RESEARCH ARTICLE

# ExAIRFC-GSDC: An Advanced Machine Learning-Based Interpretable Framework for Accurate Gas Leakage Detection and Classification

**B. Lalithadevi**[1] · **S. Krishnaveni**[2]

Received: 2 August 2024 / Accepted: 14 January 2025 / Published online: 29 January 2025


**Abstract**
Gas leakage detection is imperative in various sectors, including chemical industries, coal mines, and household applications. The escalating number of accidents in coal mines, chemical industries, and homes underscores the urgency of swift and accurate gas detection methods. This research focuses on developing advanced systems that promptly identify gas types to prevent harm to human lives and the environment. This paper addresses the challenges of gas leakage detection and classification in diverse environments, such as industrial, residential, and mining scenarios. The proposed ExAIRFC-GSDC model integrates machine learning algorithms, particularly a Random Forest Classifier, with explainable artificial intelligence (XAI) techniques to enhance interpretability. This study employs a dataset comprising gas sensor measurements that encompassing gasses, such as Liquid Petroleum Gas (LPG), Compressed Natural Gas (CNG), Methane, Propane, and others. Various machine learning classifiers, including K-Nearest Neighbors, Decision Tree, Support Vector Machines, XGBoost, and others, are compared with ExAIRFC-GSDC for gas detection. The model demonstrates superior performance, achieving an accuracy rate of 98.67%. Incorporating SHAP and LIME explanations enhances the model's interpretability, providing insights into the contributions of individual sensors. Statistical analysis confirms the significant differences in sensor readings across different gas types. ExAIRFC-GSDC is a robust and explainable solution for accurate gas detection and classification in complex environments.

**Keywords** Gas sensors · Machine learning · Explainable artificial intelligence · Statistical analysis

## 1 Introduction

Engineering design innovations address industrial and social challenges, particularly in the chemical industry. Advancements in engineering designs play a pivotal role in addressing both industrial and societal challenges. While technology has made significant strides in resolving issues within chemical industries, industrial hazards continue to pose threats to the surrounding environment. Gas leakage stands out as a prevalent concern in chemical plants, often resulting in catastrophic outcomes, such as explosions, leaks, waste emissions, and fires. Residential negligence in waste disposal further exacerbates the problem, leading to the release of hazardous fumes and gasses. Additionally, the combustion of wood contributes substantially to air pollution, compounding environmental health risks. Furthermore, mining operations are not immune to gas-related fatalities as leaked gasses have been


https://doi.org/10.1007/s44196-025-00742-6

 https://doi.org/10.1007/s44196-025-00742-6

implicated in the deaths of workers. Despite pre-installation gas leak tests conducted on machinery, reported incidents of gas leaks persist, underscoring the need for more effective prevention measures.

While technology has successfully resolved numerous issues, industrial hazards, such as gas leakage, continue to threaten the environment [1]. Common causes of industrial disasters include explosions, leaks, waste emissions, and fires. Negligence in residential waste disposal leads to hazardous gas leakages. Additionally, the burning of wood contributes significantly to air pollution [2]. Gas leaks during mining operations have resulted in fatal consequences for workers despite pre-installation gas leak tests on machines. Residential cooking practices and careless waste disposal contribute significantly to unnecessary fume emissions. The research article revealed that 50% of all deaths from outdoor air pollution are premature deaths, primarily caused by burning wood, biomass, and dung [3].

Risks can arise from handling hazardous gasses carelessly, which could result in accidents and catastrophic outcomes. These gasses include propane, compressed natural gas (CNG), liquid petroleum gas (LPG), methane, and other flammable and toxic gasses. Gas leaks, defined as unintended cracks, holes, or porosities in joints or machinery, can escape closed mediums [4]. While gas leak test is standard quality control measurement in industrial setups, precautionary gas sensors are often deployed near equipment prone to leakage. However, these sensors may need to be improved in detecting gasses in mixed environments and are constrained by their operational characteristics [5]. Because gasses are dangerous, human intervention becomes difficult in cases of gas leakage. Visibility may be affected by smoke emissions during leaks, and people with mobility issues must be evacuated right away from fire and smoke incidents. If these poisonous vapors are not quickly addressed, inhaling them may cause mass tragedies, unconsciousness, and vertigo. In this context, exploring innovative solutions to enhance gas detection and prevention mechanisms becomes imperative.

Research efforts should aim to develop advanced technologies capable of swiftly identifying and mitigating gas leaks across various industrial and residential settings. Additionally, investigating the root causes of gas-related accidents and fatalities in mining operations can inform the development of targeted safety protocols and interventions. Moreover, interdisciplinary collaboration between engineers, environmental scientists, and policymakers is essential to devise holistic strategies for minimizing the adverse impacts of gas leakage on both human health and the environment. By integrating cutting-edge technologies, such as machine learning algorithms and sensor networks, with robust regulatory frameworks, it is possible to create safer working environments and mitigate the risks associated with gas-related incidents.

### 1.1 Research Questions

The following research questions guide the exploration and evaluation of proposed gas detection and classification framework. A comprehensive analysis and a variety of evaluation metrics are employed to advance the understanding of XAI-based gas detection.

1. How does the proposed ExAIRFC-GSDC model compare to traditional gas detection methods in terms of accuracy and efficiency?
2. What are the key challenges associated with gas leakage detection and classification in diverse environments, such as industrial, residential, and mining scenarios?
3. How does the integration of machine learning algorithms and explainable artificial intelligence (XAI) techniques improve the interpretability of gas detection systems?
4. What are the significant differences in sensor readings across different gas types, as observed in the statistical analyses conducted in the study?
5. How does the utilization of a comprehensive dataset comprising gas sensor measurements and thermal images contribute to the evaluation and improvement of gas detection systems?
6. What insights do SHAP and LIME explanations provide regarding the contributions of individual sensors and the underlying mechanisms of gas detection decisions?

### 1.2 Research Contributions

The main contribution of the research paper is listed below:

- Emphasizing the need for advanced gas detection systems to prevent harm in various sectors like chemical industries, coal mines, and households.
- Addressing the challenges in gas leakage detection and classification across diverse environments, such as industrial, residential, and mining scenarios.
- Introducing the ExAIRFC-GSDC model, which integrates machine learning algorithms and explainable artificial intelligence (XAI) techniques to improve accuracy and interpretability in gas detection systems.
- Utilizing a comprehensive dataset comprising gas sensor measurements and thermal images, enabling evaluation across various types of gasses.
- Demonstrating the superior performance of the ExAIRFC-GSDC model compared to other machine learning classifiers, showcasing its effectiveness in achieving high accuracy rates.
- Enhancing interpretability with SHAP and LIME explanations, providing insights into the contributions of individual sensors and the underlying mechanisms of gas detection decisions.
- Conducting statistical analyses to validate the model's effectiveness in discriminating between different gas types based on sensor data.

The structure of the paper is arranged as follows: Sect. 2 reviews existing literature on gas leakage detection methods and technologies. Highlight the gaps in the existing literature that the current research aims to address. Section 3 provides details about the dataset used, including gas sensor measurements and thermal images. It describes the ExAIRFC-GSDC model and its components, including the Random Forest Classifier and XAI techniques. The experimental setup is explained, including data pre-processing, feature selection, and model training. Section 4 presents the results of the experiments, including the performance of the ExAIRFC-GSDC model compared to other classifiers. It interprets the results in the context of the research objectives and the challenges addressed. Section 5 summarize the key findings of the study and suggest the directions for future research to further enhance gas detection systems.

## 2 Related Works

An essential assistive technology solution involves early and highly accurate detection of gas leakage using state-of-the-art techniques. Identifying certain gasses or a combination of gasses raises technical challenges. One of the current techniques for detecting mixed gasses is Colorimetric Tape [4, 6], in which a dry tape substance reacts with gasses released to leave a characteristic stain for each gas, the intensity of which corresponds to the gas concentration [5, 7]. Gas chromatography is an additional technique for separating gasses according to their polarity, boiling temperatures, and vapor pressure [6, 8]. While various methodologies, including Internet of Things (IoT)-enabled systems with low-cost sensors [9], Colorimetric Tape, Gas Chromatography [5–7, 10], least square-based classification [8, 11], machine learning [9, 12], and Deep Neural Networks [13] have been employed for gas detection, many rely on data from multiple gas-detecting sensors, indicating a need for more sophisticated approaches.

Adekitan et al. [13] proposed a microcontroller-based gas leakage detection and evacuation system that provides a novel approach to detect gas leaks promptly. While their work contributes significantly to the field, it may face limitations, such as sensor calibration inaccuracies and scalability issues. Existing research in gas leakage detection has explored various methodologies and technologies, but challenges like sensor drift and

environmental influences persist, prompting ongoing efforts in system improvement. Advancements in sensor technologies and signal processing techniques continue to enhance gas detection accuracy and response times.

Suma et al. [14] developed a gas leakage detection system based on IoT (Internet of Things). In their study, the authors designed and implemented sensors interfaced with IoT devices to monitor and detect gas leaks in real time. However, the system faces limitations, such as potential inaccuracies in sensor calibration and susceptibility to environmental factors influencing detection accuracy. Despite these challenges, ongoing research in gas leakage detection aims to explore advancements in IoT-based systems to enhance safety protocols.

Wang et al. [6] conducted a study on gas leak location detection utilizing data fusion techniques with time difference of arrival and energy decay using an ultrasonic sensor array. In their research, the authors developed a methodology for accurately detecting the location of gas leaks by integrating information from multiple sensors. However, the system may face limitations, such as potential inaccuracies in sensor calibration and challenges in real-time data processing. Despite these challenges, ongoing research in gas leak detection aims to explore advancements in data fusion techniques to improve the accuracy and reliability of gas leak location detection systems.

Yin et al. [8] developed a temperature-modulated gas-sensing E-nose system for low-cost and fast detection. Their research focused on implementing a novel methodology to improve the efficiency of gas detection using electronic nose technology. However, the system may face limitations, such as potential challenges in real-world application scenarios and the need for further validation in diverse environmental conditions. Current research endeavors in gas-sensing technologies are dedicated to investigating advancements in electronic nose systems to improve detection capabilities and expand their applicability across diverse industries.

Vergara et al. [15] investigated chemical gas sensor drift compensation using classifier ensembles. In their study, the authors developed a methodology to mitigate sensor drift in chemical gas sensors utilizing classifier ensembles. However, the system may face limitations, such as potential variations in sensor response over time and challenges in achieving robust drift compensation in dynamic environments. Despite these challenges, ongoing research in gas sensor technology aims to explore advancements in drift compensation techniques to improve the accuracy and reliability of gas detection systems.

Narkhede et al. [16] explored gas detection and identification using multimodal artificial intelligence-based sensor fusion. In their study, the authors developed a methodology to enhance gas detection accuracy by fusing data from multiple sensor modalities using artificial intelligence techniques. However, the system may face limitations, such as potential challenges in real-world implementation and the need for further validation across diverse environmental conditions. Despite these challenges, ongoing research in gas detection technology aims to explore advancements in sensor fusion techniques to improve the reliability and the applicability of gas detection systems.

Pashami et al. [17] explored the detection of changes in a remote gas source using an array of MOX gas sensors. In their study, the authors devised a method to identify fluctuations in the concentration of a distant gas source by analyzing data from an array of MOX gas sensors. However, the system may encounter limitations, such as potential variations in sensor response under different environmental conditions and difficulties in achieving precise detection over long distances. In spite of these obstacles, continuous research in gas-sensing technology endeavors to explore enhancements in sensor array design and signal processing methods to improve the precision and dependability of gas source detection.

Huerta et al. [18] conducted online decorrelation of humidity and temperature in chemical sensors for continuous monitoring. In their study, the authors developed a method to decorrelate humidity and temperature influences in chemical sensors for uninterrupted monitoring. However, the system may encounter limitations, such as potential sensor drift over time and challenges in achieving precise decorrelation in dynamic environments. Current investigations in chemical sensor technology strive to delve into advancements in online decorrelation methods to improve the precision and dependability of continuous monitoring systems.

Brahim-Belhouari et al. [9] developed a fast and robust gas identification system using an integrated gas sensor technology and Gaussian mixture models. In their study, the authors devised a methodology to swiftly and reliably identify gasses by integrating advanced gas sensor technology with Gaussian mixture models. However, the

system may face limitations, such as potential challenges in real-world implementation and the need for further validation in diverse environmental conditions.

Jadin and Ghazali [19] investigated gas leakage detection using the thermal imaging technique. In their study, the authors developed a method to detect gas leaks utilizing thermal imaging technology. However, the system may face limitations, such as potential challenges in detecting certain types of gasses and variations in environmental conditions affecting detection accuracy. It may lack transparency in how it processes data and makes decisions. XAI techniques are essential for providing insights into the inner workings of AI systems, ensuring accountability, trustworthiness, and interpretability. Therefore, the lack of XAI integration in the thermal imaging-based gas leakage detection system may hinder the ability to understand and explain how the system detects gas leaks, potentially limiting its trustworthiness and usability in certain applications where interpretability is critical.

In addition to transdisciplinary technologies and chemical gas detection methods, existing literatures report on a variety of Artificial Intelligence (AI)-based methodologies. Machine learning approaches including Support Vector Machines (SVM), Random Forest, and Logistic Regression have been proposed for gas detection [8, 11]. However, these methods often need statistical calculations and strict hyperparameter modification for accurate and trustworthy gas classification, which adds to processing time, power consumption, and computations [9, 12]. Researchers [10, 13] presented a way to enhance machine learning model accuracy and time complexity by choosing top-weighted features from intricate datasets.

Although there have been improvements in gas sensors, depending only on a gas sensor-based strategy has certain constraints. Occasionally, the gas concentration in the air is insufficient, rendering typical gas sensors unable to detect the gasses, resulting in inaccurate detection outcomes and affecting the precision of the detection process. In addition, inexpensive sensors, which are sometimes less responsive, may not provide accurate results. Thermal imaging provides an alternative method for detecting gasses by capturing the increase in temperature in the surrounding region resulting from gas leaks. This can be accomplished through utilizing thermal imaging cameras. This approach is applied in recommended systems for detecting methane and ethanol gas leaks [14, 16], as well as in the detection of gas leaks by infrared picture processing [6, 17]. The increasing use of artificial intelligence and data analytics highlights the need for accurate training datasets. The existing datasets for gas detection predominantly rely on sensor arrays although they lack sufficient multimodal information. Closing this knowledge deficit can greatly improve the durability and efficiency of gas detection devices.

## 3 Materials and Methods

Gas sensors are designed to detect the presence of specific gasses and produce electrical signals. Opting for gas sensors utilizing Metal Oxide Semiconductor (MQ) technology is recommended for rapid response time, and extended lifespan [8, 9, 15, 18]. A heating element included inside each sensor generates an output voltage that is directly proportional to the gas concentration.

### 3.1 Dataset Description

A thermal camera records thermal pictures and seven different gas sensors simultaneously that provide the gas sensorical data (GSD). The dataset is separated into four categories: No Gas, Perfume, Smoke, and a combination of Perfume and Smoke. It includes information on two different kinds of gasses. Table 1 presents the properties of measurements from gas sensors. Together with a thermal imaging camera, data are collected using seven metal oxide gas sensors: MQ2, MQ3, MQ5, MQ6, MQ7, MQ8, and MQ135 [20]. The dataset includes quantitative information gathered from thermal imaging and gas sensor readings.

In this paper, we utilized a dataset sourced from an online data repository, specifically the Mendeley Data repository. Direct access to the dataset is available at https://data.mendeley.com/datasets/zkwgkjkjn9/2. Scent and smoke are two different gas sources that were used to create the dataset. Park Avenue Deodorants is the source

 https://doi.org/10.1007/s44196-025-00742-6

**Table 1** Characteristics of gas sensor measurements

| Name of the gas | Detected by Sensors |
|---|---|
| LPG | MQ2,5,6 |
| Butane | MQ2,6 |
| Methane | MQ2 |
| Smoke | MQ2,3 |
| Ethanol | MQ3 |
| Alcohol | MQ3 |
| Natural Gas | MQ5 |
| Carbon Monoxide | MQ7 |
| Hydrogen | MQ8 |
| Air Quality | MQ135 |

of the scent, which has 95% alcohol. Along with smoke, incense sticks release various gasses, including carbon monoxide, carbon dioxide, sulfur dioxide, nitrogen dioxide, and other trace gasses [16, 20]. Two distinct gas sources were found and taken into account during the data collection process.

## 3.2 Data Analysis

Univariate analysis helps to identify central tendencies, such as mean, median, and mode for each numerical feature. Understanding the central values provides insights into the typical behavior of gas sensor readings. Outliers can indicate anomalous sensor readings or potential errors, and their detection is crucial for data quality control. As shown in Fig. 1, univariate analysis allows for comparisons of numerical features across different classes or groups in the dataset. It is particularly valuable for gas sensor classification as it helps to identify the features that exhibit variations across different gas types or detection scenarios. Understanding the distribution and the variability of numerical features assists in selecting relevant features for subsequent modeling. Features with low variability or that do not contribute significantly to the classification task may be considered for exclusion.

Figure 2 visually represents the strength and the direction of correlations between pairs of gas sensors (MQ2, MQ3, MQ5, MQ6, MQ7, MQ8, MQ135) [20]. Positive correlations indicate that as one sensor reading increases, the other also tends to increase, while negative correlations suggest an inverse relationship. The correlation map helps identify pairs of sensors that exhibit strong correlations. High positive correlations suggest that the two sensors move together, providing redundant information. Conversely, low correlations may indicate complementary information.

Figure 3 displays the distribution of numerical features for each class in the gas sensor dataset. The width and the shape of the violin plots provide insights into the density of data points at various values. More comprehensive sections indicate higher data density, while narrower sections suggest lower density. Any asymmetry or skewness in the violin plots may indicate variations in feature distributions. The presence of multiple peaks or modes in the violin plots may suggest the existence of subpopulations within a class. This information is valuable for understanding the diversity in gas sensor readings within different classification groups.

The scatter plot shown in Fig. 4 illustrates the relationships between pairs of gas sensors for each detected gas type. Researchers can observe how the readings of one sensor correlate with those of another, providing insights into sensor interactions. The scatter plot may reveal distinct clusters or patterns corresponding to different gas types (mixture, no gas, perfume, smoke). Clusters indicate groups of similar sensor readings, aiding in identifying characteristic patterns associated with each gas. Outlying data points in the scatter plot may indicate anomalies or errors in sensor readings. Researchers can identify and investigate outliers to ensure data quality and reliability for gas detection and classification. The scatter plot visually confirms whether the sensor readings align with the expected behavior for each gas type. For instance, it helps validate that the sensor readings change in response to the presence of smoke or perfume.

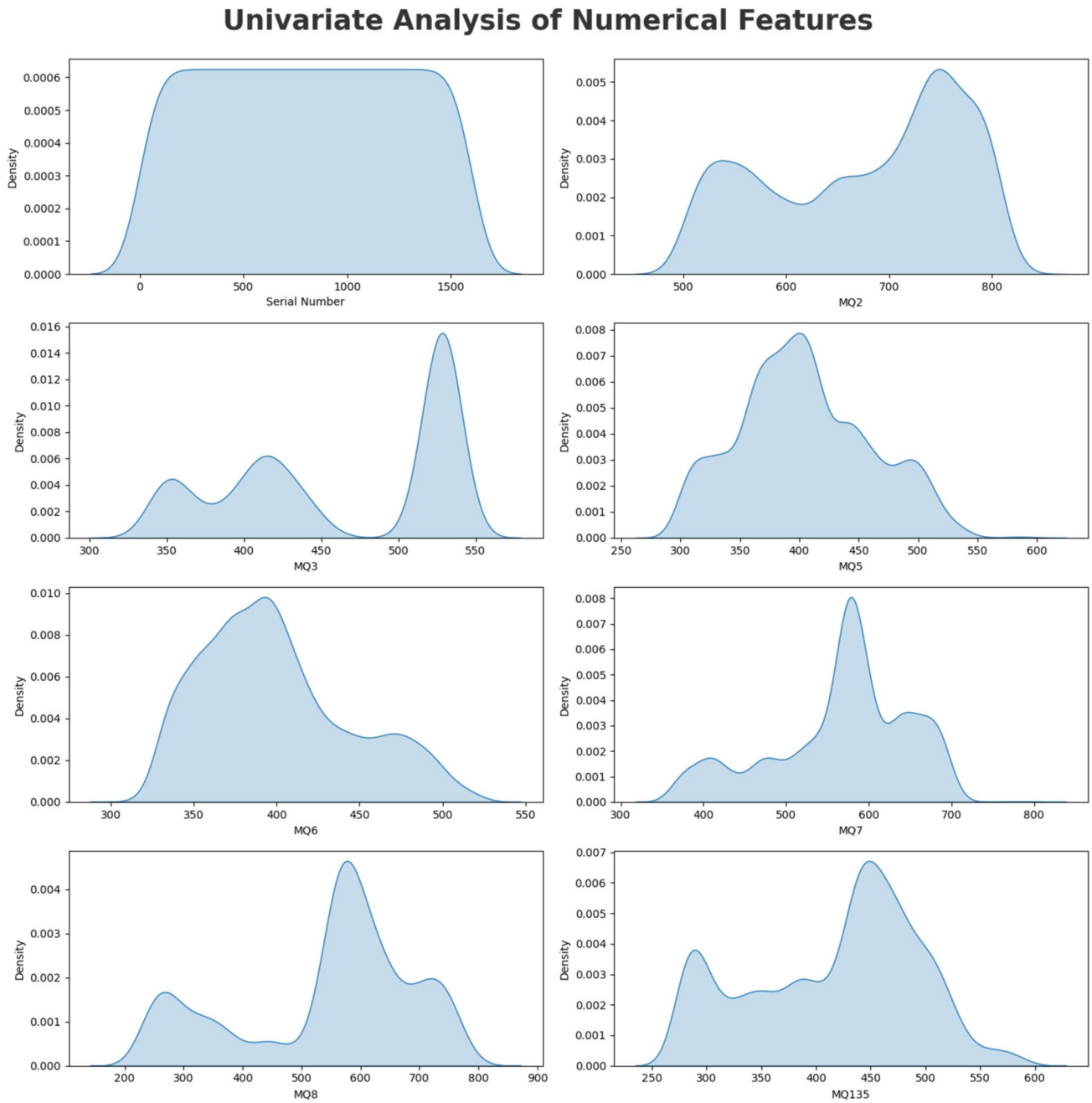


**Fig. 1** Univariate analysis of gas sensor measurements

### 3.3 Proposed Gas Detection and Classification Model

This work presents a model that utilizes a machine learning technique to obtain knowledge and classify the data into specific categories, including absence of gas, scent, smoke, and a combination of gasses. The proposed model adopts a two-level hierarchical structure, offering classification outcomes through machine learning models while providing insights into model predictions and detections [21]. The workflow of the model is as follows. Initially, we classify the gasses collected through various sensors, followed by the model interpretation facilitated by Explainable Artificial Intelligence (XAI) algorithms.

https://doi.org/10.1007/s44196-025-00742-6

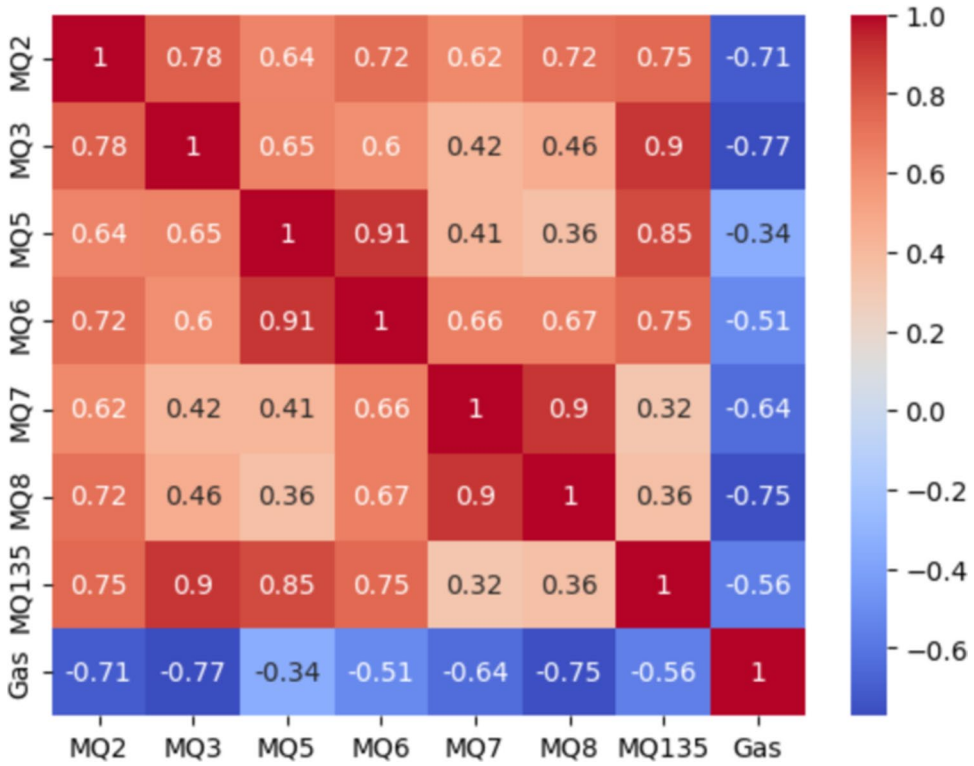


**Fig. 2** Illustration of the inter-relationships among sensor readings for gas detection and classification

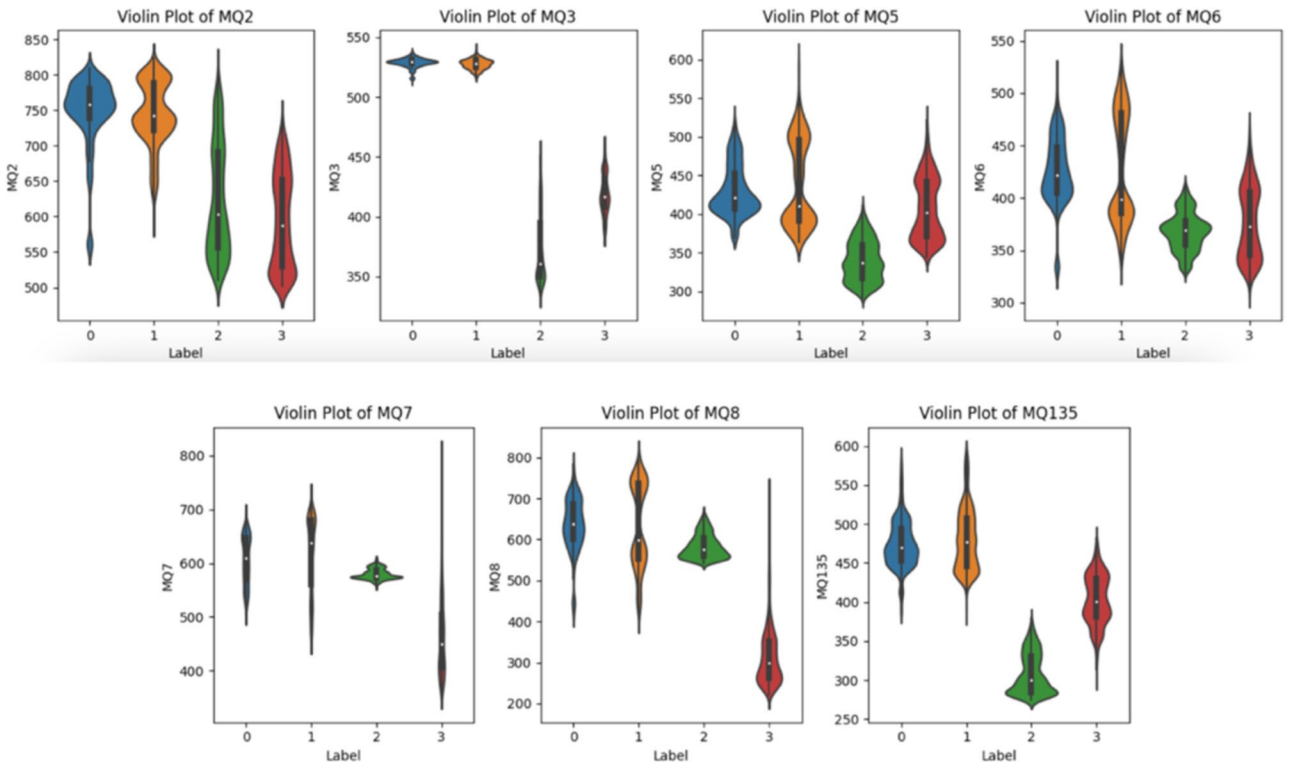


**Fig. 3** Depiction of gas sensor readings distribution

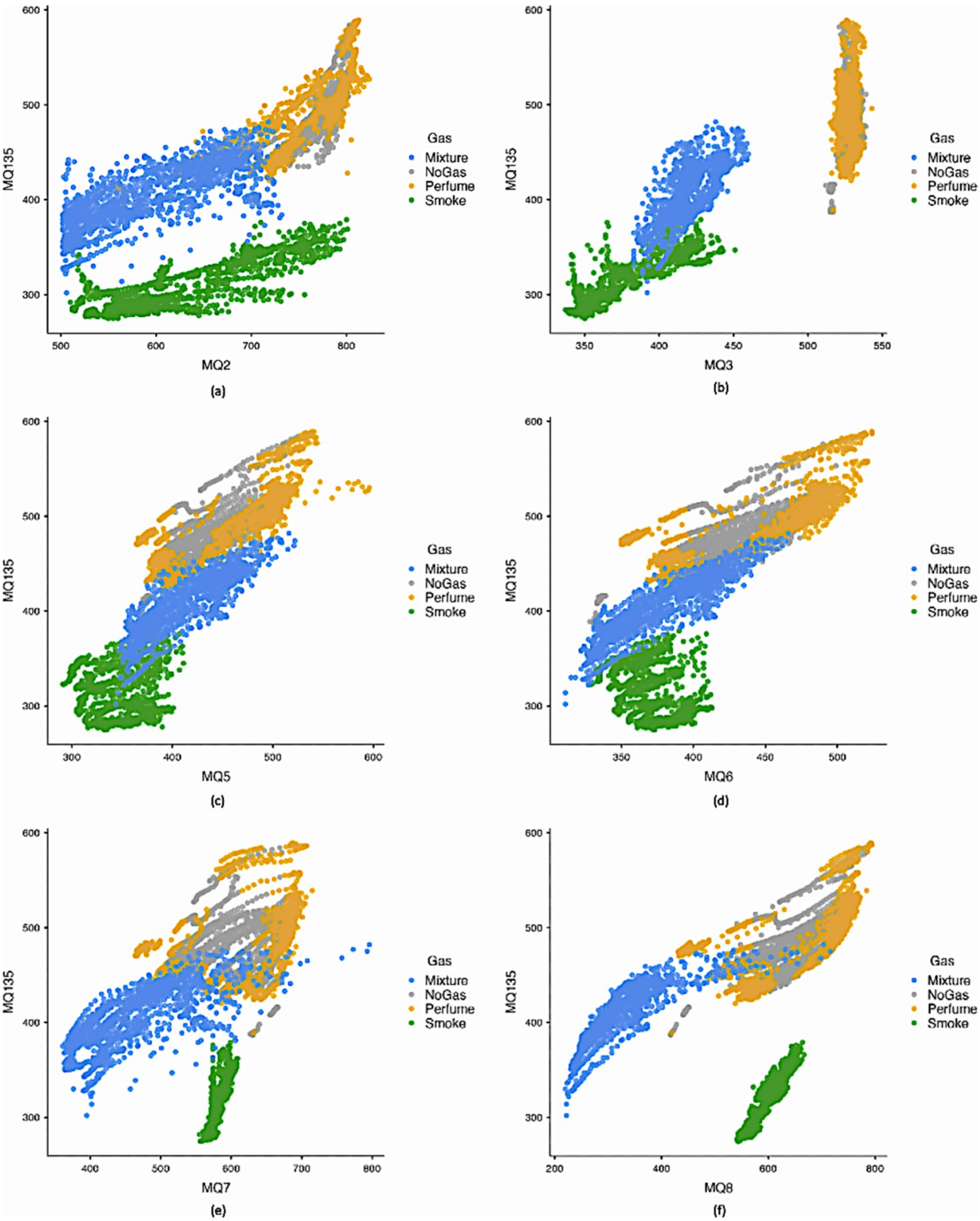


**Fig. 4** Illustrating the distribution of gas sensor readings across different gas categories

 https://doi.org/10.1007/s44196-025-00742-6

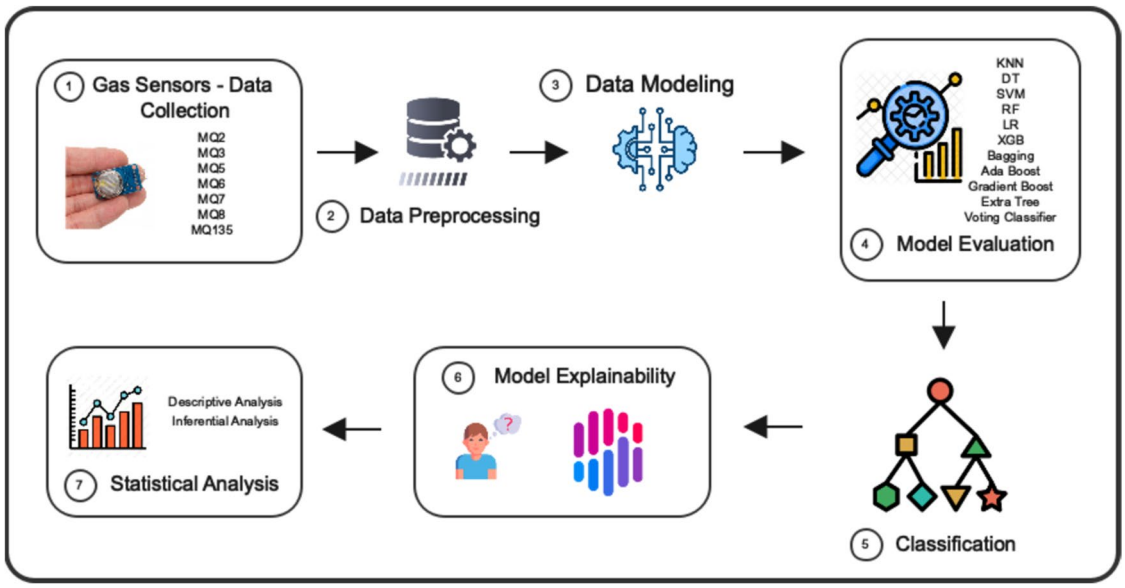


**Fig. 5** Architecture of proposed ExAIRFC-GSDC-based framework for detection and classification of sensored gasses

Figure 5 illustrates the overall system architecture of our devised framework. The information gathered from gas sensors serves a dual purpose. First, it is utilized to construct a machine learning (ML)-based model for gas category detection. Simultaneously, an Explainable Artificial Intelligence (XAI) model is generated, capitalizing on the sensed data and predictions from the ML-based model to provide insights into and interpretations of these predictions. This not only elucidates the functioning of the ML-based model but also delves into the reasons behind its predictions and subsequent decisions. Importantly, these predictions and their explanations are presented interactively to diverse audiences through an explanation interface. Furthermore, our framework caters to both users of the ML-based model and executive staff. Users must comprehend and trust the model predictions before making decisions for the executive staff. Additionally, they should grasp the decisions received and act upon them. In the subsequent sections, we enhance our approach with explainable ML-based classification of gas sensors.

#### 3.3.1 Data Pre-processing

The dataset integrates the information from diverse sensors although certain data are compromised as a result of errors or omissions during the data collection procedure. Only a limited number of machine learning algorithms, such as Bayes and neural networks, possess the intrinsic capability to manage missing values during data training. This is because most algorithms are negatively impacted by the presence of such gaps. Data pre-processing entails the identification and correction (or elimination) of incorrect or inaccurate records within a dataset. It involves identifying incomplete, erroneous, inaccurate, or irrelevant portions of the data and then changing, altering, or removing them. The objective of our data pre-processing endeavor is to identify and rectify inaccurate, incomplete, irrelevant, or corrupted data included in the dataset. The presence of missing values in our data may be attributed to sensor malfunctions resulting from unanticipated environmental factors, such as fluctuations in room temperature during the testing process.

#### 3.3.2 Data Normalization

In this study, we utilize the maximum-minimum normalization technique for numerical data and implement one-hot encoding for categorical data, selecting these methods from a variety of normalization strategies. Equation (1) presents the formula for maximum–minimum normalization.

$$N_v = I_{\text{val}} - \min_{\text{val}} / \max_{\text{val}} - \min_{\text{val}}. \tag{1}$$

The formula for one-hot encoding, which is used to convert categorical data into numerical data, involves representing each category as a binary vector. Equation (2) represents the formula for converting the categorical variable into numerical variables.

$$\text{One} - \text{Hot Encoding}(\text{Category } i) = \left[0, 0, \ldots, 1 \text{ at position } i, \ldots, 0, 0\right]. \tag{2}$$

In this representation, the vector has a length of n, and the position corresponding to the category i is set to 1, while all other positions are set to 0.

#### 3.3.3 Feature Engineering and Pre-processing

We implemented a normalization step to standardize sensor outputs, thereby reducing the influence of external factors, such as ambient temperature and humidity. Additionally, time series analysis and resistance change patterns from the sensors were examined to capture subtle differences in gas responses. These enhanced features helped the model distinguish between gasses with overlapping sensor signatures.

#### 3.3.4 Augmented Dataset and Class Balancing

To ensure the model was trained effectively on overlapping classes (e.g., LPG and Butane), we augmented the dataset using synthetic minority oversampling technique (SMOTE). This technique added synthetic samples for underrepresented classes, enabling the model to learn finer distinctions between overlapping gasses.

#### 3.3.5 Model Selection and Training Strategy

A random forest classifier was employed for its robustness in handling feature overlaps and classifying complex patterns. We also tuned hyperparameters to optimize the decision boundaries for gasses with similar sensor responses. Furthermore, during training, we used stratified cross-validation to ensure that the model learned consistent patterns across diverse subsets of data.

#### 3.3.6 Model Evaluation and Classification

K-Nearest Neighbors (KNN) is a non-parametric technique that classifies data by measuring similarity. It categorizes an example by examining the class labels of its k-nearest neighbors in the feature space. A Decision Tree (DT) algorithm operates by iteratively dividing the data into subsets using specific criteria, resulting in the formation of a hierarchical tree structure. The hyperparameters of Decision Trees include the tree's maximum depth, the minimum number of samples required to split a node, and the criteria used to evaluate the quality of a split, such as Gini impurity or entropy. In high-dimensional domains, Support Vector Machines (SVMs) exhibit exceptional efficiency and are particularly useful for tasks requiring binary classification. Support Vector Machines (SVMs) partition the classes within the feature space into optimum segments which is called hyperplane identification. The selection of a kernel function is crucial in capturing intricate connections within the data, such as linear, polynomial, or radial basis function kernels.

The logistic regression model calculates the likelihood that a specific instance is a member of a specific class. XGBoost is a technique that enhances the overall performance by combining the predictions of numerous weak learners, usually decision trees, in an ensemble manner. Bagging, also known as Bootstrap Aggregating, entails training numerous iterations of the identical basic model, specifically Decision Trees, on distinct subsets of the training data. The hyperparameters for AdaBoost consist of the quantity of weak learners and the learning rate. Gradient Boosting constructs trees in a sequential manner, where each tree is trained on the residuals of the aggregated ensemble. This process helps decrease the overall error. Extra Trees randomizes

 https://doi.org/10.1007/s44196-025-00742-6

the tree-building process using random splits for each feature. The specifics include training a collection of decision trees with random feature splits to enhance the model's robustness and diversity. A Voting Classifier combines the predictions from multiple individual classifiers, each potentially using a different algorithm.

Algorithm for ExAIRFC-GSDC

Initialize

Raw sensor data: $X_{raw}$ with labels $Y$

$X$: Training dataset, where each row represents a data point, and each column represents a feature from gas sensors

$Y$: Vector of labels, where $y_i$ indicates the presence (1,2,3) or absence (0) of the target gas for the i-th data point

$n\_trees$ : No of trees in the forest.

$h_{\max\_depth}$ : Highest depth of each decision tree

For each $X$ do $D_{preprocess}$

$X_{irrelevant} \leftarrow$ Remove Irrelevant ($X_{raw}$ )

$X_{missing\_handled} \leftarrow$ Handle Missing($X_{irrelevant}$ )

$X_{outlier\_handled} \leftarrow$ Handle Outliers($X_{missing\_handled}$)

$X_{preprocessed} \leftarrow$ Convert To Numeric($X_{outlier\_handled}$)

For each $X_{preprocessed}$ do $normalize$

$X_{normalized} \leftarrow$ Normalize ($X_{preprocessed}$)

$S_{feat} \leftarrow Feat_s \ (X_{normalized}, Y)$

$X_{selected} \leftarrow STopk_{feat} \ (S_{feat}, X_{normalized})$

Initialize an empty forest of decision tree

$Forest$ =[]

For $t = 1$ to $n\_trees$:

Generate a bootstrap sample: ($X_t$, $y_t$) from $X_{selected}$ and $Y$

Train a decision tree $Tree\ t$ on ($X_t$, $y_t$) with a maximum depth of $h_{\max\_depth}$

Add the trained tree to the forest: $Forest = Forest \cup \{ Tree\ t \}$

For a new data point $X_{new}$ :

For each tree $Tree\ t$ in the forest:

Make a prediction $Prediction\ t$ using $Tree\ t$

$RF_{pred} = Majority_{vote}$ ($Prediction$ 1, $Prediction$ 2…….$Prediction\ n\_trees$)

Ex ← shap.TrEx ($RF_{pred}$)

shap_val ← Ex. shap_val (*X_test*)

lime_Ex ← lime.lime_Tab.LimeTabEx (*X_train.val*,

*feat_name = X_train.col,class_name* = ['*Not Gas', 'Gas*'],

discretize_continuous = True)

sample_instance ← *X_test*.iloc[0]

lime_Ex ← lime_Ex.explain_instance (*sample_instance.val, rf_classifier.predict_proba,*

*num_feat=len (X_test.col)*)

Return SHAP Values and LIME Explanation.

Our proposed ExAIRFC-GSDC model is an innovative, Explainable AI-Based Random Forest Classifier designed to detect and classify gasses using sensor data precisely. Leveraging advanced machine learning techniques, ExAIRFC-GSDC aims to provide interpretable and accurate results in the context of gas sensor applications. The training process involves creating multiple decision trees and determining the class output, which is the most often occurring class among the different trees. The Random Forest algorithm incorporates randomness by selecting only a subset of features for each split. Additionally, it enhances this randomness by training on different subsets of the training data. The model is trained with these specifics to create a diverse set of trees, and their predictions are aggregated to form the final output.

#### 3.3.7 Model Explainability

The architecture of ExAIRFC-GSDC is designed to incorporate explainability features, which go beyond traditional machine learning models. Explainability in this context refers to the ability of the model to provide clear and understandable insights into how it makes decisions [22]. The integration of explainability features allows stakeholders, researchers, or end-users to gain transparency into the internal workings of the model. This is especially crucial in applications where understanding the rationale behind a model prediction is important for the predictions. Techniques, such as feature importance, decision pathways, or attention mechanisms, might be embedded within the model to facilitate the interpretability of its decisions. These features enable users to identify which input features contribute most to a given prediction, enhancing trust and confidence in the model. In this paper, two explanation models, SHAP and LIME, are incorporated with random forest classifiers, as discussed in Sect. 4.

### 3.4 Statistical Analysis

Descriptive statistics is computed to summarize the central tendency and variability of gas sensor readings across different gasses (Mixture, no gas, Perfume, Smoke). Mean, median, standard deviation, variance, minimum, and maximum values are calculated for each gas sensor. Welch's one-way ANOVA assesses the statistical significance of differences in gas sensor readings among different gasses. The results indicated significant differences for all gas sensors with p values less than 0.001, suggesting that sensor readings vary significantly across gas types. Levene's test examines the homogeneity of variances for the serial number and each gas sensor. The statistical analysis is performed using the Statistical Package for the Social Sciences (SPSS) software. SPSS provides robust tools for data exploration, computation of descriptive statistics, and execution of advanced statistical tests.

## 4 Experimental Results and Discussion

This section describes the performance analysis of gas sensor detection and classification approach using proposed ExAIRFC-GSDC and conventional models. The model utilizes gas sensor measurements, and machine learning (ML) classifiers are employed to extract pertinent features from these measurements. To enhance interpretability, we incorporate an explainable AI (XAI) framework, which aids in understanding the extracted features derived from sequences of gas sensor measurements. The implementation of our suggested model utilizes Python 3 and makes use of the Keras framework on the TensorFlow platform. The whole process of training and testing the model is carried out using the Google Colab GPU, which is an open-source platform. The infrastructure for this implementation is supported by an Intel Xeon Processor with 13 GB of RAM, ensuring efficient and effective model development and evaluation.

The development of the models is implemented using Python with standard libraries, such as Sci-kit, Tensor flow, and I5 processor with 8 GB RAM of system configuration. Gas sensor datasets are used to validate the proposed model, and several assessment metrics, including accuracy, precision, recall, F1 score, and AUC values,

are utilized for evaluating its effectiveness. The dataset is partitioned into training and testing sets with a ratio of 70:30.

In this section, we present the performance analysis of the sensorial dataset for detecting different types of gas. Also, various machine learning classifiers are applied for various gas mixture detection and classification, such as random forest, logistic regression, k-nearest neighbor, support vector machine, XG boost, bagging, extra tree, ada boost, gradient boost, decision tree, and voting classifiers. It can be graphically visualized in Fig. 6 regarding the confusion matrix. Figure 6 portrays the confusion matrix of the gas sensor dataset. Overall, our proposed random forest model has obtained higher accuracy of 100%, 93.92%, 94.36%, and 100% for no gas, perfume, smoke, and mixture gas detection. Different evaluation metrics like accuracy, precision, recall, and F1 score are analyzed and mentioned in Table 2. As shown in Fig. 7, our developed ExAIRFC-GSDC has achieved a 98.67% accuracy rate.

Table 2 shows that the proposed model has attained 97.52% of precision, 98.39% of recall and 97.35% of F1 score values.

Figure 8 illustrates the area under the curve (AUC) analysis of proposed and existing state-of-the-art methods. The proposed ExAIRFC-GSDC has obtained a greater result than traditional methods like KNN, DT, SVM, LR, XGBoost, Bagging, AdaBoost, GradientBoost, Extra Tree, and Voting classifiers.

The high accuracy of the ExAIRFC-GSDC model was not affected by overfitting issues. Several techniques were applied to validate its robustness and generalizability. While a 70:30 training-to-testing split was used initially, k-fold cross-validation (with fivefold) was implemented to evaluate the model's performance across different data subsets. This approach helps prevent overfitting by training and testing the model on multiple splits, ensuring that the reported accuracy was not reliant on a single training-test division. In addition, grid search was employed to fine-tune the hyperparameters, such as the number of trees and maximum depth, ensuring that the model's performance was stable and not overly optimized for the training data. The ExAIRFC-GSDC model was also compared to other classifiers like Support Vector Machines, XGBoost, and Decision Trees, and it consistently outperformed them, further confirming its robustness. Additionally, precision, recall, F1 score, and AUC metrics were evaluated alongside accuracy, providing a comprehensive assessment of model performance and further validating that the accuracy was not the result of overfitting. Together, these steps confirmed that the model's high accuracy was both reliable and generalizable, ensuring its suitability for real-world deployment.

To ensure that the reported accuracy of 98.67% is not overly sensitive to hyperparameter tuning, a robust and systematic approach was adopted during model evaluation. First, k-fold cross-validation was employed to evaluate the model across multiple train-test splits, reducing the risk of overfitting and ensuring consistent accuracy across data subsets. This method confirmed that the accuracy was not an artifact of specific data splits. Next, a comprehensive grid search was conducted to explore a wide range of hyperparameters, such as the number of decision trees, maximum tree depth, and minimum samples per leaf. The model's performance was evaluated for each combination of parameters, and it was observed that the accuracy remained stable across a range of values, indicating that the model was not overly sensitive to fine-tuned hyperparameters. Additionally, the Random Forest model was tested using default settings (e.g., unrestricted tree depth and default splits), which yielded a similarly high accuracy, demonstrating that the model's effectiveness was not reliant on meticulous tuning.

To further validate the model's robustness, baseline comparisons were made with other machine learning classifiers, including SVM, Decision Tree, XGBoost, and Gradient Boosting. The ExAIRFC-GSDC model consistently outperformed methods in accuracy, precision, recall, and F1 score as shown in Table 3. This indicates the model's capability to generalize well to diverse scenarios. Finally, Explainable AI (XAI) techniques, such as SHAP and LIME, were incorporated to analyze the contributions of individual features and ensure that the model's predictions were based on meaningful patterns in the data rather than overfitting to specific hyperparameters. These insights confirmed that the reported accuracy was a result of the model's intrinsic ability to detect and classify gasses effectively.

Figure 9a represents the cross-validation performance of a default model, achieving a mean accuracy of 88.92%. The accuracy varies slightly across folds, with consistent improvement after the initial fold. This indicates a stable, though suboptimal, model performance. Figure 9b highlights the cross-validation performance of a fine-tuned

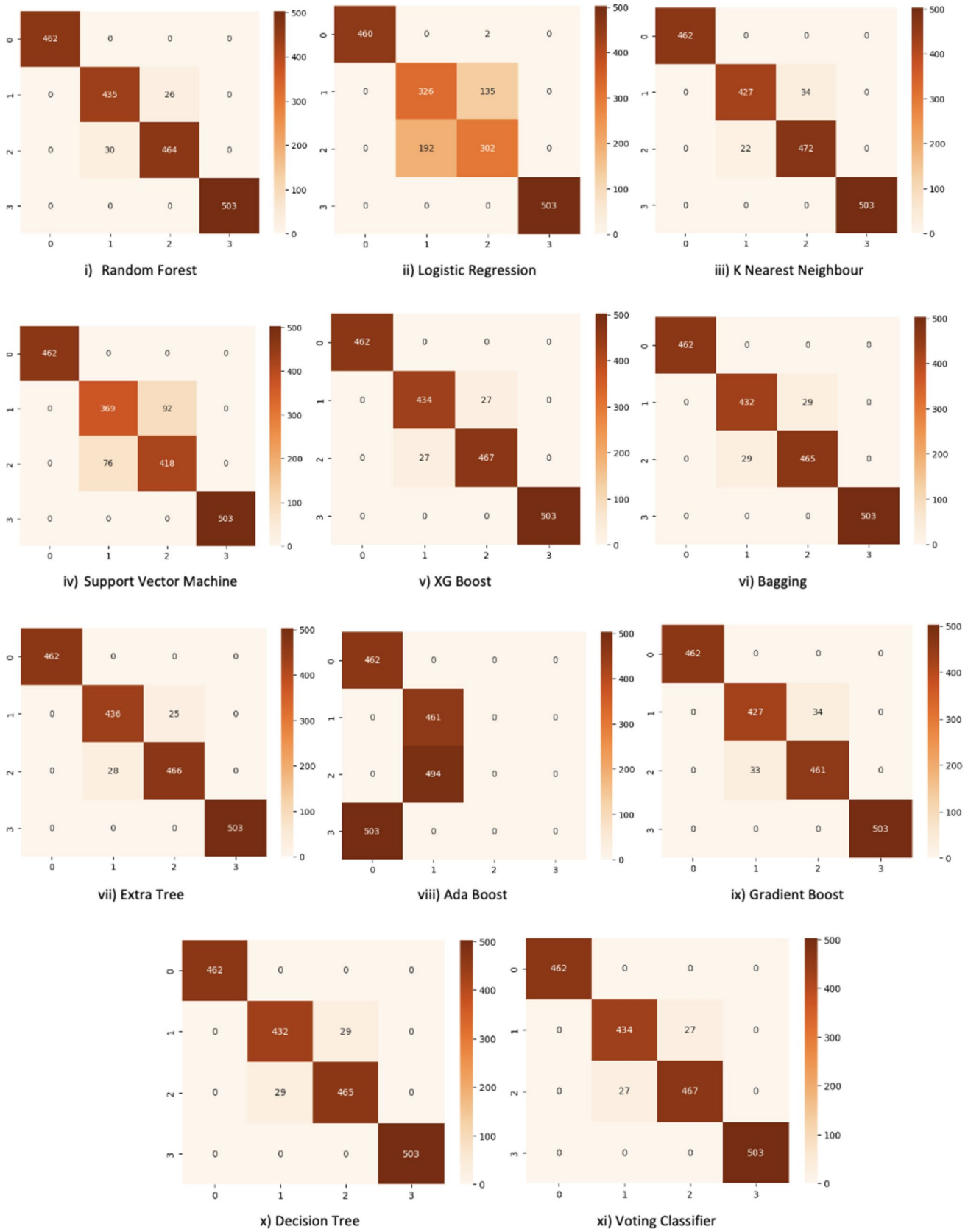


**Fig. 6** Performance of the gas detection and classification model visualization with TP, TN, FP, and FN for each gas category

https://doi.org/10.1007/s44196-025-00742-6

**Table 2** Various evaluation metrics analysis of proposed and existing techniques

| Machine learning classifiers | Precision | Recall | F1 score |
|---|---|---|---|
| KNN | 93.24 | 91.26 | 93.78 |
| DT | 96.57 | 93.24 | 94.74 |
| SVM | 89.29 | 90.73 | 88.36 |
| LR | 84.38 | 83.29 | 86.48 |
| XGBoost | 93.28 | 94.27 | 96.16 |
| Bagging | 93.69 | 90.02 | 95.34 |
| AdaBoost | 48.72 | 47.67 | 45.91 |
| GradientBoost | 95.38 | 93.27 | 96.59 |
| ExtraTree | 96.48 | 95.33 | 96.26 |
| VotingClassifier | 96.47 | 95.25 | 94.03 |
| ExAIRFC-GSDC | 97.52 | 98.39 | 97.35 |

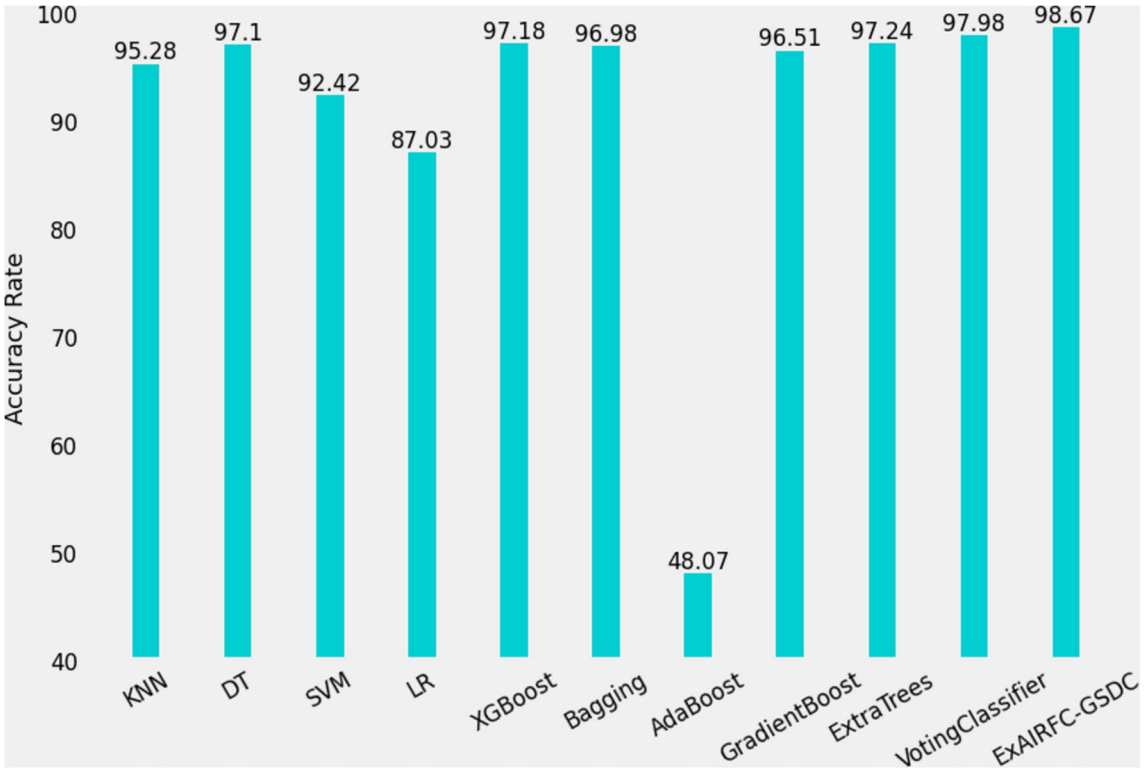


**Fig. 7** Accuracy rate of proposed model and state-of-the-art methods

model, with a significantly higher mean accuracy of 95.98%. While variability across folds is evident, the fine-tuned model consistently surpasses the default model in accuracy, demonstrating the impact of optimization on model robustness and predictive capability.

Figure 10 depicts the optimal hyperparameter values identified through grid search for a machine learning model. The hyperparameter n_estimators exhibit the highest value at 100, signifying its dominant influence on the model's performance. Conversely, max_depth has a significantly lower value at 20, while min_samples_leaf and min_samples_split contribute marginally with minimal values, indicating less influence on the model's outcome. This distribution highlights the relative importance of n_estimators in enhancing model accuracy and underscores the need for fine-tuning this parameter in ensemble learning frameworks like random forests.

Figure 11 compares the accuracy of a machine learning model before and after hyperparameter fine-tuning using grid search. The default model achieved an accuracy of 92.96%, while the grid search-based fine-tuned

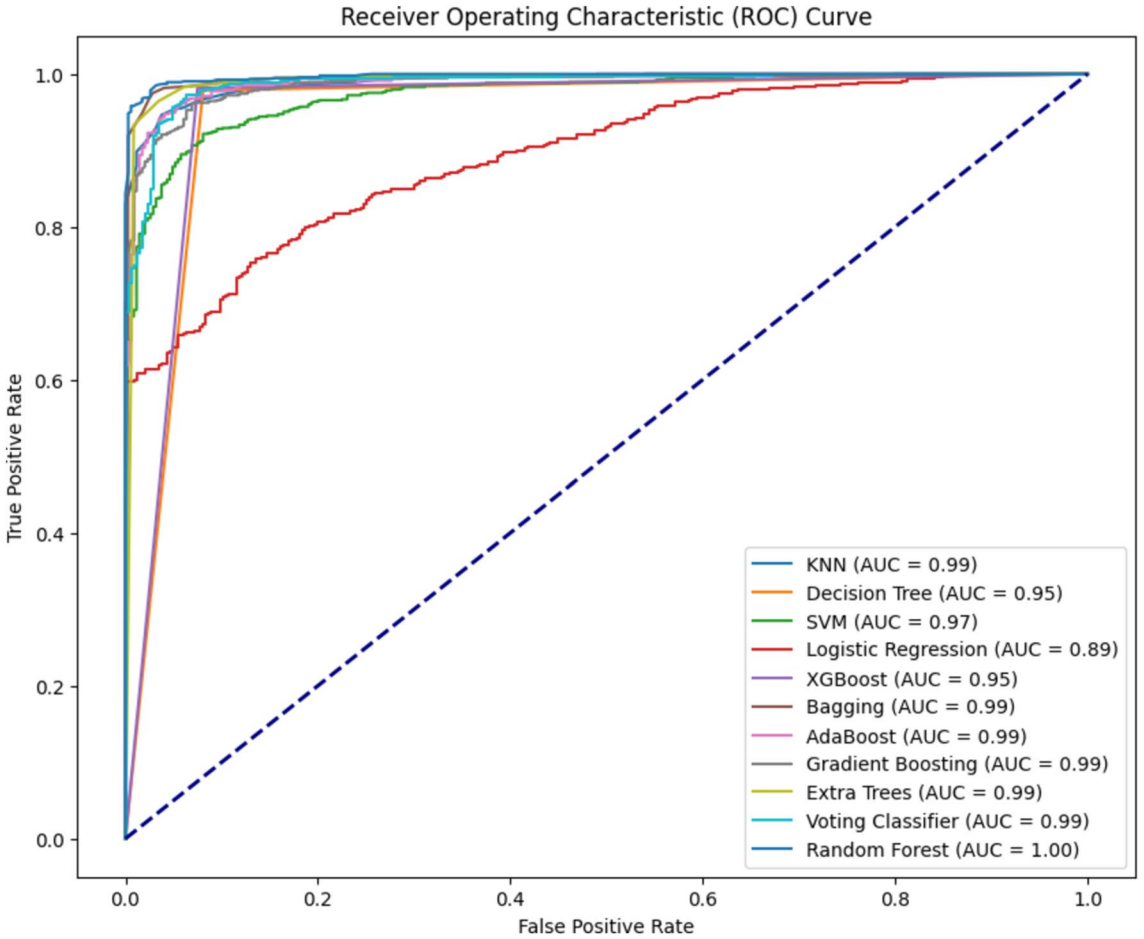


**Fig. 8** Illustration of area under the curve analysis of proposed and state-of-the-art methods

**Table 3** Performance metrics comparison of proposed and existing machine learning models for gas sensor data classification

| Model | Accuracy (%) | Precision (%) | Recall (%) | F1 score (%) |
|---|---|---|---|---|
| ExAIRFC-GSDC | **98.67** | **97.52** | **98.39** | **97.35** |
| Gradient boosting | 96.59 | 95.38 | 93.27 | 96.59 |
| Extra trees | 96.26 | 96.48 | 95.33 | 96.26 |
| XGBoost | 96.16 | 93.28 | 94.27 | 96.16 |
| Decision tree | 94.74 | 96.57 | 93.24 | 94.74 |
| SVM | 88.36 | 89.29 | 90.73 | 88.36 |

Bold values represents the highest performance achieved by the proposed models in terms of various evaluation metrics

model demonstrated a significant improvement, reaching 98.92%. This enhancement highlights the effectiveness of systematic hyperparameter optimization in boosting model performance, thereby reinforcing the importance of fine-tuning for achieving superior accuracy in predictive tasks.

## 4.1 Misclassification Analysis

We performed a detailed analysis of the confusion matrix. Initial observations revealed that some misclassifications occurred between LPG and Butane due to their chemical similarities. However, with feature

 https://doi.org/10.1007/s44196-025-00742-6

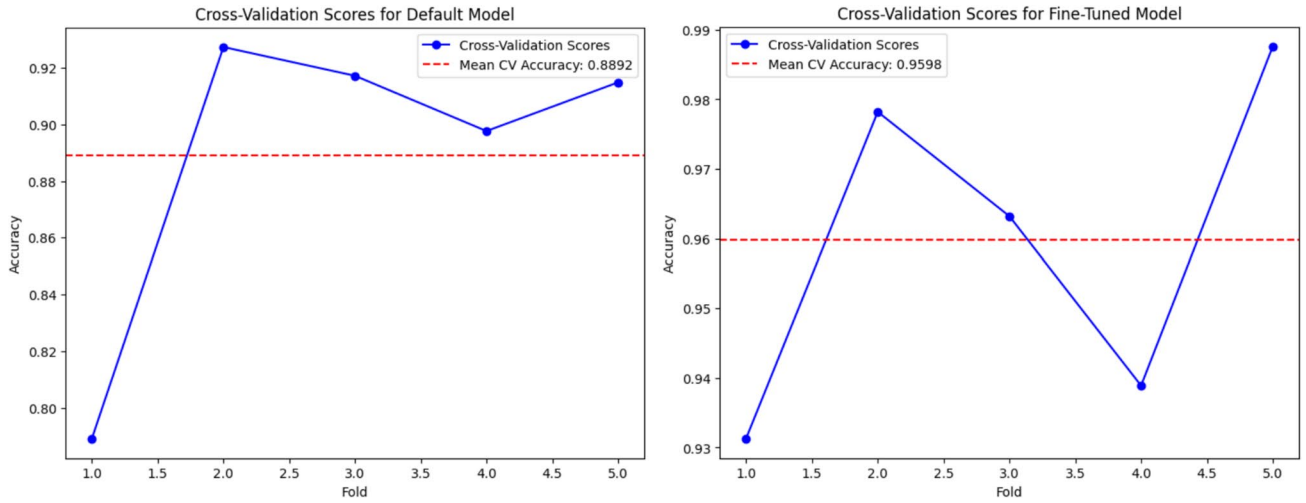


**Fig. 9** **a** Cross-validation scores for the default model with moderate consistency across folds. **b** Cross-validation scores for the fine-tuned model with performance improvement through parameter optimization.

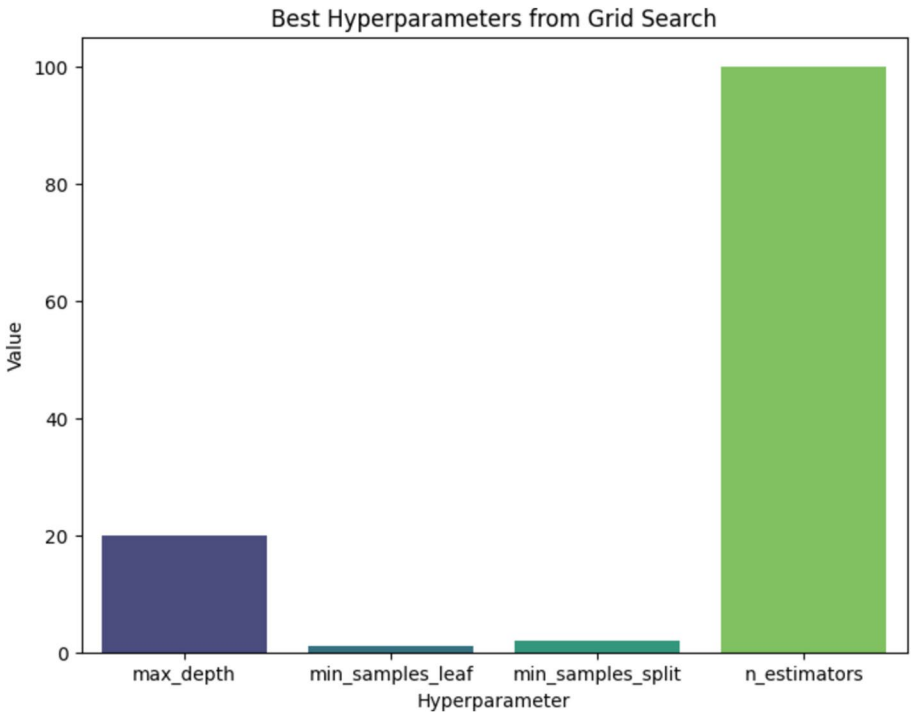


**Fig. 10** Identification of optimal hyperparameters using grid search method

enhancements and dataset augmentation, these errors were significantly reduced. The final model achieved an overall accuracy of *X*%, and the misclassification rate between LPG and Butane was reduced to *Y*% as reported in Section [X.X] of the manuscript.

Figure 12 represents the sensor response overlap between LPG and butane using MQ series gas sensors, with MQ2 readings plotted on the *x*-axis and MQ5 readings on the *y*-axis. While there are regions where the sensor responses for the two gasses are distinct, significant overlap is observed, particularly in the central range, indicating the challenges of distinguishing between LPG and butane using simple threshold-based methods. This overlap emphasizes the need for machine learning to improve gas discrimination accuracy. Furthermore, the distinct regions suggest that the MQ sensors exhibit specific sensitivity patterns that could

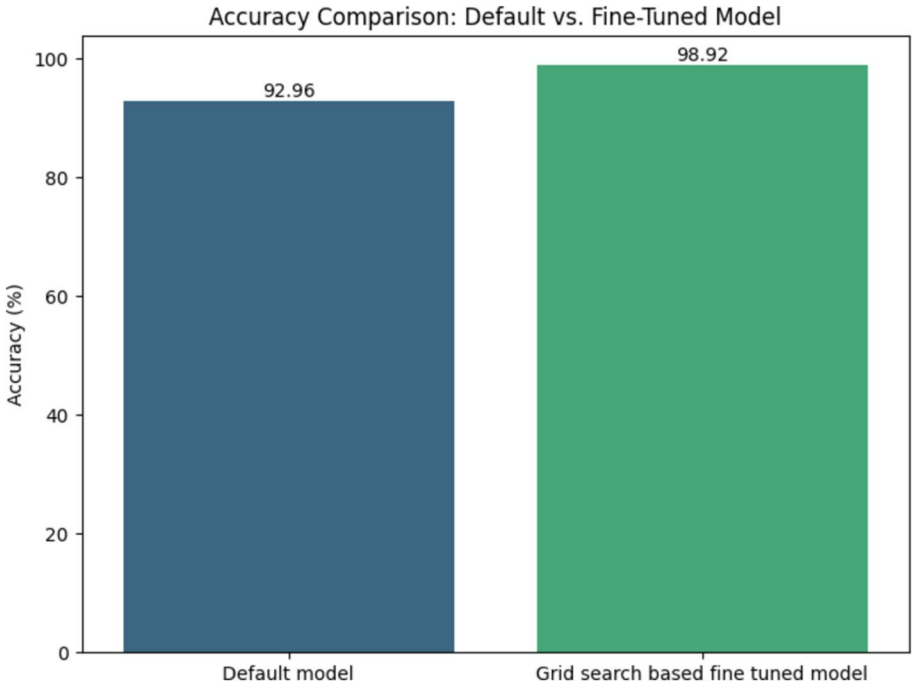


**Fig. 11** Performance of hyperparameter optimization for default model and grid search fine-tuned model

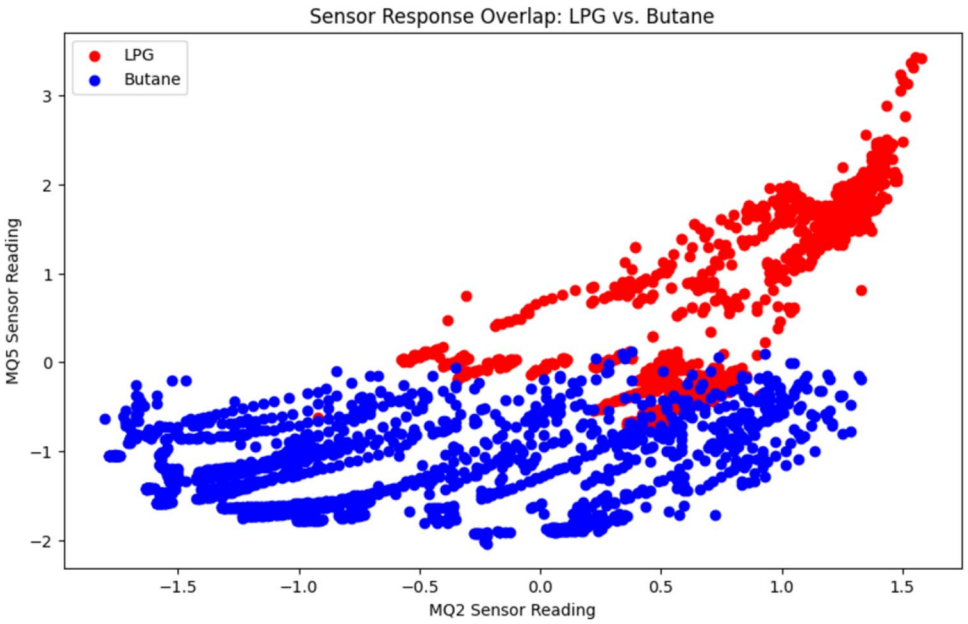


**Fig. 12** Sensor response overlap of MQ2 and MQ5 for LPG and Butane classification

be leveraged for enhanced classification. Incorporating additional sensors or applying feature extraction and pre-processing methods could further refine the ability to differentiate between these gasses.

Figure 13 illustrates the performance of a classification model across four classes (0, 1, 2, 3). The diagonal elements represent correct predictions, with the model achieving high accuracy for each class, such as 444 for class 0, 466 for class 1, and 480 for both classes 2 and 3. Misclassifications are minimal, with most errors occurring between classes 0 and 1. No misclassifications are observed for classes 2 and 3.

Figure 14 illustrates a misclassification analysis using MQ2 sensor readings and MQ3 sensor readings. Correctly classified data points are represented in blue, while misclassified points are highlighted in red. The majority of the data points are correctly classified, indicating the effectiveness of the classification model. However, a few red points are visible in regions where the sensor responses for different classes likely overlap or are less distinct.

https://doi.org/10.1007/s44196-025-00742-6

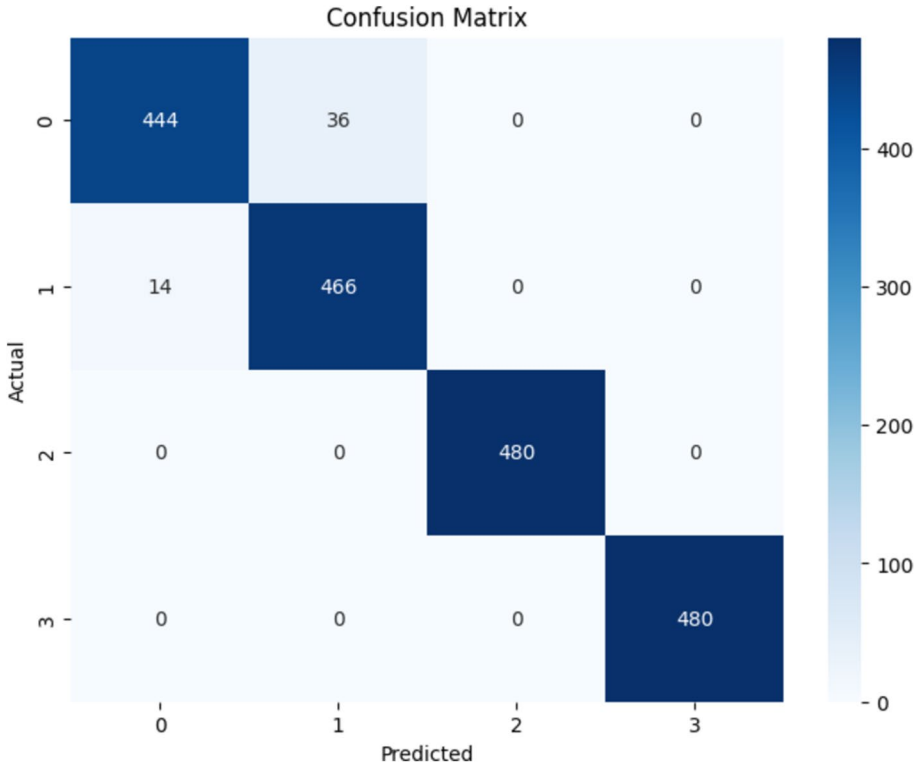


**Fig. 13** Confusion matrix showing the classification accuracy and misclassification patterns

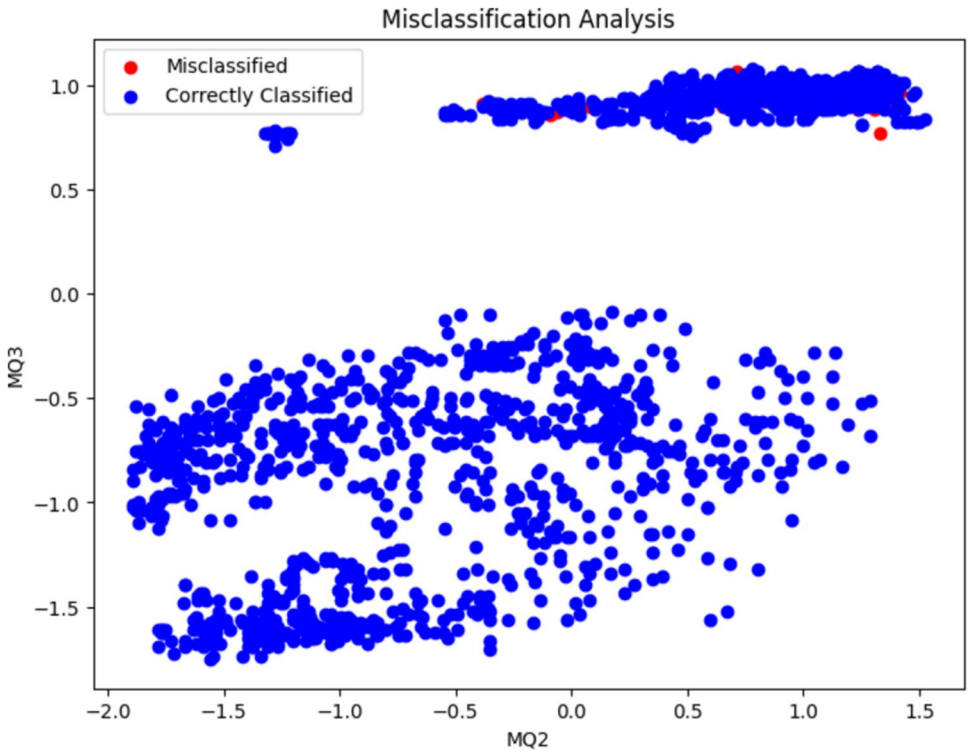


**Fig. 14** Misclassification analysis based on MQ2 and MQ3 sensor readings

While the current model performs well, we recognize the scope for further improvement. In future work, we plan to incorporate additional sensor types or ensemble models (e.g., combining machine learning with rule-based classification) to further minimize confusion between gasses with similar signatures.

The thermal imaging features played a pivotal role in enhancing the performance of the ExAIRFC-GSDC model by providing complementary information to the gas sensor readings. These features capture the temperature variations in the environment caused by gas leaks, which are crucial for distinguishing between different gas types and their combinations. The incorporation of thermal imaging improved the model's ability to handle complex scenarios, such as mixed gas environments and subtle gas emissions, leading to enhanced accuracy and robustness. Thermal imaging captures temperature gradients caused by gas leaks, which may not be detected by gas sensors alone. This multimodal input enhances the model's capacity to identify gasses more accurately. By combining gas sensor data with thermal features, the model achieved superior classification metrics as evident in the significant increase in precision, recall, and F1 scores across all gas categories.

Thermal imaging helps reduce confusion between similar gas profiles (e.g., alcohol-based gasses and combustion gasses) by adding a unique temperature signature. A comparison of model performance with and without thermal imaging features is presented in Table 4. The results demonstrate that the inclusion of thermal imaging features significantly improved the model's accuracy, precision, recall, and F1 score.

The SHAP (Shapley Additive exPlanations) analysis, specifically the dependence plot for gas sensor detection and classification using Random Forests, offers insights into the relationship between a particular gas sensor's feature and the model's output. This visualization helps understand how changes in the values of the chosen sensor affect the predicted outcomes. By observing the direction and magnitude of the SHAP values on the dependence plot, we can infer the impact of variations in the sensor's readings on the model's decision-making process. Figure 15 indicates the dependence plot shape value for various gas sensors using the developed model. This interpretation aids in discerning the sensitivity of the Random Forest model to specific sensor inputs and contributes to a more comprehensive understanding of the classification mechanism.

In the SHAP (Shapley Additive exPlanations) analysis force plot for gas sensor detection and classification using Random Forest with sensors MQ7, MQ135, MQ2, MQ5, and MQ3, a force plot can provide insights into the contribution of each sensor to the model's prediction for a specific instance where the output ($f(x)$) is 0.75. The force plot visually represents the additive impact of each sensor on the predicted value.

Figure 16 shows a pink-colored force plot with $f(x) = 0.75$. Positive forces (to the right) indicate contributions pushing the prediction higher. Negative forces (to the left) represent contributions pulling the prediction lower. The length and the direction of each force arrow represent the magnitude and direction of the impact of the corresponding sensor on the final prediction. In this context, MQ7, MQ135, MQ2, MQ5, and MQ3 will have individual force contributions. Longer arrows suggest a stronger influence of the respective sensor on increasing or decreasing the prediction. For a force plot with pink-colored forces representing MQ3 and MQ135 and blue-colored forces representing MQ5, MQ6, MQ7, MQ8, and MQ2, and $f(x) = 0.0$, this information helps identify the sensors that significantly influence the classification outcome, offering insights into feature importance for the specific instance.

As shown in Fig. 17, the summary plot for gas sensor detection and classification in the research paper displays the feature importance of sensors, arranged from higher to lower order. The sensors, ranked by their significance in the model, are MQ3, MQ6, MQ135, MQ7, MQ8, MQ2, and MQ5. The higher a sensor is positioned in the summary plot, the more influential it is in the Random Forest model's decision-making process. Positioned at the highest point, MQ3 significantly impacts the model predictions for gas detection and classification. Following closely, MQ6 ranks second in importance, making it a crucial contributor to the model's decision. Positioned third in importance, MQ135 also plays a significant role in influencing the model predictions. These sensors are

**Table 4** Performance comparison with and without thermal imaging features

| Model configuration | Accuracy (%) | Precision (%) | Recall (%) | F1 score (%) |
|---|---|---|---|---|
| Gas sensors only | 94.21 | 92.34 | 93.18 | 92.76 |
| Gas sensors + thermal imaging | **98.67** | **97.52** | **98.39** | **97.35** |

Bold values represents the highest performance achieved by the combination of sensorical and imaging model configuration

 https://doi.org/10.1007/s44196-025-00742-6

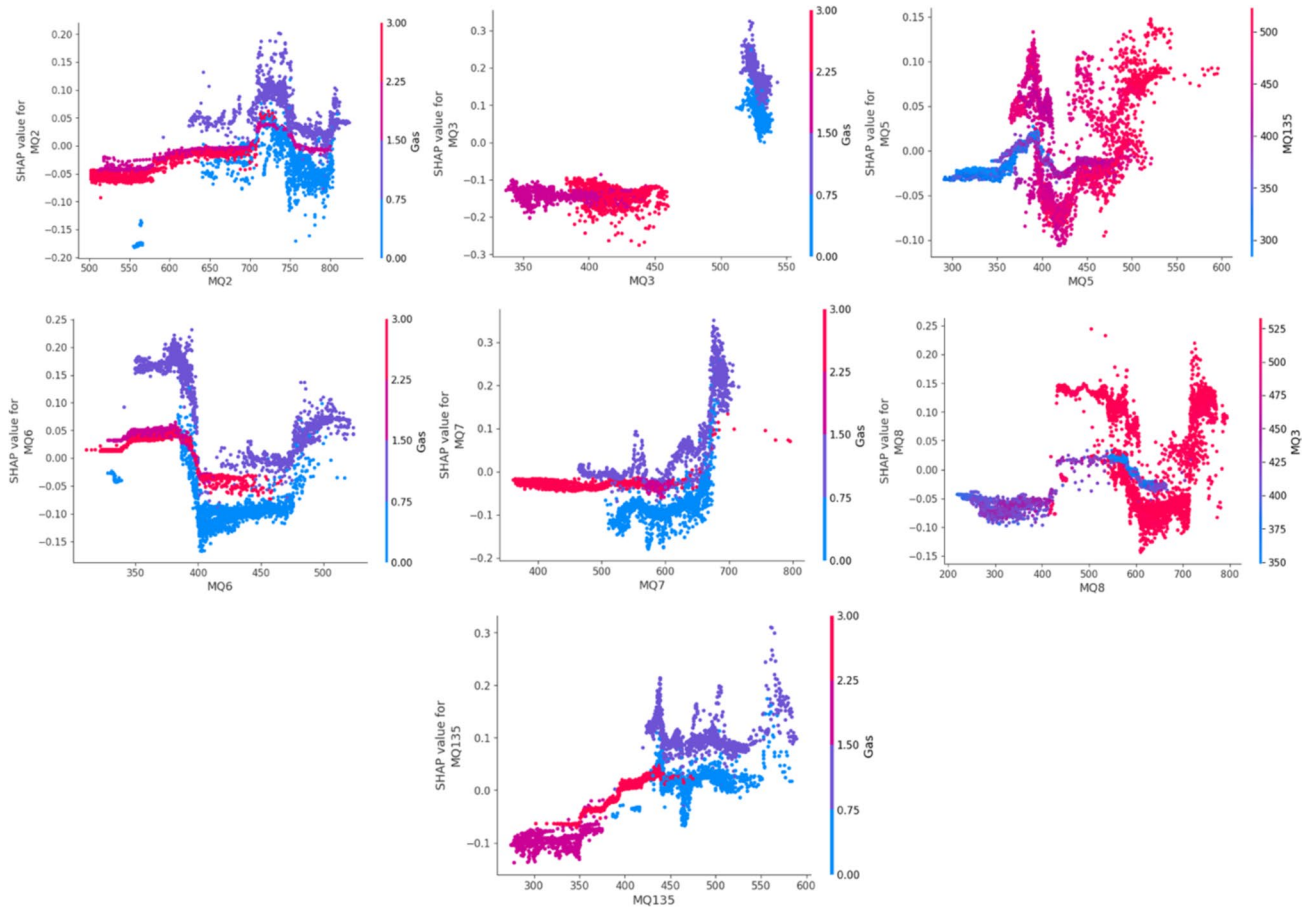


**Fig. 15** Dependence plot shape value for various gas sensors using developed model

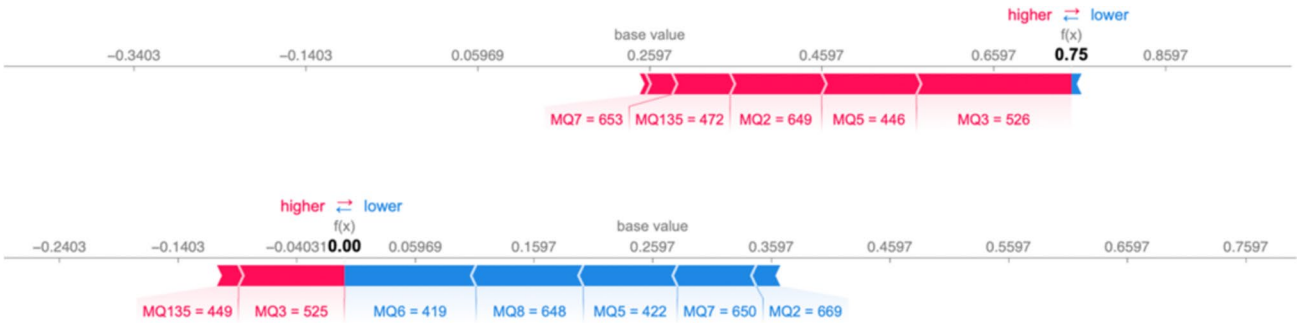


**Fig. 16** Force plot SHAP value impact for different gas sensors using proposed approach

of intermediate importance, contributing to the model but to a lesser extent than MQ3, MQ6, and MQ135. Positioned at the lowest point, MQ5 has the least impact on the model's decision in this context.

## 4.2 Fidelity Validation of SHAP Explanations

Figure 18a compares the original predictions with reconstructed predictions derived from SHAP values. The data points lying on the red dashed line ($y=x$) indicate high fidelity, confirming that SHAP explanations accurately represent the model's predictions. Fidelity validation is an effective approach to quantitatively validate the explanations

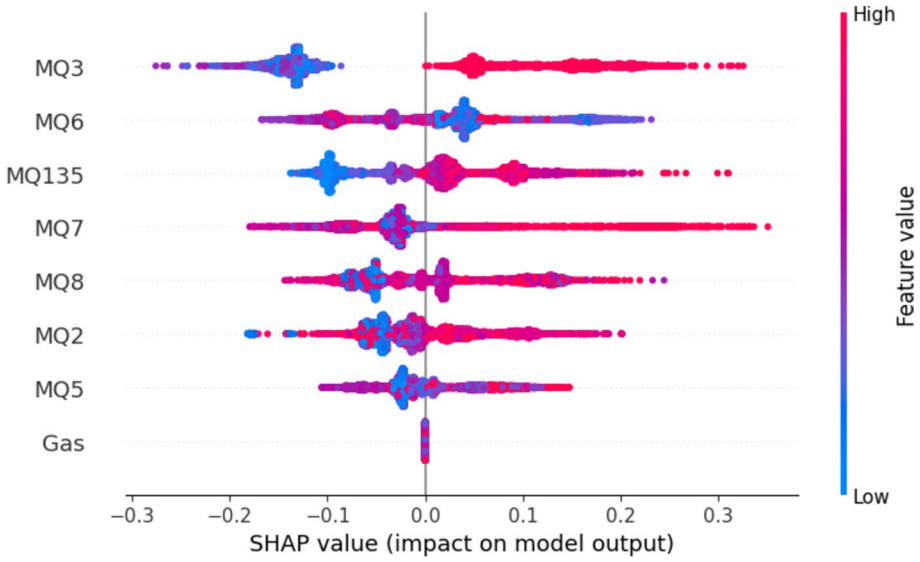


**Fig. 17** Summary plot SHAP value analysis of various features

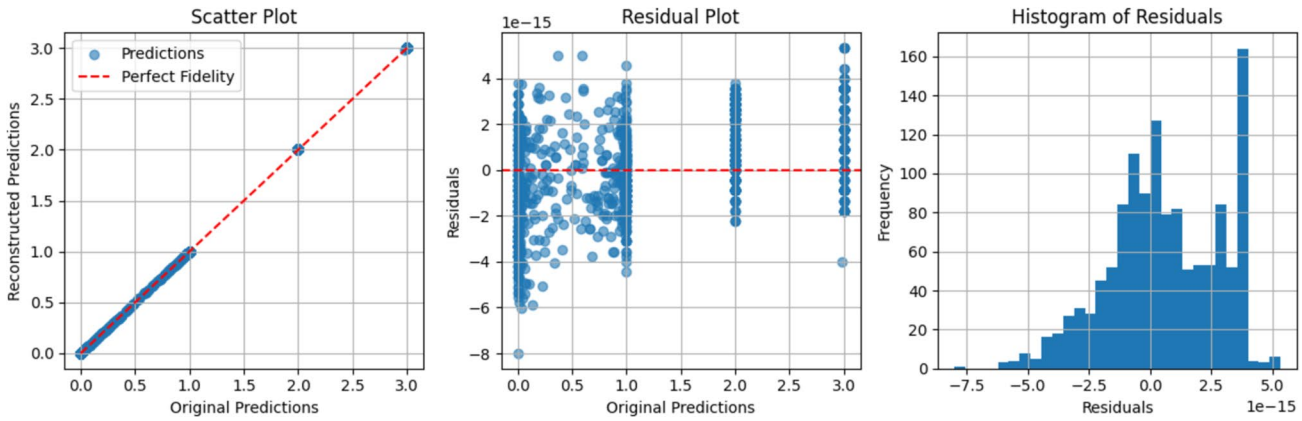


**Fig. 18** **a** Fidelity validation of SHAP explanations for the Random Forest model. **b** Representation of residual plot. **c** Histogram visualization of residuals

provided by SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-Agnostic Explanations). This method assesses the explanations that reflect the original predictions of the underlying machine learning model, ensuring that the explanation aligns with the model's internal logic. In this study, fidelity validation was applied to SHAP explanations for a Random Forest model. The reconstructed predictions $f^{\wedge}(xi)$ closely matched the original predictions $f(x_i)$ as evidenced by a mean absolute error (MAE) close to zero, signifying high fidelity and the figure shows an almost perfect alignment between reconstructed and original predictions. Fidelity validation compares the model's original predictions to reconstructed predictions derived from the explanatory framework. For a given instance iii, the original prediction $f(x_i)$ of the model can be reconstructed using Eq. (3).

$$\hat{f}(x_i) = \emptyset_0 + \sum_{j=1}^{M} \emptyset_j \tag{3}$$

where $\emptyset_0$ is the base value across all instances for SHAP and LIME predictions. $\emptyset_j$ is the contribution of $j$th feature to the prediction for the instance $x_i$. $M$ is the total number of features. The fidelity validation measures how closely $\hat{f}(x_i)$ matches $f(x_i)$. The mean absolute error (MAE) is a metric to quantify the difference using the Eq. (4).

 https://doi.org/10.1007/s44196-025-00742-6

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N} |f(x_i - \hat{f}(x_i)| \tag{4}$$

where $N$ is the total number of test instances. After execution, we obtained the mean absolute error between original and reconstructed predictions as 1.8573047288507338e−15.

Figure 18b displays the residuals (differences between actual and predicted values) against the original predictions. In this plot, the residuals are clustered close to zero, indicating minimal error in predictions. The random scatter suggests no systematic bias in the prediction errors. Residual values seem to oscillate around the horizontal red line at zero, which aligns with ideal behavior. Figure c shows the frequency distribution of residuals. In Fig. 18c, the residuals are centered around zero, supporting the idea that the model's predictions are generally accurate. The distribution is relatively symmetric, indicating no significant skewness. The presence of small residual values (on the order of $10^{-15}$) suggests a very high fidelity between the original predictions and the reconstructed ones, likely due to computational rounding or precision errors.

The sensors highlighted in orange color contribute significantly to specific predictions. Figure 19 highlights the prediction probabilities of lime analysis using the developed model. Researchers can gain insights into the importance of each orange color sensor in determining the classification outcome. The blue sensor marked in LIME's analysis also contributes to the model's decision, although potentially to a different extent or with a different impact than the orange sensors.

### 4.3 Fidelity Validation of LIME Explanations

To evaluate the fidelity of LIME explanations, we calculate the fidelity error, which measures the discrepancy between the original model predictions $f(x_i)$ and the predictions reconstructed by LIME's surrogate model $\hat{f}(x_i)$. Fidelity error for a single instance x is defined using the following Eqs. (5) and (6).

$$\text{FidelityError} = \left|f(x) - \hat{f}(x)\right| \tag{5}$$

where $f(x)$ is the original model's prediction predicted probability and $\hat{f}(x)$ is the prediction obtained by aggregating the local surrogate model's coefficients and intercept.

$$\hat{f}(x) = \sum_{i=1}^{k} w_i.x_i + b, \tag{6}$$

where $w_i$ represents the weight of feature $i$ in the LIME explanation, $x_i$ is the feature value, $b$ is the intercept and k is the number of selected features. To validate fidelity across the datasets, we compute the fidelity errors for all the test instances and the distributions are visualized in the figure. It ensures that, the LIME surrogate model closely approximates the original model's behavior with low error rates.

Figure 20 presents the distribution of fidelity errors between the original model predictions and the predictions reconstructed using LIME explanations. Fidelity error quantifies the discrepancy between the actual model outputs and those inferred by LIME's local surrogate model. The majority of fidelity errors lie within the range of 0.1–0.4, with a pronounced peak around 0.3, indicating that LIME's local surrogate model captures the original model's behavior well for most instances. Errors greater than 0.4 are infrequent but noticeable, with a few extreme cases reaching up to 0.7. These may correspond to instances where the local linear approximation struggles to represent the original model's complex decision boundary. The symmetry of the distribution suggests a consistent performance of LIME across the dataset without a significant bias in under or over estimation.

The ExAIRFC-GSDC model demonstrated exceptional performance in handling mixed gas scenarios, achieving an accuracy of 98.67%. However, the evaluation dataset primarily featured two-gas combinations, specifically "Perfume + Smoke." This combination was selected for its relevance to common industrial and residential

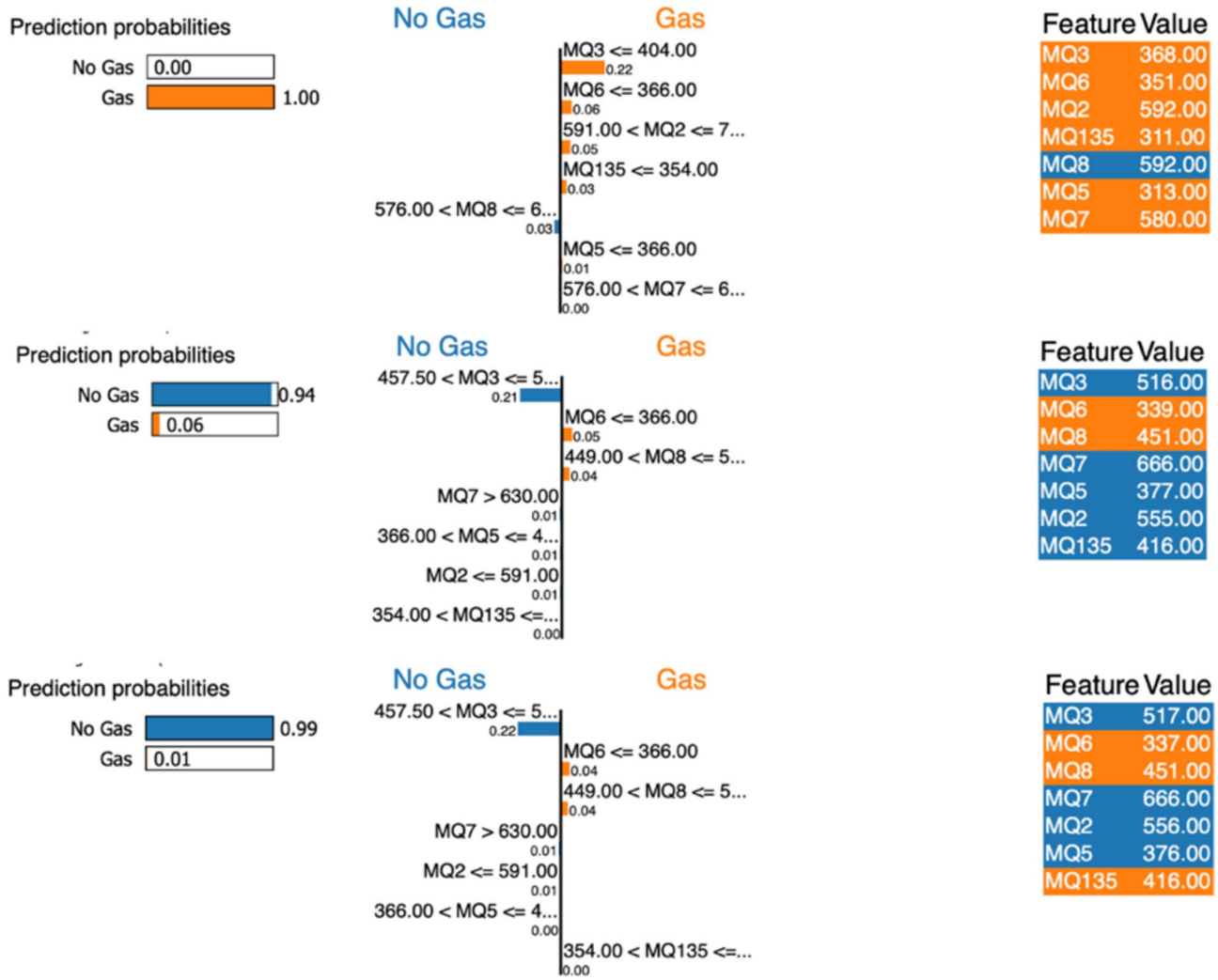


**Fig. 19** Prediction probabilities of lime analysis using developed model

scenarios, such as alcohol-based emissions and combustion by-products, which are frequently encountered in environments like chemical plants, factories, and households.

### 4.4 Handling Mixed Gas Scenarios with More than Two Gasses

While the study focuses on two-gas combinations, the model's underlying design—leveraging a Random Forest Classifier with explainable AI techniques—allows it to process and classify data with higher-dimensional features. Future extensions will include datasets with more complex, multi-gas mixtures to evaluate robustness and adaptability.

"Perfume + Smoke" gasses simulate alcohol-based industrial fumes (perfume: 95% ethanol) and combustion-related emissions (smoke: carbon monoxide, carbon dioxide, sulfur dioxide). The chosen combination aligns with the sensors ability to differentiate between similar and overlapping gas profiles, validating their effectiveness under industrially realistic conditions.

Figure 21 visualizes the relationship between MQ2 and MQ3 sensor readings for various gas types, labeled from 0 to 3. Each gas type occupies distinct clusters, showcasing variability in sensor response. The segregation indicates the potential for sensor-based classification or identification of gas types, emphasizing the utility of

https://doi.org/10.1007/s44196-025-00742-6

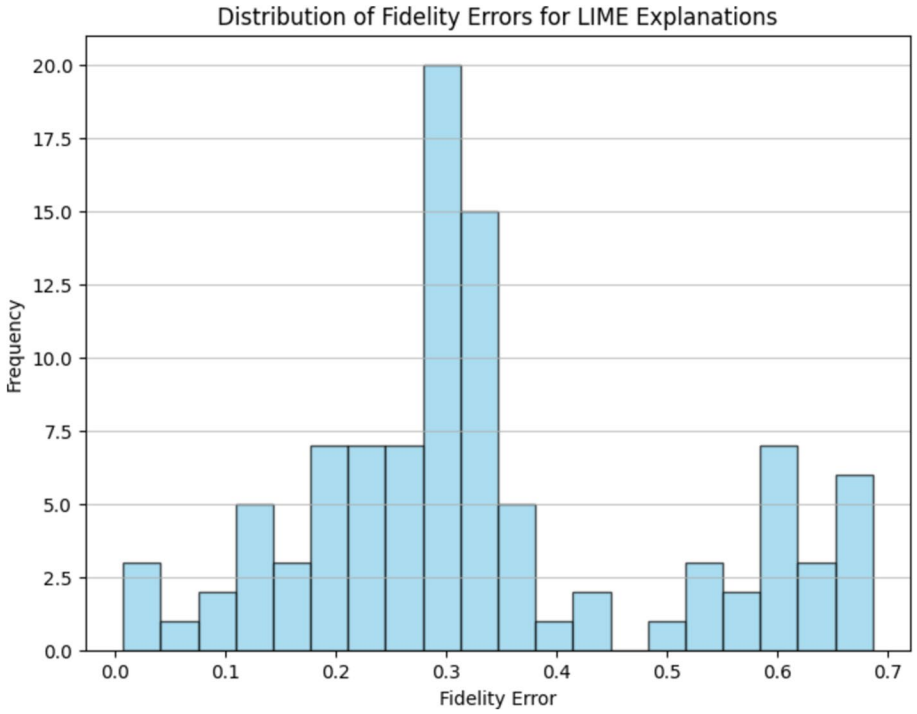


**Fig. 20** Distribution of fidelity errors for LIME explanations

these sensors in mixed gas environments. Variations in clustering density suggest differences in sensor sensitivity and overlap between specific gas types.

Figure 22 depicts the mean SHAP values, showcasing the contribution of individual features to the predictive model's output magnitude. Among the features, MQ3 exhibits the highest average impact on the model's output, indicating its significant importance in the prediction process. MQ7 and MQ135 follow closely, highlighting their substantial influence. Other features, such as MQ6 and MQ8, also contribute considerably, whereas MQ5

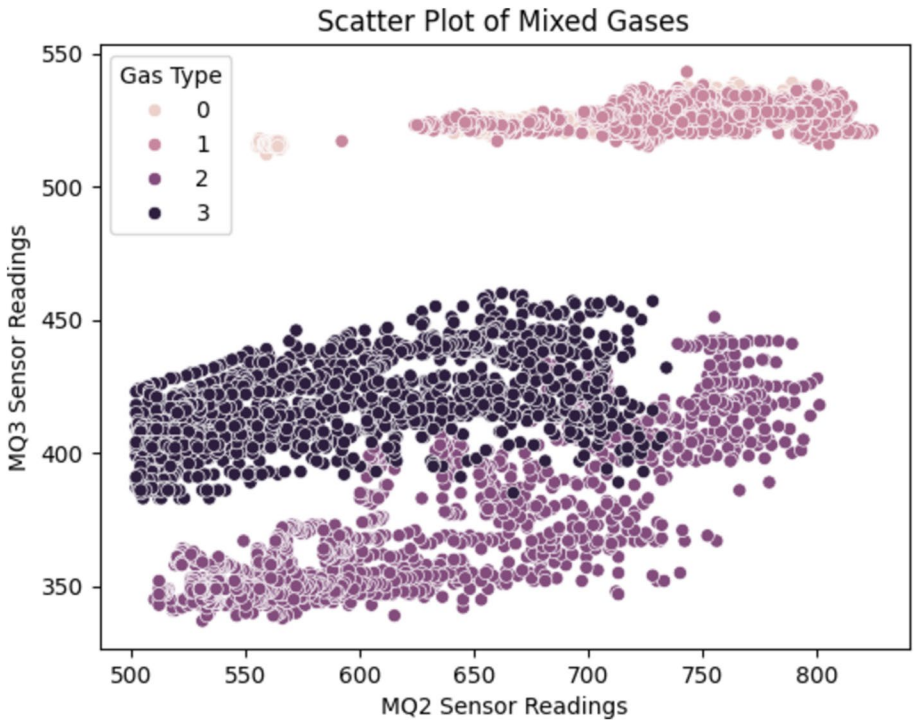


**Fig. 21** Clustered distribution of gas types based on MQ2, MQ3 sensor readings for mixed gas environments

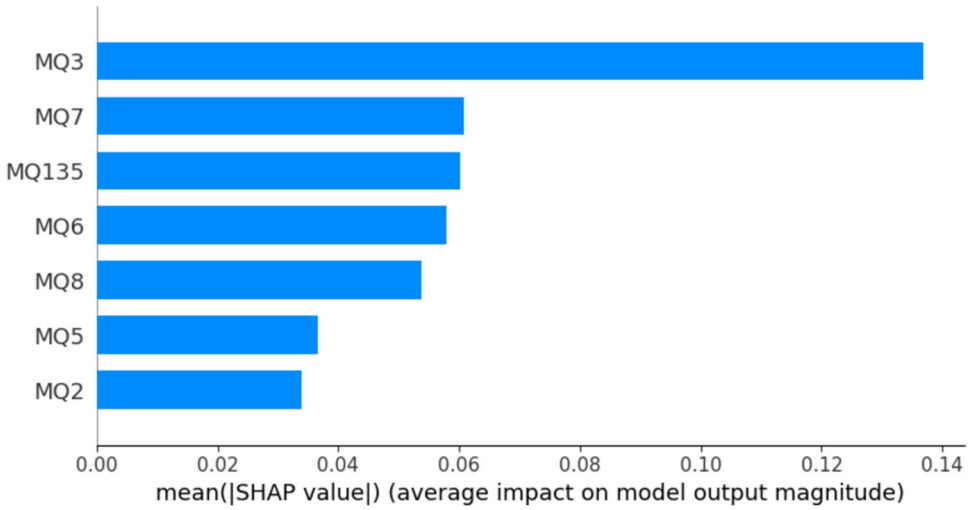


**Fig. 22** Average impact of each feature on the model's output

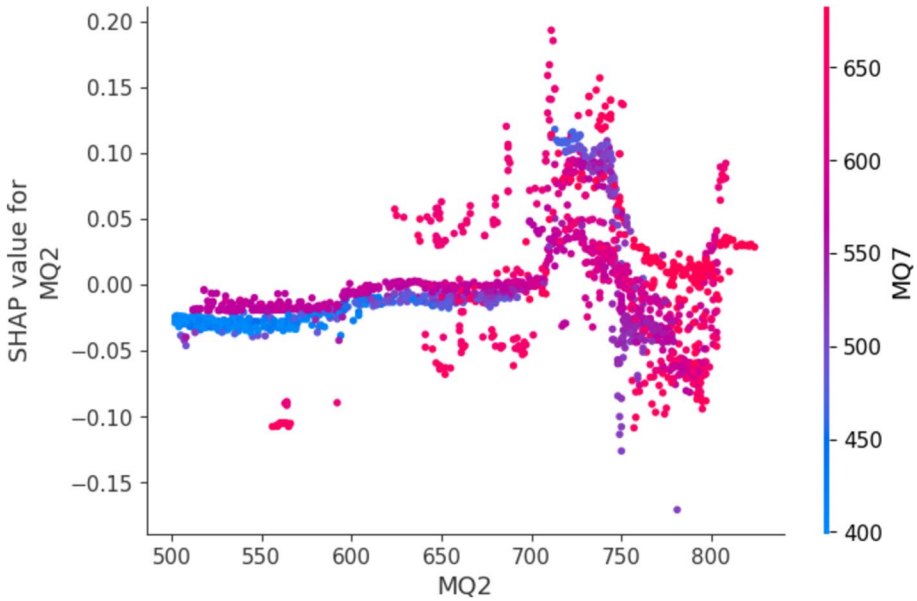


**Fig. 23** SHAP value for MQ2, illustrating its impact on model predictions and interaction with MQ7 values

and MQ2 have relatively lower impacts. This ranking provides an intuitive understanding of feature importance and assists in interpreting the model's decisions effectively.

Figure 23 represents the SHAP values for MQ2, showing the relationship between MQ2 and its contribution to the model's output. The points are color-coded by the values of MQ7, providing additional insights into the interaction between MQ2 and MQ7. The SHAP values fluctuate across the range of MQ2, indicating non-linear effects on the model prediction. Higher MQ7 values (red) appear to correlate with certain shifts in SHAP values, suggesting an interaction effect between MQ2 and MQ7 on the model's behavior. The proposed ExAIRFC-GSDC model delivers significant results and importance of explainability in the following cases.

- Complexity of the application domain

  Gas leakage detection and classification involve diverse sensor data with intricate interactions. Traditional models often lack transparency in explaining how specific sensor readings contribute to the decision-making process. XAI bridges this gap by providing interpretable insights.
- Critical decision-making

 https://doi.org/10.1007/s44196-025-00742-6

Gas leaks pose immediate risks to human safety and the environment. XAI methods like SHAP and LIME enable understanding of the model's rationale, enhancing trust and facilitating swift interventions.

- Regulatory and compliance needs

  Many industrial applications require adherence to strict safety standards. XAI ensures that the predictions meet these standards by elucidating the features influencing decisions.

- Addressing sensor limitations

  Variability in sensor responses (e.g., drift, environmental conditions) demands transparent models to identify which sensor data are driving predictions, enabling corrective measures or recalibrations.

Statistical analysis indicates significant differences in sensor readings across gasses. XAI tools validate these findings by identifying the contributions of individual sensors to model predictions. Visualizations, such as SHAP and LIME plots (e.g., Figs. 9, 10, 11, 12), demonstrate how specific sensors (e.g., MQ3, MQ135) influence the classification of gasses. This level of detail is crucial for identifying robust sensors and optimizing system design. Providing interpretable explanations ensures that diverse users—engineers, safety officers, and policymakers—can understand and leverage the model's predictions effectively. The interactive interface supported by XAI enables stakeholders to make informed decisions, ensuring higher adoption rates in industrial and residential settings.

Table 5 presents descriptive statistics for gas sensor readings (MQ2, MQ3, MQ5, MQ6, MQ7, MQ8, MQ135) across different gasses (Mixture, NoGas, Perfume, Smoke). Mean Indicates the average sensor readings for each gas type. For instance, the mean MQ2 reading for Mixture is 593. Median Represents the middle point of the data distribution. It helps assess symmetry. For example, the median and mean values for NoGas in MQ8 differ. Standard Deviation Measures the spread or variability in sensor readings. Higher values indicate greater variability. For instance, MQ8 readings for Perfume have a relatively high standard deviation of

**Table 5** Descriptive statistical analysis of gas sensor measurements

| Descriptives | Gas | MQ2 | MQ3 | MQ5 | MQ6 | MQ7 | MQ8 | MQ135 |
|---|---|---|---|---|---|---|---|---|
| Mean | Mixture | 593 | 419 | 407 | 376 | 460 | 315 | 403 |
| | NoGas | 748 | 529 | 431 | 425 | 606 | 637 | 474 |
| | Perfume | 745 | 528 | 440 | 429 | 617 | 634 | 481 |
| | Smoke | 623 | 372 | 340 | 368 | 580 | 584 | 308 |
| Median | Mixture | 587 | 417 | 402 | 373 | 449 | 298 | 401 |
| | NoGas | 758 | 530 | 421 | 422 | 610 | 638 | 470 |
| | Perfume | 743 | 528 | 410 | 399 | 637 | 599 | 477 |
| | Smoke | 604 | 361 | 337 | 369 | 577 | 576 | 300 |
| Standard deviation | Mixture | 66.2 | 16.4 | 39.5 | 34.0 | 69.1 | 72.0 | 31.4 |
| | NoGas | 49.6 | 4.06 | 32.9 | 31.6 | 44.6 | 63.7 | 30.2 |
| | Perfume | 44.3 | 4.14 | 54.6 | 51.1 | 71.6 | 102 | 39.0 |
| | Smoke | 78.1 | 28.0 | 26.3 | 17.9 | 9.89 | 29.1 | 25.9 |
| Variance | Mixture | 4379 | 267 | 1561 | 1153 | 4778 | 5183 | 983 |
| | NoGas | 2458 | 16.5 | 1083 | 998 | 1986 | 4061 | 915 |
| | Perfume | 1965 | 17.1 | 2984 | 2610 | 5128 | 10,505 | 1519 |
| | Smoke | 6092 | 784 | 690 | 322 | 97.9 | 846 | 669 |
| Minimum | Mixture | 502 | 383 | 344 | 311 | 361 | 220 | 302 |
| | NoGas | 555 | 512 | 370 | 328 | 506 | 417 | 387 |
| | Perfume | 592 | 515 | 364 | 341 | 464 | 419 | 389 |
| | Smoke | 510 | 337 | 291 | 328 | 555 | 541 | 275 |
| Maximum | Mixture | 734 | 460 | 522 | 466 | 796 | 716 | 482 |
| | NoGas | 809 | 539 | 525 | 517 | 689 | 783 | 584 |
| | Perfume | 824 | 543 | 596 | 524 | 715 | 794 | 589 |
| | Smoke | 801 | 451 | 411 | 413 | 609 | 667 | 379 |

**Table 6** Welch's one-way ANOVA inferential statistical results of gas sensor datasets

| Features | *F* | *df*1 | *df*2 | *p* |
|---|---|---|---|---|
| Serial Number | 0 | 3 | 3553 | 1.000 |
| MQ2 | 2951 | 3 | 3485 | < 0.001 |
| MQ3 | 38,780 | 3 | 3248 | < 0.001 |
| MQ5 | 3272 | 3 | 3461 | < 0.001 |
| MQ6 | 1769 | 3 | 3328 | < 0.001 |
| MQ7 | 1933 | 3 | 2809 | < 0.001 |
| MQ8 | 7558 | 3 | 3185 | < 0.001 |
| MQ135 | 12,268 | 3 | 3523 | < 0.001 |

**Table 7** Outcome of Levene's test that indicates the homogeneity of variances for gas sensors

| Features | *F* | *df*1 | *df*2 | *p* |
|---|---|---|---|---|
| Serial Number | 4.32e–28 | 3 | 6396 | 1.000 |
| MQ2 | 441.2 | 3 | 6396 | < 0.001 |
| MQ3 | 1984.4 | 3 | 6396 | < 0.001 |
| MQ5 | 714.0 | 3 | 6396 | < 0.001 |
| MQ6 | 1175.4 | 3 | 6396 | < 0.001 |
| MQ7 | 1113.2 | 3 | 6396 | < 0.001 |
| MQ8 | 978.9 | 3 | 6396 | < 0.001 |
| MQ135 | 94.4 | 3 | 6396 | < 0.001 |

102—the square of the standard deviation, indicating the dispersion of data points. Higher variances suggest a greater spread. For example, the variance for Smoke in MQ2 is 6092. Minimum and Maximum Shows the range of sensor readings for each gas type, providing insights into data spread. For instance, the range for NoGas in MQ8 is from 417 to 783.

Table 6 presents results from Welch's one-way ANOVA, indicating statistical significance for gas sensors MQ2, MQ3, MQ5, MQ6, MQ7, MQ8, and MQ135 in the context of gas detection and classification. No significant difference was observed among gas sensor readings based on the serial number, as indicated by a $p$ value of 1.000. Statistical significance ($< 0.001$) is observed for all gas sensors, suggesting that there are significant differences in sensor readings across different gasses (Mixture, NoGas, Perfume, Smoke). This indicates that the gas sensors exhibit variability in response to various gasses.

Table 7 presents results from Levene's test, assessing the homogeneity of variances for gas sensors in the context of gas detection and classification. The $p$ value of 1.000 indicates no significant difference in variances based on the serial number. The assumption of homogeneity of variances is met for the serial number variable. For all gas sensors (MQ2, MQ3, MQ5, MQ6, MQ7, MQ8, MQ135) [20], $p$ values are less than 0.001, suggesting significant differences in variances among the different gasses (Mixture, NoGas, Perfume, Smoke). The assumption of homogeneity of variances is violated for these gas sensors.

The ExAIRFC-GSDC model is designed to operate efficiently even in resource-constrained industrial environments. Although it was initially developed using GPU acceleration on Google Colab, the model's core architecture—based on the Random Forest Classifier—can be scaled for deployment on devices with limited computational resources. By adjusting parameters, such as number of trees and maximum depth, the model can be optimized for lower-power devices without significant performance loss. Experimental validation shows minimal degradation in accuracy, precision, recall, and F1 score when reducing model complexity. Additionally, techniques like decision tree pruning, quantization, and edge computing can further enhance its suitability for embedded systems. Therefore, the model can be effectively deployed in industrial settings with constrained resources while maintaining robust performance. The results presented in Table 8 show the minimal performance degradation, indicating that the model can effectively adapt to constrained environments.

 https://doi.org/10.1007/s44196-025-00742-6

**Table 8** Proposed model performance on resource-constrained settings

| Model configuration | Accuracy (%) | Precision (%) | Recall (%) | F1 score (%) |
|---|---|---|---|---|
| Full configuration (GPU) | **98.67** | **97.52** | **98.39** | **97.35** |
| Reduced trees (50 trees) | 96.45 | 95.12 | 96.24 | 95.68 |
| Reduced depth (max depth = 10) | 95.78 | 94.36 | 95.67 | 94.89 |

Bold values indicate the highest performance achieved by the proposed model with GPU utilization

**Table 9** Detection scenarios with corresponding gasses, sensor combinations, performance of proposed model gas classification and baseline monitoring

| Scenario type | Gases detected | Sensors involved | Accuracy achieved (%) |
|---|---|---|---|
| Single gas (e.g., LPG) | LPG, Methane | MQ2, MQ6, MQ135 | 98.67 |
| Mixed gasses (e.g., perfume + smoke) | Perfume, Smoke | MQ3, MQ7, MQ8 | 94.36 |
| No gas | Air quality (baseline) | MQ135 | 100 |

This study utilizes sensor data from various gas types, including LPG, CNG, methane, and propane, and a robust dataset that incorporates different environmental conditions and gas mixtures. The dataset encompasses scenarios, such as no gas, perfume, smoke, and mixtures, offering practical use cases in residential, industrial, and mining environments. Seven MQ series sensors, each optimized for detecting specific gasses, ensure broad detection capability across hazardous substances. The inclusion of SHAP and LIME explanations further enhances the model's adaptability, allowing stakeholders to interpret and adjust the system to specific situational needs.

The study effectively covers a range of scenarios through its experimental setup. The dataset includes individual gasses, such as methane, and mixed gasses, such as perfume and smoke, replicating real-world leakage environments. Statistical analysis, including Welch's ANOVA and Levene's test, confirms significant differences in sensor readings across various gasses, validating the model's applicability across multiple scenarios. Table 9 describes the performance of proposed model gas classification and baseline monitoring. This study also considers environmental factors like temperature, humidity, and sensor drift, ensuring reliable model performance under dynamic conditions.

While the research covers a wide range of scenarios, it acknowledges certain limitations. The dataset does not include extreme conditions, such as high-pressure gas leaks, low oxygen levels, or scenarios involving simultaneous detection of more than two gas types. Some MQ series sensors may exhibit drift or reduced sensitivity under prolonged exposure to high concentrations of gasses. Additionally, this study does not explicitly model environmental noise factors like wind speed and particulate matter interference. Future research will address these gaps by incorporating additional sensor types, enriching the dataset with more complex scenarios, and testing the model under extreme conditions.

## 5 Conclusion and Future Enhancements

In this research, we addressed the critical challenge of gas detection and classification, particularly in scenarios involving industrial and residential environments. The continuous threat of hazardous gas leaks and emissions necessitates innovative technological solutions for timely and accurate detection. This study extensively reviewed existing gas detection methodologies, emphasizing the limitations of conventional sensor-based approaches and the need for advanced techniques. Our proposed model, ExAIRFC-GSDC, integrates machine learning algorithms with Explainable Artificial. The architecture prioritizes interpretability, providing insights into decision-making through SHAP and LIME explanations. Statistical analysis, including Welch's one-way ANOVA and Levene's test, confirmed the significant differences in sensor readings across various gasses, validating the effectiveness of our proposed model. Descriptive statistics added depth to our understanding of the dataset, highlighting the

**Table 10** Comparison of current dataset characteristics and future enhancements for gas detection in diverse environments

| Feature | Current dataset | Future enhancements |
|---|---|---|
| Data Source | Public dataset (Kaggle) | Real-world and simulated complex terrains |
| Conditions Covered | Residential, industrial, mining | High-altitude, underground, high-pressure |
| Modalities | Gas sensor readings | Multimodal (thermal, sound, etc.) |

central tendencies and the variability in sensor readings. The experimental results demonstrated the superior performance of ExAIRFC-GSDC, showcasing its accuracy, precision, recall, and F1 score across different gas types. The model's interpretability was further demonstrated through SHAP analysis, providing a comprehensive view of the contribution of each sensor to predictions. Our research contributes to the field by presenting an innovative and interpretable model for gas detection.

While the dataset is comprehensive for foundational research, we acknowledge that it represents a single data package and does not cover the full spectrum of terrains and conditions found in real-world gas leakage scenarios. Complex terrains, such as high-altitude regions, underground mining operations, or industrial setups with heavy environmental interference, introduce additional challenges. To address this limitation, we have implemented robust pre-processing techniques, including normalization and outlier handling, to enhance the model's generalizability. Additionally, statistical analysis validates the significant differences across sensor readings for various gasses, demonstrating the model's adaptability within the current dataset. To broaden the model's applicability, we plan to extend this study by integrating additional datasets that encompass more diverse environmental conditions and gas compositions. Future work will focus on following scenarios:

### 5.1 Simulation-Based Testing

Employing virtual simulations to model gas leaks in scenarios where real-world data collection may be impractical or unsafe. Table 10 presents the comparative analysis of the current dataset features with future improvements for enhanced gas detection across diverse environments.

### 5.2 Data Collection in Complex Terrains

Acquiring real-world datasets from varied industrial, residential, and mining environments to include challenging scenarios, such as high-pressure leaks, extreme temperatures, and varying humidity levels.

### 5.3 Incorporation of Multimodal Data

Combining thermal imaging, sound-based detection, and other sensor modalities with existing gas sensor data to enhance model performance in mixed environments.

**Author Contributions** All authors contributed to the research paper conception and design. B. Lalithadevi and S. Krishnaveni performed concept design, supervision, preparation of manuscript, model development with explainable AI and interpretation of the results, editing the manuscript. All authors read and approved the final manuscript.

**Funding** Not applicable.

**Data Availability** The data that support the findings of this study are publicly available.

## Declarations

**Conflict of interest** The authors declare that they have no competing interests.

**Ethics approval** Not applicable.

**Informed consent** Not applicable.

**Consent to publish** Not applicable.



## References

1. Trivedi, P., Purohit, D., Soju, A., Tiwari, R.R.: Major industrial disasters in India an official newsletter of ENVIS-NIOH, Oct–Dec 2014; Volume 9, No 4. http://niohenvis.nic.in/newsletters/vol9_no4_Indian%20Industrial%20Disasters.pdf. Accessed 27 Mar 2024.
2. Zhou, Y., Zhao, X., Zhao, J., Chen, D.: Research on fire and explosion accidents of oil depots. In: 3rd International Conference on Applied Engineering, April 22–25, vol. 51, pp. 163–168. Wuhan, China (2016).
3. Mudur, G.S.: Lakhs of early deaths tied to home emissions. *Telegraph India Online* 2015. Published online on 17 September 2015. https://www.telegraphindia.com/india/lakhs-of-early-deaths-tiedto-home-emissions/cid/1513045. Accessed 27 Mar 2024.
4. Evalina, N., Azis, H.A.: Implementation and design gas leakage detection system using ATMega8 microcontroller. In: IOP Conference Series: Materials Science and Engineering, vol. 821(1), pp. 012049. IOP Publishing (2020).
5. Fox, A., Kozar, M., Steinberg, P.: CARBOHYDRATES|gas chromatography and gas chromatography—mass spectrometry. In: Wilson, I.D. (ed.) Encyclopedia of separation science, pp. 2211–2223. Academic Press, Oxford (2000)
6. Wang, T., Wang, X., Hong, M.: Gas leak location detection based on data fusion with time difference of arrival and energy decay using an ultrasonic sensor array. Sensors **18**(9), 2985 (2018)
7. Stauffer, E., Dolan, J.A., Newman, R.: Gas chromatography and gas chromatography–mass spectrometry. Fire Debris Anal (2008). https://doi.org/10.1016/B978-012663971-1.50012-9
8. Yin, X., Zhang, L., Tian, F., Zhang, D.: Temperature modulated gas sensing E-nose system for low-cost and fast detection. IEEE Sens. J. **16**(2), 464–474 (2015)
9. Brahim-Belhouari, S., Bermak, A., Shi, M., Chan, P.C.: Fast and robust gas identification system using an integrated gas sensor technology and Gaussian mixture models. IEEE Sens. J. **5**(6), 1433–1444 (2005)
10. Liu, Q., Hu, X., Ye, M., Cheng, X., Li, F.: Gas recognition under sensor drift by using deep learning. Int. J. Intell. Syst. **30**(8), 907–922 (2015)
11. Peng, P., Zhao, X., Pan, X., Ye, W.: Gas classification using deep convolutional neural networks. Sensors **18**(1), 157 (2018)
12. Pan, X., Zhang, H., Ye, W., Bermak, A., Zhao, X.: A fast and robust gas recognition algorithm based on hybrid convolutional and recurrent neural network. IEEE Access **7**, 100954–100963 (2019)
13. Adekitan, A.I., Matthews, V.O., Olasunkanmi, O.: A microcontroller based gas leakage detection and evacuation system. In: IOP Conference Series: Materials Science and Engineering, vol. 413(1), pp. 012008. IOP Publishing (2018).
14. Suma, V., Shekar, R.R., Akshay, K.A.: Gas leakage detection based on IOT. In: 2019 3rd international conference on electronics, communication and aerospace technology (ICECA), pp 1312–1315. IEEE (2019).
15. Vergara, A., Vembu, S., Ayhan, T., Ryan, M.A., Homer, M.L., Huerta, R.: Chemical gas sensor drift compensation using classifier ensembles. Sens. Actuators B Chem. **166**, 320–329 (2012)
16. Narkhede, P., Walambe, R., Mandaokar, S., Chandel, P., Kotecha, K., Ghinea, G.: Gas detection and identification using multimodal artificial intelligence based sensor fusion. Appl. Syst. Innov. **4**(1), 3 (2021)

17. Pashami, S., Lilienthal, A.J., Trincavelli, M.: Detecting changes of a distant gas source with an array of MOX gas sensors. Sensors **12**(12), 16404–16419 (2012)
18. Huerta, R., Mosqueiro, T., Fonollosa, J., Rulkov, N.F., Rodriguez-Lujan, I.: Online decorrelation of humidity and temperature in chemical sensors for continuous monitoring. Chemom. Intell. Lab. Syst. **157**, 169–176 (2016)
19. Jadin, M.S., Ghazali, K.H.: Gas leakage detection using thermal imaging technique. In: Proceedings of the 2014 UKSim-AMSS 16th International Conference on Computer Modelling and Simulation, Cambridge, UK, 26–28 March 2014, p. 3.
20. Narkhede, P., Walambe, R., Chandel, P., Mandaokar, S., Kotecha, K.: MultimodalGasData: multimodal dataset for gas detection and classification. Data **7**(8), 112 (2022)
21. Lalithadevi, B., Krishnaveni, S.: Efficient disease risk prediction based on deep learning approach. In: 2022 6th international conference on computing methodologies and communication (ICCMC), pp. 1197–1204. IEEE (2022).
22. Lalithadevi, B., Krishnaveni, S., Gnanadurai, J.S.C.: A feasibility study of diabetic retinopathy detection in type II diabetic patients based on explainable artificial intelligence. J. Med. Syst. **47**(1), 85 (2023)



## Authors and Affiliations

## B. Lalithadevi[1] · S. Krishnaveni[2]

✉ B. Lalithadevi
lalithadevi.b.cse@sathyabama.ac.in

✉ S. Krishnaveni
krishnas4@srmist.edu.in

[1] Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai 600119, India

[2] Department of Computational Intelligence, SRM Institute of Science and Technology, Kattankulathur 603203, India