Chapter 5

An IoT-Enabled E-Nose for Remote Detection and Monitoring of Airborne Pollution Hazards Using LoRa Network Protocol

Detection and monitoring airborne hazards using e-noses has been lifesaving and prevented accidents in real-field scenarios. E-noses generate unique signature patterns for various volatile organic compounds (VOCs) and, by leveraging artificial intelligence, detect the presence of various VOCs, gases and smokes on site. Widespread monitoring of airborne hazards across many remote locations is possible by creating a network of gas sensors using Internet connectivity, which consumes significant power. Long Range (LoRa)-based wireless networks do not require Internet connectivity while operate independently. Therefore, we propose a Networked Intelligent Gas Sensor System (N-IGSS) using LoRa low-power wide area networking protocol for real-time air-borne pollution hazard detection and monitoring. We have developed a gas sensor node by using an array of seven cross-selective tin-oxide-based Metal-Oxide Semiconductors (MOX) gas sensor elements interfaced with a low-power microcontroller and a LoRa module. Experimentally, we expose the sensor node to six classes of VOCs as released by burning samples of Tobacco, Paints, Carpets, Alcohol and Incense Sticks. Using the proposed two-stage analysis space transformation approach, the captured dataset was first pre-processed using the Standardized Linear Discriminant Analysis (SLDA) method. Four different classifiers viz. AdaBoost, XGBoost, Random Forest (RF) and Multi-Layer Perceptron (MLP) were then trained and tested in the SLDA transformation space. The proposed N-IGSS achieves 'all correct' identification of 30 unknown test samples with a low mean squared error (MSE) of 1.42 × 10-4 over a distance of 590 m.

5.1. Introduction

The World Health Organization (WHO) reports that Indoor Air Pollution (IAP) harms 3.8 million people every year [170]. It is caused by occupant activities such as cooking, smoking, using electronic equipment, using consumer items, or

emissions from construction materials inside houses or buildings. Carbon monoxide (CO), Volatile Organic Compounds (VOCs), Particulate Patter (PM), aerosol and other biological pollutants are such harmful pollutants that are found in the indoor ambience of buildings [171]. In addition, it is reported that poor Indoor Air Quality (IAQ) may harm people's health by increasing disorders that are associated with air borne pollution present in indoor ambiences [92], [172]. Thus, air-borne pollution detection and monitoring the presence of various hazardous VOCs, gases, and smoke are essential to mitigate poor air quality in the indoor ambience.

Low-Power Wide Area Network (LPWAN) technologies have attracted considerable interest and played a significant role in popularizing the Internet of Things (IoT) based applications. Depending on the deployment area, most LPWANs operate in the unlicensed Industrial, Scientific, and Medical (ISM) bands at 169, 433, 868/915 MHz, and 2.4 GHz [173]. A LoRa-based network link is a wireless communication technology that uses long-range, low-power radio frequency signals to transmit data over a long distance. These LPWAN technologies include NB-IoT, Sigfox, and LoRa. It has been demonstrated that LoRa offers more advantages in ultralong transmission, high stability, ultralow power, and low cost [101]. LoRa-based network links are designed to provide reliable and cost-effective communication for IoT devices in various applications, including smart cities, industrial automation, agriculture, and environmental monitoring [173].

LoRa-based network links have two main components, i.e., LoRa nodes and LoRa gateways. The LoRa nodes are small, low-power devices that can be embedded in IoT sensors and devices to transmit data wirelessly. The LoRa gateways act as a bridge between the LoRa nodes and the internet, receiving data from the nodes and forwarding it to a cloud-based server or an IoT platform. The characteristics of several network technologies for the IoT are shown in Table 5.1. This study introduces the LoRa WAN protocol-based pollution hazard detection and monitoring system, comprising a microcontroller and up to eight sensors and operating remotely. Further, advanced two-stage analysis space transformation-based approaches have also been used to enhance the performance and efficiency of the proposed Networked Intelligent Gas Sensor System (N-IGSS) [49], [148] – [149].

Table 5.1. Topology and frequency of the communication technologies [101], [174]-[175].

Network Technologies	Topology	Coverage Range	Power Consumption	Radio Frequency	Tx- Rx Data Size	Limitations/ Advantages
BLE	Ad-hoc	10-100 m	15-30 mA per packet	2.4 GHz-2.4835 GHz	1-3 Mbps	Short range
Wi-Fi	Star	50-100 m	2 to 20 watts	2.4 GHz – 5 GHz	1-9608 Mbps	Short distance, High battery power
ZigBee	Mesh	10-100 m	150 mA	868.3 MHz, 902-928 MHz	20-250 kbps	Short distance, Maintenance cost too much
Sigfox	Star	20-25 km	78 mA	862-928 MHz	100 bps	High module cost, High battery power
LoRa	Star/ Mesh	10-20 km	32 mA	433 MHz, 860-1020 MHz	290bps – 50 kbps	More extended range, low power battery

The literature review shows a growing interest in environmental monitoring using IoT and cloud-based technologies. Numerous environmental factors, including temperature, humidity, light, noise level, CO, and NO_2, have been monitored indoors and outdoors at a university campus using an environmental monitoring system [89]. This system includes the Raspberry Pi and MICS-4514 sensors. An IoT-based indoor environmental quality (IEQ) assessment system employing an Arduino Uno module and economical sensors like the DHT22 was developed in an experimental investigation [90]. An IoT-based measurement system with simple pressure, temperature, humidity, and gas sensors (BME680, DHT22, and MQ5) interfaced with an Arduino microcontroller has also been reported [91]. In a similar experiment, Kureshi et al. used temperature, humidity, a VOCs sensor (BME680), and PM 2.5 and PM 10 sensors (SDS011) for air quality index (AQI) calculation. They have used this information to increase awareness about indoor air pollution hazards to human health [92]. Numerous gases, including CO_2, NO_2, ethanol, methane, and propane, were identified using the gas sensors MICS-6814 in the IoT-based system [93] developed for real-time IAQ monitoring.

The emission investigation, including particles and benzene generated by different incense sticks and sparklers, has been reported by Werner et al. [94] utilized indoors, they had a significantly unfavorable impact on the air quality. The indoor AQI is further explored by using multiple low-cost PM sensors for effective compliance with the U.S. Environmental Protection Agency (EPA) [95]. The variation in the concentration ranges of CO, CO2, Total Volatile Organic Compounds (TVOCs), and PM 2.5 over a period of 24 hours is monitored in indoor environments by Capua et al., and a decision-making algorithm is used for mitigating the IAQ [96]. Rebecca et al. have described the particles and gases released by burning Arabian incense and looked at the reactions of in vitro human lung cells to incense smoke [97]. A study on mosquito-repellent incense was conducted by Wang et al. It generates excessive gaseous and particle pollutants, as permitted under GB3095-2012 rules and the WHO advice [98]. Further, green IoT for eco-friendly discussed by [177] and Opportunistic routing (OR) is widely adopted in Wireless Sensor Networks (WSNs) running asynchronous duty-cycled MAC protocols by [178].

In the recent literature, the researchers have only considered AQI and specific sensors to detect and monitor a limited number of pollution gases and PM concentrations for evaluating indoor air quality. The attention towards various VOCs, gases, and smoke in indoor environments is a research gap. Further, an electronic nose (e-nose) is a device that mimics the olfactory system of humans to identify and analyze various VOCs, gases and smoke present in the environment. E-noses have been widely used for detecting and monitoring air pollution hazards, including industrial emissions, traffic emissions, and indoor air pollutants. The proposed approach for pollution hazard detection and monitoring using e-noses is to provide real-time monitoring of air-borne pollutants at a centrally located RDPS. The data collected at the RDPS is analyzed by leveraging artificial intelligence-based algorithms and models to identify the classes of pollutants of interest [49],[148]-[149]. Pollution hazard detection networks are designed to detect, monitor, and predict the class and level of environmental pollution. Such networks can be categorized based on the sensors, spatial coverage, and data transmission methods. In addition, people worldwide are utilizing the Internet and GSM mobile-based networks which require high power and paid subscriptions [101].

In this work, we propose a cyber-physical system (CPS) suitable networked intelligent gas sensor system (N-IGSS) using a LoRa network link-based IoT platform which can be operated in private places like in industrial scenarios, mines, etc., and it can operate well even when there is no Internet or GSM mobile-based network connectivity. To the author's knowledge, this is the first work to have integrated the LoRa network link to develop a N-IGSS. The proposed N-IGSS can be used in many practical LoRa-based applications, such as hospital management, air quality monitoring, agriculture, smart cities, and industrial management. Therefore, the proposed N-IGSS will gain economic, social and environmental impacts.

5.1.1 Motivation and Contributions

We are motivated by providing a low-power, low-cost solution for widely monitoring airborne hazards across remote locations. The proposed solution addresses the challenge of monitoring airborne hazards using Internet connectivity, which consumes significant power. Furthermore, the proposed system can operate independently by utilizing LoRa-based wireless networks that do not require Internet connectivity, thus reducing power consumption and providing extended battery life. The proposed architecture has been planned as a CPS suitable network of E-Noses as shown in Figure 5.1.

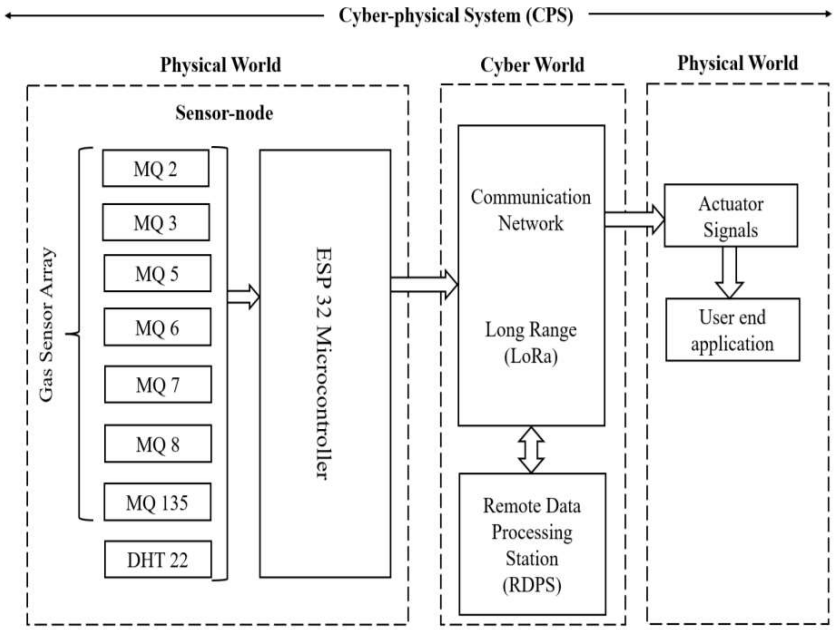


Figure 5.1 Architecture of CPS suitable network of E-Noses using LoRa Networking Protocol

The proposed N-IGSS comprises an array of low-cost MOX-based gas sensor elements integrated as the gas sensor node. In this experiment, we have considered six types of VOCs, gases and smoke released by burning tobacco, paint, carpet, incense sticks, and alcohol. Our sensor node is placed indoors to capture and transmit sensor responses on the LoRa network link to the remotely placed LoRa gateway (remote data processing station) for further analysis and predictions for real-time operations. We have processed the sensor response data at the RDPS. The block schematic diagram of our proposed N-IGSS is shown in Figure 5.2. In the proposed N-IGSS, an array of seven-element Tin Oxide based gas sensor elements, along with one digital temperature and humidity sensor, has been used for detecting six classes of VOCs, gases and smokes. Respective signature patterns are analyzed using our proposed two-stage analysis space transformation method, which closely mimics the human olfactory system [99]. The utilities of this work are highlighted as follows:

1. A N-IGSS is proposed to detect and monitor air-borne pollution hazards in indoor ambiences at RDPS.

2. For the first time, a Lora WAN networking link protocol is used for real-time networked operations of e-noses.

3. The proposed N-IGSS has been designed using a two-stage analysis space transformation method to ensure the classifier models deliver high performance.

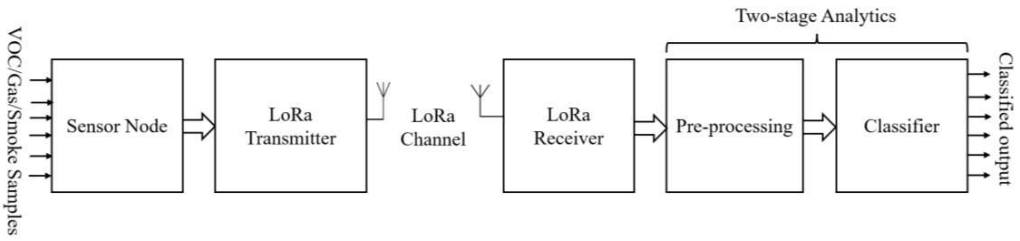


Figure 5.2. Block Schematic diagram of the Networked Intelligent Gas Sensor System (N-IGSS).

5.2. Materials and Methods

In this section, we present the proposed method of developing an N-IGSS, which has been integrated with the LoRa networking link protocol for hazardous

pollution detection in real-time at an RDPS using artificial intelligence-based algorithms and models.

5.2.1. The contextual background of N-IGSS

The proposed system comprises three parts, as shown in Figure 5.3 (a) – (c):

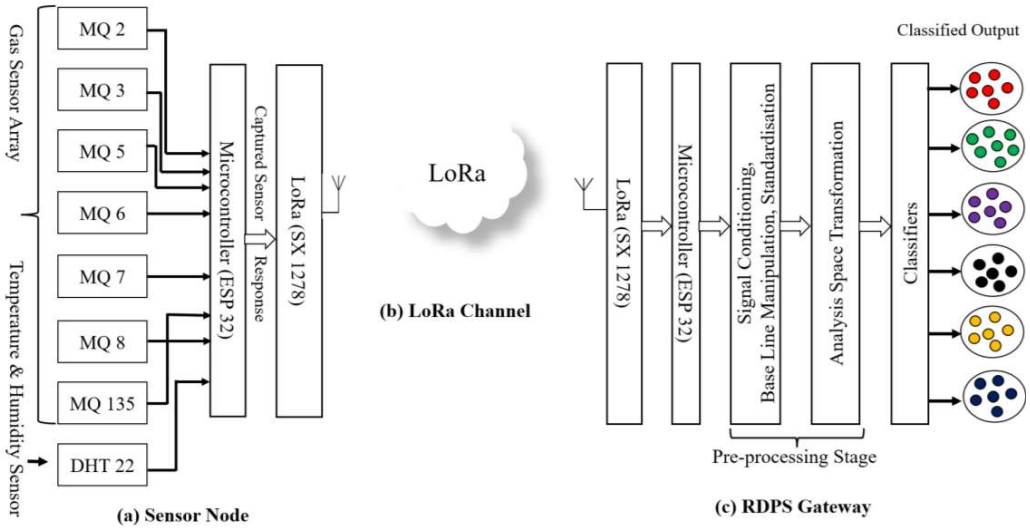


Figure 5.3 (a)-(c). The proposed architecture of networked intelligent gas sensor system (N-IGSS).

Physical Gas Sensor Node (E-Nose): The sensor node is the on-site part of the proposed N-IGSS, which consists of an array of seven tin-oxide (MOX)-based gas sensor elements viz. MQ 2, MQ 3, MQ 5, MQ 6, MQ 7, MQ 8, MQ 135 and one temperature and humidity sensor viz. DHT 22 interfaced with a microcontroller and a LoRa module, as shown in Figure 5.3 (a). These MOX sensors are cross-sensitive and respond to multiple VOCs, gases, and smokes with different sensitivities and generate unique signature patterns for various VOCs, gases, and smokes in different types of pollutants.

The LoRa link: The LoRa link is a wireless communication technology designed to enable long-range, low-power communication between devices. The LoRa link operates in the unlicensed radio spectrum, which means it can be used without needing a license or paying subscription fees. The main function of the LoRa link is to provide a reliable and efficient communication link for IoT devices. It is

particularly well-suited for applications that require low data rates and long-range communication, as shown in Figure 5.3 (b).

RDPS: We receive the real-time transmission of the sensor node responses at the other end of the LoRa link using the LoRa gateway. LoRa is a transceiver module, and we can use the same module as a transmitter and receiver according to the configuration code. At this RDPS, we have processed the gas sensor array responses using our proposed two-stage space transformation method to achieve high performance. The captured gas sensor array responses are first pre-processed in this method to transform the raw sensor responses into a suitable analysis space using the Standardized Linear Discriminant Analysis (SLDA) method. In the second stage, different types of classifiers, viz., AdaBoost, XGBoost, RF, and MLP, are designed in the SLDA analysis space to efficiently classify the samples of the six classes of VOCs, gases, and smokes, as shown in Figure 5.3 (c).

5.2.2. The Gas Sensor Node Prototype

The gas sensor node consists of seven tin-oxide-based MOX based gas sensor elements and one temperature and humidity sensor, which generates real-time signature patterns of the considered VOCs, gases, and smoke being released while burning considered types of pollutant materials. The sensor node is placed indoors to transmit sensor responses to the analyte VOCs, gases, and smoke using the LoRa network link protocol. These wirelessly transmitted sensor array responses are received at an RDPS (LoRa link gateway). The electrical characteristics of the components used for the sensor node prototype development are shown in Table 5.2.

Table 5.2. Electrical Characteristics of the Components Used in Device Prototype.

List of components	Input Voltage	Power Ratings
LoRa Module (SX1278)	3.3V	Tx:93 mA, Rx:12.15 mA, standby:1.6mA
ESP 32 Microcontroller	5V	130mA
ESP 32 GPIO pins	3.3V	40mA
MQ Sensor	5V	150mA
DHT-22	5V	2.5mA

The connection diagrams of the sensors and the LoRa transmitter and receiver with the microcontroller are shown in Figures 5.4 (a) and (b).

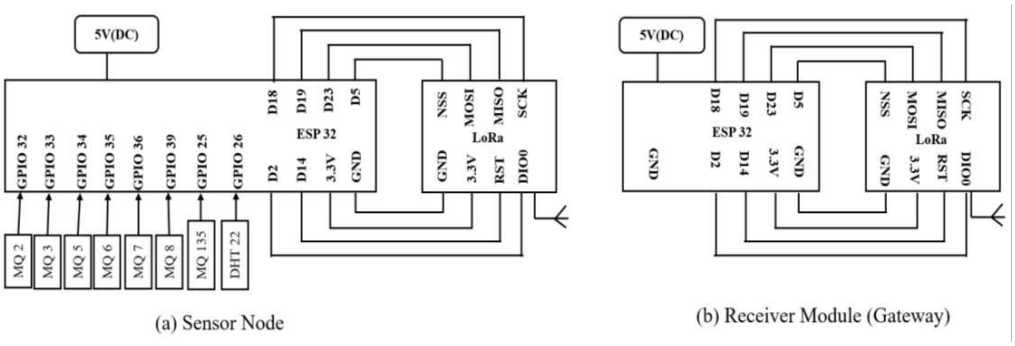
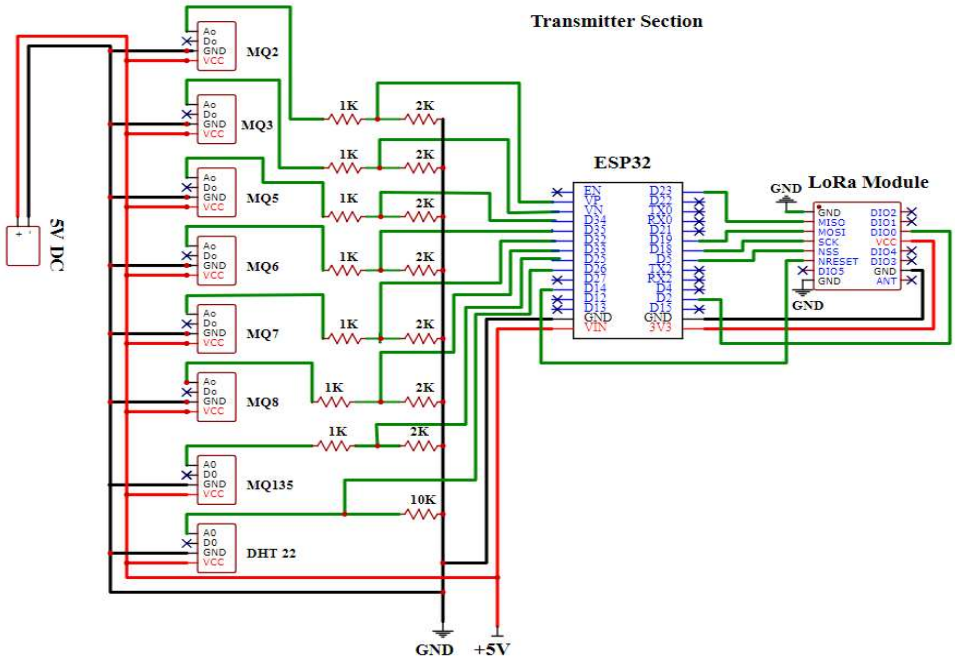
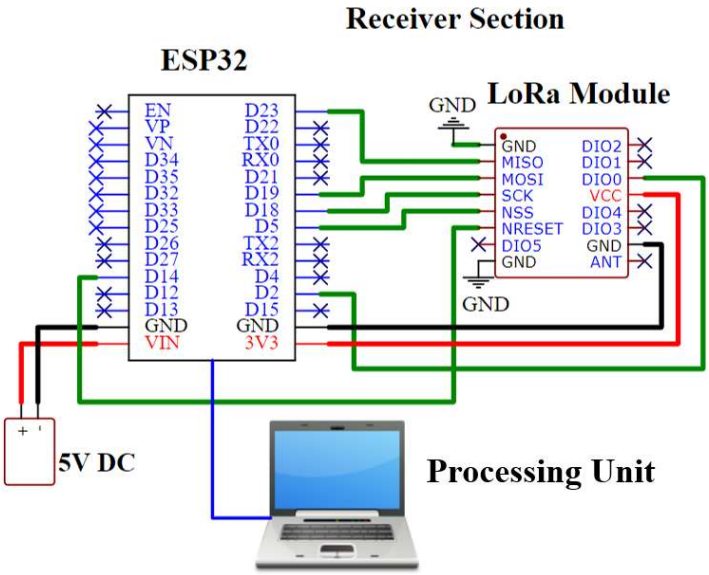


Figure 5.4. Connection Diagram of (a) Sensor Node (Transmitter), (b) Gateway (Receiver Module), as interfaced with the microcontroller and LoRa module.

We have interfaced the sensors and LoRa modules with respective microcontrollers for creating the sensor node and LoRa receiver gateway of the proposed N-IGSS. It communicates with the sensors and transmits the data through Serial Peripheral Interface (SPI) to the LoRa transceiver. The LoRa module SX1278 is a 137 MHz to 525 MHz Long Range Low Power transceiver module which can be used both as a transmitter and a receiver. The sensor node is powered using a 5V (DC) supply. The LoRa module (SX1278) is supplied 3.3V by voltage dividing the microcontroller's 5V supply. The Basic Circuit Diagram of N-IGSS is depicted in Figure 5.5 (a) Sensor Node and (b) Receiver Gateway at RDPS.



(a) Sensor Node

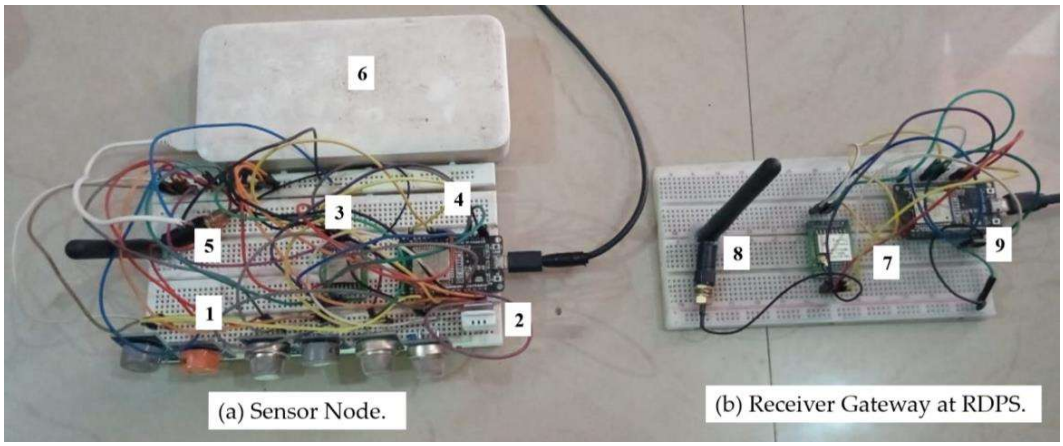


(b) Receiver Gateway at RDPS

Figure 5.5. Basic Circuit Diagram of N-IGSS (a) Sensor Node and (b) Receiver Gateway at RDPS.

LoRa consists of an antenna that transmits the sensor values through the microcontroller to the receiver node via radio waves. In our case, we used a 915 MHz

antenna in the ISM band designated for Asia. The physical view of fabricated N-IGSS for pollution hazard detection and monitoring is shown in Figures 5.6 (a) and (b).



1: Gas Sensor Array; 2: Temperature & humidity Sensor; 3: LoRa Module (Tx); 4: Microcontroller (Tx); 5: Antenna (Tx); 6: Power supply (DC, 5V); 7: LoRa Module (Rx); 8: Antenna (Rx); 9: Microcontroller (Rx)

Figure 5.6. Physical Prototype of N-IGSS (a) Sensor Node, (b) Receiver Gateway at RDPS.

The LoRa transceiver at the receiver gateway is coded exclusively to function as the gateway. This gateway can be connected to the internet to upload real-time sensor responses to the cloud. Multiple nodes are also used to create a more effective and larger network of IGSS.

5.2.3. Experimental setup

In ambient conditions, we conducted our experiments in a closed, semi-furnished room of 114 square meters (1,226.64 square feet). We deployed the N-IGSS sensor node at a fixed height of 2 feet and captured real-time responses of the sensor array for the analyte smokes of tobacco, paints, carpets, alcohol, and incense sticks. In our experiments, we have employed the SX1278 LoRa module, which can only operate on the 868 MHz (European ISM) and 915 MHz frequencies (American ISM). To adhere to Federal Communications Commission (FCC) rules, we selected the 915 MHz bands.

Initially, the gas sensor system is purged with ambient air to obtain the baseline signature pattern under steady-state conditions of the sensor array responses for ten minutes. Once the sensor array responses become steady, the array is exposed to one of the considered classes of VOCs, gases, or smokes for twenty minutes at a rate of 15 samples per minute. After twenty minutes of exposure, the gas sensor array is again purged with ambient air until the sensor responses return to the baseline signature patterns. This process is repeated for all the considered classes of VOCs, gases, and smokes, and the raw sensor array response dataset is captured in its totality.

In this experiment, the sensor node was first heated up for 15 minutes while keeping the gas chamber closed, and the steady state of the sensor responses was achieved. Consequently, the gas chamber inlet opened for the next 20 minutes, and samples of the considered analyte were captured at an RDPS. Then the analyte exposure was stopped, and for the next 30 minutes, the gas chamber was purged with fresh ambient air to ensure that the sensor responses returned to the baseline conditions. The same procedure was repeated to capture samples of the other analytes, as considered.

Accordingly, each experimental phase continues for 65 minutes, raw sensor responses are captured, and the same process is repeated in all experiments. Therefore, the experiment took 390 minutes (65 minutes × 6 types). All sensor responses return to baseline responses throughout the experiment, and no sensor poisoning occurs. During this period, 1800 samples were captured at a sampling rate of 15 samples per minute. Further details of the dataset and the samples collected are given in Table 5.3.

Table 5.3. The Distribution of Samples.

Raw Materials	Sampling Time (mins)	Total Samples	Training Samples	Testing Samples	Class
Ambient air	20	300	295	5	1
Tobacco	20	300	295	5	2
Paints	20	300	295	5	3
Carpet	20	300	295	5	4
Incense	20	300	295	5	5
Alcohol	20	300	295	5	6
Total	120	1800	1770	30	

The samples belong to the six classes of VOCs, gases, and smoke released by burning different pollutants in indoor ambient conditions. The dataset consists of 300 samples of each class: ambient air, tobacco, paints, carpet, incense sticks, and alcohol. The captured dataset was then segregated into training and testing datasets. The training dataset consisted of 295 × 6 = 1770 samples, and 5 × 6 = 30 samples for testing purposes. Finally, the testing dataset was separated from the training or validation of the classifiers at any stage.

A prototype LoRa network has been implemented and deployed in typical urban environments for air-borne pollution hazard detection and monitoring and to evaluate the efficacy of the proposed LoRa network. In the experiments, the LoRa node is gradually moved away from the LoRa gateway's location while sending uplink packets, and the received signal strength is measured continuously. A LoRa node transmits the data packet every 4 seconds, received at the LoRa gateway. Since LoRa Gateway can communicate across channels, we received the sensor responses correctly at the remote data processing station.

5.2.4. Contextual Background of Analysis Space Transformation

The proposed N-IGSS is based on performance enhancement approaches by designing suitable classifiers in the analysis space transformation domain. It has been reported that a classifier performs better when it is trained in a suitable transformation space where the data shows well-separated clusters with good inter-cluster separation [8]. In the following subsections, we have described the processes in greater detail:

5.2.4.1. Data Pre-processing

This work uses the Standardized Linear Discriminant Analysis (SLDA) space transformation method. The SLDA transforms the raw data into respective space transformation domains [4], [49], [100], [151] – [152], [176]. The first three principal components contain 98.88% information, using which we have obtained the 3-D scatter plot, which shows well-separated clusters with good inter-cluster separation [148], as shown in Figure 5.7.

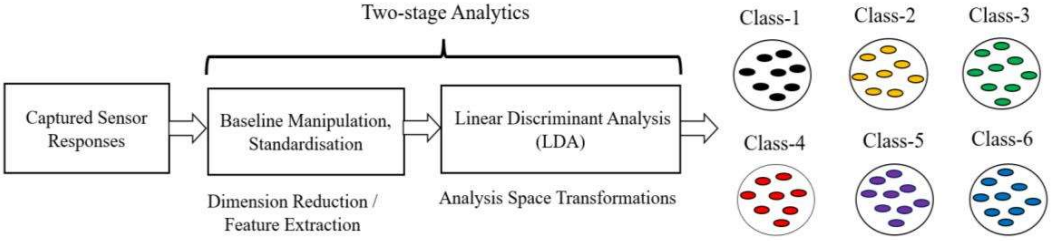


Figure 5.7. Block Schematic Diagram for obtaining 3D Scatter Plot.

Standardized Linear Discriminant Analysis (SLDA) is a statistical method that can be used for dimensionality reduction. It is a linear transformation process on the input data, which transforms the raw sensor responses into new components to maximize the separation between different classes.

The process of SLDA space transformation method has been described in the following steps:

- Calculate the mean vector, mi (i=1,2,3…) of each class of a dataset
- Scatter matrix within the class

$$S_W=\sum_{i=1}^{c} S_i \tag{47}$$

Where S_W= data points within each class deviate from their respective class

$$S_i = \sum_{x\in D_i}^{n}(x - m_i)(x - m_i)^T \tag{48}$$

Where S_i = scatter matrix of each class, x = data point, mi = mean vector and

T= transpose matrix

$$m_i = \frac{1}{n_i}\sum_{x \in x_i}^{n} x_i \tag{49}$$

- Calculate the covariance matrix by adding the scaling factor (1/(N-1)) to the within-class scatter matrix,

$$\sum_i = \frac{1}{N_i-1}\sum_{x\in D_i}^{n}(x - m_i)(x - m_i)^T \tag{50}$$

$$\text{And } S_W = \sum_{i=1}^{c}(N_i - 1)\sum_i \tag{51}$$

Where N_i= sample size of the VOC class (here = 300 × 6), Now we can drop (Ni - 1) because all classes have equal sample size

- Scatter matrix between each class (S_B)

$$S_B = \sum_{i=1}^{c} N_i(m_i - m)(m_i - m)^T \quad (52)$$

Where m = overall mean, mi= sample mean and Ni= sample size of the respective class

- Compute the eigenvectors and eigenvalues

$$A = S_W^{-1}S_B \quad (53)$$

$$AV = \lambda V \quad (54)$$

Where λ = eigenvalue and V = eigen vector of same eigen value

- Project the data onto the new subspace

$$Y = X \times W \quad (55)$$

Where X = n dimensions matrix representation the n samples and Y = Transmitted n × k dimensional samples in the new subspace

We have pre-processed the raw sensor response dataset in this experiment using the Standardized Linear Discriminant Analysis (SLDA) space transformation method. The SLDA transforms the raw data into respective space transformation domains [49], [98] – [100]. The first three principal components contain 98.88% information, using which we have obtained the 3-D scatter plot, which shows well-separated clusters with good inter-cluster separation, as shown in Figure 5.8. It can be observed that the clusters belonging to the six classes of VOCs/gases/odors are shown clearly in the SLDA-transformed dataset.

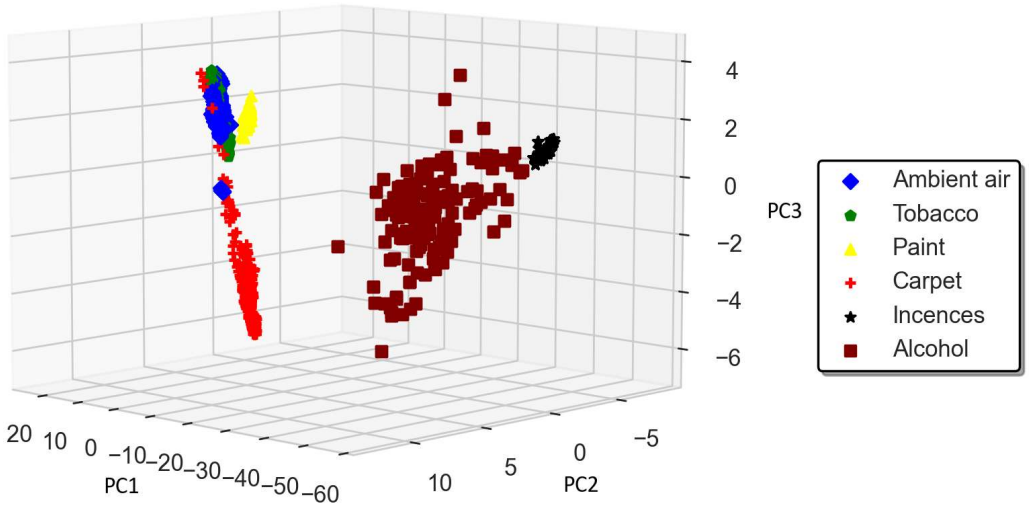


Figure 5.8. 3D Scatter Plot of SLDA transformed dataset.

5.2.4.2. Design of the Classifiers

The raw sensor responses were first transformed into the SLDA domain during pre-processing. SLDA is an effective method for feature enhancement. We have used all seven principal components (PCs) to train and test the classifier used in the IGSS without any information loss. In the SLDA transformation space, better-performing classifiers can now be designed. The transformed dataset consists of 1800 samples, with each vector consisting of seven elements. The transformed dataset was then segregated into two parts, i.e., the training and testing datasets, consisting of 1770 and 30 samples in the SLDA transformed domain, respectively. Further, we have used four popular classifiers for the experimentation, viz., AdaBoost, XGBoost, RF, and MLP. AdaBoost and XGBoost are good for boosting, RF for handling large datasets, and MLP for handling complex non-linear relationships. The process of the classifier is shown in Figure 5.9.

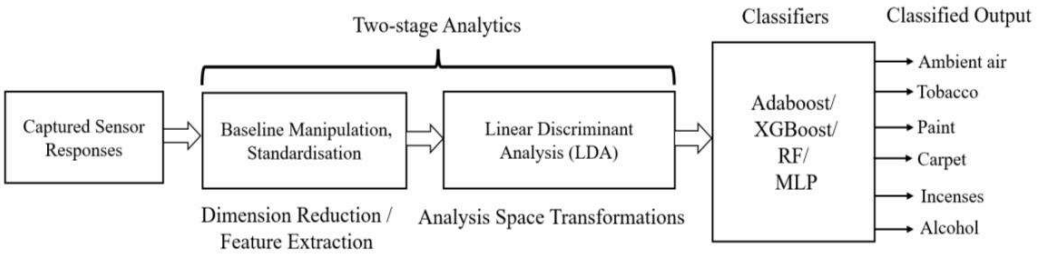


Figure 5.9. Block Schematic Diagram of Proposed Classifiers.

The AdaBoost algorithm is a boosting algorithm that combines multiple weak classifiers to form a robust classifier. The design of an AdaBoost classifier involves selecting several hyperparameters, such as the number of estimators, the learning rate, and the random state, to optimize the model's performance. The number of estimators is the number of weak classifiers to combine to form a robust classifier. The learning rate is a hyperparameter that controls the contribution of each weak classifier to the final model. The random state is a seed value that ensures the reproducibility of the results. In this case, we set the random state to 1. The AdaBoost Classifier has been shown in Figure 5.10.

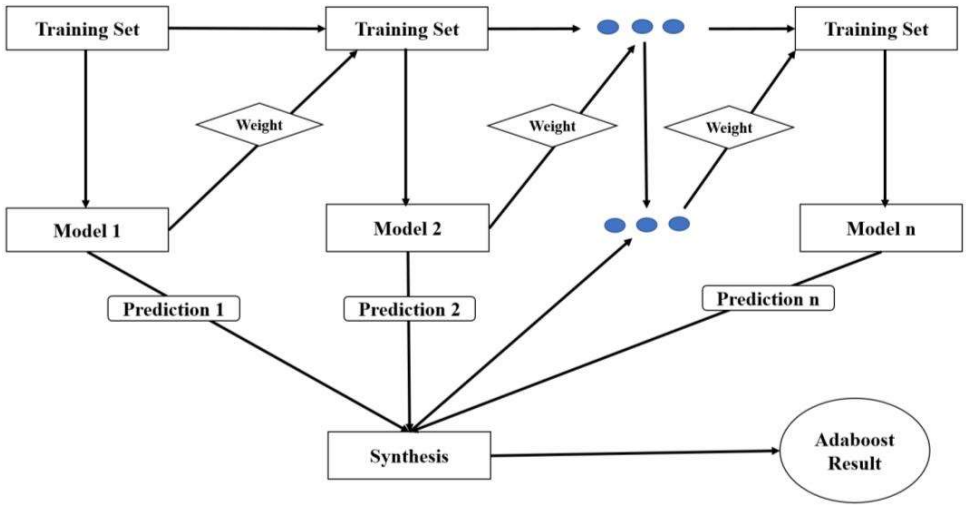


Figure 5.10. AdaBoost Classification Model.

The design of an XGBoost classifier involves selecting several hyperparameters, such as the learning rate, number of estimators, maximum depth of the tree, minimum child weight, gamma, regularization alpha, number of threads, and cross-validation. The learning rate is a hyperparameter that controls the contribution of each tree to the final model. The number of estimators is the number of decision trees to create in the ensemble. The XGBoost Classification Model has been shown in Figure 5.11.

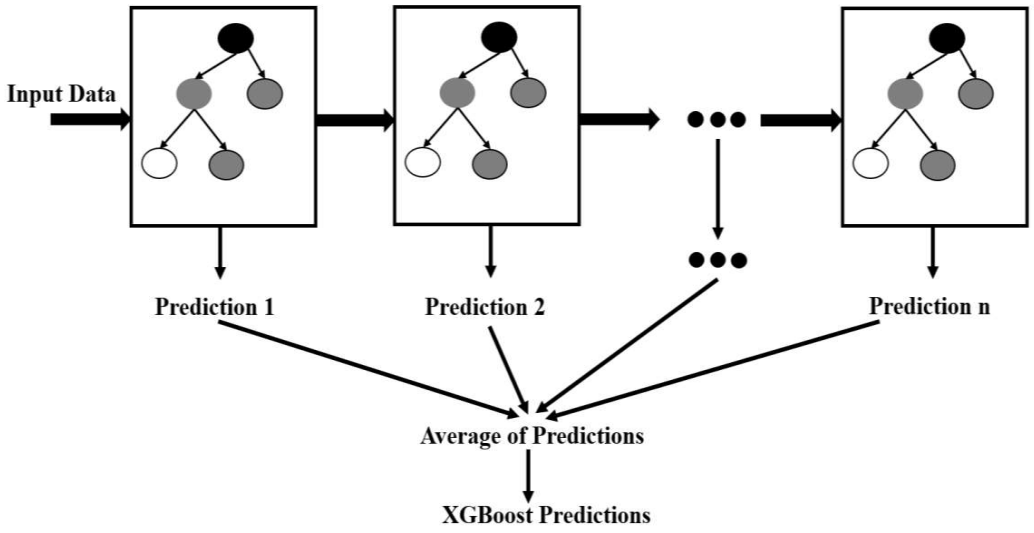


Figure 5.11. XGBoost Classification Model.

Random Forest builds many decision trees and combines their outputs to make predictions. The design of a Random Forest classifier involves selecting several hyperparameters, such as the number of estimators, criterion, learning rate, random state, and cross-validation. The number of estimators is the number of decision trees to create in the forest. The criterion is a hyperparameter that determines the function used to measure the quality of a split. The random state is a hyperparameter that controls the randomness of the algorithm. The Random Forest Classification Model has been shown in Figure 5.12.

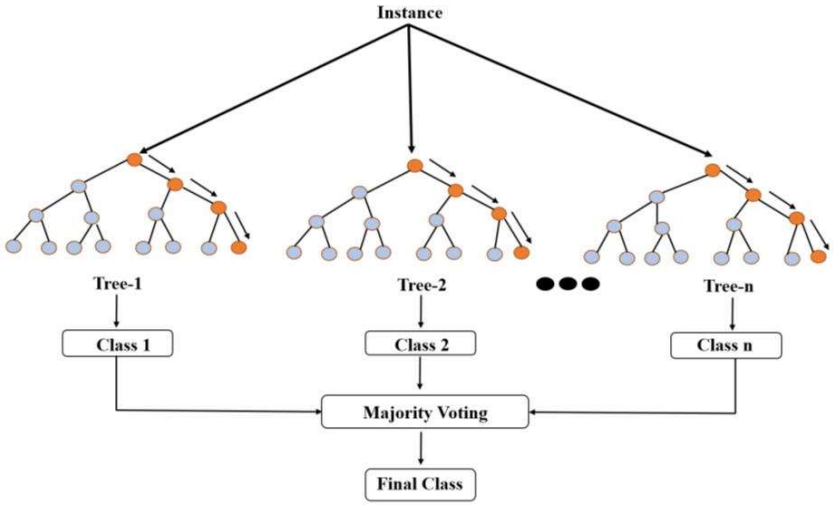


Figure 5.12. Random Forest Classification Model.

This MLP classifier has a relatively small number of hidden layers and a moderate batch size, which can make it computationally efficient and suitable for medium-sized datasets. Using the ReLU activation function and adaptive learning rates can help improve the model's performance, while using cross-validation can help reduce overfitting and improve generalization. Reduce overfitting and improve generalization. The ReLU activation function is defined as f(x) = max (0, x), which means that the output of the activation function is the maximum between the input value and zero. Adam adapts the learning rate of each weight during training based on the historical gradient information, which can lead to faster convergence and better performance. Adaptive learning rates can lead to faster convergence and better performance than fixed learning rates. An epoch is defined as one passing through the entire training dataset. The number of iterations required for the model to converge may be less than 100. The proposed MLP Classification Model has been shown in Figure 5.13.

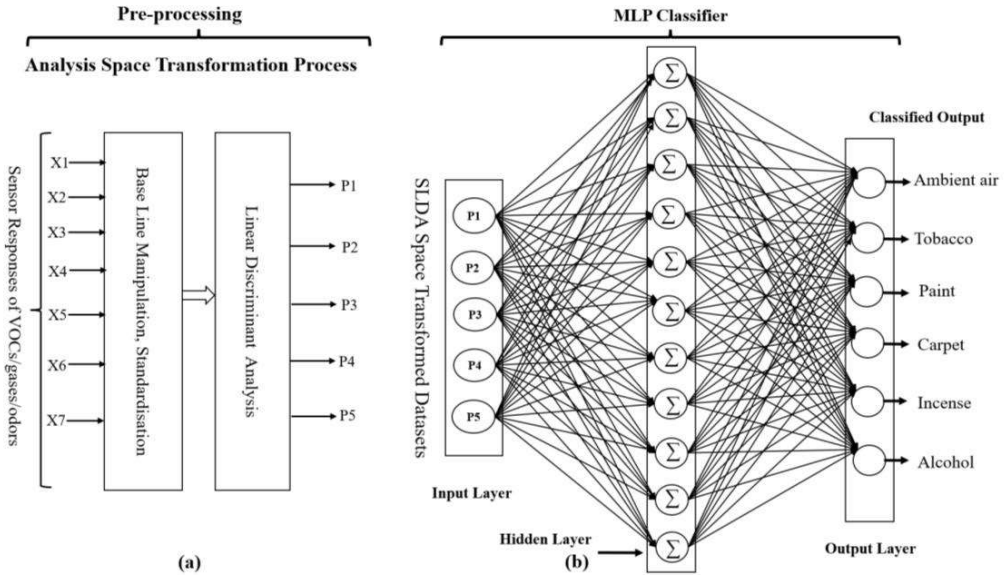


Figure 5.13. MLP Classification Model.

The performance parameters taken while designing the classifiers: AdaBoost, XGBoost, RF and MLP have been presented in Table 5.4.

Table 5.4. Parameters tunning of AdaBoost, XGBoost, RF and MLP Classifier.

Classifiers	Parameters
AdaBoost	N_estimators:0.5, learning rate: 50, random_state:1, cv=5
XGBoost	Learning_rate:0.1, n_estimators:1000, max_depth:4, min_child_weight:6, gamma=0, reg_alpha:0.005, nthread:4, cv=5
RF	N_estimators:100, criterion: Gini, random_state:1, cv=5
MLP	Hidden layer sizes =11, activation function: ReLU, solver: adam, batch size:100, learning rate: adaptive, max iteration: 100, cv=5

5.3. Results and Discussion

AdaBoost, XGBoost, RF, and MLP classifiers have been designed and tested in this experiment. The MLP trained using SLDA transformed data achieved the 'all correct' classification accuracy over the 30 test samples taken from the six classes of VOCs, gases, and smokes. The proposed IGSS has been aimed at being portable, easy to use, and affordable in real-time scenarios.

5.3.1. The LoRa Network Link Performance

In this experiment, we tested the LoRa channel performance in the indoor ambience with several wall and roof obstructions. The signal strength for the LoRa networking link is measured as the received signal strength indicator (RSSI) in decibels (dB). It observed that the RSSI is about -60 to -80 dB when the transmitted signal becomes erratic. However, the received signal exhibits errors as the RSSI decreases with increasing distance. The signal is completely lost when the RSSI becomes -164dB. Accordingly, the maximum experimental range achieved for the LoRa network link for indoor operations for N-IGSS has been 590 meters.

5.3.2. Performance of the Proposed N-IGSS for air-borne Pollution Hazard Detection

In this experiment, we have trained and tested four different classifiers. Among these, the AdaBoost, XGBoost and random forest (RF) classifiers have achieved 96.67%, 96.67% and 96.67% classification accuracy, respectively, for the 30 unknown

test samples. On the other hand, the MLP classifier achieved 100% classification accuracy for all 30 test samples, as considered. While classifying the considered VOCs, gases and smokes, we have compared the prediction error for each test sample. The error ranges between 1.42 × 10-4 to 1.24 × 10-2, with an average mean-squared error (MSE) of 2.66 × 10-3 for 30 unknown test samples. The classification performance of the best-trained classifier is shown in Table 5.5 and Figure 5.13. Further, the classification performance of the MLP classifier, trained and tested in the SLDA domain using 30 unknown test samples of the considered VOCs, gases, and smokes, has been shown in Figure 5.14.

Table 5.5. Performance of SLDA transformed data of different Classifiers.

Classifiers	Accuracy (%)
AdaBoost	96.67
XGBoost	96.67
RF	96.67
MLP	100

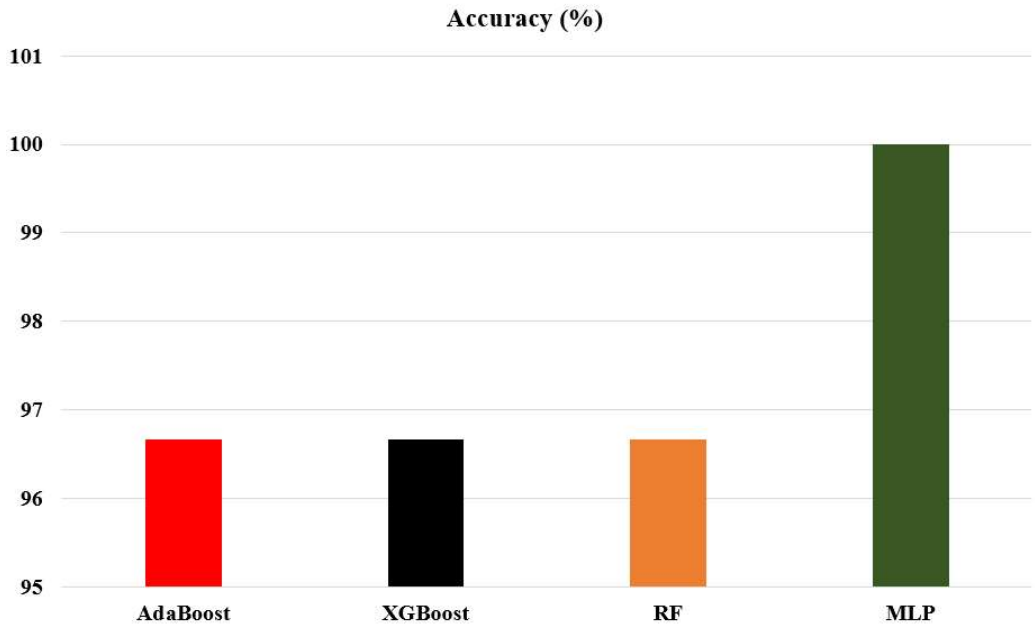


Figure 5.14. Performance of AdaBoost, XGBoost, RF and MLP Classifier.

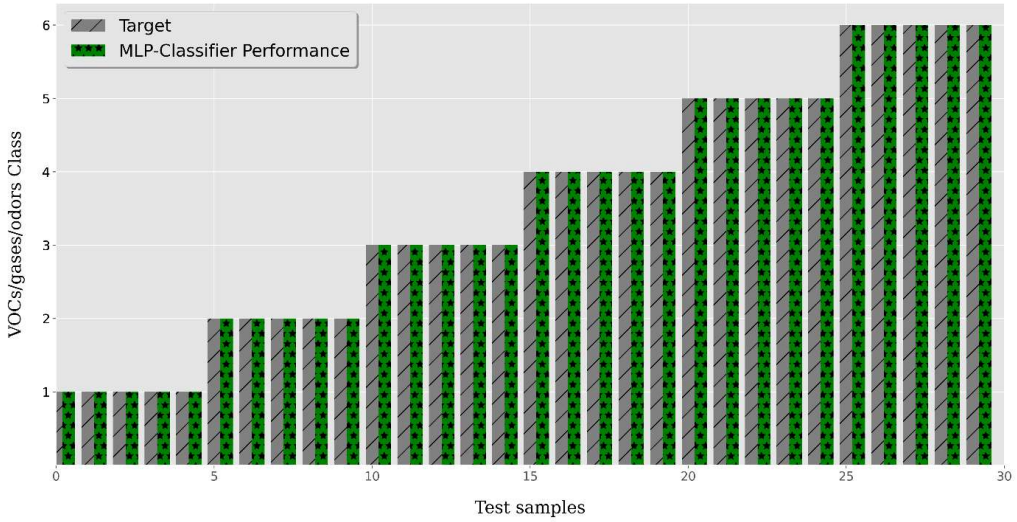


Figure 5.15. Performance of MLP Classifier.

For further clarity, the confusion matrix of the classification performance of the MLP classifier has also been shown in Figure 5.16.

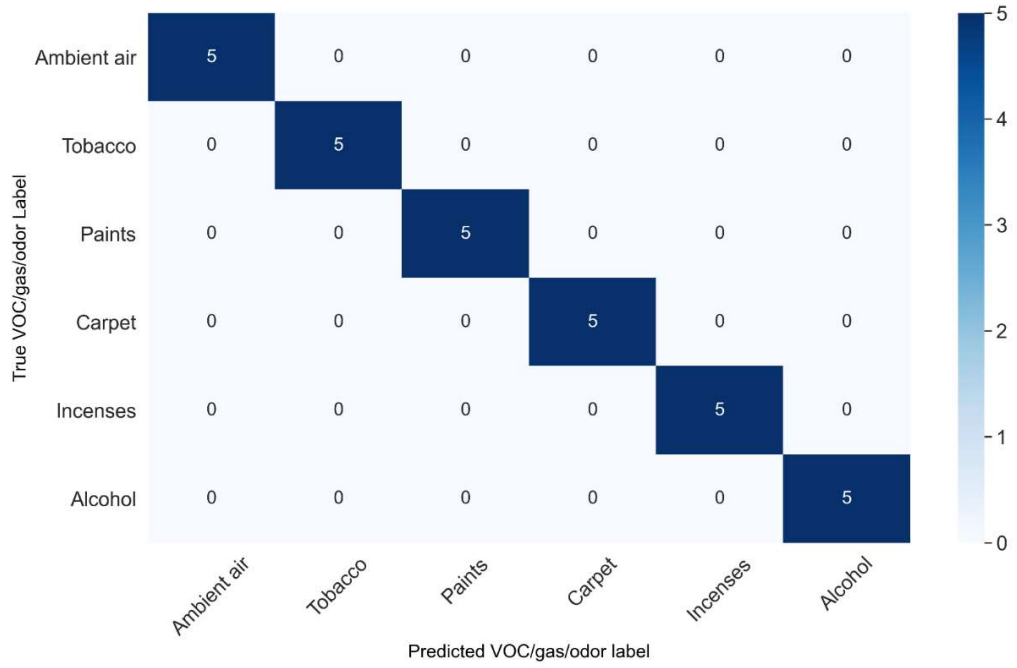


Figure 5.16. Confusion Matrix of MLP Classifier.

The classification accuracy has been calculated using popular error matrices, i.e., Mean Squared Error (MSE) and Mean Absolute Error (MAE). The MLP classifier has been the best-performing classifier. The difference between the actual and predicted values for the six classes of VOCs, gases, and smokes has been evaluated using MSE and MAE as the performance parameters, as shown in Table 5.6. For the sake of further clarity, the confusion matrix of the classification performance of the MLP classifier has been shown in Figure 5.16, which shows 'all correct' classification of the considered 30 unknown samples taken from the testing dataset and not used for training the classifier models in the SLDA transformation domain.

Table 5.6. Performance of MSE and MAE of LDA and proposed method

Class	MSE		MAE	
	LDA	Proposed Method	LDA	Proposed Method
Ambient air	4.62×10^{-4}	1.42×10^{-4}	1.003×10^{-2}	7.53×10^{-2}
Tobacco	5.38×10^{-4}	4.51×10^{-4}	1.74×10^{-2}	1.41×10^{-2}
Paints	3.39×10^{-4}	5.05×10^{-4}	1.44×10^{-2}	1.73×10^{-2}
Carpet	6.53×10^{-3}	1.53×10^{-3}	3.02×10^{-2}	2.29×10^{-2}
Incense	1.02×10^{-3}	8.94×10^{-4}	2.39×10^{-2}	2.21×10^{-2}
Alcohol	1.65×10^{-2}	1.24×10^{-2}	8.82×10^{-2}	8.37×10^{-2}