International Journal of Computing and Digital Systems
2025, VOL. 17, NO. 1, 1–17

http://dx.doi.org/10.12785/ijcds/1571033021

Real-Time Gas Monitoring and Anomaly Detection in Petroleum Industry Using IoT and Machine Learning

Sonali Antad[1], Virat Giri[2], Bhushan Bachewar[1], Shreya Barsude[1], Aayush Gadiya[1] and Harsh Badagandi[1]

[1]*Computer Engineering, Vishwakarma Institute of Technology, Pune, India*
[2]*Computer Science and Engineering, Sanjay Ghodawat Institute, Kolhapur, India*

Received 25 May 2024, Revised 7 December 2024, Accepted 4 January 2025

Abstract: Significant safety challenges are encountered in the petroleum industry due to the presence of toxic gases, which pose serious health risks and potential hazards. To address this issue, we have developed an innovative solution with the help of the Internet of Things (IoT) and Machine Learning. This solution includes an IoT-driven bot equipped with ESP8266 and Raspberry Pi Pico W microcontrollers, to monitor and detect toxic gases in real-time. The bot integrates sensors, MQ-2, MQ-3, and MQ-135, which detect harmful gases such as methane, propane, alcohol, and ammonia. The ESP8266 has Wi-Fi capabilities, allowing the bot to connect to the internet and transmit data, while the Raspberry Pi Pico W handles sensor data processing. Controlled via the Blynk IoT application, this setup enables remote operation and real-time monitoring. As the bot navigates through the petroleum facility, it collects gas concentration data, which is sent to a Google Spreadsheet for storage and analysis. This data is analyzed using machine learning techniques, specifically Isolation Forest and One-Class Support Vector Machine (One Class SVM), which assist in identifying anomalies. These algorithms analyze the data to identify unusual patterns or spikes in gas concentrations, indicating potential leaks or hazardous conditions. Upon detecting anomalies, the system triggers alerts to notify personnel, enabling prompt action to prevent risks. This system enhances safety by providing continuous monitoring and demonstrates the potential of IoT and machine learning to revolutionize workplace safety in high-risk environments, significantly improving safety protocols and protecting both workers and the environment, particularly in the petroleum and oil industries.

Keywords: Petroleum Industry, Internet of Things(IoT), Anomaly Detection, Isolation Forest, One-Class SVM, Petroleum Industry, ESP8266, Raspberry Pi Pico, MQ Sensors

1. INTRODUCTION

The petroleum industry, which has a vital role in the global energy supply, poses significant dangers to workers due to the presence of hazardous gases released during refining and drilling processes. These gases, whether emitted from damaged equipment or as byproducts of operations, pose severe health and safety risks, causing respiratory irritation, skin damage, and long-term health complications like chronic respiratory issues. In extreme cases, exposure to these gases can lead to life-threatening conditions, including asphyxiation. Uncontrolled gas releases can have catastrophic incidents such as fires, explosions, and widespread equipment damage. These challenges underscore the immediate need for real-time gas monitoring systems to protect personnel and minimize risks.

There have been incidents such as the Texas City Refinery Explosion, Mexico City Gas Explosion, and many more which have affected the lives of many people. Some improvements have been made in the industries, such as adding sensors to detect gases, but they are static and do not predict if the machinery is in good condition or not. Additionally, workers need to physically go to the site to check if there is any machine defective, which is very risky.

Current safety measures, such as personal protective equipment (PPE) and stationary monitoring systems, are essential but insufficient on their own due to their static nature and the dynamic environments in industrial sites. Despite worker training and regular inspections, the risk of gas exposure remains constant.

This highlights the need for a system to monitor gas levels throughout the industry, detect any anomalies, and quickly alert industry workers. This will reduce the need for human intervention in risky environments and enable timely preventive action to reduce future risks.

The primary objective of this research is to develop an

E-mail address: sonali.antad@vit.edu, virat.giri@gmail.com, bhushan.bachewar221@vit.edu

innovative, IoT-enabled robotic system that addresses critical safety challenges in the petroleum industry by providing dynamic gas monitoring, real-time anomaly detection, and proactive alerts. Unlike traditional static monitoring systems, which are limited in their ability to adapt to the dynamic and hazardous nature of industrial environments, this system offers a mobile, sensor-equipped solution capable of navigating the entire facility. By leveraging IoT technology and advanced machine learning algorithms such as the One-Class Support Vector Machine (SVM), the system ensures continuous, comprehensive monitoring and enhances safety measures.

This study introduces an integrated approach that combines real-time data acquisition from gas sensors with machine learning-based anomaly detection to identify irregularities in gas levels. The system's ability to detect and communicate potential hazards promptly reduces reliance on manual inspections in high-risk environments, thereby minimizing exposure to dangerous conditions for workers. Additionally, the integration of the Blynk IoT application allows remote control and monitoring, ensuring that even hard-to-reach or hazardous areas are effectively covered.

Our contributions extend beyond the prototype stage, offering insights into scalable and cost-effective implementations of similar systems in industrial settings. By demonstrating the effectiveness of using widely available sensors like MQ-2, MQ-3, and MQ-135 for initial testing, the research highlights the potential for adopting advanced sensors tailored to specific industry needs. Furthermore, the comparative analysis of machine learning models employed in this study provides a robust framework for selecting suitable algorithms for real-time anomaly detection in dynamic environments.

In response to these challenges, our project introduces a solution that uses the Internet of Things (IoT) to provide safety monitoring in the petroleum industry. We have developed an IoT-driven robotic system equipped with sensors, including MQ-2, MQ-3, and MQ-135 sensors. Unlike static sensors, our mobile solution offers dynamic monitoring by navigating the entire industrial environment and continuously scanning for gas leaks and fluctuations in gas levels.

The MQ-2 sensor is capable of detecting smoke, butane, propane, methane, alcohol, hydrogen, and liquefied natural gas (LNG). The MQ-3 sensor can identify the presence of benzene, methane (CH_4), hexane, and carbon monoxide (CO). Additionally, the MQ-135 sensor detects ammonia (NH_3), benzene (C_6H_6), carbon dioxide (CO_2), and other harmful gases and smoke. These sensors transmit real-time data to Google Spreadsheets, ensuring continuous and comprehensive environmental monitoring.

Table I gives information about the gases detected by the MQ sensors and their effect if released in the petroleum industry.

We have employed these MQ sensors since we have just made a prototype, but for industry deployment, we can employ electrochemical sensors, catalytic bead sensors, infrared gas sensors, etc. These are generalized sensors, but for each industry according to its work and emission of gases, we have to deploy specific sensors. Overall anomaly detection and its communication to the workers remain the same; just sensors can be deployed differently according to each industry's need.

To enhance the effectiveness of our monitoring system, we employ machine learning models. The models with which we worked are Isolation Forest and One-Class Support Vector Machine. After testing and comparing both models, we finally employed One-Class Support Vector Machine (SVM) algorithms for real-time anomaly detection. The detailed analysis of the machine learning model and its comparison will be explained in the paper along with other details. By continuously analyzing the collected data, these models can identify irregularities in gas levels, providing timely alerts to operators. This enables proactive risk mitigation measures, significantly reducing the likelihood of accidents and enhancing overall safety.

The integration of the IoT Blynk application allows the operators to remotely control the robotic system, directing it to any desired location for targeted monitoring. This remote control capability ensures that even the most hard-to-reach areas of the industrial site are effectively monitored.

Overall, our project demonstrates the integration of IoT technology, gas sensors, and machine learning algorithms to address critical safety challenges in high-risk industrial environments. By providing real-time, dynamic monitoring and timely alerts, our system provides situational awareness, reduces risks to personnel, and promotes safer and more efficient operations in the petroleum industry. Our system provides safety protocols, particularly for workers inspecting turbomachines in hazardous environments.

The environment in petroleum industries has the risk of gas-related incidents and a constant threat to worker safety. Traditional gas monitoring approaches often rely on periodic inspections and manual data analysis, which may not be sufficient to detect potential hazards in real-time. However, by integrating dynamic gas monitoring capabilities with anomaly detection algorithms, our solution enables continuous and proactive monitoring of gas concentrations, ensuring timely identification of anomalies or hazardous conditions.

By using machine learning, the system can analyze complex patterns and trends in gas data, enabling it to identify subtle deviations indicative of the hazards. This approach to anomaly detection allows operators to take preemptive measures to prevent risks and ensure the well-being of workers operating in the environments.

Our bot differs from the existing system in terms of

TABLE I. Gas And Sensor Information Table

Sensor	Gases Detected	Effect in Petroleum Industry
MQ2	Methane (CH4)	Explosive; poses fire and explosion hazards, leading to facility damage, injuries, and loss of life.
	Liquefied Petroleum Gas (LPG)	Highly flammable; increases the risk of fires and explosions, causing equipment damage and endangering personnel.
	Carbon Monoxide	Poisoning; can lead to asphyxiation and impaired cognitive function.
MQ135	Benzene	Highly flammable; increases the risk of fires and explosions, causing equipment damage and endangering personnel.
	Toluene	Highly flammable; poses fire and explosion hazards, leading to facility damage, injuries, and loss of life.
MQ3	Benzene	Carcinogenic; poses long-term health risks to workers, including increased cancer risk and respiratory illnesses.
	Carbon Monoxide (CO)	Poisoning; can lead to asphyxiation and impaired cognitive function.

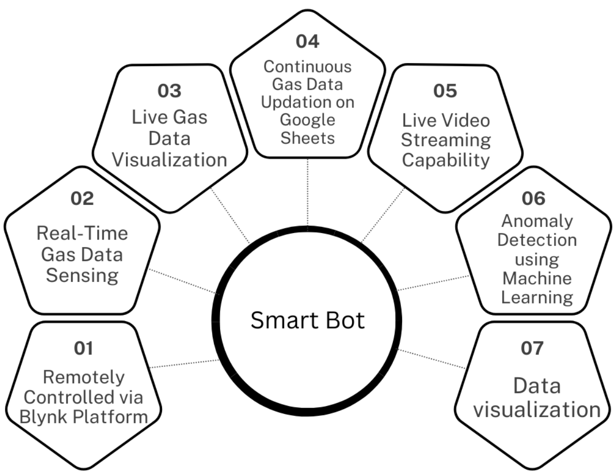

Figure 1. Features of the project

controlling IoT from anywhere and its visual presentation and description of gas in every area. Also, live video streaming, anomaly detection, and a low-cost setup make deploying it easier and cheaper.

Figure 1 showcases the features of the bot, illustrating its capabilities and functionalities in facilitating dynamic gas monitoring and anomaly detection.

The bot provides a solution for enhancing safety protocols in industrial settings through a combination of sensors, data analysis algorithms, and real-time monitoring capabilities. From detecting gas leaks to identifying anomalies in gas concentrations, the bot serves as a crucial tool in safeguarding the health and safety of workers and preventing potential accidents or incidents.

2. LITERATURE SURVEY

The literature survey reveals a growing body of research focused on IoT-based gas monitoring systems and machine learning-driven anomaly detection techniques. Previous studies have extensively investigated the development and implementation of IoT technologies for gas monitoring applications across industries. Below are a few notable examples.

The paper [1] introduces an IoT-based environmental monitoring rover designed to address the limitations of traditional monitoring methods. The rover incorporates a suite of sensors, including DS18B20 temperature sensors, DHT11 humidity sensors, MQ-135 air quality sensors, and HC-SR04 ultrasonic sensors for obstacle detection. It transmits data wirelessly to a mobile application, enabling real-time monitoring and control. The compact and cost-effective design makes it suitable for deployment in confined spaces and hazardous environments. 1 However, the current battery life is limited, and the rover lacks the capability to detect ionizing radiation. Future improvements could include autonomous movement capabilities and the integration of additional sensors for comprehensive environmental monitoring.

The paper[2] focuses on developing an IoT-based gas leakage detection and alert system leveraging the MQ2 gas sensor and the ESP8266 microcontroller. The system is designed to detect hazardous gases like methane and carbon monoxide and immediately notify users through real-time alerts via the Blynk application on their smartphones. To enhance safety, the system includes features such as a buzzer and LED indicators for local alarms, along with automatic activation of an exhaust fan to mitigate the detected gas concentration. The study emphasizes a cost-effective approach, ensuring that the solution is accessible for both domestic and industrial environments. Furthermore, it demonstrates the practical use of IoT in real-time monitoring, offering an efficient response mechanism to potential gas leakage incidents, ultimately reducing risks and improving safety

The research paper [3] proposes an IoT-based wireless sensor network system to address the pressing issue of pipeline vandalism and leakage in Nigeria's oil and gas industry. The system employs sensors such as pressure transducers, thermocouples, accelerometers, and gas detectors to monitor critical parameters. Data is transmitted wirelessly to a central server for analysis and decision-making, enabling timely responses to potential incidents. Machine learning algorithms like Support Vector Machines (SVM), Random Forest, and XGBoost can be integrated to enhance anomaly detection and predictive capabilities. These algorithms have shown promising results in similar applications, with accuracies ranging from 85% to 95%, depending on the specific dataset and model configuration. The system's effectiveness also depends on factors such as sensor selection, deployment strategies, communication protocols, and data analytics techniques. While the proposed system holds promise, further research is needed to address challenges such as battery life, sensor reliability, and security vulnerabilities to ensure its long-term viability.

Paper [4] proposes an IoT-based gas and smoke detection system utilizing an ESP8266 NodeMCU microcontroller and an MQ-2 gas sensor. The MQ-2 sensor, sensitive to various gases including LPG, propane, and butane, measures gas concentration by detecting changes in electrical resistance. When a gas leak is detected, the sensor's resistance increases, triggering an alarm. The NodeMCU processes the sensor data and sends alerts via SMS and a mobile app. The system's effectiveness depends on factors like sensor calibration, threshold setting, and communication protocols. While the paper provides a basic framework, further research can explore advanced sensor technologies, machine learning algorithms for pattern recognition, and integration with smart home systems for comprehensive safety solutions.

The research paper [5] presents an IoT-based gas detection system utilizing an ESP32 microcontroller and a combination of gas sensors (MQ-2, MQ-6, and MQ-9) to detect combustible gases like propane, butane, and carbon monoxide. The system operates by continuously monitoring the surrounding air and analyzing the sensor data. When gas concentrations exceed predefined thresholds, the system triggers multiple alerts: an audible alarm, visual alerts on an LCD display, and SMS notifications sent via the Blynk platform. Additionally, the system transmits real-time gas concentration data to a Thinkspeak server for remote monitoring and analysis. The effectiveness of the system relies on factors such as sensor calibration, threshold setting, and communication protocols. Future research could explore advanced sensor technologies, machine learning algorithms for enhanced pattern recognition, and integration with smart home systems for comprehensive safety solutions.

The research paper [6] delves into applying machine-learning techniques for detecting leaks in oil and gas pipelines. The authors propose using five machine learning algorithms—SVM, KNN, Random Forest, Gradient Boosting, and Decision Tree—to analyze operational data such as temperature, pressure, and flow rate. The dataset used was preprocessed to ensure data quality and consistency. After training and tuning the models, SVM emerged as the top performer, achieving an accuracy of 97.4%. This indicates that the model can effectively distinguish between normal and anomalous pipeline conditions, potentially leading to early detection of leaks and preventing environmental disasters. The study highlights the potential of machine learning in enhancing pipeline safety and reliability.

The paper [7]conducted a comparative study on anomaly detection in oil-producing wells using one-class classifiers in a multivariate time series dataset. Their research evaluated the performance of various classifiers for detecting anomalies in complex, high-dimensional data specific to the petroleum industry. The study highlighted the effectiveness of machine learning models in identifying operational anomalies, emphasizing the critical need for accurate, real-time detection methods to enhance operational safety and efficiency in oil production environments.

Paper [8] proposes a novel approach to anomaly detection in oil wells using one-class classifiers. The study focuses on detecting faults in naturally flowing offshore oil and subsea gas-producing wells, utilizing the publicly available 3W dataset.The authors compared the performance of various classifiers, including Isolation Forest, One-class Support Vector Machine (OCSVM), Local Outlier Factor (LOF), Elliptical Envelope, and Autoencoder with feed-forward and LSTM architectures. The LOF classifier consistently outperformed other techniques, demonstrating its effectiveness in identifying anomalies in both simulated and real-world scenarios. The research highlights the importance of feature extraction and the use of statistical tests to validate the results. The findings contribute to improving the reliability and efficiency of oil well operations by enabling early detection and prevention of anomalies, ultimately leading to significant cost savings and reduced environmental impact.

The paper [9] the development of an IoT-based LPG gas detection and warning system to mitigate the risks associated with LPG leaks in domestic and industrial settings.

Similar research has focused on using microcontroller-based devices with sensors like the MQ-6 to detect leaks, combined with GPS for location tracking and Arduino for automating alerts. These systems transmit real-time warnings via smartphones, emails, and alarms, enabling timely responses. Previous studies highlight the efficiency, cost-effectiveness, and automation of IoT-driven gas detectors, which offer reliable, real-time monitoring without the need for human intervention.

The paper [10] reviews mobile detection and alarming systems for hazardous gases and volatile chemicals, focusing on their application in laboratories and industrial environments. Recognizing the critical need for early detection to ensure safety, the study examines papers from January 2010 to August 2021, selecting 42 out of 236 papers identified from databases like Web of Science, IEEE, and Scopus. The review discusses various gas detection technologies, including sensor types (catalytic bead, MEMS, MOX), sensor specifications (response time, precision), processor types (microcontrollers, PLCs), and communication technologies (Bluetooth, Wi-Fi, ZigBee, LoRa). The goal is to highlight advancements in sensors, processors, and communication for improving hazardous gas detection and alarm systems.

[11] reports the development of autonomous robots equipped with gas sensors for detecting hazardous gases in real-time. The robots, often used in disaster management, are equipped with a variety of sensors to detect gases like carbon dioxide, liquefied petroleum gas, and alcohol vapors, while also identifying ambient gases. Navigation systems with collision avoidance help the robots to safely traverse uneven terrains. A GPS module maps the location of detected gases, and information is transmitted remotely for further analysis.

The paper [12] presents a mobile robot car for toxic gas detection and plant security. Unlike stationary systems, the robot uses a toxic gas detector linked to a microcontroller, which sends real-time gas concentration data to a server. Two motors control both the camera's rotation and the robot's movement, enhancing patrol coverage. A wireless module and movable power unit extend its operation, providing efficient gas monitoring and protecting workers from exposure.

[13] introduces a graphene-based nanosensor functionalized with copper phthalocyanine for the detection of ammonia and phosphine at room temperature. The sensor, combined with machine learning, demonstrated high sensitivity and specificity, especially for ammonia at ultralow concentrations (100 ppb, 100% accuracy). The system operates efficiently without the need for high temperatures and is capable of detecting multiple gases, offering a low-power, highly selective solution for industrial gas monitoring.

[14] proposes a bio-inspired anomaly detection method for gas sensors based on spiking neural networks. By detecting rapid changes rather than relying on magnitude thresholds, the method minimizes false positives caused by sensor drift. Tested with chemicals like surgical spirits and isobutanol, the approach successfully identifies gas anomalies. Implemented on FPGA hardware, the method is compact, energy-efficient, and suitable for low-cost applications, such as safety in drones and ground robots for hazardous scene detection.

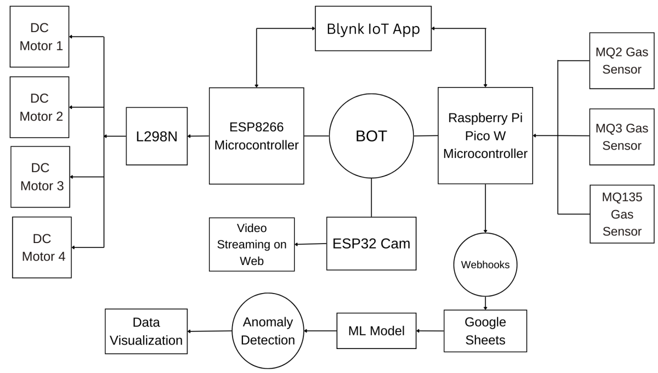

Figure 2. System Architecture of the project

[15] proposes an automatic system for LPG leakage detection and booking. It uses an MQ-5 gas sensor for high-sensitivity detection of LPG and natural gas, with real-time alerts sent via SMS and a GSM module. The system also triggers an alarm if gas levels exceed safe limits. Additionally, a weight sensor monitors LPG levels in the cylinder, offering accurate detection and quick response, enhancing existing industrial safety models.

3. METHODOLOGY

Figure 2 shows the system architecture of the project. The system integrates several components, including microcontrollers, gas sensors, data transmission modules, and machine learning algorithms. The IoT bot navigates through the facility, collecting real-time data, and operates autonomously or under remote control using the Blynk IoT application.

We have used two microcontrollers, the ESP8266 and the Raspberry Pi Pico W. The ESP8266 is responsible for providing Wi-Fi connectivity and enabling communication between the bot and the cloud. The Raspberry Pi Pico W manages the collection of sensor data from the attached gas sensors, which include the MQ-2, MQ-3, and MQ-135 sensors. These sensors detect the presence and concentration of gases such as methane, propane, alcohol, and ammonia. It transmits sensor data in real-time to a Google Spreadsheet, for data storage and further processing. Together, these microcontrollers enable the bot to continuously monitor the environment and ensure real-time data processing and transmission.

The gas sensors connected to the Raspberry Pi Pico W are the key components for real-time data collection. The sensors provide raw analog signals that represent the concentration of various gases in the environment. The Raspberry Pi Pico W microcontroller processes these signals, converting the analog readings into digital data for transmission. This data includes gas concentration levels that are calibrated based on specific threshold values for each sensor.

For use in the petroleum industry, we can use industry-specific sensors on Raspberry Pi Pico W. The MQ sensors are currently used on the robot for prototype and development purposes. As the sensors vary from industry to industry and machine to machine, we need to identify all the gases that may be emitted in a specific industry before deploying and then using the appropriate gas-specific sensors. For anomaly detection, the collected gas data is analyzed using machine learning algorithms, particularly Isolation Forest and One-Class SVM. These algorithms are selected for their ability to effectively identify anomalies in high-dimensional data, such as unexpected spikes in gas concentrations that could signal a leak or a hazardous situation.

Isolation Forest is an unsupervised algorithm designed to identify anomalies by randomly selecting a feature, such as gas concentration, and splitting the data at random thresholds. Anomalies are more easily isolated compared to normal observations, which makes them stand out. The model is trained using historical data from the sensors to understand typical gas concentration patterns. When real-time data strays from these established patterns, the algorithm marks it as an anomaly.

The One-Class SVM method is an unsupervised anomaly detection algorithm that identifies a region in feature space (gas concentrations) where most of the data points are located. Any points that fall outside this region are considered anomalies. The One-Class SVM is trained on normal gas concentration data, and it uses kernel functions to classify new sensor readings as either normal or anomalous.

The algorithms are trained using historical sensor data collected over some time. Key hyperparameters, such as the contamination rate for the Isolation Forest are kept at 0.05 and the kernel type for One-Class SVM which we used were polynomial and radial basis function (RBF), where RBF gave more efficient results. During operation, both algorithms continuously analyze incoming sensor data to detect anomalies, such as sudden increases in toxic gas levels. When an anomaly is identified, the system triggers alerts, notifying personnel of potential risks. For training the gas detection model, 10,000 data entries were collected from the MQ sensors (MQ-2, MQ-3, MQ-135) deployed in a normal environment. These data entries were gathered using the Raspberry Pi and sent directly to Google Sheets for storage and further processing.

The first step in preprocessing the dataset was to clean it by removing rows with missing values. The number of miss-

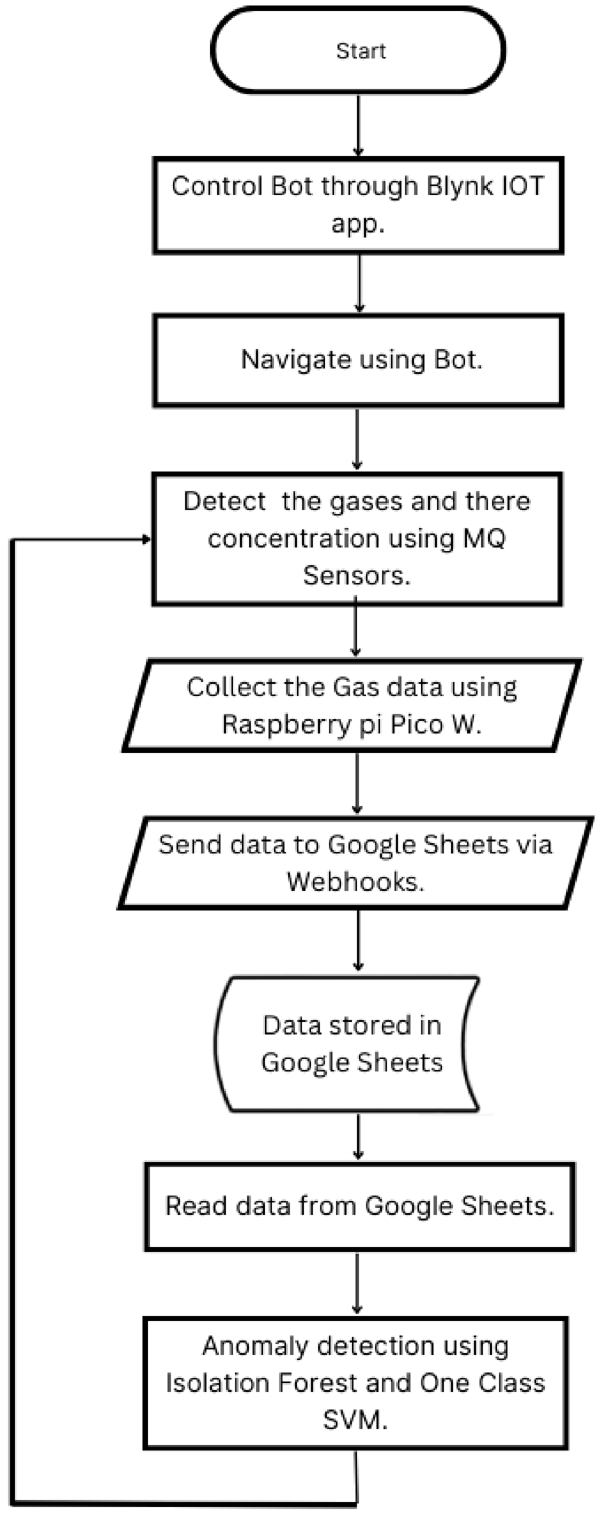

Figure 3. Flowchart of the proposed system

ing entries was minimal, so the decision was made to delete those rows rather than imputing values. After cleaning, we standardized the data using standardization, ensuring that all the sensor readings were on the same scale. Standardization helped to eliminate biases in the model due to varying scales in gas concentration values. After preprocessing, we input the data into machine learning models, specifically Isolation Forest and One-Class SVM, for training. These models were designed to identify normal gas concentration patterns, which helps in detecting anomalies during real-time operations.

Specific areas in the same environment were introduced for testing the models with smoke containing carbon and alcohol. The IoT bot was navigated through these spots to evaluate whether it could successfully detect anomalies, indicating abnormal gas levels. The system effectively identified these anomalies, confirming that the bot and the machine learning model could detect hazardous gases, validating its accuracy in real-world conditions.

In One-Class SVM, anomalies are identified by establishing a decision boundary that distinguishes the majority of normal data from the origin in a high-dimensional space. While training, the algorithm employs a kernel function to transform the data into this space and determines a boundary that includes most normal data points while leaving outliers aside.New data points are classified based on their position relative to this boundary: points inside are considered normal, while those outside are classified as anomalies. The model's effectiveness depends on parameters such as the choice of kernel and the "nu" parameter, which determines the expected fraction of outliers. In our case, the kernel we have used is the Radial Basis Function and nu is 0.9.

In Isolation Forest, anomalies are detected by isolating data points through the use of random decision trees. The algorithm constructs several trees by randomly choosing features and split values. Data points that can be easily separated from the bulk of the data, shown by shorter path lengths in the trees, are marked as anomalies. Conversely, points with longer path lengths are deemed normal. This approach effectively identifies outliers by taking advantage of their isolation characteristics.

Figure 3 shows the flowchart of the proposed system illustrates the various components and their interactions, as well as the overall flow of the system.

The ESP8266 is well-suited for SmartBot due to its low cost, built-in Wi-Fi, and efficient performance in controlling the bot via the Blynk app for remote operation. While alternatives like the ESP32 offer more power and features, the ESP8266 provides sufficient capabilities for SmartBot's tasks, such as motor control and data transmission, without added complexity.

The Raspberry Pi Pico W takes care of sending sensor data to Google Sheets. Its Wi-Fi capability and efficient power use make it ideal for SmartBot. Compared to alternatives like the Raspberry Pi Zero 2 W, the Pico W is more cost-effective and perfectly capable of managing real-time data collection.

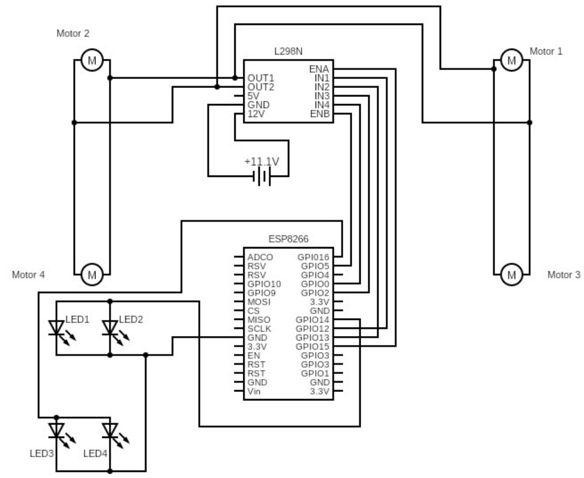

Figure 4. Connection between ESP8266, L298N motor driver and LEDs

For machine learning, the One-Class SVM is ideal for detecting gas anomalies since it excels in unsupervised learning when only "normal" data is available, making it perfect for rare hazardous events. While the Isolation Forest is faster and more efficient for sparse anomalies, One-Class SVM's precision in identifying subtle gas-related anomalies is critical for safety in this application.

4. WORKING PRINCIPLE

Figure 4 shows the hardware connection between ESP8266, L298N motor driver, and LEDs.

In the project setup, the L298N H-Bridge Motor Driver facilitates control over four motors, with two motors interconnected on each side, their output pins paralleled before connecting to the respective output pins of the L298N driver for synchronous movement. The ESP8266 IoT microcontroller, serving as the central control unit, communicates with the L298N driver through six input pins, with EN-A pin of L298N connected to D8, EN-B connected to D1, IN-1 connected to D6, IN-2 connected to D7, IN-3 connected to D4, and IN-4 connected to D3. This setup allows for remote control and monitoring via a mobile app through integration with the Blynk IoT cloud platform. Users send commands through Blynk, interpreted by the ESP8266 microcontroller, which then dispatches corresponding signals to the L298N driver, dictating motor movement and direction. The interconnected system offers seamless, remote-controlled operation, with the ESP8266's VIN connected to an 11.1-volt LiPo battery for power supply.

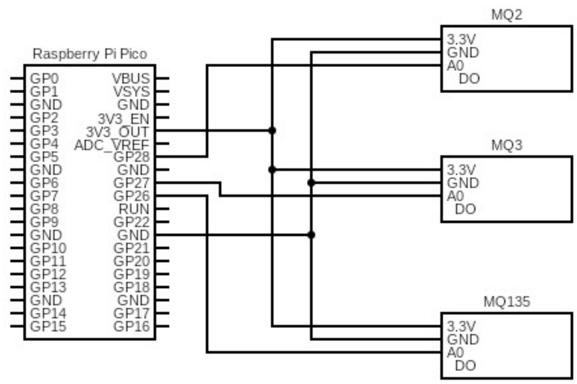

Figure 5. Connection between MQ Sensors and Raspberry Pi Pico W

The hardware setup for the connection between the gas sensors MQ2, MQ3, and MQ135 and Raspberry Pi Pico is given in Figure 5

In the sensor setup, the MQ-2 sensor is linked to GPIO 26, the MQ-3 sensor to GPIO 27, and the MQ-135 sensor to GPIO 28 on the Raspberry Pi Pico W microcontroller board. Each sensor's signal pin is connected to its respective GPIO pin, enabling the Pico W to read analog signals from the sensors for gas concentration measurements. Furthermore, each sensor includes a GND pin and a VCC pin. The GND pin connects to the power supply pin of the Raspberry Pi Pico W, while the VCC pin powers the sensors, thus completing the circuit. This configuration allows the Raspberry Pi Pico W to work with the MQ sensors and collect data for various applications, such as air quality monitoring and gas detection systems.

Data collection is made easier with advanced gas sensors such as the MQ-2, MQ-3, and MQ-135, which are integrated into the IoT-driven bot. As the bot navigates environments, these sensors continuously monitor gas concentrations in real-time, detecting a range of gases prevalent in the petroleum industry. Once collected, the gas data is transmitted to Google Spreadsheets through webhooks, ensuring centralized storage and accessibility for further analysis.

Webhooks are a method of communication between two different applications or services over the web. They allow real-time data to be sent from one application to another whenever a particular event occurs. The basic idea is that instead of a developer periodically requesting data from an application and waiting for a response (polling), the application itself will push data to a specific URL (the webhook) as soon as an event happens.

In the system's operational framework, anomaly detection is a critical component facilitated by advanced algorithms such as Isolation Forest and One-Class SVM. Isolation Forest efficiently identifies anomalies within the gas data by isolating them in sparse regions of the dataset, making it particularly effective for detecting unusual patterns or outliers. Similarly, One-Class SVM distinguishes normal data points from anomalies by constructing a boundary around the majority of the data, thereby enabling the identification of abnormal instances. Through the integration of these algorithms, the system can accurately detect deviations from expected behavior, providing operators with timely alerts to mitigate potential risks in hazardous environments within the petroleum industry.

A. *Isolation Forest*

Isolation Forests (IF) present a distinctive approach to anomaly detection by using decision trees, akin to Random Forests, but with a significant departure in their methodology. Unlike Random Forests, which are supervised models requiring pre-defined labels for classification, IF operates as an unsupervised model, eliminating the need for labeled data to identify anomalies within a dataset. Instead, IF focuses on isolating outliers or anomalies within the data points themselves.

The key principle underlying IF is to isolate anomalies by exploiting their inherent properties, such as being sparse or distant from the majority of data points. By doing so, IF efficiently distinguishes anomalies from normal data points without relying on the definition of what constitutes "normal." This approach allows IF to identify anomalies more effectively, as anomalies typically exhibit distinct characteristics, such as shorter tree paths, compared to normal data points.

One of the main advantages of IF lies in its ability to operate with shallow decision trees. Unlike traditional decision trees, which may require deeper splits to accurately classify data points into multiple classes, IF's focus on isolation means that trees within the forest do not need to be deep. As a result, IF requires less memory and computational resources, making it more efficient, particularly when dealing with high-dimensional datasets.

Mathematical Formula for anomaly score[7]:

$$S(x, m) = 2^{-E\times\frac{h(x)}{c\times m}}$$

Here, m is the number of points and x is the data point.

B. *One Class SVM*

One-Class Support Vector Machines (One-Class SVMs) offer a specialized approach to anomaly detection, tailored for scenarios where the objective is to identify outliers and novel data points within a single class. In contrast to traditional Support Vector Machines (SVMs), which are typically applied to binary classification tasks,One-Class SVMs function by exclusively training on data points from a single class, referred to as the target class.This distinctive characteristic enables One-Class SVMs to excel in scenarios where only one class of data is available during the training phase, which is common in anomaly detection applications.

The fundamental goal of a One-Class SVM is to learn a boundary or decision function within the feature space that effectively encapsulates the target class. This boundary serves as a representation of the normal behavior exhibited by the data points belonging to the target class. By delineating this boundary, the One-Class SVM aims to differentiate between normal data points and potential outliers or anomalies that lie outside of this boundary.

The exclusive focus on one class during training enables One-Class SVMs to capture the inherent characteristics and patterns of the target class with precision. This focused learning approach empowers the One-Class SVM to identify deviations from the established norms with a high degree of accuracy, thereby distinguishing outliers and anomalies effectively. Consequently, One-Class SVMs are well-suited for applications where the detection of anomalies within a specific class of data is paramount, such as identifying fraudulent transactions in financial systems or detecting anomalies in sensor data from industrial machinery.

Mathematical Formula for anomaly score[7]:

$$\min_{\omega,p,\xi}\left(\frac{1}{2}||\omega||^2_{-p} + \frac{1}{v_n}\sum_{i=1}^{n}\xi_i\right)$$

In support vector machines (SVMs), the separating hyperplane is represented by a weight vector (ω). The position of the hyperplane relative to the origin is determined by the offset (ρ) along the normal vector (ω). Slack variables (ξ_i) are associated with each data point (i) and allow for a soft margin, penalizing deviations from the margin. These variables indicate how much a data point violates the margin or falls on the incorrect side of the hyperplane. A hyperparameter (v) regulates the balance between maximizing the margin and minimizing the number of data points within the margin or on the wrong side of the hyperplane. The objective is to minimize the squared norm of the weight vector ($||\omega||^2$).

5. IMPLEMENTATION

A. Bot Controlling through the Blynk IoT App

When input is received through the Blynk IoT platform, it is sent via virtual pins to the ESP8266 microcontroller. This microcontroller acts as the main processing unit, handling the incoming data from the Blynk platform. After processing, the ESP8266 sends digital signals to the L298N motor driver, which is crucial for controlling the motors' direction. Depending on the commands received, the L298N motor driver changes the polarity of the electrical signals sent to the motors, determining if they rotate clockwise or counterclockwise. This coordinated system enables precise control and manipulation of the motors' movements based on user input from the Blynk IoT platform, whether accessed through a mobile app or a web interface.

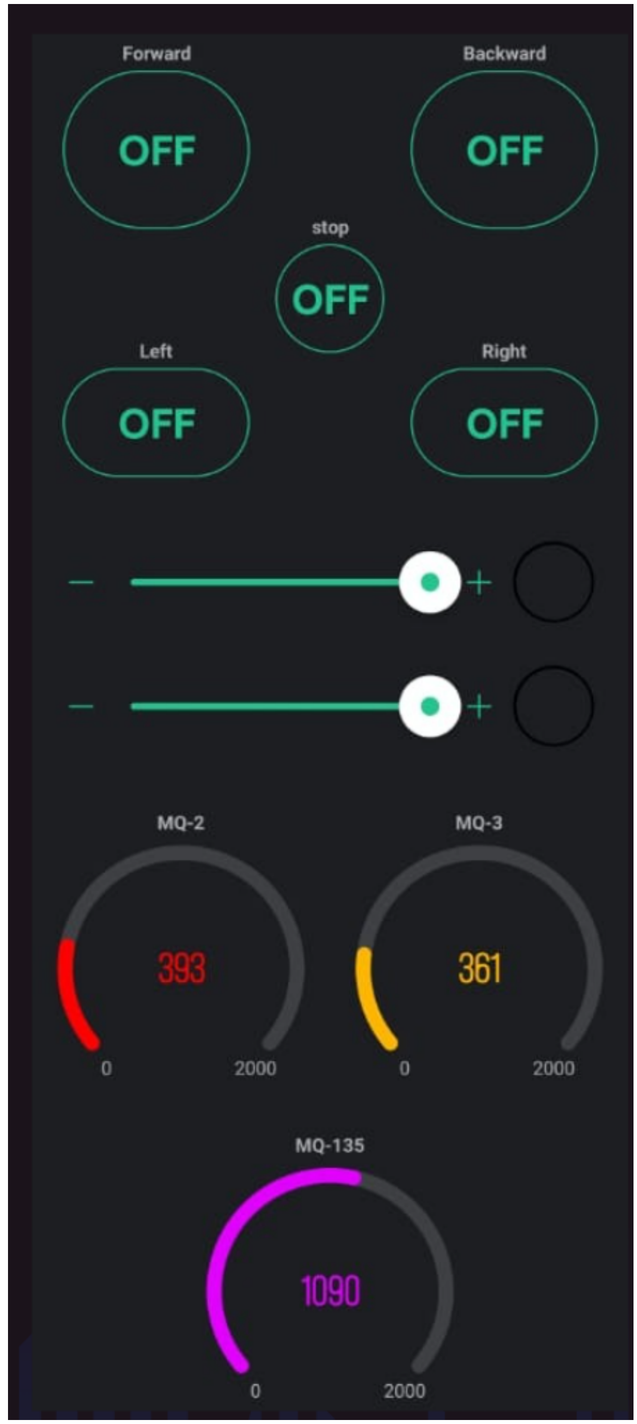

Figure 6. Blynk IoT app interface

Figure 6 shows the Blynk IoT app interface for controlling the Bot.

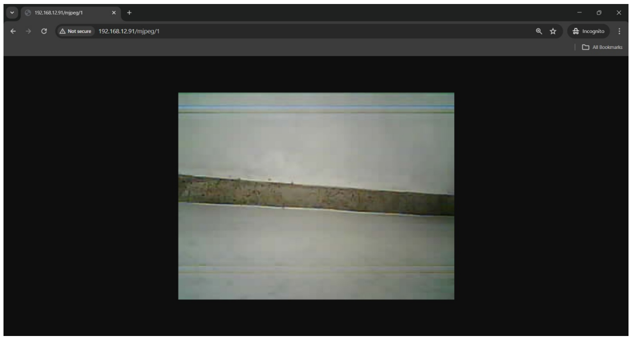

Figure 7. Live Video Streaming of Bot's Surroundings on the Web

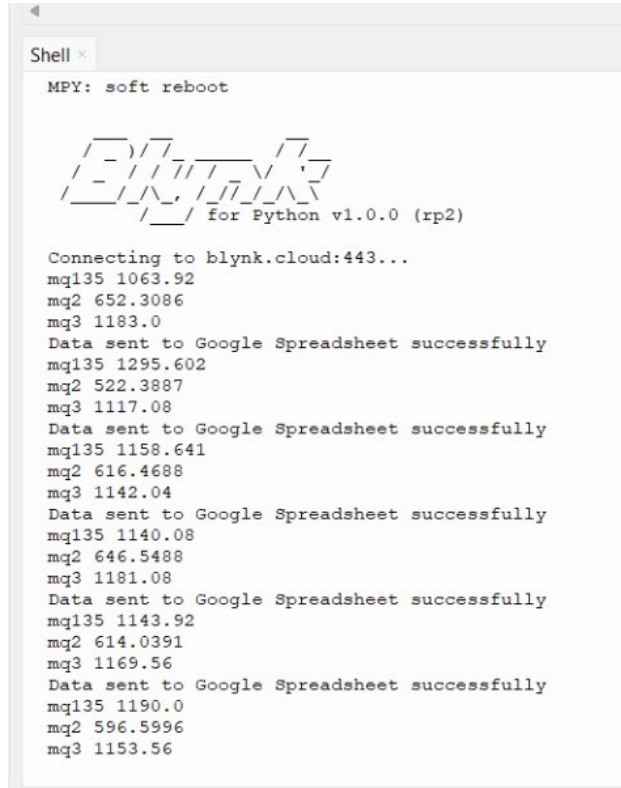

Figure 8. Raspberry Pi Pico collection of gas data

B. *Live Video Streaming*

The bot is furnished with an ESP32 CAM module, which provides live video streaming capability and facilitates immediate broadcasting of video content over the web. This feature allows operators to gain visual insights into the bot's surroundings, regardless of its deployment location. Whether navigating diverse environments or remote locations, the ESP32 CAM module ensures immediate access to live video feeds, enhancing situational awareness and operational control.

The live video streaming shows video transmission by the ESP32 Cam module connected to the internet as shown in Figure 7.

C. *Gas Data Collection through MQ Sensors*

The gas data collected by Raspberry Pi Pico from the MQ sensors is printed on the Raspberry Pi Pico's terminal and is shown in Figure 8 .

The gas sensors (MQ2, MQ3, and MQ135), integrated with the Raspberry Pi Pico W, are utilized to collect real-time data on gas concentrations within the environment. These sensors possess the ability to identify a range of gases that are essential for safety within the petroleum sector, including methane, liquefied petroleum gas (LPG), carbon monoxide, benzene, and toluene. The Raspberry Pi Pico W, serving as the system's microcontroller, interfaces with these sensors using GPIO pins or I2C/SPI communication protocols, enabling seamless data retrieval. Each sensor's readings are gathered either individually or in parallel, allowing continuous monitoring of gas levels in real-time. The selection of the Raspberry Pi Pico W is justified due to its compatibility with the sensors, low power consumption, and ability to manage multiple data streams in a dynamic industrial environment efficiently.

D. *Gas Data Transfer to Google Spreadsheet through Webhooks*

After the gas data is collected, the Raspberry Pi Pico W transmits it to Google Spreadsheets in real time through webhooks. Webhooks allow the microcontroller to communicate with external services like Google Sheets by sending HTTP requests containing the sensor data. The Raspberry Pi Pico W is programmed to send the gas concentration readings to a predefined webhook URL linked to a specific Google Spreadsheet. Each HTTP request includes the latest gas data, which Google Spreadsheets processes and automatically inserts into a new row. This approach ensures a consistent and centralized stream of real-time information, facilitating access for analysis, visualization, and subsequent selection processes. The use of webhooks ensures efficient, automated, and real-time data transfer, eliminating the need for manual intervention while maintaining an up-to-date log of gas concentrations in the monitored environment.

Figure 9 shows the MQ sensors' data collection on Google Spreadsheets.

E. *Machine Learning Models and their Evaluation*

Following the continuous transmission of gas data to Google Spreadsheets, the system initiates its anomaly detection process by employing unsupervised machine learning models, notably the Isolation Forest and One-Class SVM algorithms. These models are integrated into the system to autonomously analyze the collected gas data and identify any irregular patterns or outliers that may indicate potential anomalies or hazardous conditions.

The Isolation Forest algorithm is utilized as a key component in anomaly detection due to effectively identifying

date	MQ135	MQ3	MQ2		
5/4/2024 5:02:06	906.4414	287.36	889.0098		
5/4/2024 5:02:10	855.2402	337.92	915.7285		
5/4/2024 5:02:14	836.041	294.4	879.4082		
5/4/2024 5:02:18	833.4414	388.48	839.2188		
5/4/2024 5:02:22	704.8008	337.92	868.3398		
5/4/2024 5:02:26	1397.281	310.4	834.4199		
5/4/2024 5:02:30	811.041	360.96	880.0195		
5/4/2024 5:02:34	1160.602	381.4399	966.6387		
5/4/2024 5:02:38	1519.723	369.92	1598.949		
5/4/2024 5:02:42	1618.922	420.48	699.5098		
5/4/2024 5:02:46	1450.563	154.8799	426.0098		
5/4/2024 5:02:50	1464.602	360.96	514.6797		
5/4/2024 5:02:54	1255.961	349.4399	558.8398		
5/4/2024 5:02:58	1236.76	310.4	590.3594		
5/4/2024 5:03:02	880.1602	310.4	641.8789		
5/4/2024 5:03:06	915.4004	333.4399	495.1895		
5/4/2024 5:03:10	837.9609	280.96	561.2402		

Figure 9. Gas Data on Google Spreadsheet

and isolating anomalies within a dataset. This algorithm operates by constructing many decision trees, each of which randomly selects features and splits data points. By doing so, anomalies are often isolated into shorter paths within the trees, making them stand out as distinct anomalies. This inherent property of the Isolation Forest algorithm enables efficient detection of anomalies within the gas data, facilitating timely intervention and risk mitigation measures.

The line plot in Figure 10, Figure 11, and Figure 12 illustrates the time series data of three sensors: MQ-2, MQ-5, and MQ-135. Indexes along the x-axis track chronological data points, while sensor readings in parts per million (ppm) are displayed on the y-axis. The blue line depicts the sensor readings' trends and fluctuations over time, providing insights into gas concentration levels. Anomalies, detected using the Isolation Forest algorithm incorporating a contamination parameter of 0.05, are highlighted with red dots, the red dots depict a drastic change in normal reading, enabling easy identification of deviations from normal behavior.

One-Class SVM creates a boundary around most data points, anomalies are outside the boundary. Both models excel in detecting anomalies within the gas data without requiring labeled training data, making them suitable for unsupervised anomaly detection tasks. The system autonomously identifies anomalies within the gas data, providing operators with timely alerts when changes from regular patterns are detected. This approach helps operators to swiftly respond to potential hazards and mitigate risks, ensuring the safety of personnel and facilities in high-risk environments within the petroleum industry.

A One-Class SVM model is trained on each sensor's data to learn the underlying patterns. Predictions from the model are then used to identify anomalies, which are highlighted in the plot. Anomalies are typically represented by data points that fall outside the normal range of values or exhibit unusual behavior compared to the majority of the data.

The visualizations in Figure 13, Figure 14, and Figure 15

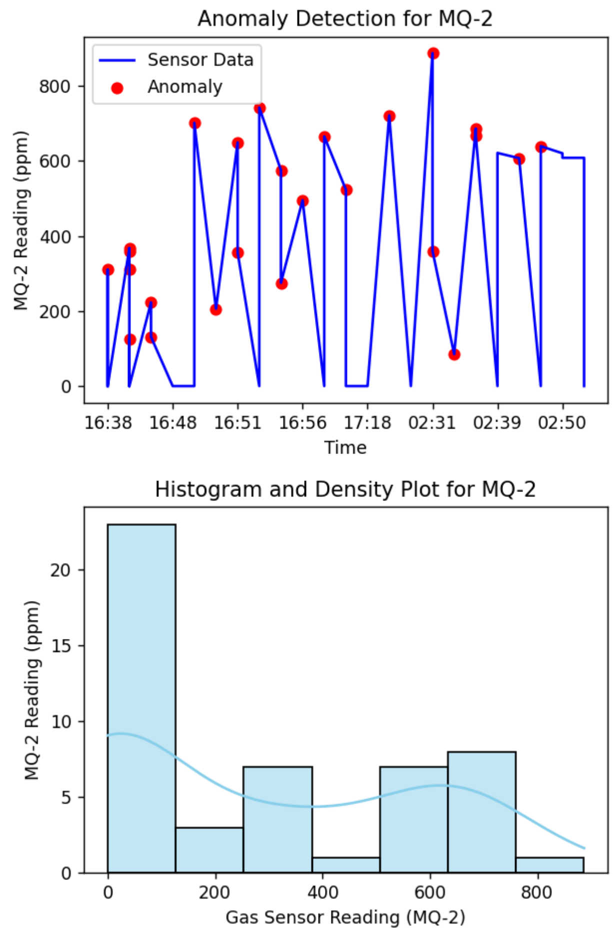

Figure 10. Isolation Forest visualization on gas data of MQ-2 sensor

depict sensor data over time, specifically focusing on three types of sensors: MQ135, MQ2, and MQ3. Each sensor is represented in its subplot within a single figure. Each subplot has a time-based x-axis and a y-axis that displays the sensor values.

We have evaluated the models based on their detection rate performance. The detection rate, also known as True Positive Rate or Recall, measures the proportion of actual anomalies that the anomaly detection model correctly identifies. It evaluates how effectively the model identifies true anomalies among all the actual anomalies present in the dataset.

Mathematical Formula for detection rate:

$$Detection\ Rate = \frac{True\ Positives\ (TP)}{True\ Positives\ (TP) + False\ Negatives\ (FN)}$$

A high Detection Rate indicates that the model is effective at identifying most of the actual anomalies. While a low Detection Rate Suggests that the model misses many

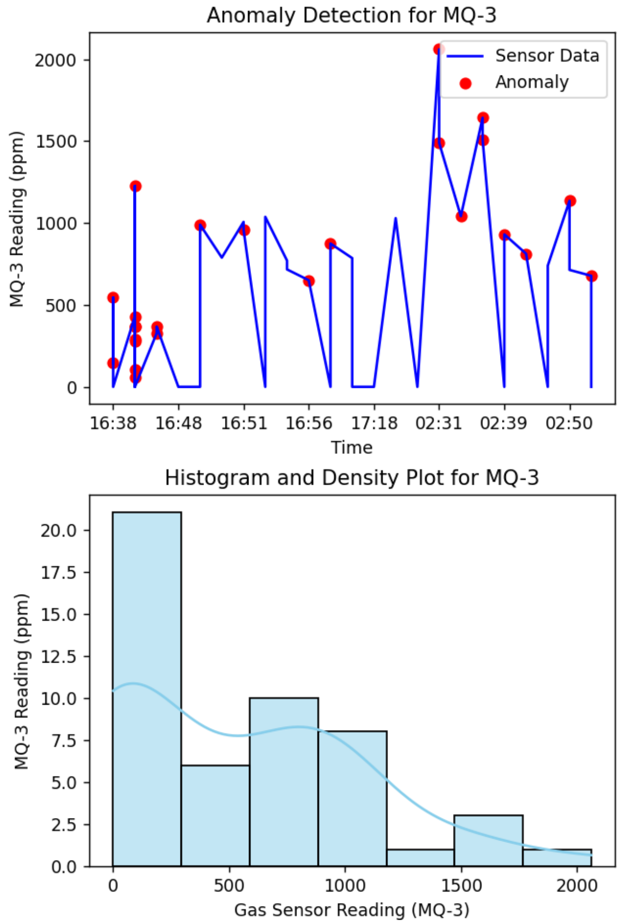

Figure 11. Isolation Forest visualization on gas data of MQ-3 sensor

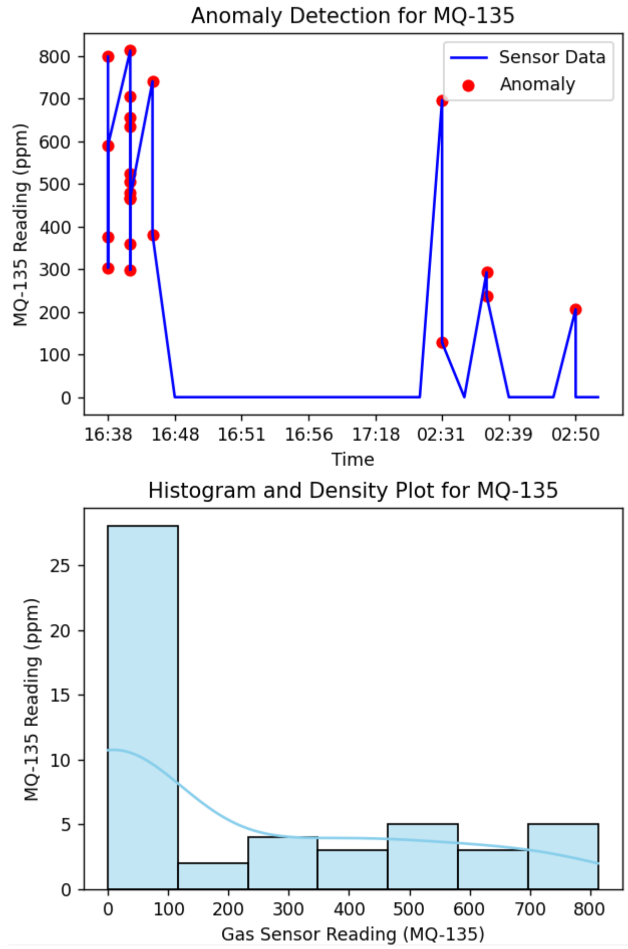

Figure 12. Isolation Forest visualization on gas data of MQ-135 sensor

anomalies, classifying them as normal may be problematic in applications where detecting anomalies is critical.

In the evaluation of classification models, including anomaly detection, True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) are critical metrics. True Positives (TP) refer to the number of instances that are correctly predicted as anomalies and are indeed anomalies in reality. True Negatives (TN) represent the number of instances correctly identified as normal and are actually normal. False Positives (FP) occur when the model incorrectly classifies normal instances as anomalies, while False Negatives (FN) are instances where the model fails to detect actual anomalies, classifying them as normal. These metrics are essential for understanding the performance of the model, as they help assess how well the model identifies true anomalies and avoids false alarms.

Table II shows the values of TP, TN, FP, and FN given by both models and the Detection Rate.

From the visualizations and table, we conclude to use On Class SVM on our project since it is more efficient.

F. Website Working

The gas monitoring system integrates the MERN stack (MongoDB, Express.js, React.js, Node.js) with Flask to create a dynamic and efficient platform for real-time gas monitoring and anomaly detection. The React.js frontend provides the user interface, displaying real-time data in the form of gas gauges and live video streaming. The video feed is pulled from an ESP32 camera and displayed to the user for continuous monitoring of the environment. Additionally, React presents graphical representations of gas concentration data, showing trends and highlighting any detected anomalies.

On the backend, Express.js handles the interaction between various system components. It retrieves gas readings from a Google Spreadsheet and sends them to the Flask server. Flask, in turn, uses a Random Forest machine learning model to analyze the gas data for potential anomalies. By comparing the incoming gas readings with pre-trained patterns, Flask identifies any unusual behavior, such as leaks or hazardous gas concentrations. Once the analysis is com-

TABLE II. Evaluation of both the models

Metric	One Class SVM	Isolation Forest
True Positives (TP)	50	45
True Negatives (TN)	80	85
False Positives (FP)	20	15
False Negatives (FN)	10	15
Detection Rate	83.3%	75%

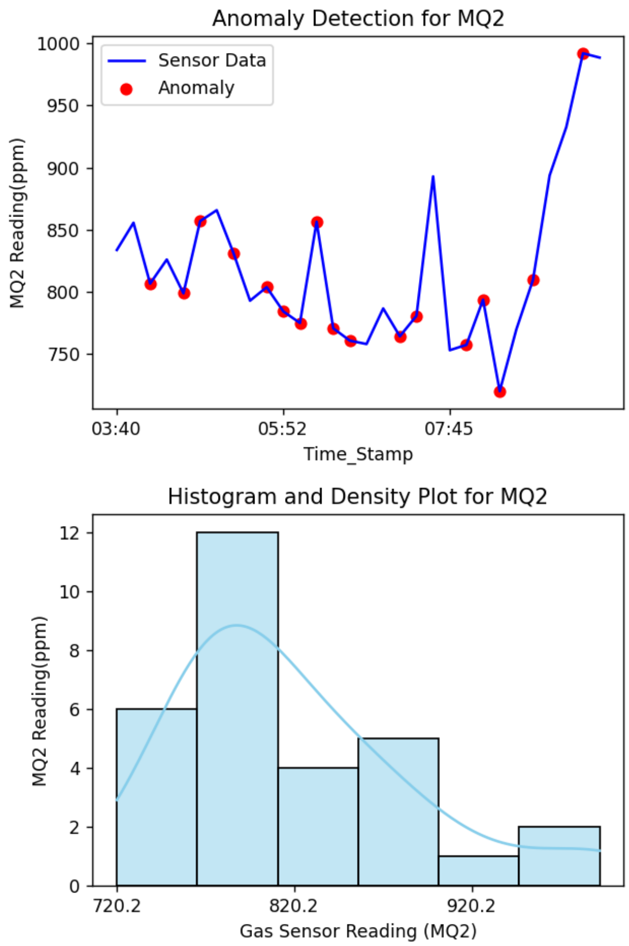

Figure 13. One Class SVM visualization on gas data of MQ-2 sensor

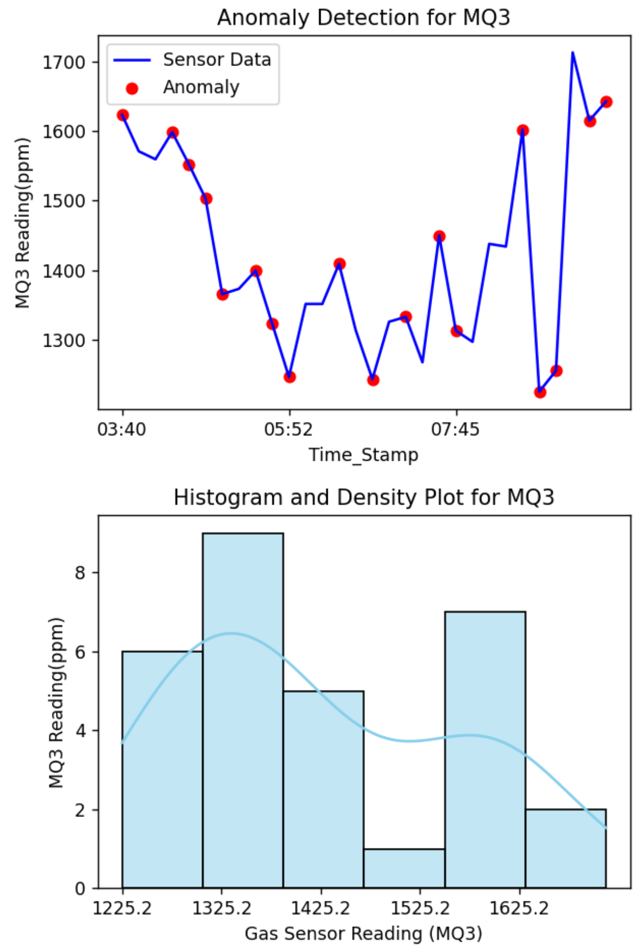

Figure 14. One Class SVM visualization on gas data of MQ-3 sensor

plete, Flask sends the results of both the gas readings and any identified anomalies back to Express.js, which forwards this information to the React frontend for display. This workflow enables real-time monitoring, anomaly detection, and visual representation of the gas readings, making the system both interactive and efficient for users.

Figure 16 shows the website interface of our project.

6. Discussion

A. *Practical Implementation*

We can deploy as many bots as required in each industry. The IoT and machine learning-based anomaly detection system offers significant benefits to the petroleum industry over traditional methods. Unlike manual checks or stationary detectors, this system enables real-time, automated monitoring of hazardous gases through mobile IoT bots. Controlled via the Blynk IoT application and equipped with sensors, these bots continuously collect data across facilities. Using the machine learning algorithm One-Class SVM, the system provides accurate anomaly detection, minimizing false alerts.

In a real-world deployment, this system enhances safety by quickly identifying gas leaks and toxic emissions, reduc-

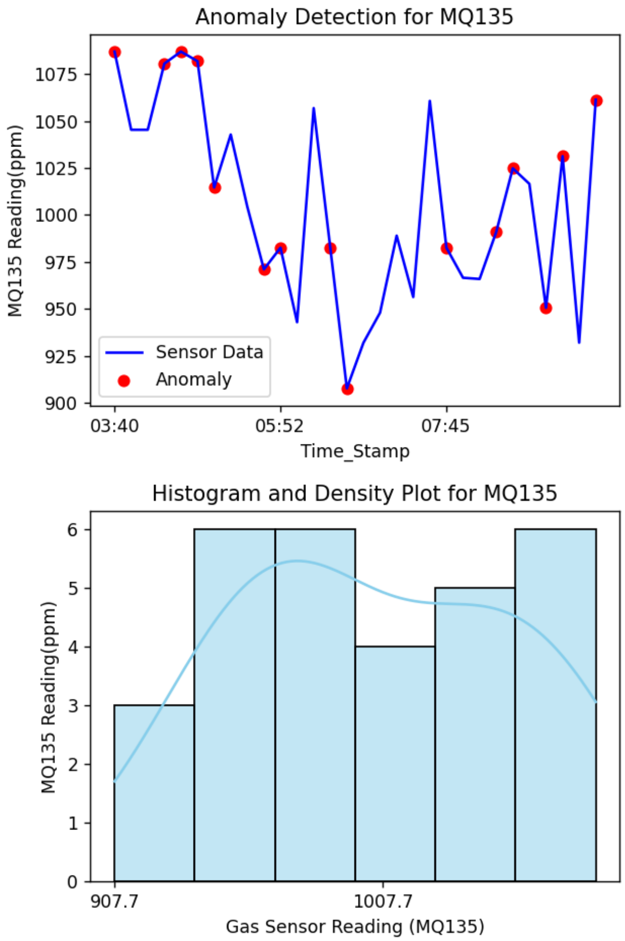

Figure 15. One Class SVM visualization on gas data of MQ-135 sensor

Figure 16. Website showing the graph, gauges and live video streaming coming from the Bot

ing human exposure, and ensuring faster response times. Its ability to provide continuous monitoring, data storage, and intelligent analysis offers a scalable, cost-effective solution that improves workplace safety and regulatory compliance.

The results from the implementation of the IoT-based gas detection system demonstrate its practical potential in real-world petroleum industry scenarios. Unlike traditional static gas detection systems that are limited in scope, the mobile nature of the bot allows for dynamic coverage of the entire facility, including hard-to-reach areas. This increases the likelihood of early detection of gas leaks or hazardous gas concentrations, potentially preventing catastrophic incidents such as fires, explosions, and asphyxiation.

B. Advantages

The system is capable of detecting anomalies in gas concentration data even when levels remain below the typical alarm thresholds used in conventional systems. This enables a more proactive risk mitigation approach, allowing operators to take preventive actions before critical conditions are met. The system's continuous real-time monitoring and anomaly detection reduce the reliance on manual inspections and improve overall safety protocols.

The key advantage of this IoT-based system lies in its scalability, cost-effectiveness, and real-time monitoring capabilities. Compared to traditional stationary gas detection systems, which are expensive and limited in coverage, the proposed system offers mobility, allowing for wider surveillance at a fraction of the cost. Moreover, the use of low-cost hardware components such as MQ-series sensors, Raspberry Pi Pico W, and ESP8266 reduces the overall system cost without compromising performance.

In terms of remote monitoring, the Blynk IoT platform provides real-time alerts and remote control capabilities, which traditional systems lack or require proprietary (and often costly) solutions for. This feature adds significant value in scenarios where human intervention is delayed or when the facility is unmanned.

C. Prior Work and Our Advancements

The prior works in IoT-based gas monitoring systems and machine learning-driven anomaly detection provide valuable insights and solutions but also come with certain limitations that our approach addresses.

In the realm of IoT-based gas detection and environmental monitoring systems, several studies have made significant contributions, yet they still present limitations that our work aims to address. Research in this field, such as [1], [2], and [4], has primarily focused on developing standalone systems that utilize different types of gas sensors like MQ-2, MQ-6, MQ-135, and microcontrollers such as ESP8266 and ESP32 to detect hazardous gases like methane, carbon monoxide, and LPG. These systems often use simple sensors that can trigger alarms when the concentration of gases exceeds certain thresholds. For example, [2] details a system that notifies users via mobile alerts through the Blynk application when hazardous gas leaks like methane or carbon monoxide are detected. It also features additional safety mechanisms like alarms, exhaust fan activation, and buzzer indicators, which improve safety in domestic and industrial environments. While these systems have proven effective in monitoring environmental conditions, they often lack real-time video surveillance and advanced data analytics that would make them more reliable for comprehensive monitoring.

Moreover, while IoT solutions like [3] employ machine learning techniques such as Random Forest, Support Vector Machines (SVM), and K-Nearest Neighbors (KNN) for anomaly detection, the application of these algorithms has often been limited to detecting simple anomalies in sensor data. For example, [6] describes an IoT system that uses machine learning to detect leaks in oil and gas pipelines, reporting high accuracy (up to 97.4%) for early detection of potential risks. However, these studies typically focus on data analysis alone without combining multiple types of input data (e.g., sensor data and video data), which limits the scope and situational awareness that can be achieved. Furthermore, although some studies in the field, such as those in [5], [9], and [10], incorporate real-time gas detection with alerts, they still tend to rely on relatively simple systems without the integration of advanced machine learning or real-time visual analysis to support decision-making.

Our work significantly advances these previous systems by incorporating multiple features. We combine gas leakage detection with live video streaming through the ESP32-CAM module, allowing for real-time monitoring of the environment. The data from the gas sensors is processed using Flask and machine learning algorithms like Random Forest to detect anomalies, and the results are sent to a React frontend, where both the gas levels and anomalies are visualized alongside the live video feed.

One of the unique contributions of our system is the integration of a movable bot controlled remotely via the Blynk app. This functionality enables users to move the bot to different locations within the monitored area, thus enhancing the coverage and flexibility of the monitoring process. The system's ability to combine mobility, real-time video, and gas sensor data sets it apart from existing solutions that are typically static or lack video analysis.

Additionally, our system integrates Google Spreadsheets for real-time data storage, making it easy for users to access and track sensor readings. We use the MERN stack (MongoDB, Express.js, React, Node.js) for the backend and frontend communication, ensuring seamless data flow between the sensors, machine learning model, and user interface.

Once the system is equipped with industry-grade sensors, it will be ready for deployment in real-world applications, such as industrial environments or hazardous areas, where its combination of gas detection, machine learning-based anomaly detection, live video streaming, and remote control will provide a comprehensive, real-time monitoring solution. This makes our system highly adaptable and robust for various applications, from industrial gas monitoring to environmental safety.

D. Limitations

Despite its many advantages, the current system faces several challenges and limitations. The system's performance is highly dependent on the quality and accuracy of the gas sensors, which can be affected by factors such as sensitivity, calibration, warm-up time, and response speed. Inaccuracies or delays in sensor readings could reduce the system's overall reliability. Another challenge lies in the real-time processing and transmission of data. While the system provides live video, gas readings, and graphical analysis, maintaining stable data flow and timely updates can be difficult, especially in remote areas with poor internet connectivity. Delays in data transmission may hinder the real-time monitoring capabilities. Furthermore, the scalability of the system in larger or more complex environments, like vast industrial plants or extensive outdoor areas, may be limited by the bot's size, mobility, and the need for manual oversight. Expanding the system to cover larger areas may require modifications to improve its range and autonomy.

Additionally, since the system is mobile, its battery life is another constraint. Continuous operation of sensors, motors, and communication modules, such as the Blynk app and live video streaming, can quickly drain the battery, reducing the operational time and necessitating frequent recharging or battery replacement.

E. Future Scope

The future scope of this system includes several potential enhancements. One major area of development is expanding the system to detect a wider variety of hazardous gases, allowing it to be used in different industrial settings beyond petroleum. Improving sensor technology for faster response times, greater accuracy, and self-calibration will further enhance system reliability.

Another key area for future research is the integration of more advanced machine learning models, such as deep learning or ensemble methods, to improve the accuracy of anomaly detection in complex or noisy environments. Additionally, incorporating predictive analytics could allow the system to not only detect anomalies but also predict potential gas leaks or hazards, making it a proactive safety solution.

Implementing edge computing could enable real-time data processing on-site, reducing the reliance on continuous network connectivity, which is especially valuable in remote locations. Moreover, the bot itself can be made autonomous, allowing it to navigate and monitor large areas without human intervention, further improving safety and efficiency.

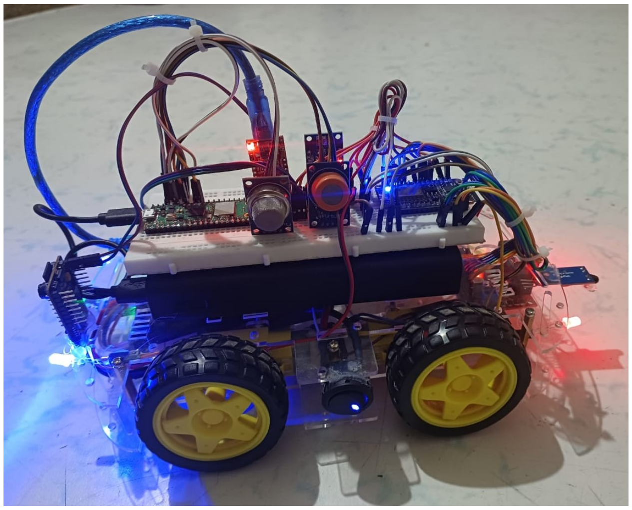

Figure 17. IoT Controlled Bot

The system could also be scaled for large industrial applications, integrated with cloud platforms for centralized monitoring, and equipped with automated maintenance features to minimize manual oversight.

These enhancements would transform the system into a more intelligent, flexible, and self-sufficient tool for improving safety and operational efficiency in high-risk industries like petroleum.

7. Conclusions

The developed gas monitoring system as shown in Figure 17 represents a significant advancement in safety monitoring technologies, particularly for high-risk industries such as petroleum. This innovative system integrates Technological advancements, encompassing the Internet of Things (IoT), machine learning algorithms, and live video streaming capabilities. The synergy of these technologies enables the system to perform real-time monitoring of hazardous gas concentrations, enhancing both proactive risk mitigation and situational awareness.

One of the key features of this system is its ability to detect anomalies using machine learning algorithms. Specifically, implementing these anomaly detection algorithms has significantly bolstered the system's capacity to identify and respond promptly to potential hazards. This timely detection is crucial as it reduces the likelihood of accidents, thereby minimizing risks to both personnel and facilities.

Moreover, the live video streaming capabilities provide an additional layer of situational awareness. By visually monitoring the environment in real-time, the system allows for more informed decision-making during critical situations, ensuring that any anomalies are detected and accurately assessed and managed.

The practical implications of this system for the petroleum industry are vast. In real-world scenarios, this system can be deployed across various facilities, ranging from refineries to drilling sites, where the risk of hazardous gas emissions is high. Unlike traditional static gas detectors, this mobile, IoT-driven solution can cover a wider area, dynamically navigating the environment to monitor gas concentrations in hard-to-reach or high-risk areas. The integration of machine learning further ensures that anomalies are detected in real time, allowing for immediate interventions that prevent accidents before they occur. Additionally, the system's live video streaming capabilities provide real-time visual data, empowering operators to respond more effectively and with greater situational awareness during emergencies.

The key benefits this system offers over traditional methods include continuous, real-time monitoring, automated anomaly detection, enhanced decision-making through live video feeds, and reduced dependence on manual inspections. This allows for more effective safety protocols, reducing the likelihood of human error and increasing overall operational efficiency. Fig. 16 depicts the prototype of the IoT-controlled bot that can be deployed in the petroleum industry.

In this paper, we present an IoT-controlled bot project designed to prevent hazards in the petroleum industry caused by oil leaks and gas spills. The project integrates machine learning, IoT, and robotics.

References

[1] S. M. M. Nurul Ayni Mat Pauzi and I. Yahya, "Low-cost environmental monitoring mini rover based on iot technology," *International Journal of Advanced Technology and Engineering Exploration*, vol. 8, 2021.

[2] M. A. N. T. A. A. Noor Kareem Jumaa, Younus Mohammed Abdulkhaleq, "Iot based gas leakage detection and alarming system using blynk platforms," *Iraqi Journal for Electrical and Electronic Engineering*, vol. 18, 2022.

[3] A. H. R. Samaila Bello, Muhammad Dikko Amadi, "Internet of things-based wireless sensor network system for early detection and prevention of vandalism/leakage on pipeline installations in the oil and gas industry in nigeria," *FUDMA Journal of Sciences (FJS)*, vol. 7, 2023.

[4] A. S. O. Ayeni, J. K., "Iot-based gas and smoke detection system using blynk application with automatic sms and alarm notifications," *University of Ibadan Journal of Science and Logics in ICT Research (UIJSLICTR)*, vol. 11, 2024.

[5] C. P. S. B. N. A. Kancharapu Chaitanya, G.G.S.V. Gayatri, "Efficient plant gas leakage detection and monitoring system based on iot integration," *Journal of Nonlinear Analysis and Optimization*, vol. 15, 2024.

[6] S. Aljameel, D. Alomari, S. Alismail, F. Khawaher, A. Alkhudhair, F. Aljubran, and R. Alzannan, "An anomaly detection model for oil and gas pipelines using machine learning," *Computation 2022*, vol. 10, 2022.

[7] W. F. Jr, K. S. Komati, and K. A. d. S. Gazolli, "Anomaly detection in oil-producing wells: a comparative study of one-class classifiers in

a multivariate time series dataset," *Journal of Petroleum Exploration and Production Technology*, 2023.

[8] K. A. d. S. G. Wander Fernandes Jr., Karin Satie Komati2, "Anomaly detection in oil-producing wells: a comparative study of one-class classifers in a multivariate time series dataset," *Journal of Petroleum Exploration and Production Technology*, 2024.

[9] T. Younas, D. Kumar, S. Mukhi, H. Z. Mufaddal, S. M. F. a. Fayyaz, M. M. a. Khan, and H. Ahmad, "Lpg gas detecting robot based on iot," *National Conference on VLSI, Embedded, and Communication & Networks*, 2020.

[10] M. F. R. Al-Okby, S. Neubert, T. Roddelkopf, and K. Thurow, "Mobile detection and alarming systems for hazardous gases and volatile chemicals in laboratories and industrial locations," *Sensors*, 2021.

[11] T. Das, D. J. Sut, V. Gupta, L. Gohain, and P. Kakoty, "A mobile robot for hazardous gas sensing," *International Conference on Computational Performance Evaluation (ComPE)*, 2020.

[12] H.-C. Cheng, M.-C. Chiu, K.-F. Zeng, and C.-M. Chiu, "A design of toxic gas detecting security robot car based on wireless path-patrol," *The 2nd International Conference on Precision Machinery and Manufacturing Technology (ICPMMT 2017)*, vol. 123, 2017.

[13] S. Huang, A. Croy, L. A. Panes-Ruiz, V. Khavrus, V. Bezugly, B. Ibarlucea, and G. Cuniberti, "Machine learning-enabled smart gas sensing platform for identification of industrial gases," *Advanced Intelligent Systems*, vol. 4, 2022.

[14] J. Liu, J. Harkin, and McDaid, "Bio-inspired anomaly detection for low-cost gas sensors," *IEEE 18th International Conference on Nanotechnology (IEEE-NANO)*, 2018.

[15] A. B. N, P. Bharath, C. G. B, S. V. Kumar, and V. G H, "Automation of lpg cylinder booking and leakge monitoring system," *International Journal of Combined Research & Development*, vol. 5, 2016.