Developmant of Gas Hazard Environment Prediction Based on Fuzzy-KNN and MQ2 Sensor in Microcontroller

Sumantri K Risandriya[1], Fardiwans Marbun[2]

[1,2]Electronic Engineering Technology, Department of Electrical Engineering, State Polytechnic of Batam, 29461, Indonesia.

[1]sumantri@polibatam.ac.id , [2]fardiwansmarbun11@gmail.com

***Abstract*— This study developed a gas hazard prediction system using the Fuzzy K-Nearest Neighbor (Fuzzy-KNN) method and an MQ-2 sensor on an Arduino Uno. The system addresses the growing risks of LPG, smoke, and methane leaks in household environments. The MQ-2 sensor was selected for its high sensitivity, while Fuzzy-KNN helps reduce data uncertainty in classifying hazard levels.The system process includes sensor data collection, Euclidean distance calculation with K = 3, fuzzy membership weighting, and classification into safe, alert, and danger categories. Outputs are provided through an LCD display and early warnings using a buzzer and LED. Testing with 46 samples achieved 87% accuracy and a response time of less than one second, demonstrating that the system can detect gas leaks in real time and has strong potential to improve home safety.**

Keywords*— *MQ-2 Sensor, Fuzzy-KNN, Gas Leak Detection, Microcontroller, Home Security System

I. INTRODUCTION

Safety in homes is often overlooked, especially regarding LPG (Liquefied Petroleum Gas) leaks. Due to its flammable nature, even small leaks can cause explosions or poisoning. Current detection methods still rely on human senses, which are often too late to provide warnings. Therefore, an intelligent system is needed to detect and predict gas leaks in real time.

Previous research has developed gas detection systems using sensors and intelligent algorithms. [1] developed a smoke detection system using fuzzy logic to automate actuators, but it only responds after dangerous thresholds are reached. Additionally, [2] demonstrated the efficacy of microcontroller-based layouts for household gas monitoring. Thus, an advanced predictive approach is required to analyze sensor patterns before hazards occur.

According to [3], combining fuzzy logic and K-Nearest Neighbor (Fuzzy-KNN) offers significant advantages for sensor-based classification. Fuzzy logic manages data uncertainty through linguistic values, while KNN classifies inputs based on historical training data. This combination provides an accurate, adaptive framework for predicting dynamic gas hazard levels.

II. METHOD

A. MQ-2 Gas Sensor

The MQ-2 sensor is an SnO_2 semiconductor-based gas sensor that detects combustible gases such as LPG, methane, hydrogen, smoke, and CO. Its fast response, high sensitivity, and low cost make it widely used for household and industrial gas leak detection [4].

According to [5] sensor output increases with closer distance to the gas source and longer exposure time, showing its ability to detect changes in gas concentration. The relationship between sensor resistance (Rs) and gas concentration (PPM) is expressed empirically:

$$\frac{Rs}{Ro} = A.(PPM)^B \quad \textbf{(1)}$$

where Ro is the resistance in clean air and A, B are calibration constants. The higher the gas concentration, the lower the resistance Rs. The MQ-2 sensor operates in a voltage divider circuit with a load resistor (RL), producing an output voltage:

$$Vout = Vcc.\frac{RL}{RL + RS} \quad \textbf{(2)}$$

and the value of Rs can be calculated from:

$$Rs = RL.(\frac{Vcc}{Vout} - 1) \quad \textbf{(3)}$$

At this step, the analog voltage value generated by the sensor can be converted into resistance. The Ro value is then compared to the Ro value to determine the gas con centration in ppm. This procedure is performed using the empirical curve specified in the datasheet, as shown in Fig. 1.

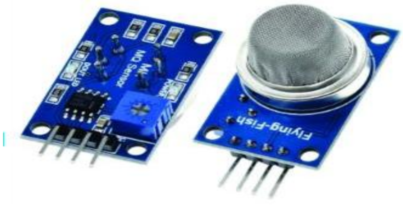


Figure 1. MQ-2 Sensor

B. Arduino Microcontroller

Arduino is an open-source microcontroller platform that is easy to program and commonly used in automation systems. Arduino Uno and Mega 2560 are popular choices due to their easy integration with gas sensors, communication modules such as SIM800L, and actuators such as buzzers and LCDs [5], [6], [7].

[6] shows that Arduino can process data from MQ-2 sensors and automatically send alerts via SMS using the SIM800L module, supporting real-time gas leak monitoring applications. Furthermore, [8] proved that microcontrollers can effectively execute local threshold automation alongside continuous sensor streaming.

C. Fuzzy-KNN

The Fuzzy-KNN method combines fuzzy logic and the K-Nearest Neighbor (K-NN) algorithm to classify uncertain data based on feature similarity. This method is suitable for gas sensor data that fluctuates and has unclear boundaries.

Fuzzy logic processes uncertainty using linguistic values such as “low,” “medium,” and “high” [9]. The Mamdani fuzzy method is commonly used due to its simple and easy-to-implement rule structure. [7] applied this method for vehicle fuel classification using sensor values and exposure time.

K-NN is a non-parametric classification algorithm that determines data classes based on the nearest neighbors in feature space. Fuzzy-KNN enhances K-NN by adding fuzzy membership weights to neighbor calculations, allowing smoother classification and better handling of sensor uncertainty. This capability makes Fuzzy-KNN suitable for real-time gas hazard prediction using MQ-2 sensor data.

According to [10], [11], the membership degree of test data (x) in class (C_i) is calculated based on the distance to neighbors and the fuzzy membership degree of training data, formulated as:

$$u_i = \frac{\sum_{j=1}^{k} u_i(x_j)\left(\frac{1}{\|x_{-} - x_j\|^{2/(m-1)}}\right)}{\sum_{j=1}^{k}\left(\frac{1}{\|x - x_j\|^{2/(m-1)}}\right)} \quad \textbf{(4)}$$

Where (u_i(x)) represents the membership degree of test data (x) to class (i), (x_j) is the (j)-th training data, and (m) is the fuzzifier parameter (typically between 1.5 and 3). The Euclidean distance formula used in K-NN is defined as follows:

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad \textbf{(5)}$$

Where (x) and (y) represent sensor readings such as CO and LPG levels. New data is compared with training data to find the three nearest neighbors ((K=3)). If the nearest neighbors do not show a clear majority class, fuzzy logic is used to determine class membership. According to [3] combining K-NN with fuzzy logic improves classification accuracy in uncertain sensor-based systems.

However, [12], [13] proposed using Minkowski distance as a generalized form of Euclidean distance, expressed as:

$$d_i = \left(\sum_{j=1}^{n} |x_{qj} - x_{ij}|^{p}\right)^{1/p} \quad \textbf{(6)}$$

The parameter determines the distance type: when (p=1), it produces Manhattan distance, and when (p=2), it produces Euclidean distance. Minkowski distance can improve classification performance for data with different distributions. [14] developed Local Mean and Global Learning Fuzzy-KNN (LMGL-FkNN), which improves classification stability by considering local and global data patterns, especially in noisy and imbalanced datasets.

Fuzzy-KNN is suitable for gas hazard prediction using MQ-2 sensors because it can handle dynamic and uncertain sensor data. Its adaptive capability enables accurate and responsive real-time gas leak prediction. In this study, the Fuzzy-KNN formulation is simplified for efficient implementation on an Arduino Uno while maintaining its fundamental principles. The distance calculation, fuzzy weighting, and membership aggregation processes are discussed in Chapter 3.3.2 (System Design).

D. Gas Leak Detection System

Gas leak detection systems based on MQ sensors and Arduino have been widely developed. Several studies have integrated these systems with IoT and communication platforms such as Telegram to provide early warnings to users [4], [9]. In several other studies, the system is equipped with an LCD, buzzer, relay, and purifiers such as exhaust fans and activated carbon filters to reduce the concentration of hazardous gases [4], [15]. Additionally, [16] emphasized that implementing automated hazard detection algorithms can drastically improve true positive identification rates in variable background environments.

E. System Flowchart

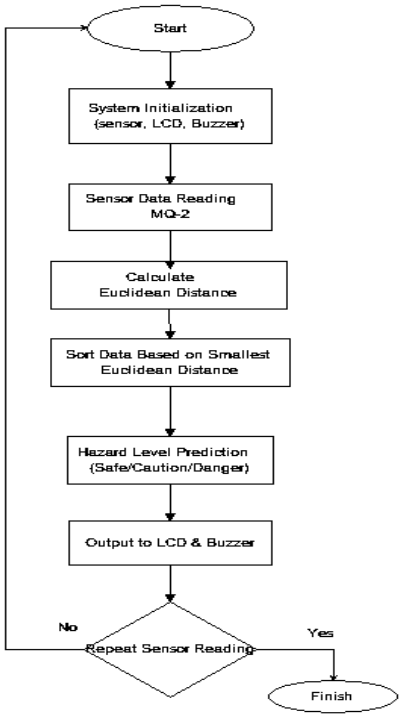
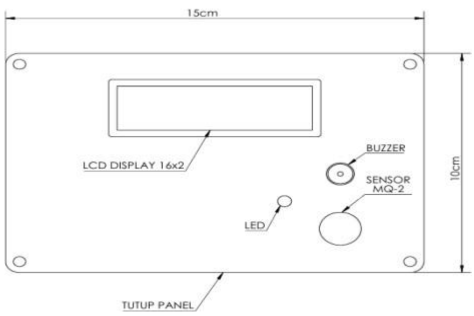


Figure 2. System Flowchart

Figure 2 shows the workflow of the gas hazard prediction system using Fuzzy-KNN. The system begins with sensor data acquisition, LCD display, and buzzer output. Sensor values are fuzzified into membership values, then Euclidean distances are calculated and sorted to determine the nearest data. Based on the results, the system classifies gas hazard levels into safe, caution, and danger categories.

F. Mechanical Drawings

The device uses a minimalist plastic box casing with a functional design. The front panel includes a 16x2 LCD for real-time gas status display, a red LED indicator, a buzzer alarm, and an MQ-2 sensor positioned for optimal gas detection. Internal components, including the Arduino Uno and supporting modules, are arranged neatly for easy maintenance. The design emphasizes safety, portability, and suitability for household applications. The mechanical design is shown in Fig 3.



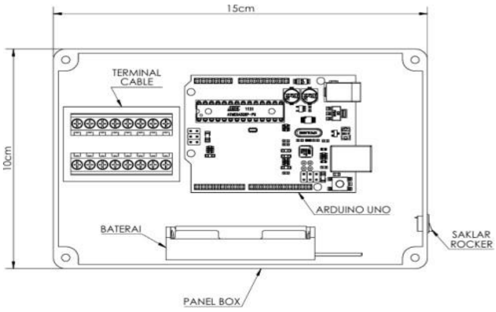


Figure 3. Mechanical Front and Interior Views

G. Electrical Design

The electrical diagram illustrates an Arduino Uno-based gas detection system. The MQ-2 sensor is connected to analog pin A0 for gas concentration measurement. The 16x2 LCD with an I2C module uses pins A4 (SDA) and A5 (SCL) to display hazard levels, while the red LED and active buzzer are connected to pins D7 and D8 as visual and audio indicators. The circuit is powered by a 5V supply from the Arduino and processes sensor data using the Fuzzy-KNN algorithm to activate indicators based on gas hazard levels, as shown in Fig 4 and Fig 5.

Figure 4. Electrical Design

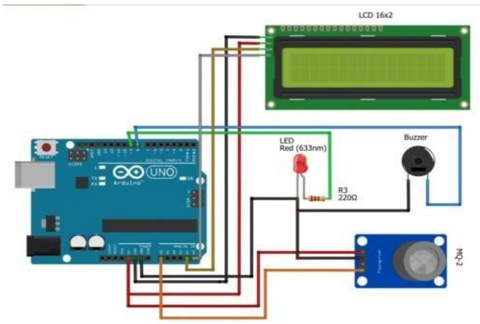
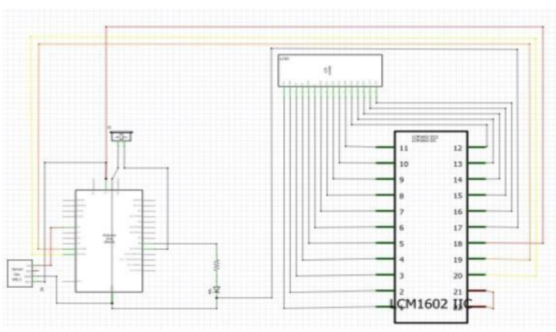


Figure 5. Schematic Design



H. System Design

At this stage, a Fuzzy-KNN-based gas prediction system was developed to classify gas hazard levels (Safe, Alert, and Danger) using the MQ-2 sensor. The analog sensor readings were converted into ppm values and processed by the Fuzzy-KNN algorithm on the Arduino Uno. The algorithm consists of three main stages:

- **Distance Calculation Stage (Distance Calculation)**

Each test data from the sensor is compared with all training data to determine the proximity between data. Distance calculation uses a simple Euclidean Distance meth od, which is formulated as:

$$d_i = |x_q - x_i| \quad (7)$$

where is the current sensor reading value (test data), while is the value in training. This formula is used in the euclideanDistance function to calculate the absolute dif ference between sensor values.

- **Neighbor Weighting Stage**

After obtaining the distance of each data point, the system assigns a fuzzy weight based on the inverse of the distance using the following formula:

$$w_i = \frac{1}{d_i} \quad (8)$$

The closer the test data is to the training data, the higher the fuzzy membership weight. This process is performed in the fuzzyKNN function to determine each neighbor's contribution to the classification result.

- **Accumulation and Class Determination Stages (Membership Aggrega tion)**

Each class (Safe, Alert, and Danger) obtains a total weight of K=3 closest neigh bors. The class with the highest total weight is selected as the final prediction result using the maximum membership rule principle, as written in the following equation:

$$c_{Pred} = arg \max_i \sum_{j=1}^{k} wij \quad (9)$$

The prediction results are displayed on a 16x2 LCD, while the red LED and buzzer provide warnings when gas concentration exceeds the danger threshold.

In this study, the Fuzzy-KNN formula is simplified for efficient implementation on the Arduino Uno. The simplification is performed by reducing the complexity of the membership function calculations and the parameter (m) used in the theoretical model. Instead, the system applies class weight calculations to determine the classification results. This approach maintains the basic principles of Fuzzy-KNN while reducing computational requirements for real-time MQ-2 gas detection. The classification process and Fuzzy-KNN calculation are shown in Figure 6.

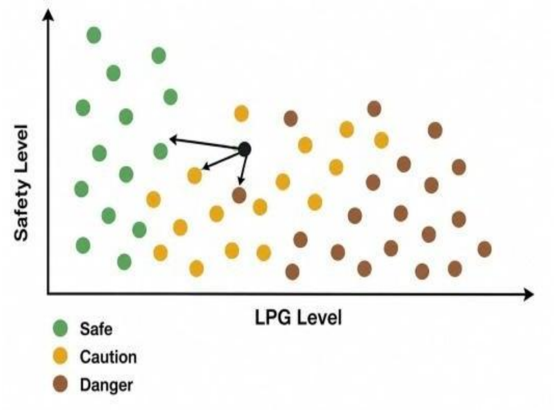


Figure 6. Illustration of Collecting 3 Neighbor Data

III. RESULTS AND DISCUSSION

A. System Design Results

The household gas hazard prediction system consists of hardware and software components. The hardware includes an MQ-2 sensor, Arduino Uno microcontroller, 16x2 LCD with I2C interface, active buzzer, and indicator LED. The MQ-2 detects flammable gases such as LPG, methane, and combustion gases, then sends sensor data to the Arduino for processing and hazard prediction. The system is designed for easy user understanding through LCD indicators ("SAFE," "ALERT," and "DANGER"), supported by LED and buzzer warnings.

B. Program Implementation Results

The Arduino IDE is used to develop the system software using C/C++ programming. The program begins with baseline calibration by storing average sensor readings in clean air as a reference to minimize detection errors. The sensor data is then converted into estimated gas concentration values in ppm using the following equation:

$$\textbf{\textit{GasPPM = (sensorValue − baseline) × 1.5}} \quad \textbf{(10)}$$

The Fuzzy-KNN algorithm is then applied by calculating the Euclidean distance between sensor input and the dataset, followed by selecting the three nearest neighbors ((K=3)). The fuzzy weight is calculated using (w=1/d), where (d) represents the neighbor distance. The class with the highest weight is selected as the prediction result, which then controls the LED and buzzer outputs according to the hazard category.

C. System Test Results

Tests were conducted to assess the system's ability to detect gases at different con centrations. The tests used household LPG gas, cigarette smoke, and paper smoke. Table I shows the test results.

TABLE I
TABLE OF TESTING THE MQ-2 SENSOR-BASED FUZZY KNN SYSTEM

No	Gas Concentration (PPM)	Sensor Voltage	LED/Buzzer Status	Prediction Classification	Description
1	0,0	0,79	OFF	SAFE	Clean air, no gases detected
2	34,5	1,55	OFF	SAFE	Very low gas concentration
3	76,5	1,14	OFF	SAFE	There are no indications of danger yet
4	94,5	1,17	OFF	SAFE	Stable sensor value in safe conditions
5	100,0	1,60	OFF	SAFE	Lower gas threshold detected
6	129,	1,63	OFF	SAFE	Gas levels increased slightly
7	150,0	1,89	OFF	SAFE	Gas detection is stable
8	193,5	1,49	OFF	SAFE	Minor fluctuations in the sesor
9	211,0	1,89	OFF	SAFE	Gas detection is stable
10	250,5	2,06	OFF	SAFE	Voltage values increased moderately
11	276,0	1,71	OFF	SAFE	Still within safe limits
12	297,0	1,85	OFF	SAFE	Almost approaching a state of alert
13	321,0	1,99	OFF	ALERT	Gas began to accumulate
14	351,0	2,30	OFF	ALERT	System detects potential gas increase
15	390,0	2,06	OFF	ALERT	Gas levels are quite high
16	433,5	2,26	OFF	ALERT	The alert threshold has been reached
17	451,5	2,41	OFF	ALERT	Early warning activated
18	498,0	2,37	OFF	ALERT	Gas is at dangerous levels
19	546,0	2,44	ON	DANGER	Alarm active, concentration exceeds safe limits
20	562,5	2,55	ON	DANGER	Hazardous gas conditions
21	620,5	2,60	ON	DANGER	High concentration gas
22	711,0	2,85	ON	DANGER	The potential for an explosion increases
23	880,5	3,12	ON	DANGER	Extremely high level of danger
24	921,0	3,30	ON	DANGER	Sensor in saturated condition
25	1170,0	4,12	ON	DANGER	Gas reaches maximum detectable level

TABLE II
SENSOR RESPONSE TIME TEST RESULTS DATA

No	Test Conditions	Response Time (s)	System Status	Remarks
1	Sudden gas exposure (LPG 1000 ppm)	0,80	LED/Buzzer ON	Quick response, active alarm
2	Moderate exposure (500 ppm LPG)	0,86	LED OFF	System stable, not yet active
3	Low exposure (LPG 200 ppm)	0,83	LED OFF	No alarm, sensor ready
4	Gas recovery	1,05	LED OFF	The sensor returns to its initial state

The response time test showed an average response of 0.83 seconds, demonstrating the system's ability to detect gas level changes and activate alarms in real time (<1 second). This result aligns with the MQ-2 sensor characteristics, which has a 0.5–1 second response time for flammable gases. The fast recovery time after gas removal indicates stable circuit performance and reliable Fuzzy-KNN processing of sensor data, as presented in Table II.

TABLE III
DATA ON THE RESULTS OF TESTING THE EFFECT
OF DISTANCE ON THE VOLTAGE OF THE MQ-2 SENSOR

No	**Distance (cm)**	**Sensor Voltage (V)**	**Description**
1	1	3,43	The sensor receives direct exposure to gas, maximum voltage
2	2	3,40	High gas concentration, slight decrease from 1 cm
3	3	2,68	Significant voltage drop due to gas diffusion
4	4	2,60	The sensor is still responding strongly to gas
5	5	2,55	Gas concentration begins to decrease
6	6	2,26	Voltage decreases as distance increases
7	7	2,21	The sensor is still able to detect gas stably
8	8	2,23	Slight fluctuations due to gas dispersion
9	9	2,02	Low voltage, gas dispersed in the air
10	10	1,50	The sensor is still detecting small amounts of gas

Based on Table 6, the MQ-2 sensor voltage decreases as the distance from the gas source increases. At 1 cm, the voltage reaches 3.43 V due to high gas exposure, while it decreases to

1.90 V at 10 cm because of gas diffusion. The significant voltage changes within 1–5 cm indicate the optimal detection range of the MQ-2 sensor, allowing the Fuzzy-KNN system to detect gas concentration changes quickly and accurately. Beyond 6 cm, the sensor maintains stable readings, showing good sensitivity and stability toward variations in distance and gas concentration.

D. Discussion

Testing of the gas hazard level prediction system based on the Fuzzy-KNN method was conducted to determine the optimal K value that produces the highest accuracy. Accuracy was calculated by comparing the system's prediction results with the actual labels of 46 test data, using the standard formula:

$$\text{Accuration } (\%) = \frac{\text{The Number of Test Data}}{\text{Number of Correct Predictions}} \times 100 \quad (4.1)$$

The test results show that the K value significantly affects Fuzzy-KNN classification performance. At K=1, the system achieved 84.8% accuracy but was sensitive to sensor fluctuations, especially near the SAFE-ALERT boundar (300–310 ppm), causing overfitting [11]. At K=2, accuracy increased to 86% due to fuzzy weighting, but classification remained unstable because of possible class ties. This confirms the use of odd K values to avoid ties in Fuzzy-KNN [14]. The best performance was obtained at K=3, achieving 87% accuracy with an average response time of 0.83 seconds. This value provided a balance between sensitivity and classification stability, making it the optimal K parameter for the system.

For K=5 and K=7, accuracy decreased to 82.6% and 78.3% due to oversmoothing, where excessive neighbors reduced the ability to distinguish local data variations [3]. Therefore, K=3 was selected as the optimal parameter, as shown in Figure 10, where accuracy increases until K=3 and decreases at higher K values.

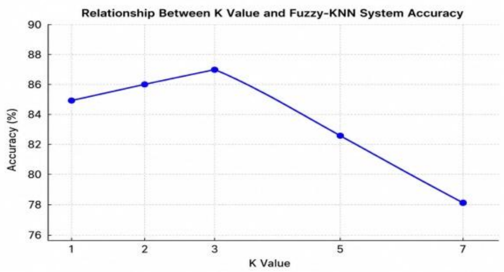


Figure 7. Relationship between K value and Fuzzy-KNN system accuracy

TABLE IV
RESULTS OF SYSTEM ACCURACY TESTING BASED ON K VALUE VARIATIONS

K Value	Correct Predictions (from 46 data points)	Accuracy (%)	System Characteristics
1	39	84,8%	Highly sensitive, overfitting
2	40	86,0%	Tends to tie, more stable
3	40	87,0%	Most optimal, sta-ble
5	38	82,6%	Oversmoothing, reduced sensitivity
7	36	78,3%	Too flat, not responsive enough

Based on Table IV, the K value directly affects system performance. Accuracy increases from K=1 and reaches the highest value at K=3, indicating optimal stability and classification capability. When K exceeds 3 (K=5 and K=7), accuracy decreases due to the oversmoothing effect. Therefore, K=3 is selected as the optimal parameter for the gas hazard prediction system.

IV. CONCLUSION

Based on the planning and testing results, the Fuzzy KNN-based gas hazard prediction system successfully classified gas conditions into three categories: safe, alert, and danger using the MQ-2 sensor. The system achieved 87% accuracy with an average response time of less than one second, enabling real-time operation on the Arduino Uno. The combination of fuzzy logic and K-NN improves handling of fluctuating sensor data, while LED and buzzer outputs function properly according to classification results. Therefore, the system is effective and reliable for early gas leak detection and warning in household environments.

V. RECOMMENDATION

This research is still limited to empirical sensor calibration and the amount of test data. Therefore, to improve the accuracy of the system, further research should per form calibration with standard gas measuring instruments and expand the amount of data sets. Additionally, the system can be developed by adding communication mod ules such as GSM or IoT so that hazard warnings can be sent remotely. Furthermore, the computational process of the Fuzzy-KNN method can be optimized by using a more advanced microcontroller.

REFERENCES

[1] D. Nurjanah, H. Handayani, en ..., "Rancang Bangun Pendeteksi dan Penetralisir Asap Rokok Dalam Ruangan Menggunakan Sensor Mq-2 dan Metode Fuzzy Logic", *... Student J. ...*, vol IV, bll 7–14, 2023.

[2] I. Muslem R, "Sistem Pendeteksi Kebocoran Gas Rumah Tangga Menggunakan Mq-2 Sensor Dan Mikrokontroler", *J. Tika*, vol 6, no 02, bll 58–64, 2021, doi: 10.51179/tika.v6i02.457.

[3] L. H. Zain, E. Setiawan, en H. Fitriyah, "Sistem Deteksi Kematangan Buah Mangga berdasarkan kandungan Gas NH3, C2H5OH dan VOCs menggunakan metode K-Nearest Neighbor (K-NN)", *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol 5(9), no 9, bll 3792–3798, 2021, [Online]. Available at: http://j-ptiik.ub.ac.id

[4] I. A. Rombang, L. B. Setyawan, en G. Dewantoro, "Perancangan Prototipe Alat Deteksi Asap Rokok dengan Sistem Purifier Menggunakan Sensor MQ-135 dan MQ-2", *Techné J. Ilm. Elektrotek.*, vol 21, no 1, bll 131–144, 2022, doi: 10.31358/techne.v21i1.312.

[5] A. Rofik, "Analisis Kebocoran Gas Lpg Dengan Menggunakan Sensor Gas Mq-2", *SinarFe7*, bll 206–208, 2021.

[6] S. W. I. Made, A. A. N. Gunawan, P. I. Ketut, en S. I. Ketut, "Design and Manufacture of LPG Gas Leak Detection Based on Arduino Uno Using MQ-2 Sensor", *Asian J. Res. Rev. Phys.*, vol 8, no 1, bll 27–32, 2024, doi: 10.9734/ajr2p/2024/v8i1155.

[7] E. Talakua, "Implementasi Sensor Gas Untuk Deteksi Jenis Bahan Bakar Kendaraan Bermotor Menggunakan Metode Fuzzy Mamdani", *Cyclotron*, vol 7, no 01, bll 76–83, 2024, doi: 10.30651/cl.v7i01.21445.

[8] N. Real-time en D. Winarti, "Sistem Monitoring Dan Deteksi Kebocoran Gas Lpg Berbasis Iot Dengan", vol 5, no 3, 2025, doi: 10.58794/jekin.v5i3.1644.

[9] Amir, Fazri, Novianda, Maulan, en Rahmat, "Sistem Pendeteksi Kebocoran Liquefied Petroleum Gas Menggunakan Metode Fuzzy Logic Mamdani Berbasis Internet of Things", *J. Teknol.*, vol 12, no 2, bll 151–158, 2020.

[10] H. Salem, M. Y. Shams, O. M. Elzeki, M. A. Elfattah, J. F. Al-amri, en S. Elnazer, "Fine-Tuning Fuzzy KNN Classifier Based on Uncertainty Membership for the Medical Diagnosis of Diabetes", *Appl. Sci.*, vol 12, no 3, bll 1–26, 2022, doi: 10.3390/app12030950.

[11] Z. Behdani, M. D.-J. of D. S. and Modeling, en undefined 2025, "Fuzzy K-Nearest Neighbor in Classification and Regression", *Jdscm.Atu.Ac.Ir*, vol 2, no 2, bll 217–244, 2024, [Online]. Available at: https://jdscm.atu.ac.ir/article_18938.html

[12] F. M. N. Akbar, "Metode KNN (K-Nearest Neighbor) untuk Menentukan Kualitas Air", *J. Tekno Kompak*, vol 18, no 1, bl 28, 2024, doi: 10.33365/jtk.v18i1.3241.

[13] M. Mailagaha Kumbure en P. Luukka, "A generalized fuzzy k-nearest neighbor regression model based on Minkowski distance", *Granul. Comput.*, vol 7, no 3, bll 657–671, 2022, doi: 10.1007/s41066-021-00288-w.

[14] A. A. Amer, S. D. Ravana, en R. A. A. Habeeb, "On enhancing data classification using local mean-based fuzzy K-nearest neighbor algorithms", *Adv. Data Anal. Classif.*, 2025, doi: 10.1007/s11634-025-00653-6.

[15] A. D. Churniawan, "Deteksi kondisi Udara dengan Sensor MQ Series dan MG Menggunakan Multi-Layer Perceptron", vol 4, no 1, bll 165–170, 2025, doi: 10.31284/p.semtik.2025-1.7012.

[16] F. Sharifonnasabi *et al.*, "Hybrid HCNN-KNN Model Enhances Age Estimation Accuracy in Orthopantomography", *Front. Public Heal.*, vol 10, no May, bll 1–14, 2022, doi: 10.3389/fpubh.2022.879418.