MACHINE LEARNING-BASED MULTI-ARRAY SENSOR SYSTEM FOR RICE CLASSIFICATION

M. Shahkhir Mozamir[1], B. Aboobaider[2] and A. Dahlan[3]

[1]Fakulti Teknologi Maklumat dan Komunikasi,
Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian
Tunggal, Melaka, Malaysia.
[2]Faculty Kecerdasan Buatan dan Keselamatan Siber, Universiti Teknikal Malaysia
Melaka, Hang Tuah Jaya, 76100 Durian
Tunggal, Melaka, Malaysia.
[3]Faculty of Computer Science, Universitas Amikom Yogyakarta, Indonesia,
Indonesia.

Corresponding Author's Email: [1]shahkhir@utem.edu.my

Article History: Received 3 November 2025; Revised 29 December 2025;
Accepted 29 December 2025

ABSTRACT: Ensuring consistent rice quality is crucial for maintaining consumer confidence and supporting standards within the rice industry. Traditional grading methods, which depend on manual visual inspection, are often subjective, time-consuming, and heavily reliant on expert judgment. The proposed system combines multiple sensors to capture essential physical and olfactory features of rice samples, while reducing power consumption through optimized data collection and processing. A Decision Tree algorithm is employed to analyze the sensor data and accurately classify rice categories. Experimental results reveal that the system achieves an 80% accuracy rate, confirming its potential as a dependable, low-power solution for real-time rice evaluation. Overall, this work demonstrates the effectiveness of integrating sensor fusion and AI-driven analysis to advance intelligent and sustainable food quality monitoring systems.

KEYWORDS: *Multi-Array Sensor, Machine Learning, Rice Classification, Decision Tree Algorithm, Smart Food Quality Monitoring*

1.0 INTRODUCTION

Rice is one of the most widely consumed staple foods globally, with numerous varieties distinguished by aroma, texture, and chemical composition. Accurate classification of rice based on its origin is critical for ensuring food authenticity, safeguarding consumer rights, and preventing fraudulent practices such as rice adulteration, which remains a persistent issue in countries like Malaysia. Conventional approaches for determining rice origin typically involve physicochemical analysis, spectroscopy, or expert sensory evaluation. While effective, these methods are often time-consuming and demand specialized expertise. Among the key features that differentiate rice varieties, aroma plays a pivotal role in influencing consumer preference and market value. However, traditional sensory evaluations depend on trained human panels, which can be inconsistent, subjective, and difficult to scale. This highlights a growing need for automated, reliable, and non-destructive methods for rice classification [1].

Machine learning algorithms are increasingly used in pattern recognition because of their ability to handle complex, high-dimensional datasets. Among these algorithms, Decision Trees (DTs) are particularly favored in embedded and real-time applications due to their simplicity, interpretability, and computational efficiency. A Decision Tree operates by creating a hierarchy of decision rules, splitting features to maximize class separation using metrics such as Information Gain or Gini Impurity. Unlike some statistical models, DTs do not assume any specific data distribution, making them well-suited for heterogeneous and non-linear data, including sensor signals. They are also ideal for low-resource platforms, as they require minimal computation and memory. Although Decision Trees can be prone to overfitting, careful feature selection and pruning strategies can mitigate this risk, making them a reliable choice for real-time classification in hardware-constrained environments [2].

Multi-array sensor technology, often known as an Electronic Nose (E-Nose), has emerged as an effective approach for classifying rice based on volatile organic compounds (VOCs). The E-Nose mimics the human sense of smell by detecting and interpreting complex odor patterns, allowing it to differentiate rice varieties according to their distinct aromatic characteristics. Advances in sensor design, signal processing, and pattern recognition algorithms have further improved the accuracy and reliability of E-Nose systems for rice classification. Compared to traditional gas chromatography techniques, the E-Nose provides rapid, real-time analysis, making it particularly suitable for large-scale rice quality monitoring in both research and industrial applications [3, 4].

Despite these advantages, several challenges remain. Factors such as sensor drift, environmental variations, and the dynamic nature of VOC emissions can affect the consistency and reliability of measurements. Additionally, post-harvest conditions including moisture content, aging, and packaging can alter odor profiles, potentially leading to incorrect classifications. Developing effective Decision Tree classifiers also requires well-balanced datasets and careful preprocessing to minimize risks of overfitting or misclassification. Consequently, optimizing Decision Tree parameters and maintaining sensor stability are critical steps in creating a reliable and reproducible system for rice origin classification [5].

The rest of this paper are arranged as follows; the related works has been described in Section 2. Section 3 will be discussed about the proposed algorithm Decision Tree (DT), and Section 4 is the experiment setup. Section 5 results and discussion and present the conclusion and potential areas of improvement for research in future.

2.0 RELATED WORKS

The application of multi-array sensor systems for classifying agricultural products has attracted considerable attention in recent years. Previous studies, including Wilson et al. [6], have shown that these sensors have rapidly advanced thanks to improvements in gas-sensing materials, signal processing, and machine learning techniques. Such developments demonstrate that volatile organic compounds (VOCs) emitted by biological products can act as distinctive chemical “fingerprints,” enabling the differentiation of agricultural goods including rice, fruits, and grains based on their unique odor profiles.

In the case of rice, multiple studies have demonstrated that multi-array sensor systems can reliably differentiate between rice varieties and quality grades. For instance, an early study titled Detection for Rice Odors and Identification of Varieties Based on Multi-Array Sensor

Technique [7] reported consistent accuracy in distinguishing rice types. Likewise, Lili Qan et al. [2] employed a multi-array sensor in combination with HS-SPME-GC-O-MS to examine the aromatic profiles of Wuchang rice, showing that the distinct sensor response patterns were capable of accurately separating varieties with varying aromatic intensities.

Recent research has further advanced rice classification by integrating multi-array sensor data with artificial intelligence to improve performance. For example, a 2022 study titled Rapid Assessment of Rice Quality Traits Using Low-Cost Digital Technologies demonstrated that combining a portable multi-array sensor with near-infrared spectroscopy and neural network algorithms enabled accurate classification of 17 commercial rice varieties [8]. This approach highlights the significant potential of AI-driven sensor fusion in making rice quality assessment faster, more precise, and scalable for industrial applications.

Despite these promising advances, several challenges persist. Sensor drift, temperature and humidity fluctuations, as well as post-harvest factors such as storage conditions and moisture content, can alter VOC emissions and compromise measurement consistency. Research on sensor stability has highlighted that long-term drift can significantly affect model reliability, emphasizing the need for recalibration and adaptive learning strategies. Similarly, studies on multi-array sensor systems for rice aging [9] underline the importance of controlled environments and optimized pattern recognition techniques to maintain consistent performance.

While many existing studies employ methods such as PCA, LDA, or neural networks for data classification, the application of Decision Tree algorithms in rice analysis using multi-array sensors remains limited. Decision Trees are valued for their simplicity and interpretability, yet they demand balanced datasets and careful parameter tuning to prevent overfitting. Moreover, ensuring sensor stability and consistent VOC detection is essential for achieving reliable predictions. This study aims to fill these gaps by investigating a Decision Tree-based multi-array sensor system that emphasizes energy efficiency and reproducible rice classification.

3.0 DECISION TREE (DT) ALGORITHM FOR RICE

CLASSIFICATION

The Decision Tree algorithm has gained increasing popularity in agricultural applications, including rice classification, due to its intuitive and effective approach to decision-making [10]. A Decision Tree mimics human reasoning by sequentially asking feature-based questions and following branches until it reaches a final classification, such as "Grade A" or "Grade B." Its transparent if-then rule structure makes it particularly attractive in contexts where explainability and interpretability are important, such as food quality assessment and product grading.

In rice classification, one of the key advantages of Decision Trees is their ability to manage a wide range of data types, including both continuous and categorical variables. For instance, Maspiyanti et al. [11] applied a Decision Tree model to classify paddy rice growth stages using hyperspectral imaging, achieving over 90% accuracy. This demonstrates that Decision Trees can perform effectively when meaningful features are extracted and properly preprocessed. Additionally, their non-parametric nature allows them to accommodate complex, non-linear decision boundaries that are common in biological and agricultural data.

Despite their advantages, several challenges arise when using Decision Trees for rice quality or origin classification. Data collected from multi-array sensors are often influenced by sensor drift, environmental variations, and temporal changes in volatile organic compounds (VOCs), which can lead to inconsistent readings and reduce the reliability of the model. Additionally, Decision Trees are prone to overfitting, especially when datasets are small or unbalanced. Without proper constraints, such as limiting tree depth or applying pruning techniques, the model may perform well on training data but fail to generalize to new samples [12].

To address these limitations, researchers have explored combining Decision Trees with ensemble or hybrid learning approaches. Algorithms like Random Forest and Gradient Boosting aggregate multiple Decision Trees to improve stability and predictive accuracy. For example, Saxena et al. [13] demonstrated that hybrid models could enhance rice variety classification by reducing sensitivity to noise while preserving interpretability. Furthermore, integrating Decision Trees with preprocessing methods such as Principal Component Analysis (PCA) or normalization can further improve classification performance, particularly when sensor data are affected by humidity or temperature fluctuations.

Beyond enhancing model robustness, Decision Trees provide valuable insights into feature importance, making them particularly suitable for use with multi-array sensor systems. Once trained, a Decision Tree can reveal which sensors or features such as aroma intensity, color values, the most significant role in differentiating rice grades. This capability is crucial for developing energy-efficient systems, as it allows the selection of the most informative sensors while minimizing unnecessary measurements. Such an approach supports the design of low-power, AI-assisted rice quality monitoring systems capable of real-time operation in industrial environments [14].

Overall, Decision Trees remain one of the most practical and interpretable algorithms for rice classification, especially in early-stage research or prototype development. Although advanced methods like deep learning may achieve higher accuracy, the Decision Tree's explainability and low computational requirements make it better suited for embedded or IoT-based sensor platforms. By combining Decision Trees with sensor fusion, data preprocessing, and pruning techniques, researchers can create robust, transparent, and efficient models for automated rice grade classification [15].

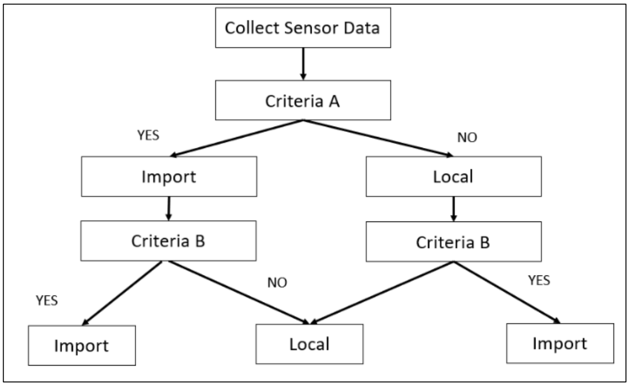

Figure 1: Logical decision tree algorithm for rice grade classification

The Figure 1 illustrates a decision-making flowchart for processing sensor data. The process begins with "Collect Sensor Data", followed by an evaluation based on Criteria A. If Criteria A is satisfied (YES), the workflow moves along the “Import” branch; if not (NO), it proceeds to the “Local” branch. This initial decision effectively separates the data based on whether it meets the first condition.

Following this, both branches undergo assessment based on Criteria B. In the Import path, meeting Criteria B (YES) maintains the final classification as Import, while failing to meet it (NO) results in Local. Conversely, in the Local path, if Criteria B is not satisfied (NO), the classification remains Local, but if it is met (YES), the outcome is Undefined.

This hierarchical and conditional structure enables a stepwise classification of sensor data, providing clear rules for labeling samples as Import, Local, or Undefined based on two sequential criteria. It illustrates how decision rules can be applied to multi-array sensor data to achieve consistent and interpretable classification outcomes.

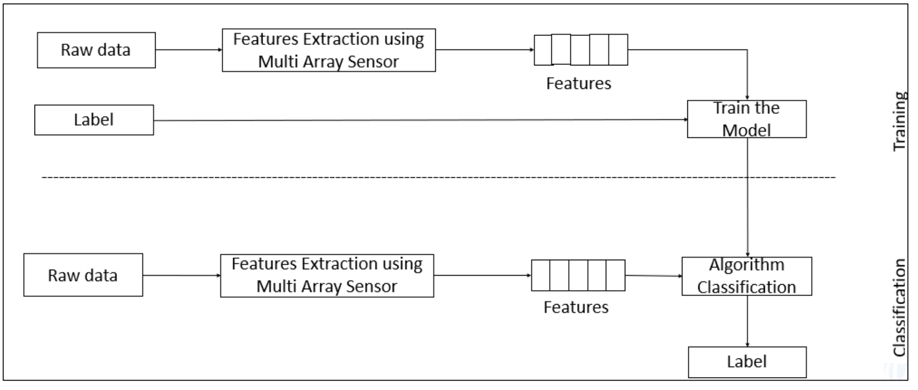

Figure 2: Training and classification process for rice grade classification

The figure 2 illustrates the workflow of a machine learning model for classification using sensor data, which is divided into two main stages: Training and Classification.

During the Training phase (upper section), raw data from a multi-array sensor undergoes feature extraction, where relevant characteristics or patterns are identified and converted into numerical features. These features, together with their corresponding labels representing known classes, are used to train the model. Through this process, the algorithm learns the relationships between the features and their labels, resulting in a trained model capable of recognizing or predicting similar patterns in new data.

In the Classification phase (lower section), new raw sensor data is processed using the same feature extraction method. The extracted features are then input into the trained model, which outputs a predicted label. Applied to rice classification, this workflow enables the system to identify and categorize different rice types or quality grades using sensor data, such as images, spectral signals, or texture patterns. The model first learns from labeled samples of known rice varieties (e.g., Basmati, Jasmine, or local types), and then, when new samples are introduced, it automatically predicts their type or quality, allowing for accurate, efficient, and automated rice classification.

4.0 EXPERIMENTAL SETUP

To evaluate the Rice Grade classification system using a Multi-Array Sensor (MAS) setup, experiments were conducted using the developed prototype, as shown in Figure 3. The prototype collects real-time sensor data from a variety of rice samples, including different grades and types. Each sample is placed within a sensing chamber, where multiple sensors such as color and odor sensors capture relevant parameters.

The collected data are processed and analyzed using a Decision Tree (DT) algorithm to determine the rice grade or category. The results are then compared with standard reference grades to assess the system's accuracy, reliability, and consistency. By integrating both visual and olfactory features, the multi-array sensor system improves classification performance and enhances the ability to distinguish between imported and local rice.

Two rice types are local and imported are used for training and classification. During the training phase, sensor data from both types are processed to extract key features, including color intensity and odor concentration, which serve as inputs to the Decision Tree model. The algorithm learns the distinguishing patterns between local and imported rice based on these features.

Once training is complete, the system enters the classification phase, where new or unknown rice samples are tested using the same MAS setup. The trained model analyzes sensor responses and automatically classifies the samples according to their similarity with the trained data. This approach allows the system to identify the origin of rice with high precision and repeatability, demonstrating its potential for automated, real-time rice grade classification.

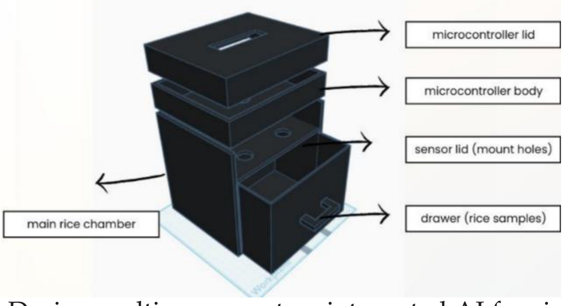

Figure 3: Design multi array system integrated AI for rice classification

Table 1: Parameter setup for rice grading using multi array sensor

Rice sample	Weight of sample	Odor sensor	Colour sensor type	Time taken for testing	Data library
2 Types (Import and Local)	*100 Gram*	*4 MQ Series Sensor*	*TCS3200*	*60 Seconds*	*4800 raw data*

5.0 RESULT AND DISCUSSION

The table below represents a confusion matrix for a classification model used to distinguish between Local and Imported rice. From the matrix, the model correctly classified 9 Local rice samples and 15 Imported rice samples. However, it misclassified 6 Local rice samples as Imported, while no Imported rice samples were misclassified as Local. These values give us:

- True Positive (TP) = 15 (Actual Imported predicted as Imported)
- True Negative (TN) = 9 (Actual Local predicted as Local)
- False Positive (FP) = 6 (Actual Local predicted as Imported)
- False Negative (FN) = 0 (Actual Imported predicted as Local)

Table 2: Results of the experiment

Metric	Local as positive	Imported as positive	Overall
Accuracy	-	-	*80%*
Precision	*100%*	*71.4%*	-
Recall	*60%*	*100%*	-
F1-Score	*75%*	*83%*	-
Specificity	*100%*	*60%*	-

The evaluation results presented in Table 2 indicate that the multi-array sensor system achieved an overall accuracy of 80% in classifying rice as either local or imported. This demonstrates that the system can reliably distinguish rice origin based on color and odor features, although some inconsistencies remain. The findings suggest that the classifier effectively captures the primary sensory differences between the two categories, validating the potential of sensor-based rice authentication.

For local rice, the system achieved 100% precision and 60% recall, resulting in an F1-score of 75%. This indicates that all samples identified as local were correctly classified, but 40% of actual local samples were misclassified as imported. The specificity of 100% confirms that the system did not falsely label imported rice as local. These results highlight that while the classifier is highly precise, it is not sufficiently sensitive to subtle characteristics of local rice, suggesting the need for improved detection of finer sensory differences.

In contrast, for imported rice, the model achieved 71.4% precision, 100% recall, and an F1-score of 83%. The system successfully identified all imported samples but occasionally misclassified some local rice as imported, as reflected by the 60% specificity. This indicates that the classifier is more sensitive to imported rice features, likely due to their stronger or more distinctive odor and color profiles. Overall, the system shows promising performance, with opportunities to enhance balance and sensitivity across both rice types through refined feature extraction or dataset augmentation.

The results also suggest that feature dominance within the dataset may favor imported rice characteristics. Imported varieties often exhibit brighter colors and more pronounced aromas, which can influence the classifier's decision boundaries, potentially causing overfitting to these features while underrepresenting subtle variations in local rice. Incorporating additional training samples or adjusting feature weighting could help improve generalization across both categories.

Future improvements could focus on sensor calibration and fusion strategies within the multi-array setup. By optimizing how color and odor sensors interact and process signals, the system could better differentiate overlapping sensory patterns. Additionally, integrating machine learning models such as Support Vector Machines (SVM) or Random Forests could further enhance classification robustness. Such enhancements would not only improve accuracy and recall for local rice but also contribute to a more reliable and scalable smart sensing platform for agricultural quality assessment and traceability.

3.0 Conclusion and Future Work

In conclusion, the multi-array sensor system integrating color and smell analysis has shown promising capability in classifying rice as either local or imported, achieving an overall accuracy of 80%. The system demonstrates high precision for local rice and excellent recall for imported rice, confirming its effectiveness in detecting distinct sensory characteristics between both categories. However, the imbalance between precision and recall across classes indicates that while the system is confident in its classifications, it still requires better sensitivity toward subtle features of local rice.

For future work, several enhancements can be explored to further improve system performance. First, expanding the training dataset with more diverse rice samples from different regions could help reduce bias and improve generalization. Second, applying advanced machine learning models such as Random Forest, SVM, or deep learning-based fusion networks can enhance pattern recognition across multi-sensor data. Finally, optimizing sensor calibration and signal preprocessing techniques will improve consistency and reliability in real-world conditions, paving the way for a robust and intelligent rice classification system suitable for industrial and agricultural applications.

ACKNOWLEDGMENTS

The authors extend their heartfelt gratitude to Universiti Teknikal Malaysia Melaka (UTeM) and Universitas Amikom for their steadfast support and encouragement throughout this research endeavour. The facilities, resources, and enriching academic environments provided by UTeM and Universitas Amikom have been pivotal to the successful completion of this work. The authors also sincerely appreciate the invaluable opportunity to grow both

academically and professionally through this experience.

REFERENCES

[1] K. Gopalakrishnan and J. D. R. Vivek, “Rice grain classification using image processing technique,” Int J Health Sci (Qassim), vol. 6, no. May, pp. 919–928, 2022, doi: 10.53730/ijhs.v6ns5.8779.
[2] L. Qian et al., “Discrimination of different varieties of rice in Wuchang area based on E-nose and HS-SPME-GC-O-MS,” Food Chem X, vol. 29, p. 102779, 2025, doi: 10.1016/j.fochx.2025.102779.
[3] R. Trihaditia et al., “Design and Development of an Electronic Nose as a Tool for Detecting the Aroma Quality of Pandanwangi Rice,” Int J Adv Life Sci Res, vol. 7, no. 2, pp. 196–203, 2024, doi: 10.31632/ijalsr.2024.v07i02.015.
[4] J. Tan and J. Xu, “Applications of electronic nose (e-nose) and electronic tongue (e-tongue) in food quality-related properties determination: A review,” Artificial Intelligence in Agriculture, vol. 4, pp. 104–115, 2020, doi: 10.1016/j.aiia.2020.06.003.
[5] Z. Ye, Y. Liu, and Q. Li, “Recent progress in smart electronic nose technologies enabled with machine learning methods,” Sensors, vol. 21, no. 22, pp. 23–26, 2021, doi: 10.3390/s21227620.
[6] A. D. Wilson, “Advances in electronic-nose technologies for the detection of volatile biomarker metabolites in the human breath,” Metabolites, vol. 5, no. 1, pp. 140–163, 2015, doi: 10.3390/metabo5010140.
[7] P. Kantakaew et al., “Influence of laminar and turbulent flow on signal response of gas sensors in electronic nose chamber for detecting rancid odor in brown rice,” Eng. Appl. Sci. Res., vol. 51, no. 2, pp. 224–234, 2024, doi: 10.14456/easr.2024.22.
[8] A. Aznan et al., “Rapid Assessment of Rice Quality Traits Using Low-Cost Digital Technologies,” Foods, vol. 11, no. 9, 2022, doi: 10.3390/foods11091181.
[9] J. R. R. Kumar and P. Chouksey, “Gas Sensor Array Drift in an E-Nose System: A Dataset for Machine Learning Applications,” IJRITCC, vol. 11, no. 6, pp. 167–171, 2023, doi: 10.17762/ijritcc.v11i6.7343.
[10] K. G. Liakos et al., “Machine Learning for Quality Control in the Food Industry: A Review,” Foods, vol. 14, no. 19, p. 3424, 2025.
[11] F. Maspiyanti et al., “Paddy Growth Stages Classification Based on Hyperspectral Image Using Feature Selection Approach,” 2015.
[12] Y. J. Chew et al., “Decision Tree Pruning with Privacy-Preserving Strategies,” Electronics, vol. 14, no. 15, 2025, doi: 10.3390/electronics14153139.
[13] P. Tiwari et al., “A Survey of Localization Methods and Techniques,” 2015.
[14] E. Elbasi et al., “Optimizing Agricultural Data Analysis Techniques,” Applied Sciences, vol. 14, p. 8018, 2024.
[15] K. G. Liakos et al., “Machine learning in agriculture: A review,” Sensors, vol. 18, no. 8, pp. 1–29, 2018.
[16] A. Çifci and İ. Kırbaş, “Fusion of machine learning and explainable AI for enhanced rice classification,” European Food Research and Technology, vol.

251, 2025.

[17] J. D. Barea-Ramos et al., “Detection of Aroma Profile in Spanish Rice Paella during Socarrat Formation by Electronic Nose,” Chemosensors, vol. 11, no. 6, p. 342, 2023.

[18] K. Cheng et al., “Characterization and feature selection of volatile metabolites in Yangxian pigmented rice varieties through GC-MS and machine learning algorithms,” Frontiers in Nutrition, vol. 12, 2025.

[19] Z. Xiaolong et al., “Analysis of volatile substances in different freshness rice by HS-SPME-GC-MS and Heracles NEO ultra-fast gas-phase electronic nose,” China Brewing, vol. 44, no. 1, 2025.

[20] A. Jana et al., “Fragrance Measurement of Scented Rice Using Electronic Nose,” 2023.