VOL. 12. NO. 2 NOVEMBER 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v12i2.7886

JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

# IOT-CLOUD BASED FISH QUALITY IDENTIFICATION IN DISTRIBUTION USING RANDOM FOREST

**Muhammad Shodiq[1*], Agus Priyono[1], Mutsna Sa Yu Zakka[1]**

University of Muhammadiyah Lamongan, Lamongan, Indonesia[1]
https://umla.ac.id[1]
shodiqmuhammad13@gmail.com[*], aguspriyono2012@gmail.com, mtsnasyzkk@gmail.com

(*) Corresponding Author
(Responsible for the Quality of Paper Content)

The creation is distributed under the Creative Commons Attribution-NonCommercial 4.0 International License

**Submitted:** 11 December 2025, **Revised:** 14 October 2026, **Accepted:** 06 November 2026, **Published:** 13 November 2026.

***Abstract***— *The decline in fish quality during distribution is a critical issue within the fisheries supply chain, as it affects consumer safety, logistical efficiency, and economic outcomes. This study aims to develop a real-time fish quality monitoring system by integrating Internet of Things (IoT) technology, cloud computing, and the Random Forest algorithm to produce more accurate quality classifications and support decision-making throughout the distribution process. The research was conducted by collecting temperature, ammonia ($NH_3$), and hydrogen sulfide ($H_2S$) data using sensors connected to a microcontroller, with measurements transmitted continuously to a cloud-based platform. The data were analyzed through the application of a Random Forest model, which was fine-tuned by testing various parameter combinations to determine the optimal configuration. The results showed that the optimal model—with n_estimators of 50, max_depth of 30, min_samples_split of 2, and min_samples_leaf of 1— achieved a mean five-fold cross-validation accuracy of 92.6%. Feature importance analysis revealed that ammonia was the most dominant variable in determining fish quality, followed by temperature and $H_2S$. These findings demonstrate that combining IoT sensor technology and machine learning models can improve the accuracy of fish quality monitoring and increase the effectiveness of distribution management across the fisheries supply chain.*

***Keywords***: *Cloud computing, fish quality, internet of things, random forest, supply chain management.*

## INTRODUCTION

The food supply chain is an interconnected system that shapes the accessibility and availability of food resources across various countries [1], including fisheries. The fisheries sector plays a crucial role in the national economy, contributing to food security and serving as a key driver of economic growth [2], as well as providing employment opportunities in many developed and developing countries [3]. Food security is critically important on a global scale, as it delivers significant health, social, and economic benefits for human populations [4]. Lamongan Regency is one of the regions in East Java that demonstrates strong potential for strategic fisheries and marine development, achieving a capture fisheries production of 149,452 tons in 2022. The fisheries sector faces challenges in achieving long-term sustainability due to political, economic, social, technological, and legal factors [3], as well as issues related to distribution [5]. This issue is crucial, considering that the decline in fish quality during distribution is still difficult to monitor quickly and accurately. Therefore, one potential approach is to utilize data-driven technologies, such as machine learning, to objectively identify fish freshness levels. Based on these challenges, this study addresses the following research question: How can IoT sensor data and the Random Forest algorithm be utilized to accurately classify fish quality during the distribution process in real time? To answer this research question, this study aims to develop a machine learning-based classification model to support fish quality monitoring, thereby contributing to improved distribution chain efficiency and the sustainability of the fisheries sector.



Accredited Rank 2 (SINTA 2) based on Decree No. 225/E/KPT/2022 issued by Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia, dated December 7, 2022. Published by LPPM Universitas Nusa Mandiri.

VOL. 12. NO. 2 NOVEMBER 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v12i2.7886

Distribution is an essential activity and a key determinant of success in the fisheries sector [5]. Fish is one of the fresh commodities that is highly perishable [6]. The problem is that fish deterioration is caused by several activities occurring from the point of capture to the consumer, particularly within the distribution chain [7], and this deterioration affects fish prices [5] and leads to economic losses [8]. The distribution chain plays a crucial role in determining quality control processes. Distribution to consumers involves multiple intermediaries, resulting in a longer and more extensive supply chain [1], which may cause a decline in fish quality if not properly managed. Therefore, the urgency of this research lies in the need for an effective method to ensure fish quality throughout the distribution process until it reaches consumers [8] [7].

Despite the importance of fisheries sustainability, one practical challenge affecting fish distribution is the inability to continuously monitor fish quality during transportation. Conventional inspection methods are generally manual and periodic, making them less effective in detecting quality degradation in real time. Therefore, an intelligent monitoring approach capable of automatically interpreting sensor data is needed to support timely decision-making during the distribution process. Supply Chain Management (SCM) consists of handling and optimizing all aspects of the supply chain processes for goods and services. Traditional SCM faces several issues, such as distribution delays and challenges in delivering real-time information. Modern SCM benefits greatly from advancements in the Internet of Things (IoT).

IoT technology can automate supply chain processes to achieve maximum operational efficiency. IoT devices within the supply chain track and trace shipments using GPS technology [9] [10]. Cloud computing is a technology that provides secure and reliable data storage with low hardware configuration requirements, while also facilitating data sharing among users through extensive network access [11]. Furthermore, the integration of IoT into SCM improves efficiency, security, and sustainability, and supports policy formulation to ensure the long-term sustainability of the fisheries sector [3] [12] [13]. Machine learning constitutes a major development in computing technology, supporting decision-making by employing prediction and classification techniques derived from historical data [14].

Previous research developed a real-time fish freshness prediction platform using machine learning models. Various parts of grass carp were stored at different temperatures, and five freshness indicators were measured. Six machine learning models were tested, including Random Forest, and the best-performing model for each indicator was integrated into a single intelligent platform. The results showed that the platform was able to predict freshness with high accuracy ($R^2$ 0.9552–0.9734), extremely fast processing (0.06 seconds), and could classify fish as Feasible, Moderate, and Not Feasible [15]. Another study entitled Smart Technologies in Aquaculture: An Integrated IoT, AI, and Blockchain Framework for Sustainable Growth examines fish spoilage detection using Terahertz spectroscopy by measuring gases produced from microbial activity, particularly ammonia ($NH_3$) and hydrogen sulfide ($H_2S$). The findings indicate that the concentrations of these gases increase over storage time and can be utilized as indicators to determine whether fish are still suitable or unsuitable for consumption [16].

Furthermore, the study Freshness Analysis of $NH_3$, Color, pH, and $H_2S$ on Salmon Spoilage Using the PCA Method discusses fish quality assessment through the measurement of spoilage gases, specifically hydrogen sulfide ($H_2S$) and ammonia ($NH_3$), which are generated by microbial activity. An increase in the levels of these gases signifies a decline in fish quality and can be used as an indicator to determine its suitability for consumption [17]. Another study developed a real-time water quality monitoring system for freshwater shrimp ponds based on IoT, cloud computing, machine learning, and a web application.

The system measures key parameters such as pH, temperature, TDS, EC, and salinity using sensors connected to a microcontroller. The data are analyzed using regression to predict conditions for the following day and classification algorithms to determine shrimp production levels. The results demonstrated high prediction accuracy ($r^2$ = 0.94) and a production classification accuracy of 97.84% (Random Forest). The Smart Aquaculture Analytics web application enables remote monitoring, visualization of historical data, prediction, classification, and automated notifications [18].

However, previous studies have primarily focused on freshness prediction or aquaculture monitoring, while limited attention has been given to integrating IoT sensing, cloud computing, and the Random Forest algorithm into a unified framework for real-time fish quality classification during the distribution process. Therefore, this study focuses on the development of an integrated IoT-cloud-based monitoring system that employs the Random Forest algorithm to classify fish quality into Feasible, Moderate, and Not Feasible categories using temperature, $NH_3$, and $H_2S$ sensor data.

CC BY NC
Accredited Rank 2 (SINTA 2) based on Decree No. 225/E/KPT/2022 issued by Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia, dated December 7, 2022. Published by LPPM Universitas Nusa Mandiri.

The main contribution of this research lies in combining IoT sensing, cloud computing, and the Random Forest algorithm to develop a real-time decision-support system for monitoring fish quality throughout distribution. Furthermore, the proposed model is optimized through hyperparameter tuning and validated using cross-validation to obtain the most reliable classification performance. The resulting system is expected to support operational decision-making while improving distribution efficiency, with supply chain sustainability regarded as a practical implication rather than the primary objective of this study.

## MATERIALS AND METHODS

### Random Forest (RF)

Random Forest is constructed by combining the outputs of multiple decision trees, or by aggregating the predictions of several decision trees, resulting in a more robust and consistent method. Therefore, RF is classified as an ensemble method. A single decision tree often experiences overfitting to the training data, especially when the tree grows too deep. However, RF can reduce overfitting by introducing randomness through two mechanisms [19]:

1 Bootstrapping: Each tree is trained using a different randomly selected subset of the data.
2 Random Feature Selection: At each splitting process, only a subset of features is randomly selected for consideration.

Within this study, the Random Forest algorithm serves as the classification method for categorizing fish quality into three groups: Feasible, Moderate, and Not Feasible. The classification process utilizes a three-dimensional feature vector consisting of ammonia, hydrogen sulfide, and temperature readings. To ensure the model's generalizability and prevent overfitting, a 5-fold cross-validation (K-Fold 5) procedure is implemented, where the dataset is partitioned into five subsets to iteratively validate the model's performance on unseen data.

To achieve the most optimal configuration, an extensive hyperparameter tuning process was conducted. This involved evaluating various combinations of key parameters, including the number of trees, the maximum depth of each tree, the minimum number of samples required to split an internal node, and the minimum number of samples required at a leaf node. The final performance of the tuned model is not only assessed by its accuracy but also rigorously validated through Precision, Recall, and F1-Score metrics. This multi-metric evaluation ensures that the model provides a balanced and reliable detection capability across all fish quality classes, which is essential for real-time supply chain monitoring.

### Data Sources

The data collected included temperature, hydrogen sulfide ($H_2S$), and ammonia ($NH_3$) generated from fish odor and measured using an IoT-based monitoring system, as well as real-time location data obtained through a Neo-6M V2 GPS module integrated into the system to monitor the fish distribution process. The IoT monitoring system was developed using an ESP32 microcontroller integrated with a DHT11 temperature sensor, an MQ-135 gas sensor for $NH_3$ detection, and an MQ-136 gas sensor for $H_2S$ detection. Prior to data collection, the MQ-135 and MQ-136 sensors were stabilized in clean air for approximately 24 hours to allow the sensing elements to reach stable operating conditions and establish baseline responses. Due to the sensitivity of MQ-series sensors to environmental conditions and other gases, the measured $NH_3$ and $H_2S$ concentrations were treated as sensor-based indicators of fish quality deterioration rather than absolute reference measurements, with potential effects of humidity, cross-sensitivity, and measurement variability acknowledged as limitations of the sensing system.

The dataset was collected from Selaroides leptolepis (yellowstripe scad) transported in a single insulated fish box distribution conditions. Temperature was measured in degrees Celsius (°C), while $NH_3$ and $H_2S$ concentrations were recorded in parts per million (ppm) based on the sensor conversion procedure. Sensor data were automatically acquired at 1-minute sampling intervals throughout the transportation process and transmitted in real time to the IDCloudHost cloud platform for data storage and monitoring. The embedded system was developed using Arduino IDE version 2.3.6, while the Random Forest model was implemented with a fixed random seed of 42 to ensure the reproducibility of the experimental results. This approach enabled the acquisition of integrated real-time data to support continuous monitoring and analysis of fish quality throughout the distribution process.

The dataset consisted of 171 sensor observations collected at one-minute intervals during a single measurement session using one insulated fish box. During the session, individual Selaroides leptolepis (yellowstripe scad) specimens were measured in the box for several minutes, resulting in multiple sensor readings from a given fish. Measurements were obtained from fish


Accredited Rank 2 (SINTA 2) based on Decree No. 225/E/KPT/2022 issued by Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia, dated December 7, 2022. Published by LPPM Universitas Nusa Mandiri.

VOL. 12. NO. 2 NOVEMBER 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v12i2.7886

representing three freshness categories—Feasible, Moderate, and Not Feasible which were determined through organoleptic assessment of appearance, color, flesh elasticity and texture, odor, and other relevant sensory attributes. The dataset comprised 56 Feasible, 95 Moderate, and 20 Not Feasible observations. Temperature and sensor-derived values associated with $NH_3$ and $H_2S$ were used as predictive features in the Random Forest classification model, while the organoleptic assessment provided the class labels.

Therefore, the 171 observations represent sensor readings rather than 171 independent fish specimens. Because multiple readings could originate from the same fish and the measurements were collected within a single session using one box, the observations may share fish-specific and session-specific effects. The current dataset provides an initial evaluation under the experimental conditions of this measurement session and does not establish generalizability across independent measurement sessions, fish boxes, shipments, or transportation conditions. Additional evaluations involving independently identified fish samples and multiple measurement sessions are required to determine the generalizability of the developed model.

The reference labels “Feasible,” “Moderate,” and “Not Feasible” were determined through organoleptic assessment based on three main criteria: eye appearance, flesh texture, and odor. The assessment was conducted by one assessor with practical experience in fish selling, who evaluated the fish at the beginning and at the end of the distribution process. Fish with clear and normal-looking eyes, firm flesh, and a fresh odor were classified as “Feasible.” Fish showing early signs of deterioration in one or more of these attributes, while not exhibiting pronounced spoilage characteristics, were classified as “Moderate.” Fish exhibiting evident deterioration, such as cloudy or sunken eyes, soft flesh, and an unpleasant or putrid odor, were classified as “Not Feasible.”

The organoleptic assessment provided the basis for assigning the reference labels, while the $NH_3$- and $H_2S$-related sensor indicators obtained from the MQ-135 and MQ-136 sensors, respectively, were used as predictive features in the Random Forest model. However, because the assessment involved a single assessor and was conducted only at the beginning and end of distribution, the labels may be subject to observer-related bias and may not fully capture changes in fish quality between assessment points. These limitations should be considered when interpreting the model's performance [17] [21].

**Data Analysis**

Data analysis was conducted through several structured stages, beginning with centralized processing and storage in the cloud. After this initial processing, the data were analyzed using the Random Forest algorithm to classify fish quality levels. The classification results were then utilized to build a decision model capable of providing automatic and informative recommendations regarding fish conditions during the distribution process. With this approach, the study not only produces an automated and real-time fish quality monitoring system but also provides faster, more accurate, and data-driven decision-support capabilities. This contributes directly to maintaining fish quality throughout all stages of the fisheries supply chain more effectively and efficiently.

**Model Performance Evaluation**

The model developed using the training data needs to be evaluated by running it on test data that have not been previously used. If the model is able to produce accurate predictions, it can be considered a good model. The following equation represents the model performance evaluation or model accuracy [19]:

$$Accuray = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

This matrix contains four main categories: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). These four categories are highly useful for calculating the model’s accuracy. To ensure the reliability of the model and mitigate the risk of overfitting, the training process employed a k-fold cross-validation method. Specifically, a 5-fold cross-validation with a shuffling mechanism was applied to the training dataset. In this procedure, the data was randomly divided into five equal subsets (folds).

The Random Forest model was trained on four of these folds and validated on the remaining single fold, with the process repeating five times so that each fold served exactly once as the validation set. The cross-validation results, presented as accuracy, precision, recall, F1-score, and cv_std in Table 3, represent the average accuracy and the standard deviation across these five iterations, thereby confirming the stability and robustness of the model's performance under various parameter configurations. The confusion matrix was generated

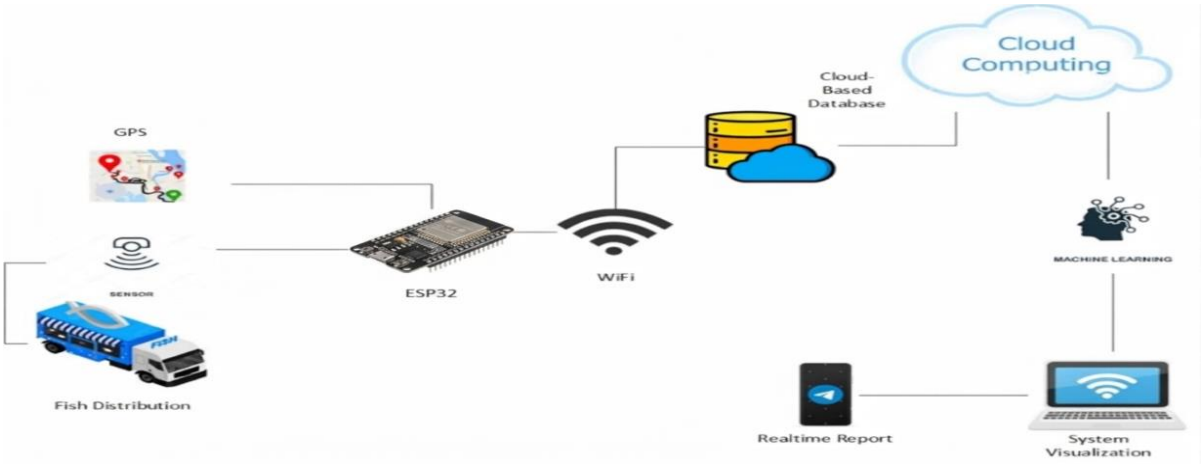
CC BY NC
Accredited Rank 2 (SINTA 2) based on Decree No. 225/E/KPT/2022 issued by Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia, dated December 7, 2022. Published by LPPM Universitas Nusa Mandiri.

from the cross-validation predictions across the available dataset. Accordingly, the evaluation reported in this study is based on five-fold cross-validation rather than on a separate independent test set. The generalizability of the model to independently collected data remains to be established.

**Research Design**

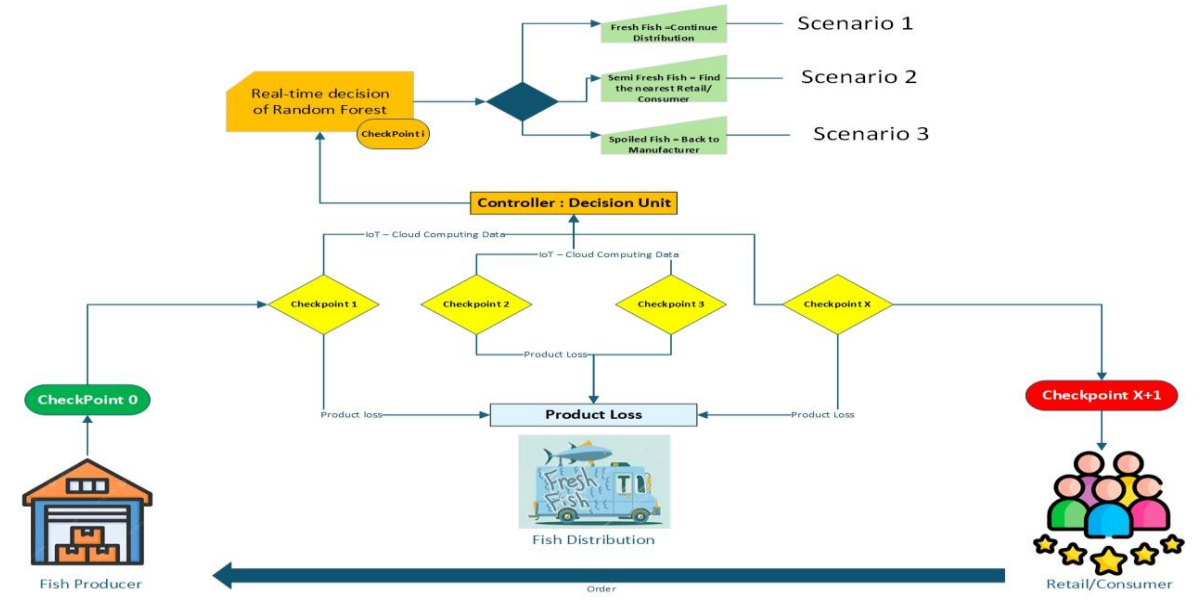
This study employs a system development design for Supply Chain Management based on IoT, Cloud Computing, and Machine Learning (Random Forest) to monitor fish quality within the distribution chain. The overview of the research is illustrated in Figure 1 (System Architecture), Figure 2 (Product and Information Flow within the System), and Figure 3 (Random Forest Data Training).

In Figure 1, it is illustrated that the transportation system used for fish delivery is equipped with sensors consisting of a temperature sensor, $H_2S$ gas sensor, and $NH_3$ gas sensor, all installed using an ESP32, which is also equipped with a GPS module. This device is then connected to a network. The network-connected device collects real-time fish odor data and stores it in the cloud. Afterward, the data are processed using the Random Forest algorithm and visualized through a system. Finally, the system sends automatic notifications via Telegram when fish quality begins to decline.




Source: (Result Research, 2025)

Figure 1. System Architecture




Source: (Result Research, 2025)

Figure 2. Product and Information Flow Within the Proposed Framework

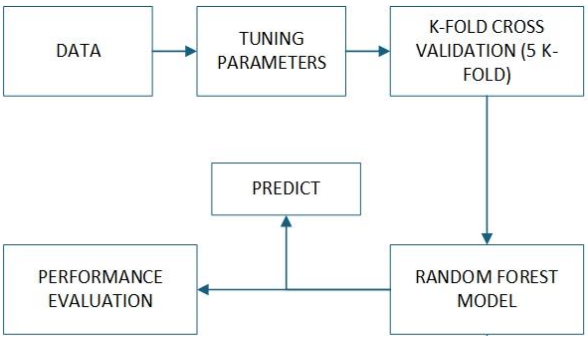


Accredited Rank 2 (SINTA 2) based on Decree No. 225/E/KPT/2022 issued by Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia, dated December 7, 2022. Published by LPPM Universitas Nusa Mandiri.

VOL. 12. NO. 2 NOVEMBER 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v12i2.7886

Figure 2 shows the fish distribution flow, starting from the producer and ending at the consumer or retail destination. Before the fish actually reaches the consumer, its quality is periodically monitored by the system at several distribution points by capturing data on ammonia and hydrogen sulfide gases, as well as the temperature of the cooler. In other words, the system continuously provides information on the fish's condition, either every minute or every hour, depending on the defined monitoring requirements. If the fish quality remains good throughout the distribution process, the system recommends continuing distribution as usual (scenario 1). However, if the system detects a decline in quality, the recommended decision is to divert distribution to the nearest retail location so the product can be sold immediately (scenario 2). Meanwhile, if the fish quality is detected as unfit for consumption, the system automatically recommends returning the fish to the producer for further processing (scenario 3). The fish quality classification process is carried out using Random Forest. Through this mechanism, the distribution process becomes more efficient while ensuring that the fish's quality is maintained throughout the journey.

Figure 3 shows the decision-making process or quality classification using a random forest. The figure illustrates the flow of the Random Forest model development and evaluation process in this study. The process begins with data collection, followed by parameter tuning to test parameter values. The data is then processed using the K-Fold Cross Validation method with 5 folds to ensure stable model performance and avoid overfitting. The training data are used to build the Random Forest model, which is an algorithm based on an ensemble of decision trees capable of performing classification or prediction tasks. Once the model is developed, it is utilized to generate predictions on both testing data and new data. The prediction results are then evaluated in the performance evaluation stage by comparing them with the actual data to assess the model's performance using metrics such as accuracy, precision, recall, and F1-score.




Source: (Result Research, 2025)

Figure 3. Random Forest Data Training

## RESULTS AND DISCUSSION

The following are the measurement results obtained from the sensors, which consist of a temperature sensor, a hydrogen sulfide ($H_2S$) gas sensor, and an ammonia ($NH_3$) sensor, as shown in Figure 4. The sensor readings collected during the experimental session were used to examine the relationship between the sensor-derived features and the assigned fish freshness categories under the experimental conditions.

**Data Sensor**

Show 10 entries Search:

| ID | Waktu | Amonia (ppm) | Sulfida (ppm) | Suhu (°C) | Prediksi | Rekomendasi |
|---|---|---|---|---|---|---|
| 1 | 2025-11-01 02:45:31 | 9.56 | 3.57 | 31.4 | Tidak Layak | Kembali ke Produsen |
| 2 | 2025-11-01 02:45:17 | 9.66 | 3.44 | 31.4 | Tidak Layak | Kembali ke Produsen |
| 3 | 2025-11-01 02:45:05 | 9.26 | 3.22 | 31.3 | Tidak Layak | Kembali ke Produsen |
| 4 | 2025-11-01 02:44:49 | 9.13 | 3.01 | 31.2 | Tidak Layak | Kembali ke Produsen |
| 5 | 2025-11-01 02:44:34 | 8.86 | 2.99 | 31.3 | Tidak Layak | Kembali ke Produsen |
| 6 | 2025-11-01 02:44:21 | 8.37 | 2.88 | 31.3 | Tidak Layak | Kembali ke Produsen |
| 7 | 2025-11-01 02:44:08 | 7.99 | 1.67 | 31.3 | Sedang | Arahkan ke Titik Terdekat : Titik Babat Babat, Jalan K.H. Ahmad Dahlan, Babat 62271, East Java, Indonesia |
| 8 | 2025-11-01 02:43:55 | 6.97 | 0.99 | 31.3 | Sedang | Arahkan ke Titik Terdekat : Titik Babat Babat, Jalan K.H. Ahmad Dahlan, Babat 62271, East Java, Indonesia |
| 9 | 2025-11-01 02:43:42 | 7.44 | 1.04 | 31.2 | Sedang | Arahkan ke Titik Terdekat : Titik Babat Babat, Jalan K.H. Ahmad Dahlan, Babat 62271, East Java, Indonesia |
| 10 | 2025-11-01 02:43:27 | 7.23 | 1.06 | 31.3 | Sedang | Arahkan ke Titik Terdekat : Titik Babat Babat, Jalan K.H. Ahmad Dahlan, Babat 62271, East Java, Indonesia |

Showing 1 to 10 of 171 entries Previous 1 2 3 4 5 … 18 Next

Source: (Result Research, 2025)

Figure 4. System Data Visualization

CC BY NC

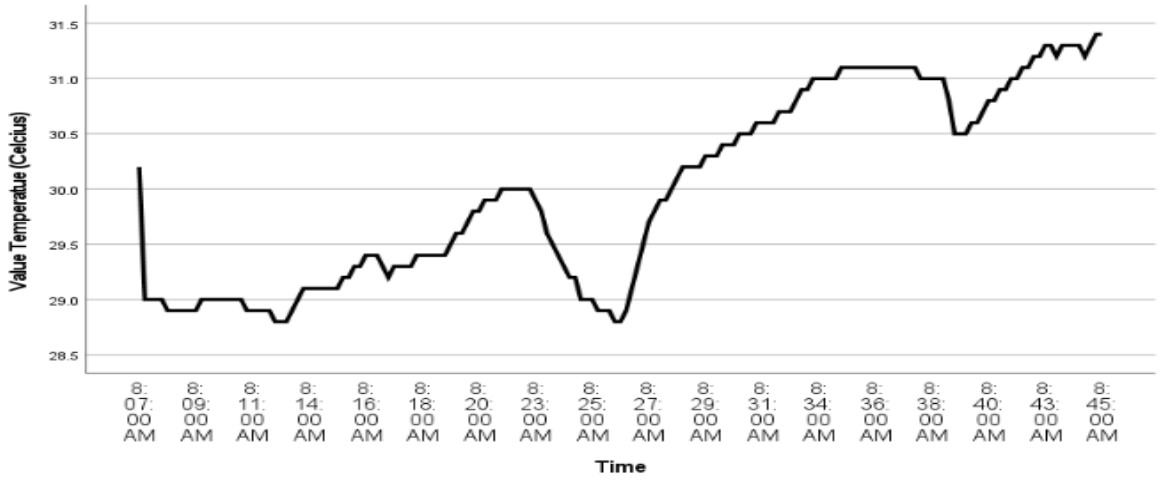
VOL. 12. NO. 2 NOVEMBER 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v12i2.7886

JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

Based on the collected data, visualizations were generated in the form of graphs for each sensor. These graphs display the patterns of temperature changes, $H_2S$ concentrations, and $NH_3$ levels over time, making it easier to clearly monitor trends in fish quality, as shown in Figures 5(a), 5(b), and 5(c). Figure 5(a) shows the temperature variation during measurement. Although the fluctuations are relatively small, temperature remains an important factor influencing microbial growth and the rate of fish quality deterioration. Therefore, temperature was included as one of the predictive features in the Random Forest classification model [15] [18].

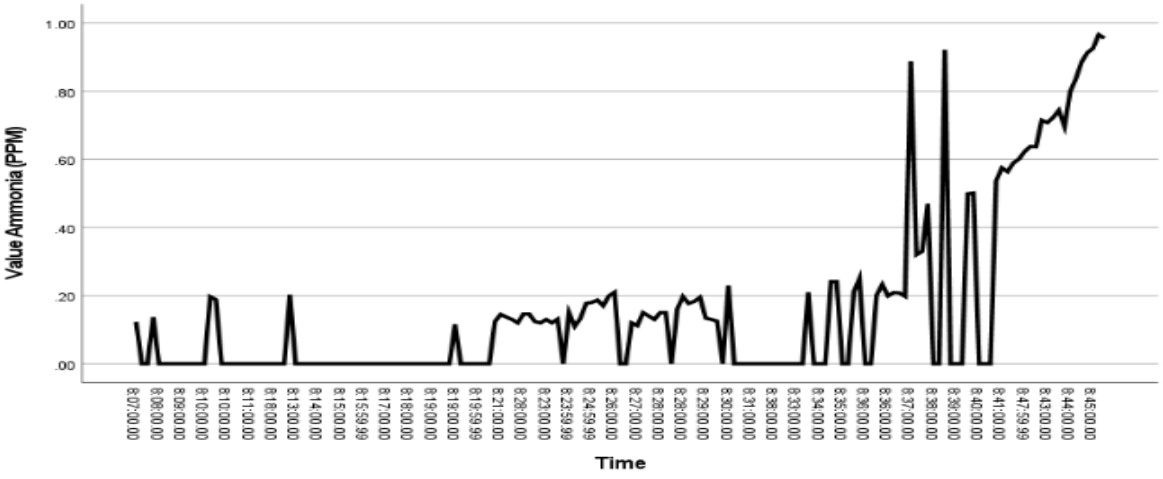
Figure 5(b) illustrates the changes in ammonia ($NH_3$) concentration during measurement. The increasing $NH_3$ concentration indicates protein decomposition caused by microbial activity, making it an effective indicator of fish spoilage. This finding is consistent with previous studies reporting $NH_3$ as a reliable parameter for assessing fish freshness [16] [17].

Figure 5(c) presents the variation in hydrogen sulfide ($H_2S$) concentration during measurement. Increasing $H_2S$ levels indicate the progression of fish spoilage and provide complementary information to temperature and $NH_3$ for improving fish quality classification. These findings are consistent with previous studies identifying $H_2S$ as an important spoilage indicator [16] [17].




Source: (Result Research, 2025)

Figure 5(a). Temperature Data Graph



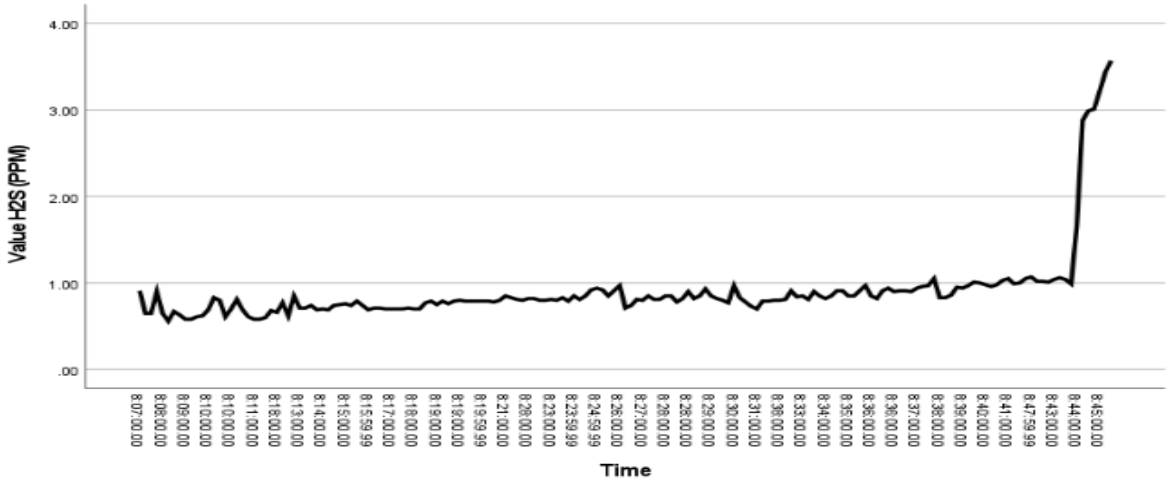

Source: (Result Research, 2025)

Figure 5(b). Ammonia Data Graph



Accredited Rank 2 (SINTA 2) based on Decree No. 225/E/KPT/2022 issued by Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia, dated December 7, 2022. Published by LPPM Universitas Nusa Mandiri.

VOL. 12. NO. 2 NOVEMBER 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v12i2.7886




Source: (Result Research, 2025)

Figure 5(c). Hydrogen Sulfide Data Graph

In this section, the original research data are presented and then analyzed using the Random Forest method. The research data are shown in Table 1. The complete dataset containing all 171 observations is provided as Supplementary Material to improve transparency and reproducibility. The structure of the collected data is summarized in Table 1.

Table 1. Research Data

| Time | $H_2S$ (PPM) | $NH_3$ (PPM) | Temperature ($^0C$) | Decision |
|---|---|---|---|---|
| 8:07 AM | 0.91 | 0.123 | 30.2 | Feasible |
| 8:07 AM | 0.65 | 0.036 | 29 | Feasible |
| 8:07 AM | 0.65 | 0.036 | 29 | Feasible |
| .... | .... | .... | .... | .... |
| 8:20 AM | 0.79 | 0.093 | 29.6 | Moderate |
| 8:20 AM | 0.79 | 0.093 | 29.7 | Moderate |
| 8:20 AM | 0.79 | 0.093 | 29.8 | Moderate |
| ... | ... | ... | .... | ... |
| 8:44 AM | 3.01 | 0.913 | 31.2 | Not Feasible |
| 8:45 AM | 3.22 | 0.926 | 31.3 | Not Feasible |
| 8:45 AM | 3.44 | 0.966 | 31.4 | Not Feasible |
| 8:45 AM | 3.57 | 0.956 | 31.4 | Not Feasible |

Source: (Result Research, 2025)

To improve the performance of the model in classifying fish quality, an optimization process was carried out using the Random Forest algorithm by testing various parameter combinations. Four main parameters were analyzed, including the maximum depth of the trees (max_depth), the minimum number of samples required at each leaf node (min_samples_leaf), the number of trees in the forest (n_estimators), and the minimum number of samples required to split a node (min_samples_split). These four parameters were tested using grid search with a parameter space as shown in Table 2.

Table 2. Test Parameters

| Parameter | Value |
|---|---|
| n_estimators (n-est) | 50, 100, 150, 200, 250, 300 |
| max_depth (md) | 5, 10, 15, 20, 25, 30 |
| min_samples_split (mss) | 2, 3, 5, 7, 9 |
| min_samples_leaf (msl) | 1, 2, 3, 4, 5, 6 |

Source: (Result Research, 2025)

This approach allows for the selection of the best model configuration that yields the most optimal accuracy in predicting fish quality, with the results presented in Table 3. Based on Table 3, the parameter combination that provides the most optimal performance in terms of both the highest accuracy and the fastest training time is n_estimators = 50, max_depth = 30, min_samples_split = 2, and min_samples_leaf = 1, with a training time of 0.639 seconds. This configuration yields the highest accuracy of 0.9259 (92.6%), while the lowest accuracy obtained from other parameter variations is 0.8963 (89.6%). The accuracy values were calculated using Equation (1), ensuring that the results are consistent with the established evaluation method.

Moreover, this optimal combination achieved robust evaluation metrics, with a Precision of 0.9297, a Recall of 0.9259, and an F1-score of 0.9245, indicating consistent classification performance across the evaluated dataset. However, these results should be interpreted within the context of the naturally imbalanced class

CC BY NC

distribution (56 Feasible, 95 Moderate, and 20 Not Feasible samples) and should not be considered sufficient evidence that the model completely mitigates class imbalance or eliminates classification bias. Future studies will employ larger and more balanced datasets, together with additional evaluation metrics, to further assess the robustness of the proposed model. The confusion matrix of the optimized Random Forest model (Figure 7) summarizes the classification results obtained through five-fold cross-validation. The matrix was constructed from the out-of-fold predictions for the 171 observations, with each observation evaluated when it was excluded from the corresponding training fold. Of the 171 observations, 159 were correctly classified, yielding a pooled accuracy of 92.98%. No separate independent test set was used to generate this confusion matrix. Therefore, the results represent cross-validation performance on the available dataset rather than performance on an independent external test set.

Table 3. Parameter Test Results

| n estimators | Max depth | Min samples split | Min samples leaf | accuracy | precision | recall | F1-score | Standard deviation | Training time |
|---|---|---|---|---|---|---|---|---|---|
| 50 | 5 | 3 | 1 | 0.9185 | 0.9218 | 0.9185 | 0.9172 | 0.0491 | 0.625 |
| 50 | 5 | 2 | 1 | 0.9185 | 0.9218 | 0.9185 | 0.9172 | 0.0491 | 0.646 |
| 50 | 5 | 5 | 1 | 0.9111 | 0.9207 | 0.9111 | 0.9113 | 0.0602 | 0.62 |
| 50 | 30 | 2 | 1 | 0.9259 | 0.9297 | 0.9259 | 0.9245 | 0.0259 | 0.639 |
| ... | .... | ... | ... | .... | ... | ... | ... | .... | .... |
| 50 | 5 | 5 | 4 | 0.8966 | 0.9117 | 0.8966 | 0.897 | 0.064 | 0.737 |
| 50 | 5 | 3 | 4 | 0.8966 | 0.9117 | 0.8966 | 0.897 | 0.064 | 0.765 |

Source: (Result Research, 2025)

Table 4. Performance Evaluation by Class

| Class | Precision | Recall | F1-score |
|---|---|---|---|
| Feasible | 0.91 | 0.89 | 0.90 |
| Moderate | 0.94 | 0.94 | 0.94 |
| Not Feasible | 0.95 | 1.00 | 0.98 |
| Macro Average | 0.93 | 0.94 | 0.94 |
| Weighted Average | 0.93 | 0.93 | 0.93 |

Source: (Result Research, 2025)

The class-specific performance of the optimized Random Forest model is summarized in Table 4. The Feasible class achieved a precision of 0.91 and a recall of 0.89, with an F1-score of 0.90. The model correctly classified 50 of 56 Feasible observations, while six Feasible observations were misclassified as Moderate. This misclassification may be attributed to overlapping sensor characteristics between the Feasible and Moderate classes during the early stages of fish quality deterioration.

The Moderate class achieved a precision, recall, and F1-score of 0.94, indicating consistent classification performance for fish with intermediate quality conditions. The model correctly classified 89 of 95 Moderate observations, while five were classified as Feasible and one as Not Feasible. The relatively high performance suggests that the selected sensor features effectively captured the characteristics associated with the Moderate quality condition.

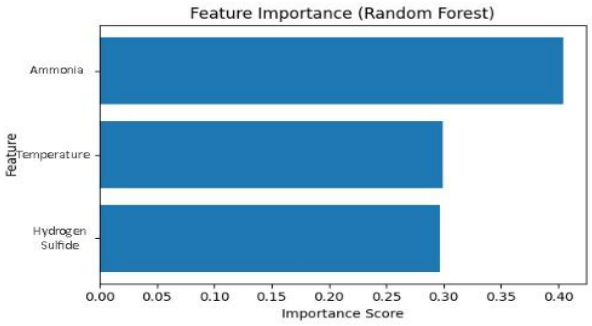
The Not Feasible class achieved a precision of 0.95, a recall of 1.00, and an F1-score of 0.98. All 20 Not Feasible observations were correctly classified, indicating that the model successfully detected all observations belonging to this class. One Moderate observation was incorrectly classified as Not Feasible, resulting in a false positive. From a food safety perspective, this type of error is less critical than failing to detect an unsafe fish sample (false negative), as it favors a more conservative decision regarding fish distribution and consumption. Overall, the model correctly classified 159 of 171 observations, resulting in an accuracy of 92.98%.

Furthermore, based on the feature importance analysis of the model, the ammonia attribute shows the most dominant influence with a value of 0.404052. The temperature attribute has a medium level of importance with a value of 0.299342, followed by the sulfide attribute, which is also in the medium importance category with a value of 0.296605, as shown in Figure 6. This shows that the ammonia variable makes the largest contribution to the fish quality classification process compared to the other two parameters. This is also supported by the research of Fish Freshness Indicator for Sensing Fish Quality during Storage [20] which states that Ammonia is very influential in fish spoilage because its levels increase with microbial activity, and this increase can be clearly


Accredited Rank 2 (SINTA 2) based on Decree No. 225/E/KPT/2022 issued by Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia, dated December 7, 2022. Published by LPPM Universitas Nusa Mandiri.

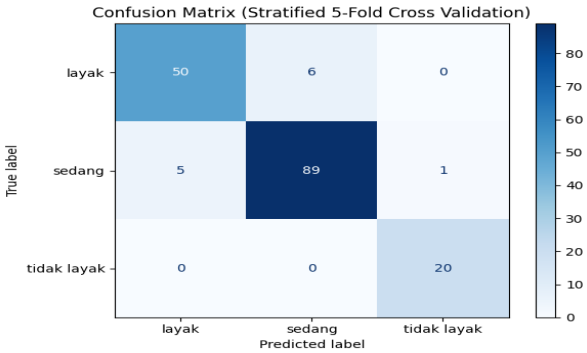
VOL. 12. NO. 2 NOVEMBER 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v12i2.7886

detected by the pH indicator which changes color from yellow to black and then purple. With a direct relationship between ammonia levels and sensor color changes, the developed indicator is effective as a fast and non-destructive tool for assessing fish freshness during storage.




Source : (Result Research, 2025)

Figure 6. Random Forest Feature Importance Graph




Source : (Result Research, 2025)

Figure 7. Confusion Matrix

Based on Figure 7, presents the confusion matrix of the optimized Random Forest model based on five-fold cross-validation predictions. Across the 171 observations, 159 were correctly classified, corresponding to a pooled accuracy of 92.98%. This value represents the accuracy calculated from the aggregated cross-validation predictions, whereas the 92.6% reported in Table 3 represents the mean cross-validation accuracy of the best hyperparameter configuration. No separate independent test set was used to generate this confusion matrix. The results of this study were also validated through Pearson correlation analysis among the variables, as shown in Table 4. This correlation analysis aims to identify the strength of the linear relationship between each pair of variables involved in the modeling. By conducting this test, a more comprehensive understanding of the interrelationships among parameters is obtained, which supports model interpretation and reinforces the findings regarding the most influential variable in determining fish quality.

Table 5. Variable Correlation

| Variable Pair | Correlation Value | Category |
|---|---|---|
| Temperature – $NH_3$ | 0.728 | Strong |
| $NH_3$ – $H_2S$ | 0.320 | Weak |
| Temperature – $H_2S$ | 0.135 | Very Weak |

Source : (Result Research, 2025)

Based on Table 5, Pearson correlation analysis was performed to examine the linear relationships among the predictor variables, namely the $NH_3$-related sensor response, temperature, and $H_2S$-related sensor response. The results showed a positive correlation between the $NH_3$-related sensor response and temperature (r = 0.728), a weaker positive correlation between temperature and the $H_2S$-related sensor response (r = 0.320), and a weak positive correlation between the $NH_3$- and $H_2S$-related sensor responses (r = 0.135). These findings describe the degree of linear association among the predictor variables but do not directly demonstrate classification accuracy or determine feature importance. The classification performance of the Random Forest model was evaluated separately using accuracy, precision, recall, F1-score, and the confusion matrix. Meanwhile, the relative importance of the predictor variables was assessed using the feature-importance scores obtained from the Random Forest model. Therefore, Pearson correlation analysis was used as supplementary information to characterize relationships among the predictors, rather than as evidence confirming the model's predictive performance or feature importance.

This research shows that $NH_3$ is the most influential feature in the Random Forest model. This result is consistent with the biological mechanism of fish spoilage, where microbial activity decomposes proteins and produces ammonia as one of the primary volatile compounds. As spoilage progresses, the $NH_3$ concentration increases, making it a direct indicator of freshness deterioration. In contrast, temperature mainly affects the rate of microbial growth, while $H_2S$ is generally produced at more advanced stages of spoilage. Therefore, $NH_3$ provides stronger discriminatory information for distinguishing fish quality categories. This finding is consistent with previous studies reporting ammonia as a reliable indicator for evaluating fish freshness and quality, supporting its use as a key feature for automated fish quality monitoring [16] [17].

From a food-quality perspective, increasing $NH_3$ concentrations indicate a reduction in fish freshness and can be used as an objective indicator

CC BY NC

of quality deterioration during distribution. Therefore, the high feature importance assigned to $NH_3$ by the Random Forest model not only improves classification performance but also supports practical decision-making for maintaining food quality and safety throughout the supply chain [16][17]. To gain a deeper understanding of the performance of the Random Forest model used, the results of this study were further compared with other studies that applied similar methods. This comparison aims to assess the consistency, advantages, and potential improvements of the approach used, as well as to provide a broader context regarding the effectiveness of the algorithm in the same field. The comparison is presented in Table 6.

Table 6. Accuracy Comparison

| Research Name | Year | Accuracy of Random Forest | |
|---|---|---|---|
| | | Highest | Lowest |
| Ahmed, F., et al. [18] | 2024 | 99,23% | 96,76% |
| Li, Q., et al. [15] | 2025 | 84% | 73% |
| Shodiq, M., et al. (this study) | 2025 | 92,6% | 89,6% |

Source : (Result Research, 2025)

Based on Table 6, a comparison of previous research results shows that the performance of the Random Forest algorithm varies significantly across studies. The study conducted by Ahmed, F., et al. (2024) reported the highest Random Forest performance, with a maximum accuracy of 99.23% and a minimum accuracy of 96.76%, indicating excellent model stability and reliability [18]. Meanwhile, the study by Li, Q., et al. (2025) showed a more moderate performance, with a highest accuracy of 84% and a lowest accuracy of 73%, reflecting challenges in data processing or the characteristics of the variables used [15]. The study by Shodiq, M., et al. (2025) achieved a highest accuracy of 92.6% and a lowest accuracy of 89.6%, representing high and relatively consistent performance. Therefore, this comparison indicates that the performance of Random Forest can be highly dependent on data quality, the types of features used, and the complexity of the problem addressed in each study.

## CONCLUSION

Based on the results of this study, the proposed IoT-Cloud-based fish quality monitoring system integrated with the Random Forest algorithm successfully classified fish quality. Across the evaluated hyperparameter configurations, the mean five-fold cross-validation accuracy ranged from 89.6% to 92.6%, with the best configuration achieving a mean accuracy of 92.6%. Separately, the confusion matrix aggregated from the cross-validation predictions showed that 159 of the 171 observations were correctly classified, corresponding to a pooled accuracy of 92.98%. These values represent different summaries of cross-validation performance and should not be interpreted as results from an independent test set. Feature importance analysis identified ammonia ($NH_3$) as the most influential variable in determining fish quality, followed by temperature and hydrogen sulfide ($H_2S$), indicating that these three variables are key indicators for monitoring fish freshness during transportation.

The satisfactory performance obtained in this study is supported by the characteristics of the collected dataset, which consists of three sensor features with relatively low noise levels. However, the dataset represents a single distribution scenario with a naturally imbalanced class distribution; therefore, the reported results should be interpreted within the context of the current experimental conditions and should not be considered sufficient evidence of the model's generalization capability across broader distribution environments. Future research will focus on gathering larger and more heterogeneous datasets across various distribution scenarios, as well as examining additional machine learning algorithms and evaluation metrics to provide a more comprehensive evaluation of the reliability and generalization capability of the developed monitoring system.

## DECLARATION OF AI USE

The authors declare that generative artificial intelligence (AI) tools were used in the preparation of this manuscript. Google Gemini was used through Google Colab to assist with Python code generation, data analysis, and data visualization. The authors reviewed and verified the results and take full responsibility for the accuracy, originality, and final content of the manuscript.

## AUTHOR CONTRIBUTIONS

This journal uses the Contributor Roles Taxonomy (CRediT) to identify individual author contributions to the research and manuscript preparation.

| Author | Contribution Roles |
|---|---|
| Muhammad Shodiq | Conceptualization, Investigation, Methodology, Software, Validation, Formal Analysis, Project Administration, Writing – Original Draft, Writing – Review & Editing. |



Accredited Rank 2 (SINTA 2) based on Decree No. 225/E/KPT/2022 issued by Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia, dated December 7, 2022. Published by LPPM Universitas Nusa Mandiri.

VOL. 12. NO. 2 NOVEMBER 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v12i2.7886

| Author | Contribution Roles |
|---|---|
| Agus Priyono | Investigation, Methodology, Software, Validation, Writing – Original Draft, Writing – Review & Editing. |
| Mutsna Sa Yu Zakka | Investigation. |

**ACKNOWLEDGMENTS**

The authors would like to express their sincere gratitude to the Pelabuhan Perikanan Nusantara (PPN) Brondong for granting permission and providing support for the research activities conducted at the port as the research observation site. The authors also appreciate the assistance, cooperation, and valuable information provided by the staff during the research process.

**FUNDING**

This research was funded by the Ministry of Higher Education, Science, and Technology (Kemendiktisaintek) through the National Competitive Research Program for Novice Lecturer Research under the Master Contract dated May 28, 2025, with Master Contract Number: 128/C3/DT.05.00/PL/2025.

**CONFLICT OF INTEREST STATEMENT**

The authors declare that there are no conflicts of interest related to the research, writing, or publication of this article.

**REFERENCE**

[1] M. Ching-Pong Poo, T. Wang, and Z. Yang, "Global food supply chain resilience assessment: A case in the United Kingdom," *Transp. Res. Part A Policy Pract.*, vol. 181, no. February, p. 104018, 2024, doi: 10.1016/j.tra.2024.104018.

[2] R. M. Elzaki, "Does fish production influence the GDP and food security in Gulf Cooperation Council countries? Evidence from the dynamic panel data analysis," *Aquaculture*, vol. 578, no. August 2023, p. 740058, 2024, doi: 10.1016/j.aquaculture.2023.740058.

[3] C. Tsirimokos, I. Tzouramani, A. Liontakis, S. Mantziaris, and A. Sintori, "Investigating the interindustry dynamics of Greece's fisheries and aquaculture sector: Effects on production and employment," *Mar. Policy*, vol. 175, no. February, p. 106621, 2025, doi: 10.1016/j.marpol.2025.106621.

[4] J. C. Onwe, M. G. Ojide, M. Subhan, and D. Forgenie, "Food security in Nigeria amidst globalization, economic expansion, and population growth: A wavelet coherence and QARDL analysis," *J. Agric. Food Res.*, vol. 18, no. September, 2024, doi: 10.1016/j.jafr.2024.101413.

[5] S. Qodrunnada and H. Hafiludin, "Analisis Rantai Distribusi Ikan Hasil Tangkapan Nelayan di Tempat Pelelangan Ikan Branta Kabupaten Pamekasan," *Juv. Ilm. Kelaut. dan Perikan.*, vol. 4, no. 3, pp. 254–263, 2023, doi: 10.21107/juvenil.v4i3.21135.

[6] S. Ramírez-Rodríguez, A. Sarroca, D. Fernández-Guerrero, S. Pelosso, and L. R. Puig, "Food waste in high income countries: Spanish fish value chain as a case study," *Fish. Res.*, vol. 279, no. July, 2024, doi: 10.1016/j.fishres.2024.107134.

[7] N. N. Afiyah, I. Solihin, and E. Lubis, "The Effect of Distribution Chain on The Quality of Tuna Mackerel (Euthynus sp.) from 'PPP Blanakan' to Consumer Area," *J. Sos. Ekon. Kelaut. dan Perikan.*, vol. 14, no. 2, p. 225, 2019.

[8] M. Hasan, N. Vasker, M. M. Hossain, M. I. Bhuiyan, J. Biswas, and M. R. Ahmmad Rashid, "Framework for fish freshness detection and rotten fish removal in Bangladesh using mask R–CNN method with robotic arm and fisheye analysis," *J. Agric. Food Res.*, vol. 16, no. October 2023, 2024, doi: 10.1016/j.jafr.2024.101139.

[9] S. Taj, A. S. Imran, Z. Kastrati, S. M. Daudpota, R. A. Memon, and J. Ahmed, "IoT-based supply chain management: A systematic literature review," *Internet of Things (Netherlands)*, vol. 24, no. October, p. 100982, 2023, doi: 10.1016/j.iot.2023.100982.

[10] W. C. Tan and M. S. Sidhu, "Review of RFID and IoT integration in supply chain management," *Oper. Res. Perspect.*, vol. 9, no. February, p. 100229, 2022, doi: 10.1016/j.orp.2022.100229.

[11] S. Fei and Q. Wang, "Research on Online English Teaching Platform Based On Cloud Computing Technology," *Procedia Comput. Sci.*, vol. 243, pp. 1153–1161, 2024, doi: 10.1016/j.procs.2024.09.136.

[12] I. Vanany *et al.*, "Social-economic and environment impacts for a fish reverse supply chain: A mixed integer linear optimization approach," *Clean. Logist. Supply Chain*, vol. 13, no. November, p. 100189, 2024, doi: 10.1016/j.clscn.2024.100189.

[13] I. Zrelli and A. Rejeb, "A bibliometric analysis of IoT applications in logistics and supply

Accredited Rank 2 (SINTA 2) based on Decree No. 225/E/KPT/2022 issued by Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia, dated December 7, 2022. Published by LPPM Universitas Nusa Mandiri.

chain management," *Heliyon*, vol. 10, no. 16, p. e36578, 2024, doi: 10.1016/j.heliyon.2024.e36578.

[14] N. K. Kamila *et al.*, "Machine learning model design for high performance cloud computing & load balancing resiliency: An innovative approach," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 10, pp. 9991–10009, 2022, doi: 10.1016/j.jksuci.2022.10.001.

[15] Q. Li *et al.*, "Development of freshness monitoring models for grass carp cutting products at multiple storage temperatures and construction of a real-time prediction platform," *LWT*, vol. 235, no. August, p. 118620, 2025, doi: 10.1016/j.lwt.2025.118620.

[16] P. Jebedass, I. Chandran, H. Ahmed, P. K. Hashir, and S. Veerasingam, "Aquacultural Engineering Smart technologies in aquaculture : An integrated IoT , AI , and blockchain framework for sustainable growth," vol. 111, no. June, 2025.

[17] E. N. Arifani, S. I. Purnama, M. A. Afandi, D. Kurnianto, and N. Latifasari, "Freshness Analysis of Color , NH 3 , H 2 S , and pH on Salmon Spoilage Using the PCA Method," vol. 01021, pp. 1–10, 2025.

[18] F. Ahmed, H. I. Bijoy, H. R. Hemal, and S. Rashed, "Heliyon Smart aquaculture analytics : Enhancing shrimp farming in Bangladesh through real-time IoT monitoring and predictive machine learning analysis," *Heliyon*, vol. 10, no. 17, p. e37330, 2024, doi: 10.1016/j.heliyon.2024.e37330.

[19] I. Technology, S. Program, and U. I. Batam, "Implementation Mean Imputation And Outlier Detection For Loan Prediction Using The Random Forest Algorithm," vol. 10, no. 4, pp. 937–944, 2025, doi: 10.33480/jitk.v10i4.6437.

[20] D. Kim, S. Park, and H. Shin, "Fish Freshness Indicator for Sensing Fish Quality during Storage," 2023.

[21] Z. Zhang, Y. Sun, S. Sang, L. Jia, and C. Ou, "Emerging approach for fish freshness evaluation: Principle, application and challenges," Foods, vol. 11, no. 13, Art. no. 1897, 2022, doi: 10.3390/foods11131897.



Accredited Rank 2 (SINTA 2) based on Decree No. 225/E/KPT/2022 issued by Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia, dated December 7, 2022. Published by LPPM Universitas Nusa Mandiri.