Classification of Coconut Milk Freshness Based on VOC Analysis Using Random Forest

Rajina R Mohamed1,2, a), Razali Yaakob1, b), Tuan Dir Tengku Mohd Azahar3, c), Nor Azura Husin1, d) and Nurul Azwaani Salehuddin Haqe2, e)

*1Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, Serdang, Malaysia*

*2College of Computer Science and Information Technology, Universiti Tenaga Nasional, Kajang, Malaysia*

*3Malaysia France Institute, Universiti Kuala Lumpur, Kuala Lumpur, Malaysia*

*Corresponding author: a)rajina@uniten.edu.my*

*b)*[*razaliy@upm.edu.my*](mailto:razaliy@upm.edu.my)

*c)*[*tgazahar@unikl.edu.my*](mailto:tgazahar@unikl.edu.my)

*d)*[*n\_azura@upm.edu.my*](mailto:n_azura@upm.edu.my)

*e)ST1085157@student.uniten.edu.my*

**Abstract.** The maintenance of coconut milk freshness plays a crucial role in preserving both food safety standards and product quality. The spoilage process of coconut milk produces Volatile Organic Compounds (VOCs) which function as indicators of microbial activity. The research aims to create a dependable method for coconut milk freshness classification through Random Forest algorithm analysis of VOC data obtained from an Electronic Nose (E-Nose) using MQ2, MQ5, MG135 and MQ6 gas sensors. Random Forest was chosen because it excels at handling complex high-dimensional data through ensemble learning methods which include boot-strap aggregating and random feature selection. The Random Forest model demonstrates successful classification between coconut milk freshness levels according to the research findings which provides an interpretable solution for real-time freshness assessment. The system demonstrates effective management of uncertain and multi-dimensional sensor data which makes it suitable for practical food industry applications. The research advances sensor-based food quality assessment through its reliable spoilage detection method which enables safer and higher-quality food products.

# Introduction

Coconut milk plays a vital role of raw materials in South East Asia and tropical food types, most frequently in curries and desserts. Coconut industry is an important part of the food and agriculture sector in Malaysia and the vitality of the coconut industry sustain the domestic food manufacturers and local farmers of coconuts. Nevertheless, the coconut milk is very perishable, as it comprises of high-water content with low acidity and is sensitive to micro-organism contamination and spoilage [1]. Not only does spoiled coconut milk compromise the quality of goods and the safety of consumers, furthermore, it causes economic losses of both yield producers and retailers due to food waste. Conventional sensory based freshness testing oxides like smell and taste are subjective and inconsistent and are not sensitive to allowing detection of early spoilage. The necessity of an automated, objective and data-based freshness classification system leading to improvement of food safety and quality control is increasing. Among the most efficient methods of food freshness determination is the analysis of the Volatile Organic Compounds (VOCs). VOCs are volatile compounds that are emitted when food is spoilt. Microbial activity increases after the production of coconut milk, making this dairy product become more prone to oxidation and more permeable to the action of VOCs (ketones, alcohols, esters, and aldehydes) [2]. High-end analysis systems like the Gas Chromatography-Mass Spectrometry (GC-MS) give more comprehensive VOC profiling, but they are expensive and time consuming. A gas sensor comprising of an Electronic Nose (E-Nose) such as MQ2, MQ5, and MQ135 is a cost-effective, real-time solution as it can sense the VOC patterns of spoilage [3]. Nevertheless, it is still, rather difficult to define the freshness levels on the basis of E-Nose data, and this necessitates the application of effective machine learning algorithms that can process large, complicated and noisy data properly. In this study, a new model is proposed through the combination of the Random Forest (RF) algorithm with the purpose of improving classification in the topic of freshness of coconut milk. RF is a very efficient ensemble learning method which builds a number of decision trees, enhancing the classification accuracy rate as well as avoiding errors that can be made by the presence of a noisy data. Compared to other classical methods of classification, including Principal Component Analysis (PCA) and Support Vector Machines (SVM), RF is well known to deal with non-linear relationships and feature spaces that are high-dimensionalized [4]. It is based on the predictive ability of RF; the specific study will help create an accurate and automated system of classifying freshness of coconut milk with high confidence levels. This study has serious implications on the food industry. That is why the suggested AI-based freshness classification model is compatible with smart packaging, where you will find it in real-time with the condition of coconut-based products and inhibit the expiration of this product. In addition, the regulatory authorities can use machine learning approach towards food safety so as to maintain global standards of food quality [5]. With the growth in food technology, the AI-based freshness detecting solutions will become extremely important in enhancing the food safety, low wastage, and supply chain optimization. The study helps provide the linkage between sensor-based food monitoring and advanced machine learning, which is useful in the next round of intelligent food quality assessment systems.

**Literature Review**

Many Machine Learning algorithms are used in food quality measurement whereby RF algorithm has been seen as one of the most vigorous classifying algorithms since it is exerting high-accuracy, is noise-resistant and also has the capacity to use high-dimensional data. RF is a commanding learning technique that builds many decision tree models and combines their summits to improve the dependability of classification [6]. It has been applied to food science and specifically in detecting and classifying freshness where it has resulted in better performance relative to other classical classifiers such as k-Nearest Neighbors (k-NN) and Support Vector Machines (SVM) [1]. RF has been coupled with some spectroscopic methods to help in food authentication analysis whereby, Fourier Transform Infrared Spectroscopy (FTIR) and Raman Spectroscopy methods among others, have been employed to identify the adulteration of liquid food products, like coconut water and milk-based beverages, among others [2]. These reports prove that RF provides better results compared with linear classifiers, particularly with regard to complex food matrices and subtle changes in chemicals that prove spoilage. Also, RF together with the PCA was applied to perform the classification of food samples by means of chemical fingerprinting with an accuracy of above 90% [3]. RF has also successfully been used to sense the maturity of coconuts in the area of product classification done with coconuts and has also been used to chart the level of coconut plantations [4]. These experiments confirm that that RF will be an effective tool in working with large datasets, which have non-linear relations, and therefore it is suitable to classify coconut milk freshness. Its potential to interpret the indications of VOCs data of sensors that mimic E-Nose further contributes to its usability in the detection of food freshness like in studies that involve the use of machine learning to monitor in real-time the spoilage of foods [5]. The current methods of freshness evaluation in coconuts milk, including sensory evaluation and microbial testing are usually insufficient as they are subjective in nature and time consuming. The combination of the RF based classification implementation of an innovative result is automatization of the freshness classification using real-time sensor data. RF is much less prone to overfitting compared to conventional models thus it is more reliable to use it in practice [1]. Finally, the prior studies confirm that RF can be used in classifying food products, especially liquid food products and the analysis of the freshness of the food products. RF and E-Nose technology combination presents a fast, non-destructive and effective technique of categorizing coconut milk as far as food safety and quality control are concerned. This work is an expansion of the related works, and has incorporated the RF to enhance the level of precise classification so that the system can be fully operational as part of the real-time monitoring solution to freshness of coconut milk.

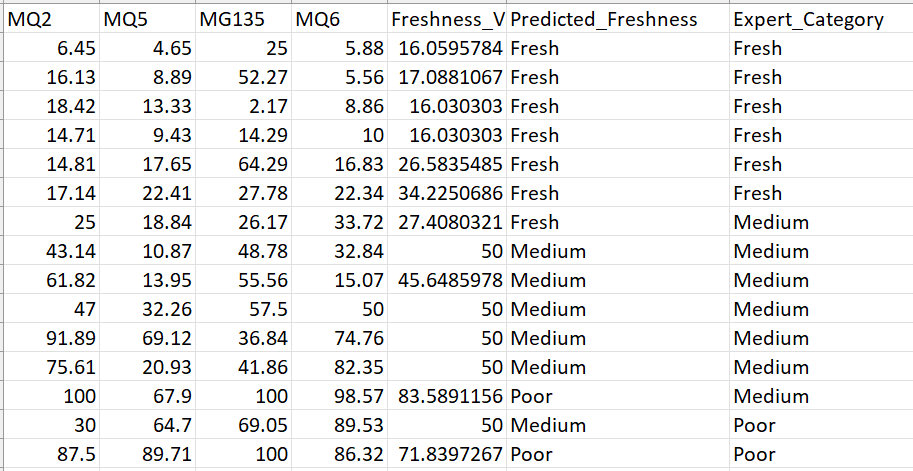
# Methodology

**Data Collection Using Electronic Nose (E-Nose)**

In order to categorize coconut milk freshness, a gas sensor of the E-Nose with four gas sensors (MQ2, MQ5, MG135, MQ6) was adopted to search VOCs released by the coconut milk sample. The E-Nose was set up in such a way that it could give real time results of the gas concentration on the sensors placed in the air above the sample. In order to be consistent at least, one part of the coconut milk was dissolved in a four-part water and kept in a sealed container at room temperature. Manual measurement was carried out by reading the sensor values every hour (at 8:00 a.m. to 6:00 p.m.), at an interval of 30 days using the values indicated on the LCD screen of the E-Nose. The duration of data collection was scheduled to be complete at 6.00 p.m. every day since earlier tests revealed that from that moment there were no significant fluctuations in gas concentration. The precision and consistency of the sensors were to be ensured by calibrating the E-Nose with a known concentration gas prior to every data collection activity. Each collected data was stored and processed in order to be analyzed.

**Hypothesis and Data Processing**

The hypothesis of this study is that the degree of VOCs concentration will exponentially grow with time, because with the development of the process of spoilage, releases of the gas will obtain higher values. The concentration of the gas was read in terms of voltage and tabulated. The data taken in 30 days gave a wide picture of development and change of the gas composition giving ability to classify coconut milk according to freshness level. The sample dataset collected by the E-Nose system is shown in Figure 1. The sensors are recorded with each value indicating a particular stage of freshness that is rated as Fresh, Medium, Poor, depending upon the concentration of the gases.

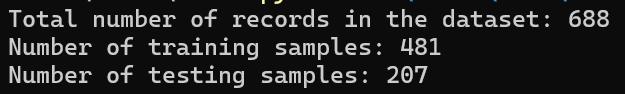


**FIGURE 1.** Sample dataset

The dataset consists of the following attributes:

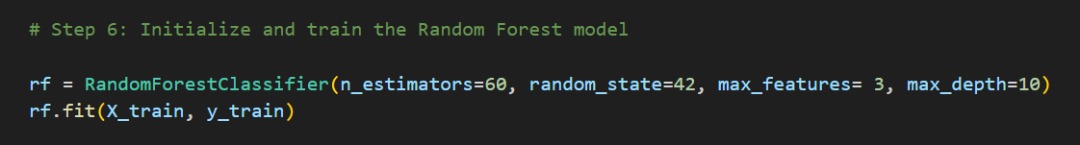
* VOC measurements from MQ2, MQ5, MQ6, and MG135 sensors.
* Freshness\_Value: A numerical score representing estimated freshness
* Predicted\_Freshness: Categorized freshness labels (Fresh = 0, Medium = 1, Poor = 2).
* Expert\_Category: The ground truth labels assigned by domain experts

To ensure optimal learning, categorical variables were converted into numeric values, and feature scaling was applied to normalize sensor readings. The dataset was split into 70% training and 30% testing, ensuring a balanced class distribution like shown in Figure 2.



**FIGURE 2.** Number of datasets

Hyperparameters are important settings in this study, which were selected wisely in order to improve the performance of the model and develop stable predictions. These settings include n\_estimators that was set at 60. That is, the model will construct 60 different decision trees and then will aggregate and average their output as a way of making its final decision, as is the case with experts voting as a team. The increased number of trees is likely to increase the accuracy and stability of the model, yet slows the model and consumes additional memory, thus 60 can be said to be a good sweet spot. random\_state was assigned to 42 to make sure that when the program is run repeatedly the results will be consistent. This is quite significant and particularly when the various models need to be tested and compared. The max feature is another important parameter, which was put to a value 3. This implies that every tree just considers 3 of the 6 available features that include MQ2, MQ5, MG135, MQ6, Freshness\_Value and Predicted\_Freshness when determines how to divide the data. The trees are made less homogeneous and, therefore more unlikely to repeat all the same errors, and whereas they are not as composed of features as possible, which then also leads to an improved model performance. And finally, max\_depth was set to 10 that will regulate the length of the attack of each tree. Excessively deep trees may simply memorize the training data and fail spectacularly when offered new data (a phenomenon of overfitting the data). The depth is set to 10 so that the trees would discover beneficial patterns without becoming too complicated. Overall, these hyperparameters were chosen manually based on understanding the dataset and experimenting with different values. While this setup wasn’t found using automated tuning methods, it still provided a good balance between accuracy, speed, and reliability. Figure 3 shows manual hyperparameters setting in this study.



**FIGURE 3.** Hyperparameters setting

**Training the Random Forest Model**

The Random Forest classifier was trained using the prepared dataset, leveraging its ensemble learning approach to improve classification robustness. The training process involved three main steps:

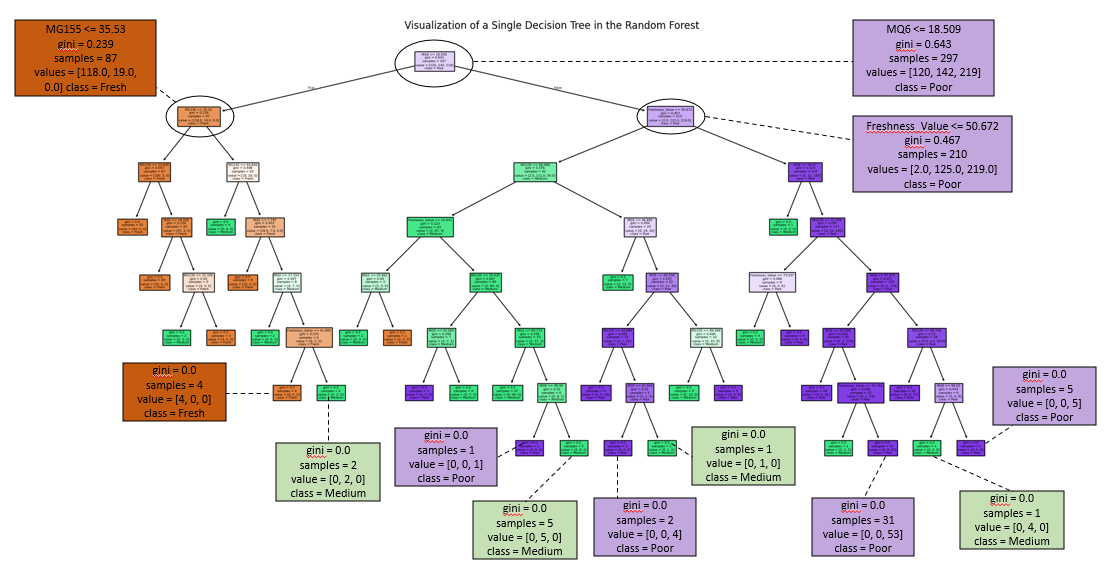
### **Bootstrap Sampling**

To meet these objectives the bootstrapping sampling technique was used to create a set of training subsets inside the dataset, which implies that each tree was trained on a distinct subset. In each boot-strap sample, some observations were put aside, whereas the rest of the samples would serve as Out-of-Bag (OOB) samples to internally validate the sample. The overall number of trees in the forest had an optimum value.

### **Decision Tree Construction**

Subsets of features were used to construct each decision tree as a way of model diversity and guarding against overfitting. Gini impurity acted as the basis on which a perfect feature threshold was chosen through the splitting process. There is an imputation method where mean values were entered where the missing values were found. Random Forest algorithm was able to assess several split points, picking the threshold that would result in minimum impurity. As an example, looking at the MQ6 as a feature, the feasible split thresholds were calculated and the Gini impurity of each of them was estimated. The last split point reduced the class impurity hence the improved classification performance.

Figure 4 shows a decision tree diagram, which is an important element of the RF model that was applied in order to classify the freshness of coconut milk in terms of sensor data. The root node that is at the top of the tree divides the data related to MQ6 sensor value and provides two main branches. Each internal node that follows in turn is a decision point at which further data-splitting is done on the basis of different sensor readings, among them being MG135, MQ5, MQ2 and Freshness\_Value. The measures of impurity of the dataset at a given stage are given by the Gini index, which is displayed in each node, and a lower value denotes purer classifications. Another indicator that is shown on the nodes is the number of samples that are taken into account on each split and distribution of classes Fresh, Medium, and Poor. When the tree branches out, it finally terminates some leaf nodes that point to final classification choices. In these nodes leaves, there are either pure or near pure classifications in which no more splits are required. The coloring of the boxes (orange-Fresh, green-Medium, and purple-Poor) allows visualizing the effect of various sensor values on the process of classification. This decision tree is not alone in the RF model and there may be several trees in this model which cooperate to make a given final decision on the classification by voting among them, and increasing the level of predictive accuracy.



**FIGURE 4.** Decision Tree Visualization

### **Majority Voting**

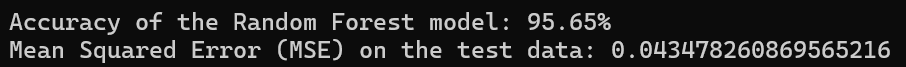
The majority voting process is determined using that formula:

|  |
| --- |
| ***y^*:** Final predicted label  Prediction made by the *B*-th decision tree in the RF  ***B*:** Total number of trees in the RF  ***mode*:** Most often prediction among all the trees |

Upon building several decision trees, the resultant final classification of a sample was based on majority vote. Every tree gave a prediction of the freshness category and the most commonly predicted category was taken as the final label. This combination method helped to enhance the stability of the classification and minimized against the overfitting of the model. RF considers the number of decision trees that had made a particular prediction, and the prevailing prediction (majority vote) is considered the final classification. Individual trees are assigned to categorize an input with learned (decision) rules, like VOCs sensor products on the freshness of coconut milk. The prediction of the various trees could differ because the trees are trained on different subsets of data. Majority voting simply means that the commonest group of the trees is the one that is used as the final output and therefore the model is resistant to noise and individual tree bias. It can be illustrated in a case where five decision trees are used to predict the freshness of a sample if three trees give the value as Fresh, the others give the value as Medium and Poor then the final classification will be given as Fresh because it has got the majority vote. As it is, this strategy makes the model more reliable, as it minimizes overfitting to any individual decision tree and also enhances generalizations as it applies to varying samples. The Majority voting ensemble should be ideal because even though a certain number of the available trees may take improper decisions, owing to minute fluctuations in the set of data, the entire model should be correct and consistent, which becomes the most convenient step in determining the freshness of coconut milk using gas sensor information.

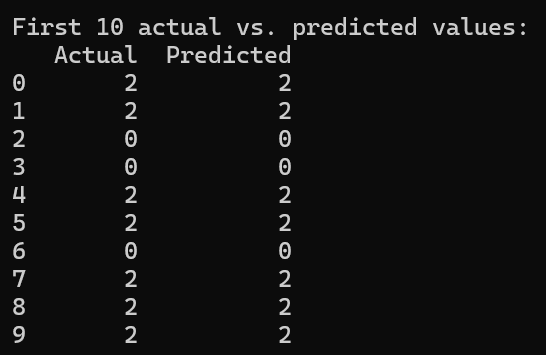
# Result and Discussion

The RF model showed excellent results that included an excellent level of accurate predictions of 95.65 in the test dataset. The level of the developed accuracy is high, which implies that the model makes the right predictions, contributing to its strength and stability. Also, the Mean Squared Error (MSE) of only 0.04347 supports the information that the prediction errors are insignificant that indicate that the model has the capacity to generalize well such as in Figure 5.



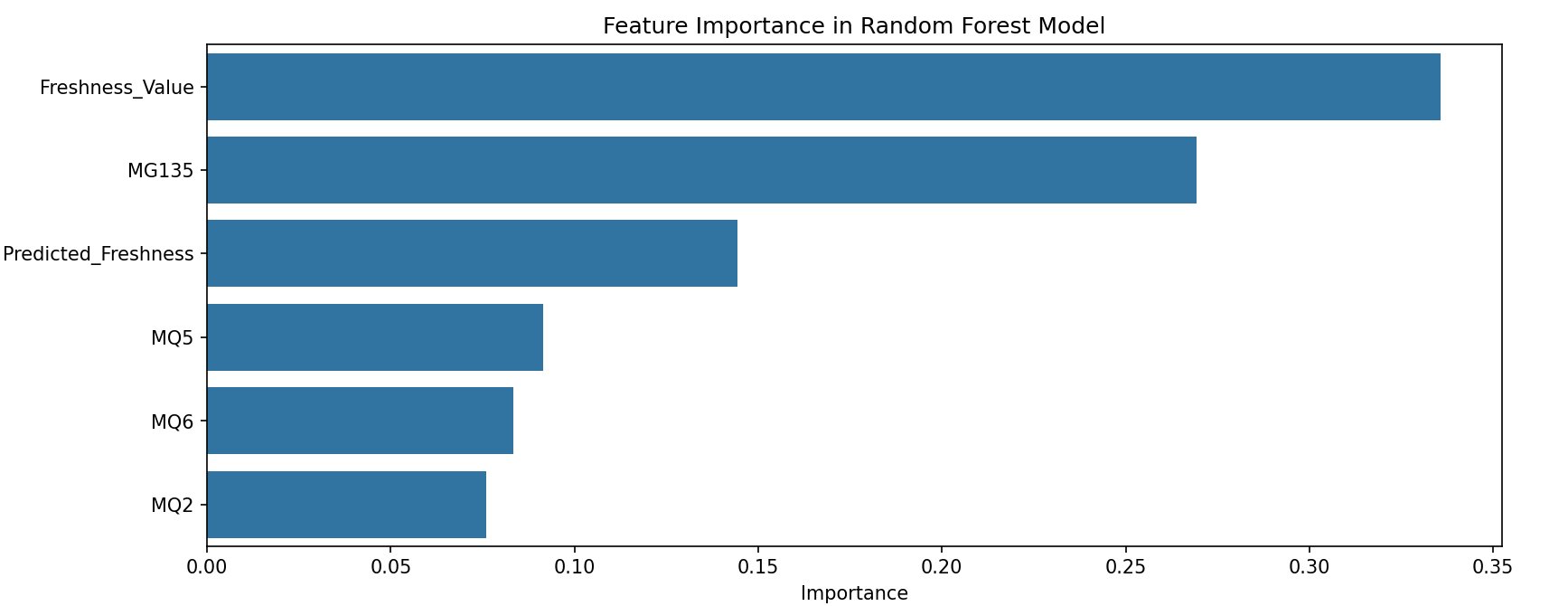
**FIGURE 5.** Accuracy and MSE result

This statement can be supported by a closer examination of the first 10 values of actual versus predicted values presented in Figure 6 whereby the predictions and the actual values coincide to each other. Such consistency implies that this model has effectively discovered the underlying patterns in the data, and thus this model is a very effective classifier.



**FIGURE 6.** First 10 actual and prediction value

The feature importance analysis of the gas sensors in the RF model as shown in Figure 7 point out to the increased contribution of certain sensors towards determining the coconut milk freshness. The MG135 sensor also becomes a major characteristic, i.e. it is sensitive to VOCs related to spoilage, possibly gases like ammonia, hydrogen sulfide, or ethanol which are released in large quantities during microbial decay. The MQ5, MQ6 and the MQ2 sensors also has a contribution in classification but with reduced significance which means that, they do detect the gases of interest but their readings can be less clear on the distinction between a fresh and spoiled sample. It could be possible that the MQ5 sensor, used to detect natural gas, methane, and Liquefied Petroleum Gas (LPG), is picking up some of the spoilage-related hydrocarbons and that the MQ6, which reacts to butane and propane, is also detecting the fertility of bacteria growth in metabolites. To a lesser extent, the MQ2 sensor is sensitive to flammable gases such as hydrogen gas and alcohol vapours so this may be due to small VOCs emission during early stages of spoilage.



**FIGURE 7.** The feature importance in Random Forest model

# Ensemble learning is also a major strength of the RF algorithm, hence minimizing variance and avoiding overfitting when compared to the single-decision classifiers such as Decision Trees. We have seen a high level of accuracy in the model, and this indicates that it has been able to generalize between various data points hence it would be adequately applicable in the real-life situations that need high accuracy in classification. These results are, nevertheless, promising; however, there is a need to evaluate how well this model performs on other datasets to see whether it can be applied in different conditions.

# In spite of the high performance of the model, there should be some limitations to be considered. To begin with, the data arrangement and assignment of classes may affect performance. The model could be biased and therefore could be inclined towards the prevailing collection of data. Further analysis in precision and recall and confusion matrix would help more information about the performance in the classes. Second, the effectiveness of the model in computation terms is an aspect to be put into consideration when using it in extensive applications. Although they have the high accuracy, the RF models can be computationally costly due to many decision trees. Efficiency could be further improved by including optimizations, i.e. feature selection and hyperparameter tuning without compromising accuracy.

# Aside from this, another study would be to consider other ensemble techniques, like Gradient Boosting and XGBoost, in order to compare their performance and determine which one is the best model. Also, it would be beneficial to implement more varied data and experiment with various conditions to prove robustness and the viability of real-life implementation of the model. On the whole, these outcomes define the efficiency of RF method, which can be used as the potent instrument of high-accuracy classification tasks.

# Conclusion

The paper proves the high accuracy of RF model and obtains 95.65 % accuracy in the model with a low level of MSE, about 0.04347, which compares favorably to the traditional classification methods. The ensemble learning mechanism used in the model makes it robust and highly reliable and predictable. Such results are consistent with earlier studies, which indicate that RF is useful in modeling non-linear relationships in data, which is challenging. On the one hand, the outcomes are very encouraging, and some possible limitations should be taken into account. This might affect generalization because the dataset composition and possible class imbalances might influence the results so additional validation on other datasets is needed. In addition, the computational efficiency will also be taken into account because RF models are quite demanding in terms of resources, especially when working with large-scale data. One should further attempt to optimize the hyperparameter tuning and seek other ensemble approaches like XGBoost or Gradient Boosting Machines (GBM) in the future research to improve performance without sacrificing computational costs. To sum everything up, we can state that the findings that were obtained in this study prove that RF is one of the most reliable and effective classifiers and, therefore, it is an adequate option to use it in high-accuracy predictive modeling tasks. The fact that it performs well in basic tasks related to classification confirms again its relevance in many other areas, such as medical diagnosis, fraud detection and financial risk detection. Since the field of machine learning is still under development, it is obvious that the future brings more improvements to the ensemble methods, expanding the horizons of predictive accuracy and model efficiency. Internet of Things (IoT) technology can easily be incorporated into the real-world application of the freshness classification system of coconut milk proposed, which is based on gas sensors and the use of RF algorithm, in order to ensure real-time and endless execution of the application. With the integration of Wireless Sensor Networks (WSN) and edge computing devices, i.e., Raspberry Pi or ESP32 microcontrollers, the use of the E-Nose system enables the real-time capture of VOCs data, process, and transmit it to cloud-based technologies so that additional analysis can be implemented. Such arrangement allows the monitoring of the freshness of coconut milk at the storage and transportation premises remotely and thus protects against spoilage and contributes to food safety. Also, it is also possible to modify machine learning models and integrate them directly with IoT gateways and perform freshness prediction on a device, minimizing the use of cloud computing resources and providing greater efficiency to the system. It is also possible to develop smart packaging where inbuilt sensors relay freshness information through RFID tags or even mobile phones and the consumer or retailer can get real time spoilage information. Moreover, the system can be extended to a wider scope of food supply chain as a method of warning distributors and retailers of the possibility of spoilage before the sale of the products to the consumers. Automated seasonal alerts in the form of SMS or IoT dashboards would mean freshness can be proactively tracked and food waste is minimized as well as improving inventories. Future changes may include AI-powered anomaly detection, in which the past freshness data is to be used to forecast spoilage patterns, allowing probing interventional methods. Such application in real life would greatly enhance the standards concerning food safety and it would minimize losses of goods that are perishable.

# References

1. T.A. Teklemariam, F. Chou, P. Kumaravel, and J. Van Buskrik, “ATR-FTIR spectroscopy and machine/deep learning models for detecting adulteration in coconut water with sugars, sugar alcohols, and artificial sweeteners,” Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy **322**, 124771 (2024).
2. T.A. Teklemariam, F. Chou, P. Kumaravel, and J. Van Buskrik, “ATR-FTIR spectroscopy and machine/deep learning models for detecting adulteration in coconut water with sugars, sugar alcohols, and artificial sweeteners,” Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 322, 124771 (2024).
3. J.A. Caladcad, S. Cabahug, M.R. Catamco, P.E. Villaceran, L. Cosgafa, K.N. Cabizares, M. Hermosilla, and E.Jr. Piedad, “Determining Philippine coconut maturity level using machine learning algorithms based on acoustic signal,” Computers and Electronics in Agriculture **172**, 105327 (2020).
4. S.V. Nithya, K. Ragunath, S. Pazhanivelan, R. Kumaraperumal, and P. Latha, “Mapping coconut plantation in Western Agro-Climatic zone using object-based classification and machine learning technique,” Plant Science Today **11**(sp4), (2024).
5. S.K. Manoharan, R.K. Megalingam, and S.P. Sreeji, “Machine Learning Based Classification of Coconut Trees Based on Tree Parameters,” in *2024 2nd World Conference on Communication & Computing*, (Institute of Electrical and Electronics Engineers, RAIPUR, India, 2024), pp. 1–6.
6. L. Breiman, “Random Forests,” Machine Learning **45**(1), 5–32 (2001).
7. M. Schonlau, and R.Y. Zou, “The random forest algorithm for statistical learning,” The Stata Journal: Promoting Communications on Statistics and Stata **20**(1), 3–29 (2020).
8. H.A. Salman, A. Kalakech, and A. Steiti, “Random Forest Algorithm Overview,” Babylonian Journal of Machine Learning **2024**, 69–79 (2024).
9. X. Wang, Y. Bouzembrak, A.O. Lansink, and H.J. Van Der Fels‐Klerx, “Application of machine learning to the monitoring and prediction of food safety: A review,” Comprehensive Reviews in Food Science and Food Safety **21**(1), 416–434 (2022).
10. X. Deng, S. Cao, and A.L. Horn, “Emerging Applications of Machine Learning in Food Safety,” Annual Review of Food Science and Technology **12**(1), 513–538 (2021).
11. K.B. Chhetri, “Applications of Artificial Intelligence and Machine Learning in Food Quality Control and Safety Assessment,” Food Engineering Reviews **16**(1), 1–21 (2024).
12. S. Sapienza, and A. Vedder, “Principle-based recommendations for big data and machine learning in food safety: the P-SAFETY model,” Artificial Intelligence and Society **38**(1), 5–20 (2023).
13. Y. Lin, J. Ma, Q. Wang, and D.-W. Sun, “Applications of machine learning techniques for enhancing nondestructive food quality and safety detection,” Critical Reviews in Food Science and Nutrition **63**(12), 1649–1669 (2023).
14. P.-K. Revelou, E. Tsakali, A. Batrinou, and I.F. Strati, “Applications of Machine Learning in Food Safety and HACCP Monitoring of Animal-Source Foods,” Foods **14**(6), 922 (2025).
15. K. Kudashkina, M.G. Corradini, P. Thirunathan, R.Y. Yada, and E.D.G. Fraser, “Artificial Intelligence technology in food safety: A behavioral approach,” Trends in Food Science & Technology **123**, 376–381 (2022).
16. I. Kumar, J. Rawat, N. Mohd, and S. Husain, “Opportunities of Artificial Intelligence and Machine Learning in the Food Industry,” Journal of Food Quality **2021**, 1–10 (2021).
17. S. He, and X. Shi, “Microbial Food Safety in China: Past, Present, and Future,” Foodborne Pathogens and Disease **18**(8), 510–518 (2021).
18. S.A. Ibrahim, R.D. Ayivi, T. Zimmerman, S.A. Siddiqui, A.B. Altemimi, H. Fidan, T. Esatbeyoglu, and R.V. Bakhshayesh, “Lactic Acid Bacteria as Antimicrobial Agents: Food Safety and Microbial Food Spoilage Prevention,” Foods **10**(12), 3131 (2021).
19. H. Chen, and Y. Feng, “Moving Forward to the Future: A Review of Microbial Food Safety Education in China,” Foodborne Pathogens and Disease **18**(8), 547–566 (2021).
20. H. Singh, S.K. Bhardwaj, M. Khatri, K.-H. Kim, and N. Bhardwaj, “UVC radiation for food safety: An emerging technology for the microbial disinfection of food products,” Chemical Engineering Journal **417**, 128084 (2021).
21. E. Teshome, S.F. Forsido, H.P.V. Rupasinghe, and E. Olika Keyata, “Potentials of Natural Preservatives to Enhance Food Safety and Shelf Life: A Review,” The Scientific World Journal **2022**, 1–11 (2022).
22. O. Misiou, and K. Koutsoumanis, “Climate change and its implications for food safety and spoilage,” Trends in Food Science & Technology **126**, 142–152 (2022).
23. Y. Zhang, J. Kan, X. Liu, F. Song, K. Zhu, N. Li, and Y. Zhang, “Chemical Components, Nutritional Value, Volatile Organic Compounds and Biological Activities In Vitro of Coconut (Cocos nucifera L.) Water with Different Maturities,” Foods **13**(6), 863 (2024).
24. Z. Cheng, M.G. O’Sullivan, S. Miao, J.P. Kerry, and K.N. Kilcawley, “Sensorial, cultural and volatile properties of milk, dairy powders, yoghurt and butter: A review,” International Journal of Dairy Technology **75**(4), 761–790 (2022).
25. A.K. Pathak, and C. Viphavakit, “A review on all-optical fiber-based VOC sensors: Heading towards the development of promising technology,” Sensors and Actuators, A: Physical **338**, 113455 (2022).
26. A. Pathak, K. Swargiary, N. Kongsawang, P. Jitpratak, N. Ajchareeyasoontorn, J. Udomkittivorakul, and C. Viphavakit, “Recent Advances in Sensing Materials Targeting Clinical Volatile Organic Compound (VOC) Biomarkers: A Review,” Biosensors **13**(1), 114 (2023).
27. Y. Feng, Y. Wang, B. Beykal, M. Qiao, Z. Xiao, and Y. Luo, “A mechanistic review on machine learning-supported detection and analysis of volatile organic compounds for food quality and safety,” Trends in Food Science & Technology **143**, 104297 (2024).