Random Forest Algorithm for Meat Classification and Microbial Population Prediction

By Salman Hanif

Random Forest Algorithm for Meat Classification and Microbial Population Prediction

Salman Hanif
School of Applied Science
Telkom University
Bandung, Indonesia
slmnhanif@stundent.telkomuniversity.ac.id

Dedy Rahman Wijaya
School of Applied Science
Telkom University
Bandung, Indonesia
dedyrw@telkomuniversity.ac.id

Pramuko Aji
School of Applied Science
Telkom University
Bandung, Indonesia
pramukoaji@telkomuniversity.ac.id

Abstract*— Meat is one of the various protein sources that needed for human body. Now, the consumption of meat has always increased from year to year due to various factors, including having high nutrition, as a source of protein, to its distribution which can be found almost everywhere. In selecting meat, many consumers do not know about the quality of the meat offered by the seller, both the duration and the preservatives used by the seller in marketing the meat. The conventional way buyers use to determine the quality of meat is by checking the smell of the meat using the nose manually. In overcoming these problems, determining the quality of the meat needed the right method in checking. By applying the Random Forest Classification and Regression method, Electronic Nose can work structured on every component needed to determine meat quality. This experiment showed that the Random Forest Classification and Regression algorithm obtained that the best parameter for classification is n estimators 72 with result mean test score is 0.954955 and the best parameter for regression is n estimators 233 with result RMSE is 0.0141 and R^2 is 0.9876.

***Keywords*—Meat, Machine Learning, Random Forest Classification and Regression, Electronic Nose**

I. Introduction

Meat is one of the top livestock products that contain protein in it [1]. The content possessed by meat, among others, is a protein that contains a high composition of amino acids. Amino acids are proteins that have benefits to support the growth and repair of muscle tissue [2]. In addition to containing amino acid protein, meat also contains energy. The energy, in essence, has the benefit of maintaining muscle mass and strengthening the immune system [3]. In addition, there is a fat content in meat that serves to keep healthy blood and muscles.

On the other side, consuming meat that is not feasible or has been rotten caused by a microbial population that is too much will harm body health [4]. This can drive consumers to suffer from obesity, headaches, and heart attacks that cause fat accumulation [5].

The cause of the sale of meat that is not fit for consumption is that the officer determines the quality of the meat by calculating or predicting the quality of the meat. Officers have difficulty determining and predicting meat quality in the existing system because the determination and prediction of meat quality are made manually [6]. For example, smelling the aroma of meat, pressing the meat to determine the texture, and seeing the shape of the meat with the naked eye to assess the quality of the meat [7]. As a result, some meat that is no longer fit for consumption will be estimated to be still done for consumption and traded by sellers, distributors, and consumers in the market.

Existing methods associated with predicting meat quality show that conventional methods directly test the meats by sniffing the meat using the nose manually to classify the meat's quality [8]. Apply the manual way makes the outcomes are not accurate therefore it is based on the nose of the human. In this experiment, we offer another method to classify meat quality. By applying the Random Forest Classification and Regression method, Electronic Nose can work structured on every component needed to determine meat quality. Besides that, Random Forest has tremendous potential of becoming a popular technique for future classifiers because its performance and accuracy has been found to be comparable with ensemble techniques bagging and boosting [9]. This method has proved its success in both regression and classification problems in recent years and is one of the best machine learning algorithms used in many different fields [10]. RF algorithm uses a simple predetermined probability to select the most important relevant attribute. Breiman formulated the RF algorithm using sample data subsets and to construct multiple decision trees by mapping random sample of feature subspaces [11].

Implementing the Electronic Nose has advantages such as the influence of other scents from outside can be minimized because the aroma produced in the sample room is brought to the sensor by utilizing the air flow through the hoses and valves that are regulated for use [12].

Arranged of this paper is: Section II describes related works. Section III describes the model of the system through this experiment. Section IV describes the result of the implementation, and Section V is showing the conclusions of this experiment.

II. Related Works

Predicting something such as meat, rice, and water with applying machine learning has been used efficiently and effectively to this day. To classify the quality and the microbial population of the meat, the dataset of the electronic nose can be applied therefore it is more compatible for Non-Destructive Testing way that produces appropriate output, including food level security, scents detection, and diseases control [13].

In research conducted [14], the Random Forest Classification to classify the Phishing Email, shows that this technique can classify the dataset merged from 200 phishing and ham emails with classification accuracy score until 99.7% and low rates of false-negative and false-positive.

In another research [15], the experiments will return the formula of Random Forests and make a study about the prediction result of the real-world and simulated the datasets for maximalize the size of trees during overfit. These experiments show that gains can be recognized by additional tuning to set tree size via restricting the split number and the size of the nodes.

In this experiment [16], the self-developed e-nose system is utilized to monitor a beef sample. The e-nose device is composed of the gas sensor array, microcontroller, and WiFi-shield. Metal-oxide semiconductor (MOS) gas sensors are assembled to Arduino microcontroller.

In this study [17], 12 types of beef samples were used, which produced twelve homogeneous data sets through an e-nose. A homogeneous data set is a data set with small differences, generated from several observations .

This paper will explain the combination between the Electronic Nose dataset and the Random Forest algorithm to predict meat quality and the microbial population in the meat. This experiment examines further the classification of meat quality and the microbial population of meat using Electronic-Nose and a lot of sensors such as MQ5 and MQ8. The result of this experiment, meat quality, and microbial population testing will be more efficient and effective for the seller and consumer.

III. Model of System

A. Dataset

The Electronic Nose outcomes is the dataset that used for this experiment.

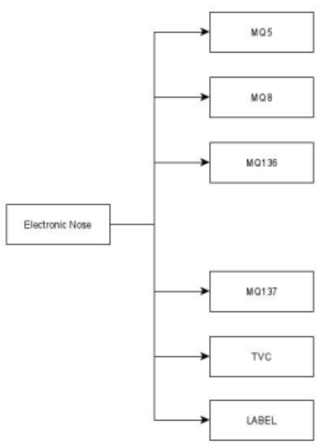

Fig. 1. Features of Dataset

Fig. 1**Error! Reference source not found.** represents the feature that is available in the dataset of Electronic Nose. There are four features (MQ5, MQ8, MQ136, and MQ137) [18]. There are two labels (TVC and Label) that output from sensors of the electronic nose [19].

The first feature is MQ5. The MQ5 is a sensor that useful for deteting gas, such as LPG, i-butane, methane, alcohol, Hydrogen, and smoke. The MQ5 sensor is commonly used for detecting gas leakage for various applications. The MQ5 gas sensor will detect the concentration of gas in the meat [20]. The second feature is MQ8. The MQ8 sensor as known as hydrogen detector is a gas sensor that has high sensitivity to hydrogen gas. This sensor also has sensitivity to alcohol, LPG gas and cooking fumes but has little sensitivity. This sensor works stably and has a long life in use [21]. The third feature is MQ136. The MQ136 gas sensors can detect physical phenomena of sulfur gas concentration. The change in air velocity in the gas causes the concentration of gas to also change, the estimation of changes in air velocity regarding gas and is expected to affect the ability to read data by the sensor [22]. The last feature is MQ137. The MQ137 sensor is devoted to detecting the presence of NH3 gas in the air. The MQ137 gas sensor has high sensitivity to Ammonia, too other organic amines. The sensor can be used to detect gas different containing Ammonia [23]. Next, there are two labels that used in this experiment. TVC is continuous label of microbial population and Label is discrete label of meat quality such as excellent, good, acceptable, and spoiled [18].

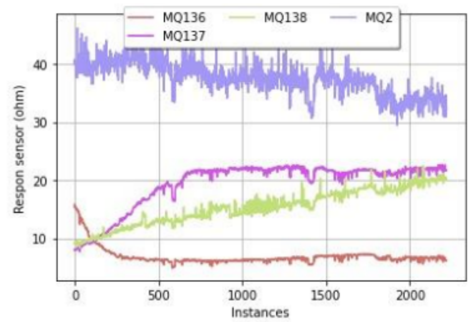

Fig. 2. Electronic Nose Signals

Fig. 2 represents the outcomes available from the dataset of Electronic Nose then can be seen in the figure 2 above. The response of sensor's can be seen in Y-axis and the instances can be seen in X-axis.

B. Proposed Method

Fig. 3 is the arrangement of the proposed method, and the description is below the figure.

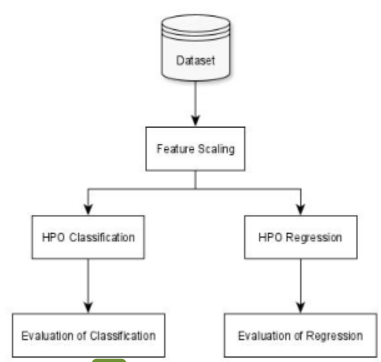

Fig. 3. Proposed Method

Fig. 3 represents the stage or process for predicting the quality of meat and microbial population using the Random Forest algorithm refers to the dataset of Electronic Nose.

1. *Feature Scaling*

Feature Scaling is the first stage of predicting process. The data will be scaled with the standard method in the future scaling stage. The standard method is a process of producing a decent average distribution of 0. This experiment uses the standard scaler or normalization [24].

Standard Scaler Formula:

$$z = \frac{x-\mu}{\sigma} \quad (1)$$

x = raw score

μ = mean

σ = standard deviation

z = standard scaler (result)

2. *Hyperparameter Optimization*

Hyperparameter Optimization is a stage using the grid search method. This method has the function to discover the best parameter for the dataset that has been selected. The best parameter must produce accurate prediction results. TABLE I represents the parameter and parameter value of hyperparameter optimization classification.

TABLE I
HYPERPARAMETER OPTIMIZATION CLASSIFICATION

Grid Search	
Parameter	Parameter Value
n_estimators	10
	17
	25
	33
	41
	48
	56
	64
	72
max_features	auto
	sqrt
max_depth	2
	4
min_samples_split	2
	5
min_samples_leaf	1
	2
bootstrap	True
	False

TABLE II represents the parameter and parameter value of hyperparameter optimization regression.

TABLE II
HYPERPARAMETER OPTIMIZATION REGRESSION

Grid Search	
Parameter	Parameter Value
n_estimators	200
	233
	266
	300
	333
max_features	auto
	sqrt
min_samples_split	2
	5
min_samples_leaf	1
	2
bootstrap	True
	False

This experiment uses a random forest algorithm, an ensemble learning method that executes by arranging many decision trees during the process of training. It is a nonparametric model described by a graph-like tree used in regression and classification role [25]. It is an extension of bagged trees. It has been primarily used for classification problems and several benchmarking studies have proven that it is one of the best machine learning techniques currently available [26]. We have used the boosted random forest because of its accurate classification performance on imbalanced [27].

$$IG(D_p, f) = I(D_p) - \frac{N_{left}}{N} I(D_{left}) - \frac{N_{right}}{N} I(D_{right})$$
(2)

f = feature split on

D_p = the parent node dataset

D_{left} = the left child node dataset

D_{right} = the right child node dataset

I = criterion of impurity

N = total number of samples

N_{left} = samples number at a left child node

N_{right} = samples number at a right child node

3. *Evaluation*

Evaluation is a stage that contains two parts, that is Root Means Square Error (RMSE) and R^2. Root Means Square Error has been used as a statistic standard metric to measure the example of implementation in quality of air and studies of climate research. [28] concluded that statistics such as the RMSE and the standard error have imprescriptible fifteen ambiguities and suggest the alternates. The RMSE formula [29].

$$RMSE = \left(\frac{\sum(y_i - \hat{y}_i)}{n}\right)^{\frac{1}{2}} \quad (3)$$

RMSE = Root Means Square Error

y = observed values

$\hat{y}$ = predicted result values

i = order data in databases

n = total data

, Besides that there is R^2. R^2 is a measuring tool that has a statistical result used to determine the implementation of the models of regression. The statistic represents the non-independent variable's score that determines the independent variable. It measures the strength of the connection between the non-independent variable's and regression models which are suitable with the scale [30].

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (4)$$

R^2 = R^2 score

y_i = I response observation

$\hat{y}_i$ = forecast response to - I

$\bar{y}$ = average

n = total data

To determine the accuracy of Performance Metrics, it is important to know the formula to calculate it. Here is The formula to find the precision, recall and accuracy [31].

$$\text{Precision} = \frac{True\,Positive}{True\,Positive + False\,Positive} \quad (5)$$

$$\text{Recall} = \frac{True\,Positive}{True\,Positive + False\,Negative} \quad (6)$$

$$\text{Accuracy} = \frac{True\,Positive + True\,Negative}{True\,Positive + True\,Negative + False\,Positive + False\,Negative} \quad (7)$$

C. Tools Used

The tool used in this experiment is Google Colab that is an executable document that can be used to store, write, and share programs that have been written in Google Drive. The hardware used in this experiment is Asus ROG GL503 Laptop. The Processor of this Laptop is Intel Core i7-8750H with NVIDIA® GeForce GTX™ 1050 Ti Graphics.

IV. Result and Discussion

This section will represent the result below in the predicted quality of meat.

A. Hyperparameter Optimization Classification

TABLE III will give information about the result of the hyperparameter by representing the grid search using the Random Forest algorithm for classification.

TABLE III
GRID SEARCH CLASSIFICATION RESULT

Parameter						**Score**
n_estimators	max_features	max_depth	min_samples_split	min_sample_leaf	bootstrap	**Mean Test Score**
10	auto	2	2	1	True	0.884685
10	auto	2	5	2	True	0.866066
16	auto	2	2	1	True	0.854054
16	auto	2	5	2	True	0.870870
17	auto	2	2	1	True	0.892492
23	auto	2	2	1	True	0.844444
23	auto	2	5	2	True	0.870870
25	auto	2	2	1	True	0.892492
30	auto	2	2	1	True	0.849849
33	auto	2	2	1	True	0.890691
37	auto	2	2	1	True	0.868468
41	auto	2	2	1	True	0.886486
44	auto	2	2	1	True	0.852852
48	sqrt	4	5	2	False	0.951952
51	auto	2	2	1	True	0.855855
56	sqrt	4	5	2	False	0.948348
58	auto	2	2	1	True	0.861861
64	sqrt	4	5	2	False	0.948348
65	auto	2	2	1	True	0.847447
72	**sqrt**	**4**	**5**	**2**	**False**	**0.954955**

The result in TABLE III represents that the best parameter to use is n_estimators = 72, max_features = sqrt, max_depth = 4, min_sample_split = 5, min_sample_leaf = 2, bootstrap = False, and Mean Test Score = 0.954955.

B. Confusion Matrix of Random Forest Classification.

This section will represent the confusion matrix of meat quality. Fig. 4 below describes the predicted label of spoiled, acceptable, good, and excellent meat.

Fig. 4. Confusion Matrix Classification

The result in Fig. 4 represents that 962 meat will be predicted to be of excellent quality. 224 of them will be spoiled.

C. Performance Metrics

TABLE IV will show the information about the precision, recall, and accuracy of the classification.

TABLE IV
PERFORMANCE METRICS

	precision	recall	f1-score	support
Excellent	1.00	0.98	0.99	225
Good	0.92	0.84	0.88	180
Acceptable	0.83	0.64	0.73	270
Spoiled	0.92	0.99	0.95	990
accuracy			0.92	1665

macro avg	0.92	0.86	0.89	1665
weighted avg	0.91	0.92	0.91	1665

The result for the performance metrics above is accuracy with f1-score 0.92.

D. Hyper Parameter Optimization Regression

TABLE V will give information about the result of the hyperparameter by representing the grid search using the Random Forest algorithm for regression.

TABLE V
GRID SEARCH REGRESSION RESULT

Parameter					**Score**	
n_estimators	max_features	min_sample_split	min_sample_leaf	bootstrap	**RMSE**	$\boldsymbol{R^2}$
200	auto	2	1	True	0.014	0.985
200	auto	2	1	True	0.016	0.983
200	sqrt	5	2	False	0.016	0.983
233	auto	2	1	True	0.039	0.967
233	sqrt	5	2	False	0.039	0.968
233	**sqrt**	**2**	**1**	**True**	**0.014**	**0.987**
266	auto	2	1	True	0.017	0.986
266	sqrt	5	2	False	0.013	0.983
266	sqrt	5	2	False	0.016	0.986
300	auto	2	1	True	0.012	0.989
300	auto	1	2	True	0.016	0.986
300	sqrt	5	2	False	0.012	0.989
333	auto	2	1	True	0.011	0.991
333	sqrt	5	2	False	0.036	0.970
333	sqrt	5	2	False	0.010	0.991

The result in TABLE V represents that the best parameter to use in n_estimator = 233 and max_feature = sqrt.

E. Graph Result of Random Forest Regression

This section will represent the confusion matrix of meat quality. Fig. 5 below described the predicted label of expired meat.

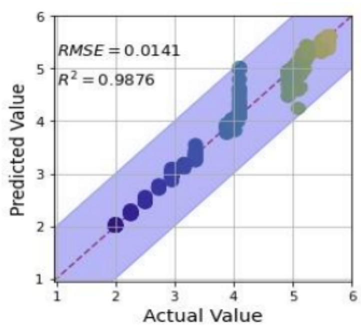

Fig. 5. Graph Result Regression

The graph result represents that the RMSE point is 0.0141 and the R^2 or R squared point is 0.9876.

V. Conclusion

The conclusion from this experiment is the Electronic Nose dataset and the Random Forest algorithm can classify the meat quality and predict the population of microbial in the meat. The confusion matrix shows the estimation of meat quality, 962 labels are excellent and 224 labels are spoiled. The result from this experiment represents that Mean Test Score = 0.954955 for classifying the meat quality, Accuracy Score = 0.92 and then RMSE = 0.0141 and R^2 = 0.9876 for predicting the population of microbial. From the classification outcomes, we can conclude the Random Forest algorithm is good to classify the excellent and spoiled meat and from the regression outcomes, we can conclude that Random Forest can be a good algorithm to predict the population of microbial because the result is accurate.

Acknowledgements

This research was funded by the Directorate General of Vocational Education, Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia through the Vocational Product Research scheme. This research was also partially funded by Telkom University.

References

[1] D. R. Wijaya, R. Sarno, E. Zulaika, and S. I. Sabila, "Development of mobile electronic nose for beef quality monitoring," in *Procedia Computer Science*, 2017, vol. 124, pp. 728–735. doi: 10.1016/j.procs.2017.12.211.

[2] M. Simonson, Y. Boirie, and C. Guillet, "Protein, amino acids and obesity treatment," *Reviews in Endocrine and Metabolic Disorders*, vol. 21, no. 3. Springer, pp. 341–353, Sep. 01, 2020. doi: 10.1007/s11154-020-09574-5.

[3] S. Banerjee, "The Essence of Indian Indigenous Knowledge in the perspective of Ayurveda, Nutrition, and Yoga", doi: 10.5281/zeno.

[4] D. Hidalgo, F. Corona, and J. M. Martín-Marroquín, "Manure biostabilization by effective microorganisms as a way to improve its agronomic value," *Biomass Conversion and Biorefinery*. Springer Science and Business Media Deutschland GmbH, 2022. doi: 10.1007/s13399-022-02428-x.

[5] J. J. DiNicolantonio, J. H. O'Keefe, and W. Wilson, "Subclinical magnesium deficiency: A principal driver of cardiovascular disease and a public health crisis," *Open Heart*, vol. 5, no. 1. BMJ Publishing Group, Feb. 01, 2018. doi: 10.1136/openhrt-2017-000668.

[6] M. Bukari *et al*., "Household factors and gestational age predict diet quality of pregnant women," *Maternal and Child Nutrition*, vol. 17, no. 3, Jul. 2021, doi: 10.1111/mcn.13145.

[7] U. S. Geletu, M. A. Usmael, Y. Y. Mummed, and A. M. Ibrahim, "Quality of Cattle Meat and Its Compositional Constituents," *Veterinary Medicine*

International, vol. 2021. Hindawi Limited, 2021. doi: 10.1155/2021/7340495.

[8] New Zealand Food Safety Authority, "A Guide to Calculating the Shelf Kife of," vol. 4, 2018.

[9] M. Rose and H. R. Hassen, "A Survey of Random Forest Pruning Techniques," Dec. 2019, pp. 99–109. doi: 10.5121/csit.2019.91808.

[10] C. M. Yeşilkanat, "Spatio-temporal estimation of the daily cases of COVID-19 in worldwide using random forest machine learning algorithm," *Chaos, Solitons & Fractals*, vol. 140, p. 110210, Nov. 2020, doi: 10.1016/j.chaos.2020.110210.

[11] V. Jackins, S. Vimal, M. Kaliappan, and M. Y. Lee, "AI-based smart prediction of clinical disease using random forest classifier and Naive Bayes," *Journal of Supercomputing*, vol. 77, no. 5, pp. 5198–5219, May 2021, doi: 10.1007/s11227-020-03481-x.

[12] X. Zou *et al*., "Design of Electronic Nose Detection System for Apple Quality Grading Based on Computational Fluid Dynamics Simulation and K-Nearest Neighbor Support Vector Machine," *Sensors*, vol. 22, no. 8, Apr. 2022, doi: 10.3390/s22082997.

[13] A. Ghosh *et al*., "Development of Electronic Nose for early spoilage detection of potato and onion during post-harvest storage." [Online]. Available: https://pubs.thesciencein.org/jmns

[14] J. Rastenis, S. Ramanauskaitė, I. Suzdalev, K. Tunaitytė, J. Janulevičius, and A. Čenys, "Multi-language spam/phishing classification by email body text: Toward automated security incident investigation," *Electronics (Switzerland)*, vol. 10, no. 6, pp. 1–10, Mar. 2021, doi: 10.3390/electronics10060668.

[15] K. ; Kirasich, T. ; Smith, and B. Sadler, "Random Forest vs Logistic Regression: Binary Classification for Heterogeneous Datasets." [Online]. Available: https://scholar.smu.edu/datasciencereviewAvailable at:https://scholar.smu.edu/datasciencereview/vol1/iss3/9http://digitalrepository.smu.edu.

[16] D. R. Wijaya, R. Sarno, and E. Zulaika, "DWTLSTM for electronic nose signal processing in beef quality monitoring," *Sensors and Actuators, B: Chemical*, vol. 326, Jan. 2021, doi: 10.1016/j.snb.2020.128931.

[17] D. R. Wijaya and F. Afianti, "Information-Theoretic Ensemble Feature Selection with Multi-Stage Aggregation for Sensor Array Optimization," *IEEE Sensors Journal*, vol. 21, no. 1, pp. 476–489, Jan. 2021, doi: 10.1109/JSEN.2020.3000756.

[18] D. R. Wijaya, R. Sarno, and E. Zulaika, "Electronic nose dataset for beef quality monitoring in uncontrolled ambient conditions," *Data in Brief*, vol. 21, pp. 2414–2420, Dec. 2018, doi: 10.1016/j.dib.2018.11.091.

[19] Z. Ma, G. Luo, K. Qin, N. Wang, and W. Niu, "Online sensor drift compensation for E-nose systems using domain adaptation and extreme learning machine," *Sensors (Switzerland)*, vol. 18, no. 3, Mar. 2018, doi: 10.3390/s18030742.

[20] A. Manhas, N. Chambyal, M. Raina, Dr. S. Dutta, and Er. P. Singh, "LPG Gas Leakage Detection Using IOT," *International Journal of Scientific Research in Science and Technology*, pp. 520–526, Aug. 2021, doi: 10.32628/cseit2174121.

[21] D. A. H. Fakra, D. A. S. Andriatoavina, N. A. M. N. Razafindralambo, K. abdallah Amarillis, and J. M. M. Andriamampianina, "A simple and low-cost integrative sensor system for methane and hydrogen measurement," *Sensors International*, vol. 1, Jan. 2020, doi: 10.1016/j.sintl.2020.100032.

[22] T. H. Nasution, A. Hizriadi, K. Tanjung, and F. Nurmayadi, "Design of Indoor Air Quality Monitoring Systems," in *2020 4rd International Conference on Electrical, Telecommunication and Computer Engineering (ELTICOM)*, Sep. 2020, pp. 238–241. doi: 10.1109/ELTICOM50775.2020.9230511.

[23] D. Hofstetter, E. Fabian, and A. G. Lorenzoni, "Ammonia generation system for poultry health research using arduino," *Sensors*, vol. 21, no. 19, Oct. 2021, doi: 10.3390/s21196664.

[24] A. Pangestu, D. R. Wijaya, E. Hernawati, and W. Hidayat, "Wrapper Feature Selection for Poverty Level Prediction Based on E-Commerce Dataset," in *2020 International Conference on Data Science and Its Applications (ICoDSA)*, Aug. 2020, pp. 1–7. doi: 10.1109/ICoDSA50139.2020.9212999.

[25] I. Reis, D. Baron, and S. Shahaf, "Probabilistic Random Forest: A Machine Learning Algorithm for Noisy Data Sets," *The Astronomical Journal*, vol. 157, no. 1, p. 16, Dec. 2018, doi: 10.3847/1538-3881/aaf101.

[26] T. Hengl, M. Nussbaum, M. N. Wright, G. B. M. Heuvelink, and B. Gräler, "Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables," *PeerJ*, vol. 2018, no. 8, 2018, doi: 10.7717/peerj.5518.

[27] C. Iwendi *et al*., "COVID-19 patient health prediction using boosted random forest algorithm," *Frontiers in Public Health*, vol. 8, Jul. 2020, doi: 10.3389/fpubh.2020.00357.

[28] A. Botchkarev, "Performance Metrics (Error Measures) in Machine Learning Regression, Forecasting and Prognostics: Properties and Typology."

[29] Universitas Telkom, IEEE Communications Society. Indonesia Chapter, and Institute of Electrical and Electronics Engineers, *The 2021 IEEE Asia Pacific Conference on Wireless & Mobile : proceedings*.

[30] F. Rustam *et al*., "COVID-19 Future Forecasting Using Supervised Machine Learning Models," *IEEE Access*, vol. 8, pp. 101489–101499, 2020, doi: 10.1109/ACCESS.2020.2997311.

[31] D. R. Wijaya, F. Afianti, A. Arifianto, D. Rahmawati, and V. S. Kodogiannis, "Ensemble machine learning approach for electronic nose signal processing," *Sens Biosensing Res*, vol. 36, Jun. 2022, doi: 10.1016/j.sbsr.2022.100495.

Random Forest Algorithm for Meat Classification and Microbial Population Prediction

ORIGINALITY REPORT

18%
SIMILARITY INDEX

PRIMARY SOURCES

1	Submitted to Telkom University Your Indexed Documents	69 words — 2%
2	www.researchgate.net Internet	61 words — 2%
3	Dedy Rahman Wijaya, Riyanarto Sarno, Enny Zulaika. "DWTLSTM for electronic nose signal processing in beef quality monitoring", Sensors and Actuators B: Chemical, 2021 Crossref	58 words — 2%
4	Submitted to Telkom University Your Indexed Documents	35 words — 1%
5	Dedy Rahman Wijaya. "Information-Theoretic Ensemble Feature Selection With Multi-Stage Aggregation For Sensor Array Optimization", IEEE Sensors Journal, 2020 Crossref	34 words — 1%
6	arxiv.org Internet	31 words — 1%
7	link.springer.com Internet	31 words — 1%

8	www.hindawi.com Internet	30 words — 1%
9	www.ijeat.org Internet	19 words — 1%
10	oamjms.eu Internet	17 words — 1%
11	Syahrizal Hanif, Dedy Rahman Wijaya, Wawa Wikusna. "K-Nearest Neighbors Algorithm for Prediction Shelf Life of Rice Based on Electronic Nose Dataset", 2021 IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob), 2021 Crossref	15 words — < 1%
12	export.arxiv.org Internet	15 words — < 1%
13	www.ncbi.nlm.nih.gov Internet	15 words — < 1%
14	Praywin Moses Dass Alex, Akash Ravikumar, Jerritta Selvaraj, Arun Sahayadhas. "Research on Human Activity Identification Based on Image Processing and Artificial Intelligence", International Journal of Engineering & Technology, 2018 Crossref	12 words — < 1%
15	repository.tudelft.nl Internet	12 words — < 1%
16	Derya Aytan-Aktug, Philip TLC Clausen, Judit Szarvas, Patrick Munk et al. "PlasmidHostFinder: Prediction of plasmid hosts using random forest", Cold Spring Harbor Laboratory, 2021 Crossref Posted Content	11 words — < 1%

17	123dok.com Internet	9 words — < 1%
18	Furqan Rustam, Aijaz Ahmad Reshi, Arif Mehmood, Saleem Ullah, Byung-Won On, Waqar Aslam, Gyu Sang Choi. "COVID-19 Future Forecasting Using Supervised Machine Learning Models", IEEE Access, 2020 Crossref	9 words — < 1%
19	Submitted to Telkom University Your Indexed Documents	9 words — < 1%
20	www.notebookcheck.net Internet	9 words — < 1%
21	Damien Ali Hamada FAKRA, Dolly Armel Sylvain ANDRIATOAVINA, Ny Aina Manitra Nany RAZAFINDRALAMBO, Kamil abdallah AMARILLIS et al. "A simple and low-cost integrative sensor system for methane and hydrogen measurement", Sensors International, 2020 Crossref	8 words — < 1%
22	Submitted to Telkom University Your Indexed Documents	8 words — < 1%
23	dedyrw.staff.telkomuniversity.ac.id Internet	8 words — < 1%
24	Dedy Rahman Wijaya, Farah Afianti, Anditya Arifianto, Dewi Rahmawati, Vassilis S. Kodogiannis. "Ensemble machine learning approach for electronic nose signal processing", Sensing and Bio-Sensing Research, 2022 Crossref	7 words — < 1%
25	Lecture Notes in Electrical Engineering, 2015. Crossref	6 words — < 1%