TheJournalish: Social and Government
http://thejournalish.com/ojs/index.php/thejournalish/index
Vol. 4 No. 5 (2023): Social and Government
DOI: https://doi.org/10.55314/tsg.v4i5.607 Hal. 168-180

Integrating Water Indicators In A Data-Driven Artificial Intelligence Model For Food Security Classification

Harun Al Azies
Study Program in Informatics Engineering, Faculty of Computer Science, Universitas Dian Nuswantoro
Correspondence Email: harun.alazies@dsn.dinus.ac.id

Abstrak: Penelitian ini menjelaskan integrasi indikator air dalam pengklasifikasian ketahanan pangan provinsi di Indonesia. Pentingnya topik ini terkait dengan hubungan yang signifikan antara air, pangan, dan energi dalam mencapai ketahanan pangan berkelanjutan. Penggunaan kecerdasan buatan, khususnya dalam bentuk algoritma XGBoost, merupakan langkah cerdas dalam mengolah data dan melakukan pengklasifikasian ketahanan pangan. Data indikator air dan cut-off point Indeks Ketahanan Pangan digunakan dalam mengembangkan model XGBoost yang bertujuan mengklasifikasikan provinsi-provinsi sebagai high food vulnerability atau "high food security. Metode penelitian melibatkan pengumpulan data, preprocessing data, serta penerapan algoritma XGBoost dengan tuning parameter. Hasil penelitian menunjukkan bahwa model yang dikembangkan memiliki akurasi sebesar 91%, dengan variabel proporsi rumah tangga yang memiliki akses terhadap sumber air minum yang aman (X1) sebagai faktor paling berpengaruh dalam pengklasifikasian ketahanan pangan. Penelitian ini bukan hanya memberikan wawasan penting terkait ketahanan pangan provinsi di Indonesia, tetapi juga menunjukkan potensi besar kecerdasan buatan dalam mengatasi permasalahan kompleks seperti ketahanan pangan. Dengan hasil yang diperoleh, dapat diperkuat argumen pentingnya penerapan teknologi kecerdasan buatan dalam mendukung kebijakan dan tindakan nyata dalam upaya mencapai ketahanan pangan yang lebih baik dan berkelanjutan.

Kata kunci: Klasifikasi Ketahanan Pangan, Kecerdasan Buatan, Indikator Air, Algoritma XGBoost, Kebijakan Berkelanjutan

Abstract: *This study explains how water indicators are used to classify regional food security in Indonesia. The significance of this topic stems from the critical link between water, food, and energy in establishing long-term food security. The employment of artificial intelligence, particularly in the form of the XGBoost algorithm, is a wise move in data processing and food security classification. The XGBoost model was developed using water indicator data and Food Security Index cut-off points to classify provinces as "high food vulnerability" or "high food security." Data collecting, data preprocessing, and the use of the XGBoost algorithm with parameter tuning are all part of the study approach. According to the research findings, the constructed model has a 91% accuracy rate, with the proportion of households with access to a source of safe drinking water (X1) being the most relevant element in defining food security. This study not only gives valuable insights into provincial food security in Indonesia but also demonstrates the tremendous potential of artificial intelligence in overcoming complicated problems like food security. With the findings, the case for using artificial intelligence technology to help policy and specific actions in attempts to achieve better and more sustainable food security can be enhanced*.

Keywords: *Food Security Classification, Artificial Intelligence, Water Indicators, XGBoost Algorithm, Sustainable Policy*

Article History :
Received; 14-09-2023; Revised; 07-10-2023; Accepted; 04-11-2023

INTRODUCTION

One of the fundamentals in the attempt to accomplish the Sustainable Development Goals (SDGs) established by the United Nations (UN) is food security (Gil et al., 2019; Pérez-Escamilla, 2017). Several concerns and elements are becoming more significant in the context of sustainable food security, such as the availability of clean water, resilience to climate change, equal access to nutrient-dense food, and agricultural productivity (Çakmakçı et al., 2023). The realization of sustainable food security is contingent upon the water nexus—the relationship between food, water, and energy (Smajgl et al., 2016). This phenomenon emphasizes the interdependence of the availability of sufficient and effective water resources and food security. In a world that is getting more interconnected and complex by the day, it is crucial to comprehend these problems and find solutions(Liu et al., 2015). However, with the rapidly growing world population and the grave challenges posed by climate change, achieving sustainable food security is challenging (Rasul, 2021). Water problems are often related to food security problems in terms of quantity and quality(Pahl-Wostl, 2019). To address the unpredictability of the food supply, water resource efficiency, protection, and improved management are essential.

As a result, by emphasizing the critical importance of water resources, our research aims to contribute to the accomplishment of long-term food security. The purpose of this study is to employ machine learning and artificial intelligence approaches to categorize the level of food security in each Indonesian region. When classifying this population, the proportion of households with access to a source of safe drinking water, the proportion of households using safely managed sanitation services, the proportion of households with open defecation, the proportion of households that have a handwashing facility with soap and water at home and the number of drinking water companies, will all be considered. This research entails a thorough grasp of the need to incorporate variables of water availability into a food security framework. Regions with a lower level of food security frequently confront major challenges in terms of access to sustainable clean water and suitable sanitation facilities. This is where this research comes in handy, identifying locations that are sensitive to food supply volatility and supporting the achievement of SDG food security targets.

Several earlier studies have looked into the topic of food security. For example, Widada et al., looked into the agricultural and economic aspects of food security in Indonesia. They discovered that food security was significantly impacted by the output of rice, corn, soybeans, chicken, and beef as well as inflation indicators like the Consumer Price Index (CPI) (Widada et al., 2017). Furthermore, clustering approaches were utilized in research by Rahayu et al., to categorize levels of food security. Based on how food security levels were classified, the research's findings yielded three clusters that offer a more in-depth understanding of regional differences in food security (Rahayu et al., 2019). Furthermore, a study conducted in Pandaan

District, Pasuruan Regency, Indonesia by Prayitno et al., integrated an examination of livelihood assets and food security indices. The findings of this study delineate the constituents of livelihood assets and identify the villages that are classified as food insecure (Prayitno et al., 2019).

Nonetheless, the previous study has not explicitly emphasized the significance of water resources and food security in Indonesia. Herein is the originality and innovation that this research delivers. A lot of water variables that may have an impact on food security can be thoroughly analyzed by us thanks to our artificial intelligence (AI) and machine learning technology (Al Azies & Herowati, 2023). To help policymakers allocate resources and efforts to promote food security in the relevant regions, we want to use our model machine to provide more accurate and timely information about the state of food security in each province in Indonesia. The novel element of this study is not only the analytical approach used but also the emphasis on the function of water in ensuring food security. This research, by taking this technique, fills a knowledge gap that has hitherto gone unnoticed in the scientific literature. The findings of this study will make an essential contribution to Indonesia's attempts to attain long-term food security.

RESEARCH METHODS

The research methodology is explained in detail in this chapter and is illustrated in Figure 1's study framework. Utilizing a dataset serves as the foundation for analysis and is where it all begins. There are two categories of data in the dataset: data on food security and data on water indicators. The secondary data used in this study was obtained from government entities. This data set contains two different types of variables: the target variable, Food Security Index Cut-off Point data received from the National Food Agency (NFA) (Anwar, 2022), and the water indicator data, variables the proportion of households with access to a source of safe drinking water (X_1), the proportion of households using safely managed sanitation services (X_2), the proportion of households with open defecation (X_3), the proportion of households that have a handwashing facility with soap and water at home (X_4) and the number of drinking water companies (X_5), collected from Statistics Indonesia. As its principal analytical technique, this study employs machine learning, or more specifically, artificial intelligence. In the machine learning framework, two types of variables are crucial. The first variable in the analysis is the category of the food security index. The Food Security Index Cut-Off Point of the National Food Agency acts as a classification reference for this study. Provinces with a score of 74 or higher on this index are considered to have "high food security," while those with a score less than 74 are considered to have "high food vulnerability." The second variable is the feature variable, which has five indications that represent the dimensions of water. These measures are used as qualities that influence. The second variable is the feature variable, which includes five indicators that reflect the dimensions of water. These factors are utilized as criteria that determine food security classification. The analysis process uses a mix of target variables and feature variables to construct a model that can predict the food security status of Indonesian regions.

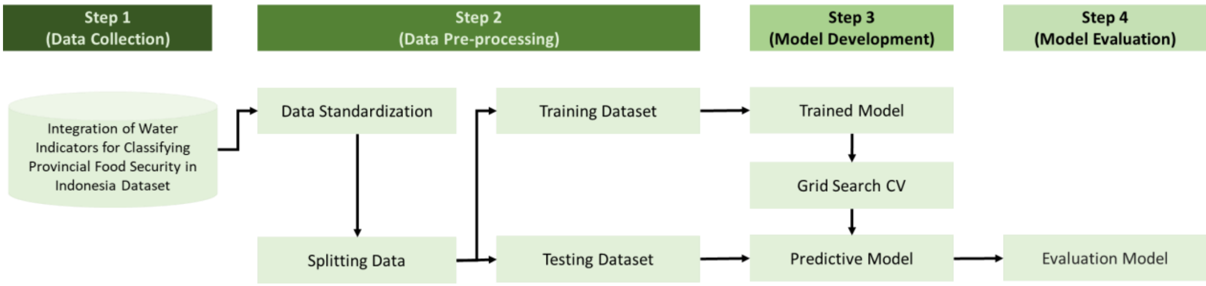

Figure 1. Research framework for food security classification models

The following stage is data preprocessing, which involves treating the data to ensure that everything is on a consistent scale(García et al., 2015; Luengo et al., 2020). This is critical for optimizing model performance. Following that, the data is split into two parts: training data and testing data(Nguyen et al., 2021). This divide is carried out in a 70:30 ratio, with training data used to train the model and testing data used to determine the model's ability to predict accurately (Muraina, 2022). The development of models is crucial to this study. The XGBoost machine learning method is used at this level to classify provinces based on their food security status. A machine learning approach called XGBoost (Extreme Gradient Boosting) is applied to regression and classification problems(Ibrahem Ahmed Osman et al., 2021). Using decision trees as the foundational model, this approach applies ensemble learning techniques. XGBoost is well known for its consistently high performance and is frequently utilized in data science contests (Obiora et al., 2021). To create a complicated model, XGBoost combines numerous relatively simple decision trees. Every decision tree is added to the model in turn, with each new tree attempting to fix the flaws of its predecessors. Because of this, XGBoost performs well when handling complicated and noisy data (Kavzoglu & Teke, 2022). Moreover, XGBoost has several features, such as the capacity to determine feature relevance, automatic handling of missing information, and strong parameter tuning capabilities with GridSearchCV or RandomizedSearchCV (Teja et al., 2021). Because of all this, XGBoost is an effective tool for developing prediction models and analyzing data (Dhaliwal et al., 2018). This procedure entails fine-tuning the algorithm parameters. This entails adjusting the algorithm settings for optimal performance. The GridSearchCV method is used to identify the best combination of parameters that results in the best model (Budholiya et al., 2022).

After the model has been developed, the following step is to evaluate it. Several performance criteria, such as accuracy, recall, and F1 score, are used to evaluate the model (Yacouby & Axman, 2020). These metrics provide an overview of the model's ability to categorize provinces as "high food vulnerability" or "high food security." The research findings also include a variable importance analysis, which indicates each variable's contribution to the model. High-importance variables have a considerable impact on food security classification. This entire procedure attempts to provide a more in-depth understanding of food security in various Indonesian provinces. As a result, our research can make an essential contribution to provincial decision-making regarding policies and resources to increase food security.

RESULTS AND DISCUSSION

This is a quantitative study that classifies the level of food security in each Indonesian region using numerical data. This study used XGBoost, a machine learning algorithm that is widely used in quantitative data analysis for categorization purposes. The goal of this research is to use water indicators that indicate the degree of food security to categorize provinces as having high food resilience or high food vulnerability. The GridSearchCV approach is used to adjust the algorithm parameters as the first stage in this research. The optimal set of parameters for the machine learning model is determined using this technique. The goal of this parameter-tuning process is to optimize the model's performance and increase the accuracy of the predictions made about the degree of food security in different regions. Training data was used in this study's parameter adjustment or parameter tuning. A set of data known as training data is used to fine-tune the XGBoost machine learning model so that it can accurately classify the state of food security. When the model's parameters are optimized, it can predict food vulnerability more accurately when applied to test data. This procedure serves as a crucial foundation for accomplishing research goals and assisting initiatives to achieve sustainable food security. 'n_estimators' with options [100, 200, 300], 'max_depth' with options [3, 4, 5], and 'learning_rate' with options [0.01, 0.1, 0.2] are among the adjusted parameters (Arifin et al., 2023). To guarantee that the XGBoost model can produce extremely accurate classification results, this is done to determine the most ideal combination of parameters. Using training data, the parameter tuning procedure entails testing with alternative combinations of parameter values, such as learning rate, max depth, and number of estimators. Finding the ideal set of parameters to create a model with the best performance is the aim of parameter tweaking. To put it another way, searching for parameters that let the model both effectively learn patterns from the training data and generalize to test data or data that has never been seen before.

Table 1. Experimental results of tuning parameters of the XGboost algorithm on training data

Learning rate	max depth	n estimators	Accuracy	Learning rate	max depth	n estimators	Accuracy
0.01	3	100	66%	0.1	3	100	59%
0.01	3	200	79%*	0.1	3	200	59%
0.01	3	300	78%	0.1	3	300	59%
0.01	4	100	66%	0.1	4	100	59%
0.01	4	200	78%	0.1	4	200	59%
0.01	4	300	78%	0.1	4	300	59%
0.01	5	100	66%	0.1	5	100	59%
0.01	5	200	78%	0.1	5	200	59%
0.01	5	300	78%	0.1	5	300	59%
0.2	3	100	59%	0.2	4	300	55%
0.2	3	200	59%	0.2	5	100	59%
0.2	3	300	59%	0.2	5	200	59%
0.2	4	100	59%	0.2	5	300	55%
0.2	4	200	59%				

*) Best parameters with highest performance

Table 1 displays the experimental results of optimizing the XGBoost algorithm settings using training data. This study focuses on three main parameters: learning rate, max depth, and n_estimators (number of estimators)(A. Gupta et al., 2020). The accuracy results are given as a percentage. During the parameter-tuning process, the ideal parameters that produce the best performance are found and are denoted with an asterisk. The best values are max depth of 3, learning rate of 0.01, and n_estimators of 200, yielding a 79% accuracy rate. This indicates that using a more complex model with a lower learning curve is the most efficient way to classify the level of food security in Indonesian regions. In this study, the optimal parameters will be discovered and test data will be used to predict food security status. Applying the model with the best parameters to test data will enable a deeper comprehension of the level of food security in every Indonesian province. This allows for the identification of the provinces with high levels of food vulnerability as well as those that have successfully established strong food security. The basis that this data will provide will be essential for formulating policies and allocating funds to regions that need more help to achieve sustainable food security.

The confusion matrix in this study's application of the best model from the previous stage using testing data is the outcome of applying the XGBoost algorithm to classify the food security status of Indonesian regions (Dharmawan et al., 2022). Figure 2 shows the degree to which these provinces may be categorized into two primary groups using the established model: "high food vulnerability" and "high food security."

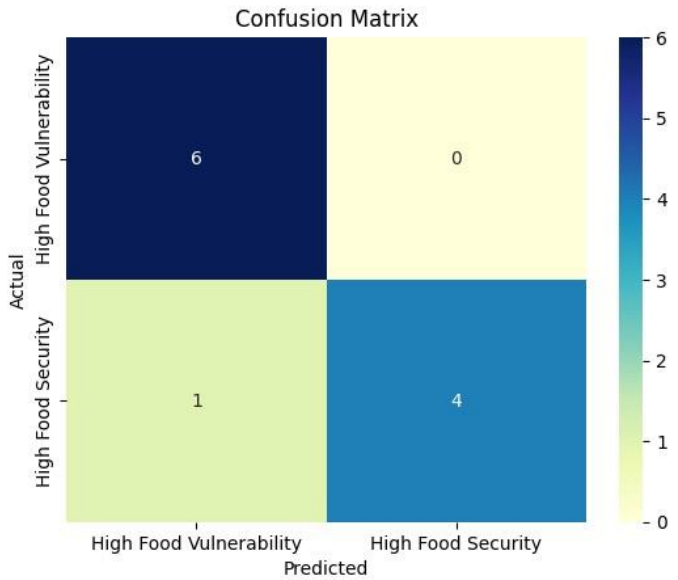

Figure 2. Confusion matrix of food security classification model

The model correctly predicted that all six provinces would be classified as "High Food Vulnerability," but the actual class is "High Food Vulnerability." As a result of the model's success in identifying provinces with the lowest levels of food security, this is a very favorable consequence. These provinces are included in the True Positive (TP) category. Furthermore, the model predicts "High Food Security" correctly in four provinces where the actual class is

"High Food Security." This illustrates that the model can identify provinces with high levels of food security. This result is also classified as a True Positive (TP). It should be noted, however, that there is one province where the actual class is "High Food Vulnerability," yet the model predicts it as "High Food Security." This is a false Negative (FN) example, showing that the model did not accurately identify the province. This confusion matrix gives an overview of how well the methodology employed in this study can categorise provinces in Indonesia into two major groups: "high food vulnerability" and "high food security." A careful examination of this matrix can assist researchers in improving models and developing more effective policies to attain longterm food security.

The investigation is then carried on by analyzing the model performance outcomes using the confusion matrix shown in Figure 2. True positive (TP), true negative (TN), false positive (FP), and false negative (FN) are the outcomes of the confusion matrix, and they are particularly significant in evaluating model performance. These results will then be utilized to determine many key performance indicators, including precision, recall, F1-score, and accuracy, as shown in Classification Report Table 3.

Table 3. Food security classification model performance report

Class	Precision	Recall	F1-score	Accuracy
High Food Vulnerability	0.86	1.00	0.92	0.91
High Food Security	1.00	0.89	0.89	

Model performance analysis based on precision value:
The precision of a model is determined by how well it can identify which provinces, out of all those designated as having "high food vulnerability," are categorized as such. Precision gauges the model's degree of positive precision(D. L. Gupta et al., 2012). For the class "High Food Vulnerability," the precision is 0.86. This indicates that of all provinces predicted to have "high food vulnerability," up to 86% do. This precision demonstrates the model's ability to detect provinces with low food security. Precision is 1.00 for the class "High Food Security" This demonstrates that, of the provinces anticipated to have "high food resilience," all of them do. The model's high precision shows that it is highly good at identifying provinces with high levels of food security.

Model performance analysis based on recall value:
The model's recall quantifies the degree to which it can accurately identify all provinces that genuinely have "high food vulnerability" as being "high food vulnerable." Recall gauges how well the algorithm can identify provinces with issues related to food security(Sofaer et al., 2019). The category "High Food Vulnerability" has a recall of 1.00. That is, the model correctly identified all provinces in the "High Food Vulnerability" group. This high recall indicates that the model does not overlook provinces with low food security. The recall rate for the "High Food Security" class was 0.80. This means that the model correctly recognized 80% of the provinces as being in the " High Food Security" category.

Model performance analysis based on F1 score:

The XGBoost model's capacity to strike a compromise between detecting the bulk of provinces with significant food vulnerability issues and accuracy in recognizing those provinces may be gauged by looking at the F1 Score(Chicco & Jurman, 2020). The F1 score is a metric that combines precision and recall into a single metric. The F1 score for the "High Food Vulnerability" class was 0.92. Selecting provinces with low food security demonstrates a fair combination of precision and recall. The F1-score for the " High Food Security" class was 0.89, demonstrating a fair balance of precision and recall in selecting regions with high food security.

Model performance analysis based on Accuracy:

The degree to which a model can accurately identify data is measured by its accuracy(Bernadó-Mansilla & Garrell-Guiu, 2003). Accuracy in this context is defined as the proportion of provinces accurately identified as either "high food security" or "high food vulnerability." The model performs better with the greater the accuracy value. The Accuracy is the model's overall success in classifying provinces. In this scenario, the accuracy is 0.91, indicating that the model successfully classified 91% of the provinces. According to the Classification Report, the model utilized was successful in classifying provinces in Indonesia into two groups, "High Food Vulnerability" and "High Food Security," with high precision and recall. The ultimate accuracy is 0.91, indicating that this model performs well in forecasting food security status in each province. This data will serve as a solid foundation for better policy decisions in Indonesia's efforts to attain long-term food security.

Following that, a variable importance analysis was performed to establish the extent to which particular variables influence food security in various Indonesian provinces (Figure 3). The findings of variable importance analysis are extremely useful in the context of decision-making and policy-building since they aid in identifying the essential elements that most influence food security. With a value of 0.741, the first variable, the proportion of households with access to a source of safe drinking water (X_1), has the highest relevance score. This high significance score demonstrates how important adequate access to safe drinking water is in determining food security. The proportion of households with open defecation (X_3) gets a significance score of 0.130. This suggests that open defecation in public is still a severe concern in these provinces in terms of food security. This variable has a major impact, and preventive interventions to limit open defecation practises could be part of a food security plan.

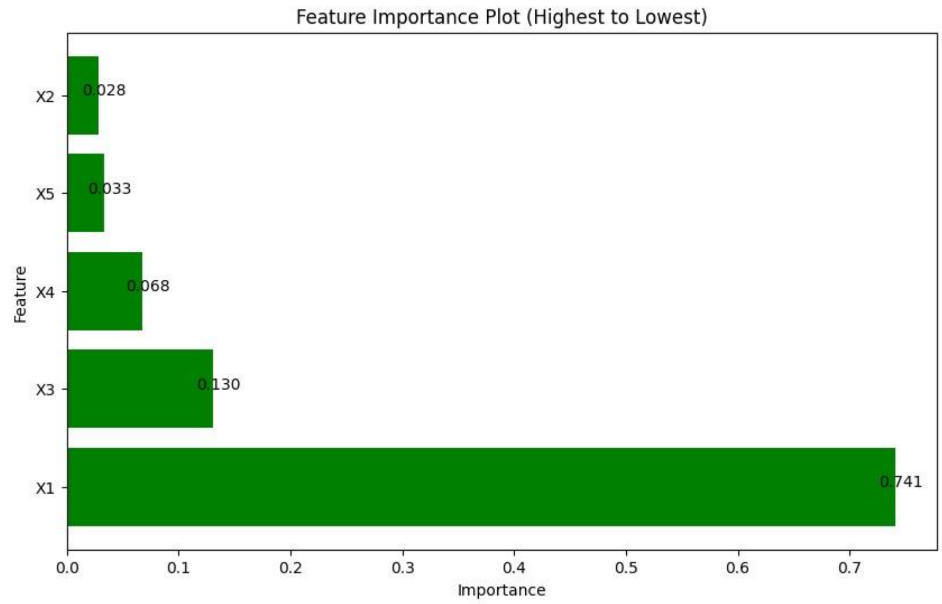

Figure 3. Level of importance of variables in the food security classification model

The third variable is the proportion of households that have a handwashing facility with soap and water at home (X_4). Even though it has a lower relevance score than the first two factors (0.068), it nevertheless makes a considerable contribution to food security. The availability of soap and water for handwashing in households is a crucial aspect of preserving cleanliness and public health. The fourth variable, with a significance value of 0.033, is the number of drinking water companies (X_5). This variable shows the availability of safe drinking water in the area. Even if it contributes less than other variables, it is nonetheless crucial in the context of food security because access to safe water sources is a requirement for supplying sufficient and healthy food.

Finally, the proportion of households using safely managed sanitation services (X_2) is the fifth variable, which has the lowest relevance score of 0.028, despite its minimal contribution, this variable is still essential since adequate sanitation improves people's health and food security. The importance of these variables can be utilized to determine priorities when developing policies and activities to increase food security at the provincial level. Variables having a high relevance score, such as access to safe drinking water (X_1) and the reduction of open defecation practices (X_3), may require further focus in attempts to promote food security. This insight will assist policymakers in allocating resources more efficiently and effectively to achieve long-term food security in all provinces.

CONCLUSION

This study emphasizes that the XGBoost machine learning model, when tuned to the best parameters, performs well in classifying the level of food security in Indonesian provinces. This model has an accuracy of roughly 91%, indicating its ability to identify areas of significant food vulnerability as well as those of good food security. Aside from that, this model has strong precision and recall for both classes, namely high food vulnerability and high food resilience, according to the classification report. The main contribution of this research is to show how water-related problems significantly affect food security. This provides a basis for better decision-making when creating suitable policies and actions, as well as an improved understanding of the complexities of food security concerns. For the government, nonprofits, and other interested parties trying to increase food security in Indonesia, the study's conclusions are helpful. By recognizing the crucial role that water plays in the context of food security, the nation's long-term goal of food security can be supported by taking deliberate steps to improve vulnerable areas' access to clean water, sanitary conditions, and hygiene. Based on the findings of this study, policy recommendations include priority access to clean water, hygiene programs, longterm monitoring and evaluation, community empowerment, long-term water resource management, and the need for additional research. It is envisaged that implementing these recommendations will boost food security in Indonesian provinces, promote community welfare, and help attain long-term food security.

DAFTAR PUSTAKA

Al Azies, H., & Herowati, W. (2023). Unravelling Income Inequality in Indonesia. Jurnal Riset Ilmu Ekonomi, 3(2), 89–100. https://doi.org/10.23969/JRIE.V3I2.63

Anwar, N. (2022). Indonesia’s Regional Food Security In Light Of The Impending Global Food Crisis. Trikonomika, 21(2), 101–110. https://doi.org/10.23969/TRIKONOMIKA.V21I2.7113

Arifin, M., Widowati, W., & Farikhin, F. (2023). Optimization of Hyperparameters in Machine Learning for Enhancing Predictions of Student Academic Performance. Ingenierie Des Systemes d’Information, 28(3), 575–582. https://doi.org/10.18280/ISI.280305

Bernadó-Mansilla, E., & Garrell-Guiu, J. M. (2003). Accuracy-Based Learning Classifier Systems:

Models, Analysis and Applications to Classification Tasks. Evolutionary Computation, 11(3), 209–238. https://doi.org/10.1162/106365603322365289

Budholiya, K., Shrivastava, S. K., & Sharma, V. (2022). An optimized XGBoost based diagnostic system for effective prediction of heart disease. Journal of King Saud University - Computer and Information Sciences, 34(7), 4514–4523. https://doi.org/10.1016/J.JKSUCI.2020.10.013

Çakmakçı, R., Salık, M. A., & Çakmakçı, S. (2023). Assessment and Principles of Environmentally

Sustainable Food and Agriculture Systems. Agriculture 2023, Vol. 13, Page 1073, 13(5), 1073. https://doi.org/10.3390/AGRICULTURE13051073

Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics, 21(1), 1–13. https://doi.org/10.1186/S12864-019-6413-7/TABLES/5

Dhaliwal, S. S., Nahid, A. Al, & Abbas, R. (2018). Effective Intrusion Detection System Using XGBoost. Information 2018, Vol. 9, Page 149, 9(7), 149. https://doi.org/10.3390/INFO9070149

Dharmawan, H., Sartono, B., Kurnia, A., Hadi, A. F., & Ramadhani, E. (2022). A study of machine learning algorithms to measure the feature importance in class-imbalance data of food insecurity cases in Indonesia. Commun. Math. Biol. Neurosci., 2022(0), Article ID 101. https://doi.org/10.28919/CMBN/7636

García, S., Luengo, J., & Herrera, F. (2015). Data preparation basic models. Intelligent Systems Reference Library, 72, 39–57. https://doi.org/10.1007/978-3-319-10247-4_3

Gil, J. D. B., Reidsma, P., Giller, K., Todman, L., Whitmore, A., & van Ittersum, M. (2019). Sustainable development goal 2: Improved targets and indicators for agriculture and food security. Ambio,

48(7), 685–698. https://doi.org/10.1007/S13280-018-1101-4/TABLES/3

Gupta, A., Sharma, S., Goyal, S., & Rashid, M. (2020). Novel XGBoost Tuned Machine Learning Model for Software Bug Prediction. Proceedings of International Conference on Intelligent Engineering and Management, ICIEM 2020, 376–380. https://doi.org/10.1109/ICIEM48762.2020.9160152

Gupta, D. L., Malviya, A. K., & Singh, S. (2012). Performance Analysis of Classification Tree Learning Algorithms. International Journal of Computer Applications, 55(6), 975–8887.

Ibrahem Ahmed Osman, A., Najah Ahmed, A., Chow, M. F., Feng Huang, Y., & El-Shafie, A. (2021). Extreme gradient boosting (Xgboost) model to predict the groundwater levels in Selangor Malaysia. Ain Shams Engineering Journal, 12(2), 1545–1556. https://doi.org/10.1016/J.ASEJ.2020.11.011

Kavzoglu, T., & Teke, A. (2022). Predictive Performances of Ensemble Machine Learning Algorithms in Landslide Susceptibility Mapping Using Random Forest, Extreme Gradient Boosting (XGBoost) and Natural Gradient Boosting (NGBoost). Arabian Journal for Science and Engineering, 47(6), 7367–7385. https://doi.org/10.1007/S13369-022-06560-8/FIGURES/12

Liu, J., Mooney, H., Hull, V., Davis, S. J., Gaskell, J., Hertel, T., Lubchenco, J., Seto, K. C., Gleick, P., Kremen, C., & Li, S. (2015). Systems integration for global sustainability. Science, 347(6225). https://doi.org/10.1126/SCIENCE.1258832/ASSET/F0BA4106-BF5F-4AE3-9D77-2E70C1A5C51F/ASSETS/GRAPHIC/347_1258832_FA.JPEG

Luengo, J., García-Gil, D., Ramírez-Gallego, S., García, S., & Herrera, F. (2020). Big Data Preprocessing: Enabling Smart Data. Big Data Preprocessing: Enabling Smart Data, 1–186. https://doi.org/10.1007/978-3-030-39105-8/COVER

Muraina, I. O. (2022). Ideal Dataset Splitting Ratios In Machine Learning Algorithms: General Concerns For Data Scientists And Data Analysts. https://www.researchgate.net/publication/358284895

Nguyen, Q. H., Ly, H. B., Ho, L. S., Al-Ansari, N., Van Le, H., Tran, V. Q., Prakash, I., & Pham, B. T. (2021). Influence of data splitting on performance of machine learning models in prediction of shear strength of soil. Mathematical Problems in Engineering, 2021. https://doi.org/10.1155/2021/4832864

Obiora, C. N., Ali, A., & Hasan, A. N. (2021). Implementing extreme gradient boosting (xgboost) algorithm in predicting solar irradiance. 2021 IEEE PES/IAS PowerAfrica, PowerAfrica 2021. https://doi.org/10.1109/POWERAFRICA52236.2021.9543159

Pahl-Wostl, C. (2019). Governance of the water-energy-food security nexus: A multi-level coordination challenge. Environmental Science & Policy, 92, 356–367. https://doi.org/10.1016/J.ENVSCI.2017.07.017

Pérez-Escamilla, R. (2017). Food Security and the 2015–2030 Sustainable Development Goals: From Human to Planetary Health: Perspectives and Opinions. Current Developments in Nutrition, 1(7). https://doi.org/10.3945/CDN.117.000513

Prayitno, G., Safitri, N., Subagiyo, A., & Riska, B. (2019). Food security index and livelihood assets of Pandaan District, Pasuruan Regency, Indonesia. IOP Conference Series: Earth and Environmental Science, 361(1), 012029. https://doi.org/10.1088/1755-1315/361/1/012029

Rahayu, R. S., Purwaningsih, Y., & Daerobi, A. (2019). Mapping Of Provincial Food Security In Indonesia Using Based Clustering Model. Jurnal Ekonomi Pembangunan: Kajian Masalah Ekonomi Dan Pembangunan, 20(1), 69–79. https://doi.org/10.23917/JEP.V20I1.7096

Rasul, G. (2021). A Framework for Addressing the Twin Challenges of COVID-19 and Climate Change for Sustainable Agriculture and Food Security in South Asia. Frontiers in Sustainable Food Systems, 5, 679037. https://doi.org/10.3389/FSUFS.2021.679037/BIBTEX

Smajgl, A., Ward, J., & Pluschke, L. (2016). The water–food–energy Nexus – Realising a new paradigm. Journal of Hydrology, 533, 533–540. https://doi.org/10.1016/J.JHYDROL.2015.12.033

Sofaer, H. R., Hoeting, J. A., & Jarnevich, C. S. (2019). The area under the precision-recall curve as a performance metric for rare binary events. Methods in Ecology and Evolution, 10(4), 565–577. https://doi.org/10.1111/2041-210X.13140

Teja, K. U. V. R., Reddy, B. P. V., Likith Preetham, A., Patil, H. Y., & Poorna Chandra, T. (2021).

Prediction of Diabetes at Early Stage with Supplementary Polynomial Features. Proceedings - 1st International Conference on Smart Technologies Communication and Robotics, STCR 2021.

https://doi.org/10.1109/STCR51658.2021.9588849

Widada, A. W., Masyhuri, M., & Mulyo, J. H. (2017). Determinant Factors of Food Security in Indonesia. Agro Ekonomi, 28(2), 205–219. https://doi.org/10.22146/JAE.26245

Yacouby, R., & Axman, D. (2020). Probabilistic Extension of Precision, Recall, and F1 Score for More Thorough Evaluation of Classification Models. 79–91. https://doi.org/10.18653/V1/2020.EVAL4NLP-1.9