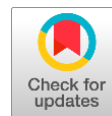


# Data mining-based analysis of factors affecting paddy farming



Iston Utama<sup>a</sup> | Budi Harsanto<sup>b</sup> | Mulyani Karmagatri<sup>a</sup> | Yusuf Iskandar<sup>c</sup> | Yulistyne Kasumaningrum<sup>b</sup> | Intan Rahmatillah<sup>a</sup>

<sup>a</sup>Entrepreneurship Department, BINUS Business School Undergraduate Program, Bina Nusantara University, Indonesia.  
<sup>b</sup>Faculty of Economics & Business, Padjadjaran University, Indonesia.  
<sup>c</sup>Department of Management & Jaya Launch Pad, Universitas Pembangunan Jaya, Indonesia.

**Abstract** This research investigates the key factors influencing the paddy rice harvest class in West Java, Indonesia, with a focus on understanding the interplay between traditional farming knowledge and modern technological solutions to support food security. It is recognize that rice is a staple for a large portion of the global population and critical to food security. This research explores how decision tree algorithms can help identify the most influential factors and support sustainable agricultural practices, especially in the context of small-scale rice farming in West Java, Indonesia. A dataset encompassing demographic variables (age, gender, household involvement), environmental conditions, and farming experience was analyzed using a decision tree model. The model's performance was validated using cross-validation, achieving an average accuracy of 71.43%. In this study reveal that gender and household size of farmers showed moderate influence, reflecting the socio-economic dynamics of family-based farming. Meanwhile, other factors such as land ownership and paddy field type had a limited impact to the paddy productions. These findings highlight the need for structured knowledge transfer programs from experienced from experience farmers to younger farmers and suggest that combining traditional farming practices with technology farming tools might can improve rice yield outcomes. Furthermore, the research applies the CRISP-DM methodology to guide data analysis and ensure alignment with agricultural goals. The integration of data mining in agriculture not only improves understanding of paddy productivity but also supports adaptive strategies in response to climate variability and demographic shifts. This research provides practical implications for farmers, researchers, and policymakers seeking to implement targeted interventions that contribute to sustainable rice production and long-term food resilience in developing regions.

**Keywords:** big data, data mining, decision tree algorithm, food security

## 1. Introduction

Global food security is a pressing concern, and rice, a staple food for a significant portion of the world's population, plays a critical role in addressing this challenge (El Bilali et al., 2020; Harsanto, 2021). The demand for rice continues to surge, as evidenced by the consumption of 520.4 million metric tons globally in 2023 (FAO, 2017; Statista, 2024). Despite global rice production exceeding consumption, reaching 780 million tons with major contributions from China, India, and Bangladesh, as presented in Picture 1, the growth in rice production has slowed in recent years, whereas demand has dramatically increased (Maraseni et al., 2018; Muthayya et al., 2014). This widening gap between supply and demand underscores the urgency of optimizing rice production to ensure food security for future generations (Figure 1).

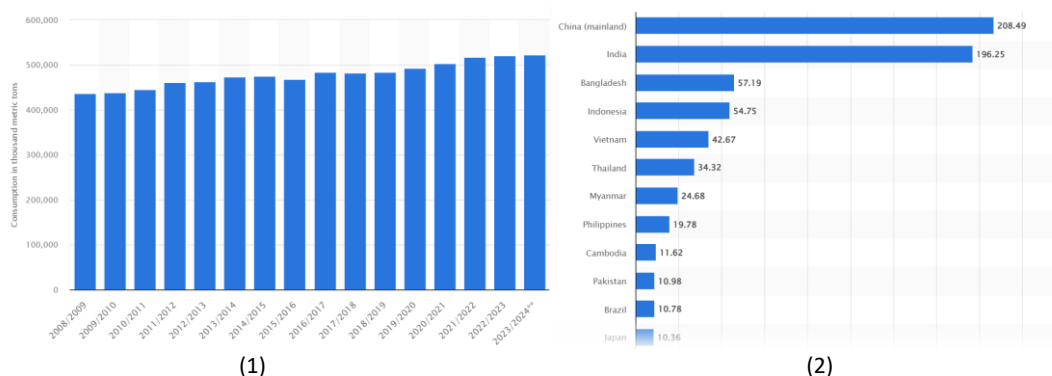


Figure 1 Global Rice Consumption (1) and Global Rice Production (2).  
Source: Statista (2024).



While various farming techniques, including modern farming methods, smart farming, and rice variety engineering, have been implemented to boost rice production, their impact in meeting the escalating demand has been limited (Alfred et al., 2021; Firdaus et al., 2020; Nawaz et al., 2022). Moreover, the reliance on traditional farming practices, while valuable in many contexts, often fails to capitalize on the potential of data-driven approaches to optimize resource allocation and enhance yields (Majumdar et al., 2017; Ngandee et al., 2021).

Data mining, a powerful tool for extracting valuable insights from large datasets, has demonstrated its potential in various industries, including agriculture (Lausch et al., 2015; Sudirman, 2023). Countries such as China, Vietnam, Thailand, and Malaysia have successfully employed data mining techniques to forecast agricultural production, mitigate risks associated with climate variability, and optimize overall productivity (Asli and Rahman, 2023; Boughton et al., 2021; Firdaus et al., 2020; Kong et al., 2021). However, Indonesia has lagged behind in adopting these technological innovations, hindering the potential to leverage data for informed decision-making in rice production (BPS, 2023; Carrasco-Torrontegui et al., 2021).

Furthermore, despite the proven effectiveness of data mining in other agricultural sectors, its application in rice production research remains limited (Fahad et al., 2019; Prabhakaran et al., 2021), with a significant proportion of existing research focusing on the impact of climate change on rice production or relying on conventional statistical approaches to analyze traditional farming practices (Firdaus et al., 2020; Mohamed et al., 2016). This gap in research underscores the need to explore the potential of data mining techniques to optimize rice production in specific contexts, such as Indonesia.

According to the description above, this research aims to address this gap by (1). A decision tree algorithm was used to identify the key factors influencing the paddy rice harvest class in West Java, Indonesia. This analysis provides insights into the complex interplay of demographic, environmental, and resource management variables affecting rice production outcomes. (2). The implications of these findings for knowledge transfer and technology integration in the context of an aging farming population are explored. Given the significant role of experienced farmers in traditional rice production systems, this research examines how data-driven approaches can be integrated with existing knowledge systems to ensure the preservation and transfer of valuable expertise to future generations. Thus, by reviewing these objectives, this research aims to contribute to a deeper understanding of how data mining techniques can be harnessed to increase rice production resilience, optimize resource utilization, and contribute to achieving food security in Indonesia.

## 2. Materials and methods

### 2.1. Data Mining and Agriculture

Data mining, a subset of artificial intelligence and machine learning, refers to extracting useful patterns from large datasets (Lausch et al., 2015; Sudirman, 2023). In the context of agriculture, it plays a pivotal role in addressing complex challenges such as crop yield prediction, soil fertility assessment, and precision farming (Majumdar et al., 2017). Agriculture, a data-rich sector, generates extensive amounts of data through various sources, such as remote sensing, weather stations, soil sampling, and modern agricultural machinery. Paddy rice production, an essential global crop, is an area where data mining has proven to be transformative in ensuring food security.

As agricultural data continues to grow exponentially, traditional methods of managing and analyzing this data have proven insufficient (Firdaus et al., 2020; Mohamed et al., 2016). Data mining offers the ability to derive actionable insights that can enhance decision-making, optimize resources, and increase productivity. Paddy rice, a staple food crop, presents numerous complexities in its production cycle, such as yield variability due to climate change, pest infestations, and soil conditions. Data mining and machine learning techniques have been employed extensively to optimize the entire paddy rice production process, from predicting yields to optimizing water usage.

According to recent studies, machine learning models such as decision trees, support vector machines (SVMs), and random forests are known for their ability to handle high-dimensional data and classify complex patterns; random forests, which combine multiple decision trees to improve prediction accuracy and robustness; and clustering algorithms, which group similar data points on the basis of shared characteristics, aiding in identifying patterns and trends (Son et al., 2020). The applications of these techniques span a wide range of agricultural challenges. For example, machine learning models have been effectively utilized to predict yields for crops such as rice on the basis of a combination of agronomic and environmental factors, empowering farmers with real-time decision-making capabilities for optimizing inputs such as fertilizers, water, and pesticides (Alfred et al., 2021; Witten and Frank, 2002). Clustering techniques have also been instrumental in grouping paddy fields with similar soil characteristics and climatic conditions, enabling the implementation of precision farming techniques that reduce resource waste and increase overall efficiency (Barbosa Júnior et al., 2024).

### 2.2. Decision Tree Algorithm in the Agricultural Sector

Decision tree algorithms are popular machine learning techniques used for both classification and regression tasks (Osei-Bryson, 2004). They work by splitting datasets into subsets based on feature values, following a tree-like structure where each node represents a decision and each branch represents an outcome (Witten and Frank, 2002). Decision trees are valued for their simplicity, interpretability, and ability to model complex relationships without requiring extensive preprocessing of data.

Their clear visual representation makes them especially suitable for users with limited technical expertise, such as farmers and agricultural practitioners. The applications of decision trees extend across various domains, including healthcare, finance, and manufacturing, where decisions based on multiple interacting factors are necessary. For example, decision trees are used in diagnosing diseases, credit scoring, and quality control in industrial processes.

In the context of agriculture, decision tree algorithms have proven particularly effective in addressing the challenges of crop management, including yield prediction, disease detection, and resource optimization. For example, decision tree models have been employed to predict paddy rice yield by analyzing multiple variables, such as soil quality, weather patterns, and nutrient content. Decision trees could outperform traditional statistical methods in predicting rice yield, providing farmers with actionable insights to optimize crop production (Saruta et al., 2013). Similarly, decision tree algorithms have been used to classify soil fertility, guiding farmers in optimal fertilizer application based on soil properties. This approach ensures the precise use of resources, resulting in increased productivity while reducing waste (Sujatha and Jaidhar, 2024). Decision trees have also been applied to detect diseases in crops, such as rice blast, enabling early interventions to minimize crop losses (Varsha et al., 2024).

Data mining can play a crucial role in documenting, analyzing, and preserving this traditional knowledge. By systematically collecting and analyzing data on traditional practices, including planting times, water management techniques, crop rotation strategies, and pest control measures, valuable insights can be extracted and shared with younger generations (AL-Falahi et al., 2023; Denashurya et al., 2023; Santiago and Buot, 2018). Data mining can also be used to validate and refine traditional methods, identifying areas for improvement or adaptation in light of evolving challenges. Furthermore, data visualization techniques and user-friendly decision support systems are pivotal in making data-driven insights accessible and actionable for farmers, effectively bridging the gap between traditional agricultural practices and modern technology. The integration of advanced technologies such as the Internet of Things (IoT), big data analytics, and artificial intelligence (AI) has transformed conventional farming into a more precise and data-driven approach. The adoption of IoT devices and sensors enables farmers to monitor various environmental variables in real time, thus optimizing resource utilization and enhancing productivity (Christanto et al., 2023; Singh et al., 2023).

The use of decision tree algorithms in agriculture continues to expand as the industry becomes more data-driven. With the integration of decision trees into precision agriculture, water management systems, and disease detection tools, farmers can make more informed decisions tailored to specific field conditions. For example, decision tree-based models have optimized irrigation schedules for water-intensive crops such as paddy rice, ensuring efficient water usage while maintaining high yields (Rahi et al., 2024). However, despite their advantages, decision trees also face challenges such as overfitting and dependency on high-quality data. As agriculture continues to evolve, hybrid models that combine decision trees with other machine learning techniques, such as random forests and neural networks, hold the potential to address these limitations, making decision trees an increasingly vital tool in sustainable agriculture.

### 2.3. CRISP-DM methods

The CRISP-DM (cross-industry standard process for data mining) methodology is one of the most widely adopted frameworks for structuring data mining projects (Schröder et al., 2021; Sharda et al., 2017). It provides a comprehensive and iterative approach to solving problems by translating business requirements into data-driven solutions. The CRISP-DM process comprises six stages (Nodeh et al., 2020; Schröder et al., 2021): business understanding, data understanding, data preparation, modeling, evaluation, and deployment. This methodology ensures that data mining tasks are well aligned with business goals and provides flexibility to revisit earlier stages based on insights gained during the project. In this research, the CRISP-DM approach was adopted to apply the decision tree algorithm for predicting agricultural outcomes, particularly in paddy rice production. The steps of CRISP-DM are explained below, as represented in Figure 2.

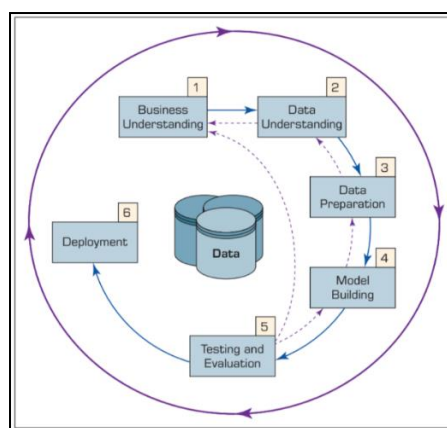


Figure 2 The CRISP-DM Procedure.

### 2.3.1. Step 1: Business Understanding

The first step in the CRISP-DM methodology involves understanding the business problem and defining the objectives that the data mining project aims to achieve. The primary goal of this research was to leverage decision tree algorithms to improve paddy rice production in West Java, Indonesia, by identifying key influencing factors, including demographic, environmental, and resource management variables. This step focused on clearly defining the problem of enhancing agricultural productivity and resilience amidst challenges such as climate variability and socioeconomic constraints. Aligning data mining objectives with practical agricultural needs such as yield prediction, disease detection, and optimal resource allocation ensures actionable insights that can support farmers, researchers, and policymakers in promoting sustainable and resilient agricultural practices.

### 2.3.2. Step 2: Data Understanding

This step involved gathering and becoming familiar with the data required for analysis. For this project, agricultural datasets were sourced from government agricultural databases and sensor data from paddy fields, encompassing historical data on paddy rice production, soil characteristics, weather patterns, and water usage. The data were assessed for completeness, identifying missing values, outliers, and inconsistencies that could affect the model's performance. This phase ensured that the relevant data were appropriately defined and contextualized to support the modeling objectives.

### 2.3.3. Step 3: Data Preparation

The data preparation phase involved cleaning, integrating, and transforming the collected datasets into a suitable format for modeling. Data cleaning addresses missing values and outliers, whereas data transformation ensures compatibility for model training. Features such as soil pH, temperature, rainfall, and pest presence were selected for their potential relevance to predicting key target variables, including rice yield and plant health. Ensuring data quality and structure at this step is critical for accurate model development.

### 2.3.4. Step 4: Modeling via the decision tree algorithm

Once the data are prepared, the modeling phase commences by employing a decision tree algorithm, which is chosen for its interpretability and suitability for complex, nonlinear relationships present in agricultural data. The model was trained to identify and split data into subsets on the basis of features such as age, years of work, and soil characteristics, creating a hierarchy of decision nodes that predict paddy harvest. Hyperparameter tuning was performed to optimize the tree depth, minimum number of samples per node, and other parameters, reducing overfitting and ensuring robust model performance. Cross-validation methods were applied to validate the model's ability to generalize effectively across different data subsets.

### 2.3.5. Step 5: Evaluation

The evaluation phase of the CRISP-DM process focuses on assessing the model's performance to ensure that it meets the defined business objectives. In this research, the decision tree model was evaluated via metrics such as accuracy, precision, recall, and the F1 score for classification tasks (such as disease detection) and the mean absolute error (MAE) and root mean square error (RMSE) for regression tasks (such as yield prediction). The evaluation demonstrated that the decision tree performed well in predicting rice yield based on soil, weather, and management variables and showed strong performance in identifying early signs of disease in rice plants. The results indicated that the decision tree model could provide reliable, actionable insights to help farmers make better decisions about irrigation, fertilization, and pest control.

### 2.3.6. Step 6: Deployment

The deployment phase focused on integrating the decision tree model into a decision support system designed for paddy farmers. This user-friendly platform allows farmers to input field-specific data and receive tailored recommendations for irrigation, fertilization, and disease prevention measures. Graphical visualizations of the decision-making process enhance comprehension by illustrating how factors such as soil moisture and weather patterns influence outcomes. The deployment aims to empower farmers with practical tools to make informed, data-driven decisions that boost productivity and resource efficiency.

### 2.3.7. Step 7: Iteration and Continuous Improvement

One of the key strengths of the CRISP-DM framework is its iterative nature. Following model deployment, real-world performance and user feedback are continuously analyzed, with opportunities to revisit earlier phases as needed. For example, additional features such as real-time weather data or pest outbreak alerts may be incorporated on the basis of farmer input,

and model parameters can be further tuned to increase prediction accuracy. Continuous monitoring and refinement ensure that the model remains adaptive, relevant, and effective in addressing evolving challenges in paddy rice production.

### 3. Results and Discussion

#### 3.1. Model Performance

The decision tree model, trained on a dataset encompassing demographic, environmental, and resource management variables related to paddy rice production in West Java, Indonesia, demonstrated reasonable predictive ability. The cross-validation scores ranged from 0.694 to 0.729, with an average accuracy of approximately 71.43%, indicating the robust performance of the predictive model being evaluated (Habibi et al., 2024; Phinzi et al., 2021). This performance, while promising, is subject to the inherent limitations of the dataset, particularly its restricted size and the presence of class imbalances. These limitations underscore the importance of future research utilizing more extensive and balanced datasets to enhance the model's generalizability and predictive accuracy.

#### 3.2. Feature Importance Analysis

The decision tree analysis, as depicted in Figure 3, revealed the relative importance of different features in predicting "Kelas Panen" (harvest class). Notably, age (Usia) emerged as the most influential factor, with an importance score of 0.372. Years of work followed closely (Lama Bekerja), with an importance score of 0.316. This prominence of age and experience aligns with previous research highlighting the pivotal role of accumulated knowledge and traditional practices in achieving successful agricultural outcomes (Srisopaporn et al., 2015; Tankou et al., 2017). Gender, with an importance score of 0.181, had a moderate impact on harvest class prediction. This finding suggests potential gender-based differences in labor division, access to resources, and decision-making authority within farming households, echoing observations made in studies exploring gender dynamics in agricultural systems (Kinkinginhoun Medagbe et al., 2020). The number of household members (Anggota Kel), with an important score of 0.132, plays a significant role in influencing productivity, primarily because of its association with labor availability during critical farming stages. This relationship is supported by various studies that highlight the impact of household composition on agricultural productivity. For example, research shows that gender norms and the allocation of labor within households can significantly affect agricultural outcomes, emphasizing the importance of understanding labor distribution among household members (Arora and Rada, 2020; Haile, 2016). Interestingly, features such as ownership status and type of paddy field exhibited zero importance in the model. This suggests that these factors, within the specific dataset analyzed, might have limited variability or a weaker relationship with yield outcomes than the more dominant influences of age, experience, and household dynamics. This finding is consistent with the previous research which emphasized that household characteristics, including the number of economically active members and farming experience, are significant determinants of agricultural productivity (Gebrehiwot, 2015; Nuno and Baker, 2021). Their study highlights how these factors can overshadow other variables, such as land ownership, which may not significantly affect productivity outcomes.

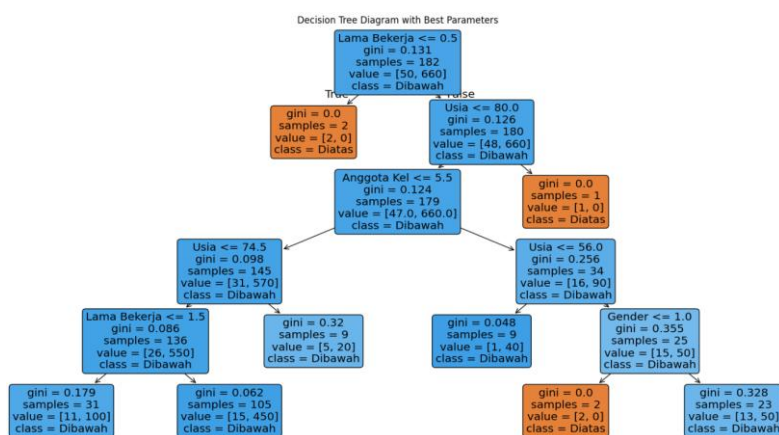


Figure 3 Decision Tree Analysis Results.

#### 3.3. Interpretation and Discussion

The predominance of age and years of work as the main predictors emphasizes the importance of accumulated knowledge and experience in traditional farming, as reported in a previous research (Srisopaporn et al., 2015). Although agricultural methods may offer technological solutions, our findings highlight that experience-driven approaches remain essential in many agricultural contexts (Tankou et al., 2017). Another aspect that has a high impact on harvest is gender and household dynamics, which reflect the economic and social factors that can also affect the productivity of paddy fields (Kinkinginhoun Medagbe et al., 2020).



Agricultural work, especially rice farming, often requires good judgment on the basis of years of practical experience. In general, older farmers accumulate a wealth of hands-on knowledge that spans multiple seasons and a variety of climatic conditions. This allows them to make better choices about agricultural practices (Dissanayake et al., 2021; Li et al., 2022). These choices may include the most suitable planting time. Proper water management techniques include crop rotation strategies and pest control measures according to the specific needs of the region (AL-Falahi et al., 2023; Chongtham et al., 2017). The relationship between traditional practices and community wisdom increases the resilience and ability to adapt to farming methods, but it also increases the need for a structured approach to transferring knowledge from senior to junior individuals. To ensure that valuable expertise is not lost over time (Denashurya et al., 2023; Santiago and Buot, 2018).

Moreover, age can reflect a deeper connection to ancestral practices and community wisdom that are handed down through generations. In regions where traditional and subsistence farming are prevalent, such as West Java and Indonesia, older farmers often act as custodians of this knowledge, providing mentorship to younger generations and preserving essential farming techniques (Carrasco-Torrentegui et al., 2021; Mukandahiro et al., 2024). This accumulated expertise is invaluable, especially in scenarios where modern technological solutions may be inaccessible or underutilized.

Another finding regarding paddy harvest also strengthens the findings of previous research studies, which show that older farmers often exhibit greater resilience and adaptability to challenges such as shifting weather patterns, soil degradation, and pest infestations. Older farmers legitimate their farming knowledge and skills to the next generation, which is closely related to their children. Their historical familiarity with both land and environmental variations gives them an edge in their ability to respond swiftly and effectively to these issues (Abdullah et al., 2018; Asli and Rahman, 2023). This resilience is likely to contribute to their ability to achieve higher yields, as reflected in the "Kelas Panen" classification. Age is also closely tied to social and economic roles within farming communities. Senior farmers often hold leadership positions within agricultural cooperatives or informal networks, giving them access to shared resources, information, and decision-making power that can influence crop productivity. Their prominent role in community-based farm initiatives often means that they can mobilize resources more effectively and rapidly adopt best practices.

Given the strong influence of age on productivity outcomes, policy interventions should prioritize knowledge transfer programs that capture and disseminate the expertise of older farmers to younger generations. This can be achieved through mentorship schemes, community-based training sessions, or digital platforms designed to document and share traditional practices. Such initiatives could ensure that the collective experience of older farmers leads to improved overall productivity and resilience across generations.

Moreover, integrating modern technological solutions with traditional practices may yield synergistic effects. For example, combining digital tools for data collection with age-old soil management techniques can help optimize resource use and improve crop yields to gain resilience for food security in Indonesia, as previous studies indicate the importance of adaptability, capabilities, and innovation for a more resilient and sustainable future, especially in light of disruptions such as the COVID-19 pandemic (Astuty et al., 2024; Harsanto et al., 2024; Santiago and Buot, 2018).

#### 4. Conclusions

This research demonstrated that age, years of work, gender, and household dynamics significantly influence "Kelas Panen" (harvest class) in paddy rice production. By leveraging data mining techniques, particularly decision trees, we identified key factors that impact productivity and yield outcomes. Age emerged as a critical predictor because of its strong correlation with accumulated experience, knowledge, and adaptive capacity in farming practices. The findings underscore the need for a balanced integration of traditional farming expertise and modern technological solutions to optimize rice production and improve food security in Indonesia.

The research also provides valuable insights for farmers, researchers, experts, and decision-makers, promoting collaboration to improve resilience, productivity, and sustainability within the agricultural sector. Nonetheless, the study faced constraints due to a limited sample size and class imbalance, which could influence the applicability of the findings. Future studies ought to utilize larger and more diverse datasets, investigate additional variables, and implement more sophisticated models to enhance comprehension and facilitate decision-making in rice production systems.

#### Acknowledgment

All the authors would like to express their gratitude to God for the completion of this article. We thank all the academic and staff members from Padjadjaran University, Bina Nusantara University, and Pembangunan Jaya University for providing invaluable facilities, educational guidance, and unwavering support.

#### Ethical considerations

This research was conducted in strict adherence to ethical principles, ensuring the protection of participants' rights and welfare. Before their participation, all participants provided informed consent. Data collection and analysis were performed with a commitment to maintaining confidentiality and anonymity.

## Conflict of interest

All the authors declare that they have no conflicts of interest.

## Funding

The research was funded by the Ministry of Education, Culture, Research, and Technology of Indonesia under the Strategic Research Collaboration (KATALIS) grand 2024 with the grand contract number: 058/E5/PG.02.00/PL.BATCH.2/2024; 1082/LL3/AL.04/2024; 186/VRRRTT/X/2024.

## References

- Abdullah, N. A. N., Karim, Z. Z. A., Hussin, S. A. S., and Khairi, S. S. M. (2018). Determination of Perception and Adaptive Capacity towards Climate Change among Paddy Farmers in State of Kelantan, Malaysia. *International Journal of Engineering & Technology*, 7(4.1), Article 4.1. <https://doi.org/10.14419/ijet.v7i4.1.28238>
- AL-Falahi, A. H., Barry, S., Gebrechorkos, S. H., Spank, U., and Bernhofer, C. (2023). Potential of Traditional Adaptation Measures in Mitigating the Impact of Climate Change. *Sustainability*, 15(21), Article 21. <https://doi.org/10.3390/su152115442>
- Alfred, R., Obit, J. H., Chin, C. P.-Y., Haviluddin, H., and Lim, Y. (2021). Towards Paddy Rice Smart Farming: A Review on Big Data, Machine Learning, and Rice Production Tasks. *IEEE Access*, 9, 50358–50380. <https://doi.org/10.1109/ACCESS.2021.3069449>
- Arora, D., and Rada, C. (2020). Gender norms and intrahousehold allocation of labor in Mozambique: A CGE application to household and agricultural economics. *Agricultural Economics*, 51(2), 259–272. <https://doi.org/10.1111/agec.12553>
- Asli, W. M. F., and Rahman, S. (2023). CONCEPTUALISING RESILIENCE LIVELIHOOD STRATEGIES OF PADDY FARMERS TOWARDS MALAYSIA NATIONAL FOOD SECURITY AGENDA. *PLANNING MALAYSIA*, 21. <https://doi.org/10.21837/pm.v21i28.1346>
- Astuty, E., Sudirman, I. D., and Aryanto, R. (2024). Sustainable resilience strategy: Unleash the micro-businesses's potential in the digitalization and sustainability era. *Cogent Business & Management*, 11(1), 2313672. <https://doi.org/10.1080/23311975.2024.2313672>
- Barbosa Júnior, M. R., Moreira, B. R. de A., Carreira, V. dos S., Brito Filho, A. L. de, Trentin, C., Souza, F. L. P. de, Tedesco, D., Setiyono, T., Flores, J. P., Ampatzidis, Y., Silva, R. P. da, and Shiratsuchi, L. S. (2024). Precision agriculture in the United States: A comprehensive meta-review inspiring further research, innovation, and adoption. *Computers and Electronics in Agriculture*, 221, 108993. <https://doi.org/10.1016/j.compag.2024.108993>
- Boughton, D., Goeb, J., Lambrecht, I., Headey, D., Takeshima, H., Mahrt, K., Masias, I., Goudet, S., Ragasa, C., Maredia, M. K., Minten, B., and Diao, X. (2021). Impacts of COVID-19 on agricultural production and food systems in late transforming Southeast Asia: The case of Myanmar. *Agricultural Systems*, 188, 103026. <https://doi.org/10.1016/j.agsy.2020.103026>
- BPS. (2023). *Ringkasan Eksekutif Luas Panen dan Produksi Padi di Indonesia 2023* (No. 05100.2317). Badan Pusat Statistik.
- Carrasco-Torrontegui, A., Gallegos-Riofrío, C. A., Delgado-Espinoza, F., and Swanson, M. (2021). Climate Change, Food Sovereignty, and Ancestral Farming Technologies in the Andes. *Current Developments in Nutrition*, 5, 54–60. <https://doi.org/10.1093/cdn/nzaa073>
- Chongtham, I. R., Bergkvist, G., Watson, C. A., Sandström, E., Bengtsson, J., and Öborn, I. (2017). Factors influencing crop rotation strategies on organic farms with different time periods since conversion to organic production. *Biological Agriculture & Horticulture*, 33(1), 14–27. <https://doi.org/10.1080/01448765.2016.1174884>
- Christanto, H. J., Sutresno, S. A., Lim, R. J., Angkur, L. V. G., Charmelita, P., Valentina, V., and Dewi, C. (2023). Design and Development of the Mobile-Based Hydroponic Planting Machine Application MyHydro. *Journal of Information Systems and Informatics*, 5(4), Article 4. <https://doi.org/10.51519/journalisi.v5i4.581>
- Denashurya, N. I., Nurliza, Dolorosa, E., Kurniati, D., and Suswati, D. (2023). Overcoming Barriers to ISPO Certification: Analyzing the Drivers of Sustainable Agricultural Adoption among Farmers. *Sustainability*, 15(23), Article 23. <https://doi.org/10.3390/su152316507>
- Dissanayake, S. P., Gunaratne, L. H. P., Sivanathewer, T., and Ginigaddara, G. a. S. (2021). Impact of Adoption of Sustainable Agricultural Practices on Household Food Security in Small-scale Paddy-Cattle Farming Systems in Anuradhapura District, Sri Lanka. *Sri Lankan Journal of Agriculture and Ecosystems*, 3(2). <https://doi.org/10.4038/sljae.v3i2.80>
- El Bilali, H., Henri Nestor Bassole, I., Dambo, L., and Berjan, S. (2020). Climate Change and Food Security. *The Journal "Agriculture and Forestry,"* 66(3). <https://doi.org/10.17707/AgricultForest.66.3.16>
- Fahad, S., Adnan, M., Noor, M., Arif, M., Alam, M., Khan, I. A., Ullah, H., Wahid, F., Mian, I. A., Jamal, Y., Basir, A., Hassan, S., Saud, S., Amanullah, Riaz, M., Wu, C., Khan, M. A., and Wang, D. (2019). Major Constraints for Global Rice Production. Woodhead Publishing. <https://doi.org/10.1016/B978-0-12-814332-2.00001-0>
- FAO. (2017). *The future of food and agriculture: Trends and challenges*. Food and Agriculture Organization of the United Nations.
- Firdaus, R. B. R., Leong Tan, M., Rahmat, S. R., and Senevi Gunaratne, M. (2020). Paddy, rice and food security in Malaysia: A review of climate change impacts. *Cogent Social Sciences*, 6(1), 1818373. <https://doi.org/10.1080/23311886.2020.1818373>
- Gebrehiwot, K. G. (2015). The impact of agricultural extension on households' welfare in Ethiopia. *International Journal of Social Economics*, 42(8), 733–748. <https://doi.org/10.1108/IJSE-05-2014-0088>
- Habibi, L. N., Matsui, T., and Tanaka, T. S. T. (2024). Critical evaluation of the effects of a cross-validation strategy and machine learning optimization on the prediction accuracy and transferability of a soybean yield prediction model using UAV-based remote sensing. *Journal of Agriculture and Food Research*, 16, 101096. <https://doi.org/10.1016/j.jafr.2024.101096>
- Haile, D. F. (2016). Factors affecting women farmers' participation in agricultural extension services for improving the production in rural district of Dendi West Shoa Zone, Ethiopia. *INTERNATIONAL JOURNAL OF HOME SCIENCE EXTENSION & COMMUNICATION MANAGEMENT*, 3(2), 59–71. <https://doi.org/10.15740/HAS/IJHSECM/3.2/59-71>
- Harsanto, B. (2021). Sustainability innovation in the agriculture sector in Indonesia: A review. *E3S Web of Conferences*, 306, 02022. <https://doi.org/10.1051/e3sconf/202130602022>
- Harsanto, B., Farras, J. I., and Firmansyah, E. A. (2024). Sustainability Innovation, Firm Competitiveness, Export Performance, and Sales Growth: A System Dynamics Approach. *ASEAN Journal on Science and Technology for Development*, 41(1). <https://doi.org/10.61931/2224-9028.1562>



- Kinkinginhoun Medagbe, F. M., Komatsu, S., Mujawamariya, G., and Saito, K. (2020). Men and Women in Rice Farming in Africa: A Cross-Country Investigation of Labor and Its Determinants. *Frontiers in Sustainable Food Systems*, 4. <https://doi.org/10.3389/fsufs.2020.00117>
- Kong, J., Yang, C., Wang, J., Wang, X., Zuo, M., Jin, X., and Lin, S. (2021). Deep-Stacking Network Approach by Multisource Data Mining for Hazardous Risk Identification in IoT-Based Intelligent Food Management Systems. *Computational Intelligence and Neuroscience*, 2021(1), 1194565. <https://doi.org/10.1155/2021/1194565>
- Lausch, A., Schmidt, A., and Tischendorf, L. (2015). Data mining and linked open data – New perspectives for data analysis in environmental research. *Ecological Modelling*, 295, 5–17. <https://doi.org/10.1016/j.ecolmodel.2014.09.018>
- Li, H., Zhou, X., Tang, M., and Guo, L. (2022). Impact of Population Aging and Renewable Energy Consumption on Agricultural Green Total Factor Productivity in Rural China: Evidence from Panel VAR Approach. *Agriculture*, 12(5), Article 5. <https://doi.org/10.3390/agriculture12050715>
- M., Sujatha., and C.d., Jaidhar. (2024). Machine learning-based approaches to enhance the soil fertility—A review. *Expert Systems with Applications*, 240, 122557. <https://doi.org/10.1016/j.eswa.2023.122557>
- Majumdar, J., Naraseyappa, S., and Ankalaki, S. (2017). Analysis of agriculture data using data mining techniques: Application of big data. *Journal of Big Data*, 4(1), 20. <https://doi.org/10.1186/s40537-017-0077-4>
- Maraseni, T. N., Deo, R. C., Qu, J., Gentle, P., and Neupane, P. R. (2018). An International Comparison of Rice Consumption Behaviours and Greenhouse Gas Emissions From Rice Production. *Journal of Cleaner Production*, 172, 2288–2300. <https://doi.org/10.1016/j.jclepro.2017.11.182>
- Mohamed, Z., Terano, R., Shamsudin, M. N., and Abd Latif, I. (2016). Paddy Farmers' Sustainability Practices in Granary Areas in Malaysia. *Resources*, 5(2), Article 2. <https://doi.org/10.3390/resources5020017>
- Mukandahiro, F., Kranjac-Berisavljevic, G., and Osei, E. A. (2024). Effect of Gender and Generation on the Adoption of Irrigation and Good Farming Practices in Bontanga and Golinga Irrigation Schemes in Northern Ghana. *International Journal of Irrigation and Agricultural Development (IJIRAD)*, 8(1), Article 1. <https://doi.org/10.47762/2024.964x.135>
- Muthayya, S., Sugimoto, J. D., Montgomery, S., and Maberly, G. F. (2014). An Overview of Global Rice Production, Supply, Trade, and Consumption. *Annals of the New York Academy of Sciences*, 1324(1), 7–14. <https://doi.org/10.1111/nyas.12540>
- Nawaz, A., Rehman, A. U., Rehman, A., Ahmad, S., Siddique, K. H. M., and Farooq, M. (2022). Increasing Sustainability for Rice Production Systems. *Journal of Cereal Science*, 103, 103400. <https://doi.org/10.1016/j.jcs.2021.103400>
- Ngandee, S., Taparugssanagorn, A., Anutariya, C., and Kuwornu, J. K. M. (2021). Assessment of rice yield prediction models based on big data analytics for better supply chain decision-making in Thailand. *International Journal of Value Chain Management*, 12(3), 221–240. <https://doi.org/10.1504/IJVC.2021.118289>
- Nodeh, M. J., Calp, M. H., and Şahin, İ. (2020). Analyzing and Processing of Supplier Database Based on the Cross-Industry Standard Process for Data Mining (CRISP-DM) Algorithm. In D. J. Hemanth. Springer International Publishing. [https://doi.org/10.1007/978-3-030-36178-5\\_44](https://doi.org/10.1007/978-3-030-36178-5_44)
- Nuno, D. B., and Baker, M. M. (2021). The Determinants of Agricultural Crop Productivity among Smallholder Households in Haramaya District, Eastern Ethiopia. *Grassroots Journal of Natural Resources*, 04(04), 146–153. <https://doi.org/10.33002/nr2581.6853.040410>
- Osei-Bryson, K.-M. (2004). Evaluation of decision trees: A multi-criteria approach. *Computers & Operations Research*, 31(11), 1933–1945. [https://doi.org/10.1016/S0305-0548\(03\)00156-4](https://doi.org/10.1016/S0305-0548(03)00156-4)
- Phinzi, K., Abriha, D., and Szabó, S. (2021). Classification Efficacy Using K-Fold Cross-Validation and Bootstrapping Resampling Techniques on the Example of Mapping Complex Gully Systems. *Remote Sensing*, 13(15), Article 15. <https://doi.org/10.3390/rs13152980>
- Prabhakaran, A., Chithra Lekshmi, K. S., and Janarthanan, G. (2021). Advancement of data mining methods for improvement of agricultural methods and productivity. Academic Press. <https://doi.org/10.1016/B978-0-12-820601-0.00010-0>
- Rahi, M., Ebrahimnejad, A., and Motameni, H. (2024). Evaluation of predicted fault tolerance based on C5.0 decision tree algorithm in irrigation system of paddy fields. *International Journal of Intelligent Computing and Cybernetics*, 17(2), 253–305. <https://doi.org/10.1108/IJIC-07-2023-0174>
- Santiago, J., and Buot, I. (2018). Conceptualizing the Socio-Ecological Resilience of the Chaya Rice Terraces, a Socio-Ecological Production Landscape in Mayoyao, Ifugao, Luzon Island, Philippines. *Journal of Marine and Island Cultures*, 7(1). <https://doi.org/10.21463/jmic.2018.07.1.07>
- Saruta, K., Hirai, Y., Tanaka, K., Inoue, E., Okayasu, T., and Mitsuoka, M. (2013). Predictive models for yield and protein content of brown rice using support vector machine. *Computers and Electronics in Agriculture*, 99, 93–100. <https://doi.org/10.1016/j.compag.2013.09.003>
- Schröder, C., Kruse, F., and Gómez, J. M. (2021). A Systematic Literature Review on Applying CRISP-DM Process Model. *Procedia Computer Science*, 181, 526–534. <https://doi.org/10.1016/j.procs.2021.01.199>
- Sharda, R., Delen, D., and Turban, E. (2017). *Business Intelligence, Analytics, and Data Science: A Managerial Perspective* (4 edition). Pearson.
- Singh, N. K., Sunitha, N. H., Tripathi, G., Saikanth, D. R. K., Sharma, A., Jose, A. E., and Mary, M. V. K. J. (2023). Impact of Digital Technologies in Agricultural Extension. *Asian Journal of Agricultural Extension, Economics & Sociology*, 41(9), 963–970. <https://doi.org/10.9734/ajaees/2023/v41i92127>
- Son, N.-T., Chen, C.-F., Chen, C.-R., Guo, H.-Y., Cheng, Y.-S., Chen, S.-L., Lin, H.-S., and Chen, S.-H. (2020). Machine learning approaches for rice crop yield predictions using time-series satellite data in Taiwan. *International Journal of Remote Sensing*, 41(20), 7868–7888. <https://doi.org/10.1080/01431161.2020.1766148>
- Srisopaporn, S., Jourdain, D., Perret, S. R., and Shivakoti, G. (2015). Adoption and continued participation in a public Good Agricultural Practices program: The case of rice farmers in the Central Plains of Thailand. *Technological Forecasting and Social Change*, 96, 242–253. <https://doi.org/10.1016/j.techfore.2015.03.016>
- Statista. (2024). *Total Global Rice Consumption*. Statista. <https://www.statista.com/statistics/255977/total-global-rice-consumption/>
- Sudirman, I. D. (2023). *Data-Driven Entrepreneur: Bisnis Berdaya Saing dengan Data Science dan RapidMiner*. Salemba Empat.
- Tankou, C. M., Snoo, G. R. D., Persoon, G., and Iongh, H. H. D. (2017). Evaluation of smallholder farming systems in the Western Highlands of Cameroon. *IOSR Journal of Engineering*, 07(01), 01–11. <https://doi.org/10.9790/30210701010111>
- Varsha, M., Basavaraju, P., and P, P. K. M. (2024). Advanced Machine Learning Techniques for Disease Management in Paddy Crop for Sustainable Rice Production: Prediction of Rice Blast Disease Severity. In *Natural Resource Management Issues in Human-Influenced Landscapes* (pp. 64–87). IGI Global. <https://doi.org/10.4018/978-1-6684-7051-0.ch004>



Witten, I. H., and Frank, E. (2002). Data mining: Practical machine learning tools and techniques with Java implementations. *ACM SIGMOD Record*, 31(1), 76–77. <https://doi.org/10.1145/507338.507355>

