Rice Farmer Registry Application and Management System Using Classification and Regression Trees Algorithm

Belen M. Tapado

College of Information and Communications Technology, Catanduanes State University, Catanduanes, Philippines
Email: bmtapado@catsu.edu.ph

***Abstract*—Global rice production surpasses half a billion tons, with Asia at the forefront. However, the Philippines continues to be a significant importer, representing six percent of global imports. The country faces considerable obstacles in domestic rice production, such as the necessity for cost-effective labor, consistent rainfall, efficient land management, irrigation infrastructure, and appropriate soil conditions. Insufficient government support intensifies these issues. To improve crop production and reduce dependence on imports, modern agricultural methods, such as smart farming technologies, are crucial. In light of these challenges, this research emphasized the vibrant role of rice farmers in the Philippines' economy. The study employed data from rice farmers registered in the "Ani at Kita: Registry System for Basic Sectors in Agriculture (RSBSA)." The Classification and Regression Trees (CART) algorithm was utilized to autonomously select qualifying farmers for incentives in each barangay. Every farmer was regarded as a distinct data point, characterized by attributes including registration status, revenue source, farming activities, household size, and involvement in beneficiary programs. The algorithm employed a hierarchical framework to classify farmers into smaller groups according to dynamic variables such as Gini impurity or entropy, finally ranking them in leaf nodes. Web-based technology was likewise employed in to the application solution to facilitate data interchange between barangays and local government bodies. The system garnered favorable feedback, achieving an average rating of 4.59 out of 5, signifying strong endorsement from respondents.**

***Keywords*—Classification and Regression Trees (CART) algorithm, rice, rice farming, rice production, local government assistance, Ani at Kita: Registry System for Basic Sectors in Agriculture (RSBSA)**

I. Introduction

In the 2019–2020 marketing year, the worldwide production of milled rice amounted to almost 500 million tons. Asia, specifically China and Southeast Asia, played a substantial role in both the production and consumption of rice. China was the leading producer, accounting for 30% of the total, with India, Bangladesh, Indonesia, Vietnam, and Thailand following suit [1]. The Philippines was noted as a significant importer, responsible for 6% of global imports.

The insufficient domestic rice production in the Philippines, as indicated by its significant imports, emphasized the necessity to improve local agriculture to satisfy the increasing demands of the growing population. Rice cultivation, ideally suited for places with cheap labor costs and abundant rainfall, necessitates meticulous field management, efficient irrigation, and favorable soil conditions to achieve maximum yields [2]. Although rice farmers play a vital part in the country's economy, they have difficulties because of insufficient government assistance [3].

To tackle these issues, the study underscored the significance of embracing contemporary agricultural methods, such as the utilization of intelligent farming technologies. The utilization of data mining principles from the datasets obtained from the accomplished RSBSA facilitated well-informed decision-making and policy formation in this context [4]. The study conducted a pilot test in the Local Government Unit (LGU) of San Andres, Catanduanes, utilizing the Decision Tree Classifier technique to determine eligible beneficiaries for government assistance in the field of rice growing.

The research presents a new use of the Classification and Regression Trees (CART) algorithm in the agricultural domain through the promotion of sustainable and data-driven methods for cultivating rice. This manner allowed for well-informed decision-making, which in turn improved government support that boosted rice production in the area, and guaranteed food security. Furthermore, its technical contributions were as follows: (a) a thorough process of compiling, evaluating, and interpreting the results of the CART algorithm to choose the worthiest rice farmer to receive government funding; and (b) the creation of an easy-to-use user interface that can be used by staff members of local government units with varying degrees of technical proficiency. Together, these contributions enhanced the field of agricultural informatics and demonstrated how machine learning algorithms, such as CART, may transform the data management of rice farmers and greatly increased agricultural output in the locale. Similarly, the implementation of the automated system by the local

Manuscript received January 11, 2024; revised May 13, 2024; accepted October 7, 2024; published March 6, 2025.

doi: 10.12720/jait.16.3.312-317

government was a vital measure in acknowledging the significance of rice farmers in driving the island's economic development. This attempt significantly enhanced the welfare of farmers and effectively diminished the country's dependence on rice imports.

II. LITERATURE REVIEW

Several articles had contributed to the inception, design and development of this study. They are as follows:

Ojo and Lloyd's research [5] on smallholder rice farmers in South-West Nigeria found that their climate change adaptation strategies are influenced by socio-economic factors, institutional factors, and locational variables. Implementing these methods increased net farm income, especially for farmers affected by droughts and floods. They recommend financial innovations in climate change education and capacity-building.

In a similar vein, a study on sustainable rice cultivation in six irrigated regions of Vietnam, Thailand, Indonesia, Myanmar, Sri Lanka, and China [6] revealed significant differences in crop productivity and financial rewards between successful and average-performing smallholder systems. The research suggested reducing greenhouse gas emissions without sacrificing crop output can be achieved through improved fertilizer and pesticide management and irrigation system optimization.

Vietnamese rice farmers are utilizing Smart Farming Technique. They have adopted the Alternate Wetting and Drying (AWD) method, to reduce water usage and greenhouse gas emissions while maintaining rice yields, and using sensors for irrigation [7].

The paper by Obit, Chin Haviluddin, and Lim [8] emphasizes the importance of integrating technologies like Big Data, Machine Learning, and Internet of Things (IoT) in smart farming, particularly in rice production, to transform traditional methods of rice farming into a new understanding of intelligence in rice precision agriculture.

Likewise, Paul and Santos' study utilized data science to improve agricultural productivity in the Philippines by utilizing a web-based tool that monitored farming activities, expenses, and weather data enabling informed decision-making. The system improved crop productivity, especially in remote regions, and boosted the nation's economy [9].

Moreover, an ensemble-based machine learning model that can combine two or more models with the CART could also be used in order to enhance the quality and novelty approach to rice cultivation [10].

The study reviewed explored sustainability, climate change adaptability, and technical innovations in rice cultivation, considering factors like geography, ecology, institutions, and socioeconomic conditions. They suggested IoT, data mining, machine learning, and AWD approach for improved crop production. Nevertheless, obstacles like infrastructure, budgetary limitations, education, training, and legislative support hinder extensive adoption.

CART algorithms, although potentially subjective, can generate flexible models for agricultural management and decision-making when integrated with precision farming. This study's application of CART enhanced data analysis and yielded practical recommendations for local government entities, underscoring technology's significance in enhancing rice farming techniques.

III. RESEARCH METHOD

This study utilized system development, survey research, and documentary analysis to examine evidence on rice farmers, farms, and production in the area. The Rapid Application Development (RAD) methodology was used for system development, while the Classification and Regression Trees (CART) Algorithm was used to prioritize farmers receiving government assistance.

A. System Architecture of the Study

Fig. 1 shows the project's System Architecture that includes a computer unit, user interface, Decision Tree Classification framework, and system users. This project gathered data from the Barangay Secretary, Municipal Agricultural Officer, and Provincial Agricultural Officer using web-based techniques. The system has three (3) user tiers: Barangay Secretary manages farmer registration data, the Municipal Agriculture Officer verifies it, and the Provincial Agriculture Officer supervises user accounts. The Model-View-Controller (MVC) [11, 12] architecture was likewise used, with the model layer managing data logic, the view layer displaying data, and the controller handling user input and interaction [13].

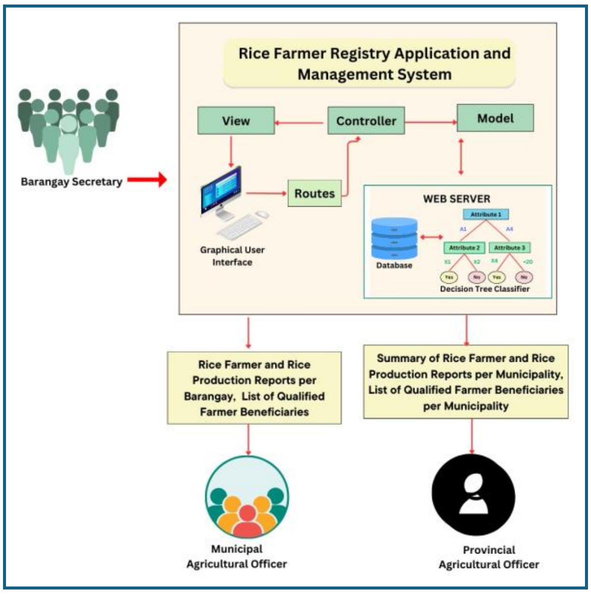


Fig. 1. System architecture of the study.

B. System Development Method

The system development of this study involves seven stages (Fig. 2) which was adopted from Rapid Application Development (RAD) methodology [14] which is recommended for projects with limited timeframes and immediate user feedback. These stages were: gathering and updating data on rice farmers, developing and improving the system, integrating the CART algorithm, implementing the system, providing user training, and evaluating the program.

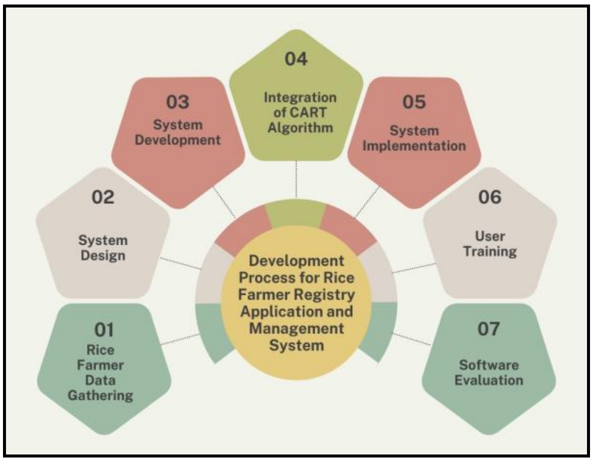


Fig. 2. System development method of the study.

C. *CART Model Construction and Evaluation*

The CART technique (Fig. 3) was used to create a decision-making model for farmers to determine their eligibility for financial aid. This non-parametric, hierarchical algorithm, used in supervised learning for classification and regression tasks, follows a tree-like pattern with root nodes, branches, internal nodes, and leaf nodes [15].

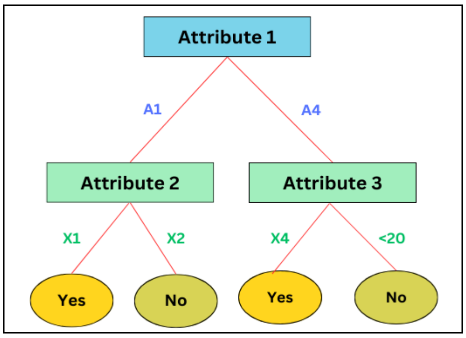


Fig. 3. CART technique.

The algorithm analyzed a five-year dataset from 2019 to 2023 of rice farmers in Barangay Manambrag, the barangay with the highest number of rice farmers in San Andres, Catanduanes, to predict their eligibility based on socio-economic and agricultural factors such as annual income, livelihood source, disability status, government assistance programs, farmer association affiliation, agrarian aid receipt, and household member count.

The dataset had 1019 entries of rice farmers, each encompassing seven variables listed in Table I. Prior to data analysis, data pre-processing measures were implemented to address missing values and homogenize the input variables. A total of 33 entries were eliminated owing to insufficient information, and the remaining data was standardized to improve model performance.

The CART method was chosen for its ability to handle classification and regression problems, making it suitable for forecasting a farmer's eligibility for government support (a classification task), by constructing decision trees by recursively partitioning the dataset based on significant variables.

TABLE I. DATASET FEATURES

Item	CART Variables		
	Variables	Variable Type	Description
1	annual_income	continuous	This refers to the head of the family's annual income
2	main_livelihood	categorical	This indicates the head of the family's primary occupation
3	Is_pwd	continuous	This indicates whether the farmer is a person with a disability (1 if yes, 0 if no)
4	is_four_ps	continuous	This indicates whether the farmer is a member of an existing assistance grant (1 if yes, 0 if no)
5	is_org_member	binary	This indicates whether the farmer is a member of an agricultural association or organization (1 if yes, 0 if no)
6	is_agrarian_beneficiary	continuous	This indicates whether the farmer is a recipient of agrarian aid (1 if yes, 0 if no);
7	is_priority	binary	This indicates whether or not a farmer should receive preference for help and subsidies (1 if yes, 0 if no)

The model uses API data to compute a priority score for assistance and trains a Decision Tree Classifier, a supervised learning algorithm, to classify targets and predict aid priorities. The model then uses these forecasts to prioritize individuals based on their likelihood of needing assistance, as shown in Table II.

TABLE II. PRIORITY SCORE CALCULATION

Item	CART Variables and Priority Calculation	
	Variables	Priority Calculation Method
1	annual_income	The lower the income, the higher the priority score (max +3 for income <5000 pesos).
2	Is_pwd	Adds +2 to the score if the person is disabled
3	is_four_ps	Subtracts 1 if they are part of the government aid program (assumes lower priority due to existing aid).
4	is_org_member	Adds 1 if they are part of an organization.
5	is_agrarian_beneficiary	Subtracts 1 for agrarian beneficiaries (assuming they are already receiving benefits).
6	Threshold	If the final score is 3 or more, the person is given a priority_for_aid of 1 (high priority); otherwise, it's 0 (low priority).

To facilitate effective model construction, the dataset was partitioned into a training set containing 70% of the data and a testing set encompassing 30% of the data. This division resulted in 691 records designated for training the CART model and 295 records assigned for evaluating its performance. The 70–30 split provided a balanced approach, allowing the model to incorporate a substantial volume of data while maintaining enough samples to evaluate its generalizability. In the training and testing division, the following notations were utilized: (a) X denotes the feature matrix (input data) excluding Name and Priority for Aid; (b) Y signifies the target variable (Priority for Aid), which the model aims to predict; (c) train_test_split() facilitates the partitioning of the data into training (70%) and testing (30%) sets. The random_state = 42 guarantees reproducibility.

TABLE III. CART TECHNIQUE NOTATIONS

Item	Metrics used for CART Technique	
	Parameter	**Best Value**
1	Criterion	"gini"
2	Max Depth	5
3	Min Samples Split	2
4	Min Samples Leaf	1
5	Max Features	None
6	Class Weight	"balanced"

The parameters used for the CART technique is shown on Table III. The Gini impurity is used for the "Criterion" parameter due to its efficiency, as it requires fewer processing resources compared to entropy while achieving comparable performance in binary classification tasks. To mitigate overfitting, particularly with a dataset of modest size, a value of 5 is chosen for the "Max Depth" parameter. This choice strikes a balance between training complexity and generalization. The "Min Samples Split" parameter determines that a node will only be divided if there are a minimum of two samples, enabling the model to capture more intricate patterns in the data. The "Min Samples Leaf" parameter is set to have a minimum of one sample per leaf node. This, along with the max depth, guarantees that the tree is not excessively intricate. When the "Max Features" parameter is set to "None", it means that the model will consider all available information at each split. Finally, the "Class Weight" parameter uses the value "balanced" to prioritize the minority class, which is crucial when there is an imbalance, such as having a larger number of "non-priority farmers."

The decision tree method was used to categorize farmers based on their dynamic properties, with each ranked according to its importance. This method, which uses factors like Gini impurity or entropy [16], automatically selects the most deserving farmers for incentives, enhancing efficiency and ensuring transparency and interpretability in the decision-making process in the Local Government Unit (LGU).

Following training, the decision tree model was evaluated on a testing set including 986 records to determine its correctness and efficacy. The model attained an accuracy of 90%, indicating that it accurately predicted the eligibility of farmers in most instances. Furthermore, precision and recall were computed, yielding values of 91 and 91, respectively, signifying a balanced capacity to identify eligible farmers (recall) while reducing the misclassification of ineligible farmers as eligible (precision).

IV. RESULT AND DISCUSSION

A. Development of the CART-Enabled Rice Farmer Registry Application and Management System

The phases of system development in this study included design, development, and employment of the CART Algorithm, implementation, user training and user evaluation of the system. The figures below are the screenshots of the developed system.

The dashboard [17] (Fig. 4) served as a centralized interface that displays visual data on rice farmers, fields, farmer's enrollment in the RSBSA, and beneficiaries in each barangay. It provides a concise overview of information. Fig. 5 depicts the manner of collecting data from farmers, in accordance with the RSBSA. Collection of data was made by the Barangay Secretary.

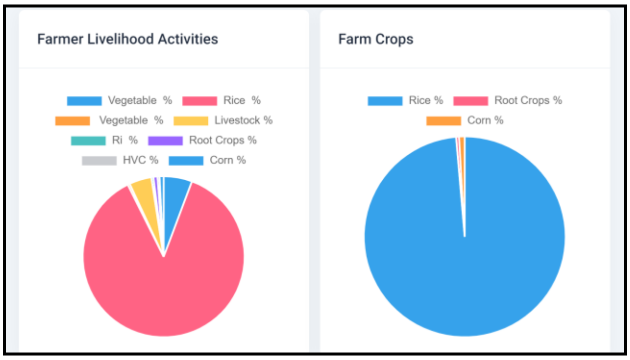


Fig. 4. Rice farmer application and management system dashboard.

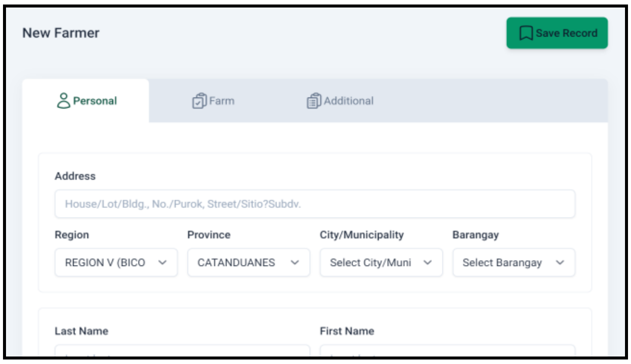


Fig. 5. Adding new farmer.

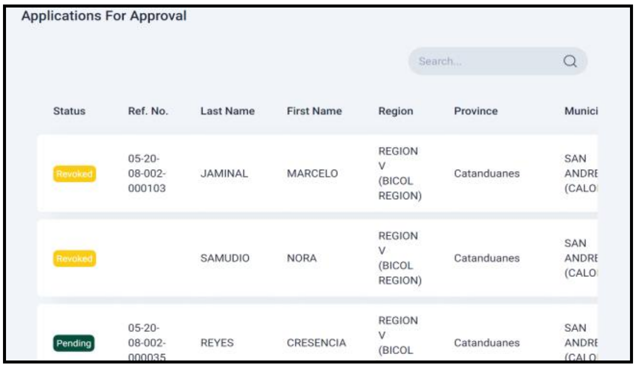
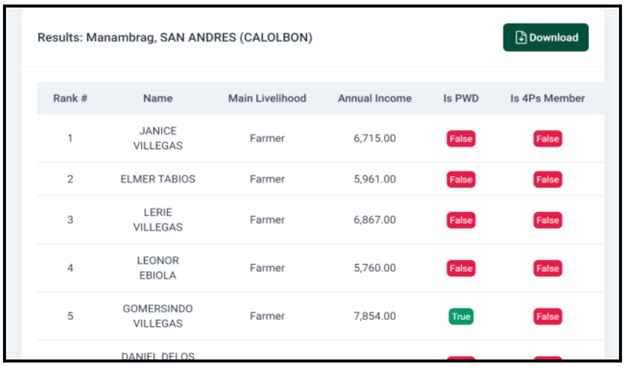


Fig. 6. Notification to the MAO relative to rice farmer registry enrolment.

Once captured by the system, a notification would be given to the MAO for approval of the farmer's application. The status of the application could be pending approval (green), revoked (yellow) and denied (red). This is shown in Fig. 6. The outcome of the process of selection and prioritization of rice farmers to receive grant is shown in Fig. 7. From the figure, parameters were set to determine the most qualified farmer to receive government support per barangay. They were ranked according to the parameter set and the most qualified farmers were in the topmost ranking list.

Fig. 7. CART parameter setting.

Figs. 8 and 9 shows the list of qualified farmers to receive the grant.



Results: Manambrag, SAN ANDRES (CALOLBON) Download

Rank #	Name	Main Livelihood	Annual Income	Is PWD	Is 4Ps Member
1	JANICE VILLEGAS	Farmer	6,715.00	False	False
2	ELMER TABIOS	Farmer	5,961.00	False	False
3	LERIE VILLEGAS	Farmer	6,867.00	False	False
4	LEONOR EBIOLA	Farmer	5,760.00	False	False
5	GOMERSINDO VILLEGAS	Farmer	7,854.00	True	False
	DANIEL DELOS				

Fig. 8. List of qualified rice farmer to receive a financial grant.

Results: Manambrag, SAN ANDRES (CALOLBON) Download

IPs Member	Is Org. Member	Is Agrarian Beneficiary	Priority Probability	Predicted Priority
False	True	False	1	1
False	True	False	1	1
False	True	False	1	1
False	True	False	1	1
False	True	False	1	1

Fig. 9. List of qualified rice farmer to receive a financial grant.

B. *Software Quality Evaluation*

Eight staff from the Local Government Unit (LGU) and ten academic members from Catanduanes State University in this country conducted a practical hands-on evaluation of the system to assess its functionality. Afterwards, the survey instrument was employed, which incorporated the software testing standards and ISO/IEC 25010 factors. These parameters encompass functional suitability, performance efficiency, compatibility, dependability, usability, security, maintainability, and portability. The survey questionnaire employed a five-point Likert Scale to evaluate each benchmark statement. The result of the evaluation of the system is depicted in Fig. 10. The system received an "Excellent" rating with a grand mean of 4.59, indicating significant benefits for the LGU community. The highest rating was for "Performance Efficiency," indicating rapid and effective system functionality. However, the lowest ranking was for "Portability," suggesting developers prioritize it over end-users. Overall, the system received high ratings in the rest of the parameters.

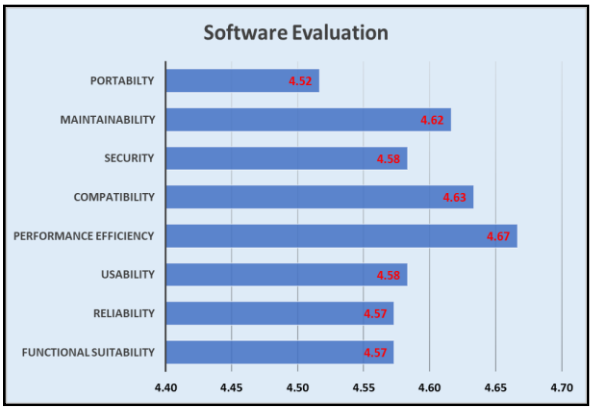


Fig. 10. Software quality evaluation.

V. Conclusion

The Rice Farmer Registry Application and Management System employed the Classification and Regression Tree (CART) methodology to create a decision-making model for evaluating farmers' eligibility for financial aid using rice farmer data collected from 2019 to 2023. The strategy effectively categorized and prioritized aid beneficiaries by analyzing socio-economic and agricultural characteristics, such as income, livelihood, and governmental support. The model, trained and assessed using a 70–30 data split, included parameters including Gini impurity, maximum depth, and minimum samples split to improve its performance and reduce overfitting. The results, evaluated by accuracy, precision, and recall criteria, confirmed the system's effectiveness in identifying and selecting competent farmers.

The system developed in this study significantly improves the efficiency of Local Government Units (LGUs) in agricultural support. With the use of International Organization for Standardization (ISO) and the International Electro-Technical Commission (IEC), 25010 software metrics, the system received an "Excellent" grade from technical experts and users with a mean score of 4.59, notably optimizing aid distribution to rice farmers. This sets a new standard for technological progress in the industry. The study represents a significant step in agricultural modernization, particularly in developing nations, providing valuable insights for farmers, policymakers, and stakeholders. Continuous enhancement and improved cooperation between local and provincial bodies are crucial for maximizing the system's effectiveness. The initiative also enhances local agricultural practices and promotes sustainable development in rural communities.

Conflict of Interest

The author declares no conflict of interest.

Funding

This research was funded by the Catanduanes State University with support through faculty-teaching equivalent (deloading) and other research grants provided to the author.

Acknowledgment

The author wishes to express her heartfelt gratitude to the Municipal Agricultural Officer and the staff of the Local Government Unit (LGU) of San Andres, Catanduanes, Philippines, for generously providing their time and valuable information, which was essential for this investigation. Sincere thanks are also extended to everyone who contributed to the success of this endeavor, from the conceptualization phase to the publication of this research. Special appreciation is given to Prof. Arejay Tabirara for his assistance in the development, training, testing, and evaluation of the CART model evaluation.

References

[1] USDA | Foreign Agricultural Service. (2024). Production—Rice. [Online]. Available: https://fas.usda.gov/data/production/commodity/0422110
[2] ROYS FARM. (Nov. 2018). Rice farming: Best business plan with 20 tips. [Online]. Available: https://www.roysfarm.com/rice-farming
[3] T. Laiprakobsup, "The policy effect of government assistance on the rice production in Southeast Asia: Comparative case studies of Thailand, Vietnam, and the Philippines," *Development Studies Research*, vol. 6, no. 1, pp. 1–12, 2019. doi: 10.1080//21665095.2019.1568900
[4] A. Twin. (February 2024). What is data mining? How it works, benefits, techniques, and examples. *Investopedia*. [Online]. Available: https://www.investopedia.com/terms/d/datamining.asp
[5] T. Ojo and Lloyd, "Determinants of climate change adaptation strategies and its impact on the net farm income of rice farmers in south-west Nigeria," *Land Use Policy*, vol. 95, pp. 103946–103946, 2020. doi: 10.1016/j.landusepol.2019.04.007
[6] K. P. Devkota, E. Pasuquin, A. Elmido-Mabilangan *et al.*, "Economic and environmental indicators of sustainable rice cultivation: A comparison across intensive irrigated rice cropping systems in six Asian countries," *Ecological Indicators*, vol. 105, pp. 199–214, 2019. doi: 10.1016/j.ecolind.2019.05.029
[7] I. Esterman. (July 2023). Vietnamese rice farmers go high-tech to anticipate a low-water future. *Mongabay Environmental News* [Online]. Available: https://news.mongabay.com/2023/07/vietnamese-rice-farmers-go-high-tech-to-anticipate-a-low-water-future/
[8] R. Alfred, J. H. Obit, C. P.-Y. Chin, H. Haviluddin, and Y. Lim, "Towards paddy rice smart farming: A review on big data, machine learning, and rice production tasks," *IEEE Access*, vol. 9, pp. 50358–50380, 2021. doi: 10.1109/access.2021.3069449
[9] E. Paul and Santos, "Development and assessment of online platform for crop production with decision support system," *The QUEST Journal of Multidisciplinary Research and Development*, vol. 1, no. 1, 2021. doi: 10.60008/thequest.v1i1.12
[10] B. Jlifi, C. Abidi, and C. Duvallet, "Beyond the use of a novel Ensemble based Random Forest-BERT Model (Ens-RF-BERT) for the sentiment analysis of the hashtag COVID19 tweets," *Social Network Analysis and Mining*, vol. 14, no. 1, 2024. doi: 10.1007/s13278-024-01240-x
[11] Z. Svirca. (May 2020). Everything you need to know about MVC architecture. *Towards Data Science* [Online]. Available: https://towardsdatascience.com/everything-you-need-to-know-about-mvc-architecture-3c827930b4c1
[12] Rick-Anderson. (June 2023). ASP.NET MVC overview. [Online]. Available: https://learn.microsoft.com/en-us/aspnet/mvc/overview/older-versions-1/overview/asp-net-mvc-overview
[13] R. Sheldon. (September 2023). Model-View-Controller (MVC). [Online]. Available: https://www.techtarget.com/whatis/definition/model-view-controller-MVC
[14] K. Casey. (Jan. 2022). Application development | rad methodology complete guide. *CODERSERA.* [Online]. Available: https://codersera.com/blog/rad-methodology/
[15] What is a decision tree. (2023). *IBM*. [Online]. Available: https://www.ibm.com/topics/decision-trees#:~:text=A%20decision%20tree%20is%20a,internal%20nodes%20and%20leaf%20nodes
[16] Anuradha and G. Gupta, "A self explanatory review of decision tree classifiers," in *Proc. International Conference on Recent Advances and Innovations in Engineering*, 2014, pp. 1–7.
[17] What is a dashboard? A complete overview. (2024). *Tableau*. [Online]. Available: https://www.tableau.com/learn/articles/dashboards/what-is#:~:text=The%20main%20use%20of%20a,relevant%20data%20in%20key%20areas