World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(Research Article)

Check for updates

# Enhancing the livelihoods of small-scale paddy farmers in sri lanka through computational approaches for inclusive and equitable agriculture

Sonal Jayawardana *, Dulan Bamunuge, Chamodh Abeygoonawardana, Kaveen Wijethunga, Dinuka Wijendra and Samitha Vidhanaarachchi

*Sri Lanka Institute of Information Technology, Department of Computer Science and Software Engineering, Malabe, Western Province, Sri Lanka.*

World Journal of Advanced Research and Reviews, 2023, 20(01), 882–893

Publication history: Received on 02 September 2023; revised on 14 October 2023; accepted on 16 October 2023

Article DOI: https://doi.org/10.30574/wjarr.2023.20.1.2087

## Abstract

Paddy cultivation, also known as rice cultivation, contributes significantly to Sri Lanka's food security and economic prosperity. Nevertheless, a number of challenges exist in the practice, such as insufficient furrow irrigation techniques, inefficient fertilizer subsidy distribution, a lack of yield prediction models, and limited access to research findings. In order to address these concerns, this research proposes an integrated system with four components. An application for furrow irrigation planning is included in the first component, utilizing support vector machine algorithms and random forest algorithms to optimize irrigation paths. Second, a priority-based fertilizer subsidy system is introduced, employing Gradient Boosting to allocate subsidies efficiently based on farmers' needs. The third component focuses on improving the dissemination of research findings by employing text summaries in order to make the findings more accessible to small-scale farmers. In the fourth component, a yield prediction model is implemented using a Random Forest algorithm that takes into account climatic and soil variables in order to forecast expected yields. The integrated system aims to enhance the efficiency, productivity, and profitability of paddy farming in Sri Lanka, providing practical solutions to challenges faced by farmers. By adopting these innovations, farmers can make informed decisions and optimize their agricultural practices, leading to sustainable rice production and economic growth.

**Keywords:** Paddy cultivation; Small-scale farmers; Fertilizer subsidy distribution; Random Forest algorithms; Gradient Boosting; Sustainable rice production

## 1. Introduction

Paddy cultivation, also known as rice cultivation, holds significant importance in Sri Lanka, serving as a vital agricultural practice. With a rich history dating back over 2,500 years, paddy farming has played a crucial role in the nation's food security and economic prosperity [1]. In 2020, paddy cultivation accounted for approximately 29% of total agricultural land use in Sri Lanka, with a cultivation area of around 818,000 hectares and a production of 3.5 million metric tons of rice [2]. The agriculture industry, including paddy cultivation, contributed about 7.7% to Sri Lanka's GDP in the same year [2].

Paddy agriculture in Sri Lanka is deeply rooted in its culture and society, connecting communities across generations. Passed down through time, the ancient practices and procedures embody accumulated wisdom. This agricultural practice requires meticulous planning and execution. Farmers prepare fields, level soil, and construct irrigation systems to ensure a steady water supply for rice plants. Transplanting young rice seedlings in flooded areas demands precision and expertise. Sri Lanka's favorable climatic conditions and diverse ecology contribute to successful paddy cultivation. Monsoonal rainfall patterns and well-maintained irrigation infrastructure create an ideal environment for rice

* Corresponding author: Sonal Jayawardana

production. The country's geographical diversity, encompassing lowlands, uplands, and hill areas, allows for the cultivation of different rice varieties tailored to specific regions and elevations. Paddy farming extends beyond large-scale commercial farms, with small-scale farmers practicing subsistence farming playing a significant role in Sri Lanka's rice output. These farmers cultivate paddy using traditional methods on small plots of land, relying on family and community support.

Paddy cultivation in Sri Lanka faces several challenges that need attention. Firstly, the absence of a proper mechanism to utilize the furrow irrigation techniques. This contributes to an inadequate and uneven distribution of water across paddy fields which results in subpar crop growth and contamination of water due to runoff [3] [4]. Secondly, the fertilizer subsidy program in the country suffers from disorganization, inefficiency, and outdated practices. The absence of a priority-based fertilizer distribution process impedes productive agriculture's progress [5] [6]. Moreover, small-scale farmers in Sri Lanka lack access to an efficient prediction model that can forecast profits and related expenses in paddy cultivation. [7] Consequently, due to the absence of an accurate and affordable yield prediction for small-scale paddy farmers in Sri Lanka; farmers struggle to make informed decisions and optimize their farming practices [8]. Furthermore, small-scale farmers in Sri Lanka lack access to the latest research findings on paddy cultivation [9]. This absence of a proper mechanism to disseminate the latest research findings to paddy farmers is primarily due to inadequate communication between farmers and research institutes and the technical complexity of the research findings. [10].

## 2. Literature review

In a comprehensive 2015 study [16], researchers explored furrow lengths and field application techniques for paddy cultivation using the Bontanga irrigation system. The surge technique demonstrated superior performance with the highest advance rate (1.26 min/m) and the shortest opportunity time (11 minutes) in Block A, while bunds had the lowest advance rate (0.92 min/m). This study effectively identified an efficient method for achieving uniform water flow. Additionally, a separate 2014 study [17] addressed the management of water resources in Sri Lanka's dry zone for paddy cultivation, emphasizing the absence of a proper mechanism for identifying irrigation paths. This lack leads to inefficient water allocation and increased water usage compared to countries utilizing advanced agricultural technologies.

An analysis of Sri Lanka's Fertilizer Subsidy Program [6] proposed policy alternatives considering economic, environmental, and socio-political factors. The analysis highlighted potential issues such as misuse, corruption, and dependency while presenting options for maintaining, removing, or reforming the subsidy program through targeted or gradual reduction approaches. Another investigation [5] revealed that farmers maintain optimal fertilizer application despite challenges, with short-term price and policy changes having minimal impact on fertilizer demand. The study recommended a phased-out subsidy scheme, accompanied by infrastructure investments and improved marketing efficiency, for long-term productivity. Additionally, an analysis [13] demonstrated the positive effects of fertilizer use on paddy production in Sri Lanka, underscoring the significance of sustainable fertilizer policies given domestic and international market implications. It emphasized the need to evaluate experimental policies related to land, fertilizer, and inputs over time.

Researchers from Rice Research and Development Institute (RRDI) have identified barriers to disseminating research outputs to farmers, including a lack of extension staff, weak research/extension links, inadequate training, and insufficient extension programs [9]. Small-scale farmers in Sri Lanka encounter difficulties in accessing up-to-date research findings due to limited communication between farmers and research institutes, hindering the utilization of valuable knowledge [10]. Improved communication channels are necessary, as small-scale farmers primarily rely on informal networks and personal contacts for information [18]. Technical research papers pose challenges for small-scale farmers, highlighting the importance of presenting findings in a more accessible manner. Valuable sources of agricultural information for paddy farmers include agricultural offices, agriculture instructors, nearby farmers, and innovative farmers [20].

Several studies focused on paddy yield and farming information in Sri Lanka. A 2021 study [21] developed crop-weather models based on weather indices to predict paddy yield, with the Random Forest model proving to be the most accurate. However, non-climatic factors such as soil fertility and market demand were not considered in this study. Another study from 2020 [22] investigated the relationship between paddy yield and climatic parameters using Artificial Neural Networks, finding the Levenberg-Marquardt algorithm to be the most effective, although further research is needed for improved accuracy. Additionally, in 2014 [23], researchers addressed the challenge of accessing farming information by creating a mobile-based solution with a knowledge repository tailored to the specific needs of farmers at different stages of the farming life cycle, aiming to enhance information accessibility for farmers in Sri Lanka.

## 3. Methodology

The overall system aims to address challenges in paddy cultivation in Sri Lanka through four research components. The first component develops a mobile application for efficient furrow irrigation planning. The second component enhances the fertilizer subsidy program by introducing a priority-based distribution process. The third component improves research dissemination to small-scale farmers, enabling informed decision-making. The fourth component develops a mobile-based solution to predict farmers' expected yield, considering climatic and non-climatic variables. Together, these components aim to enhance the efficiency, productivity, and profitability of paddy farming in Sri Lanka.

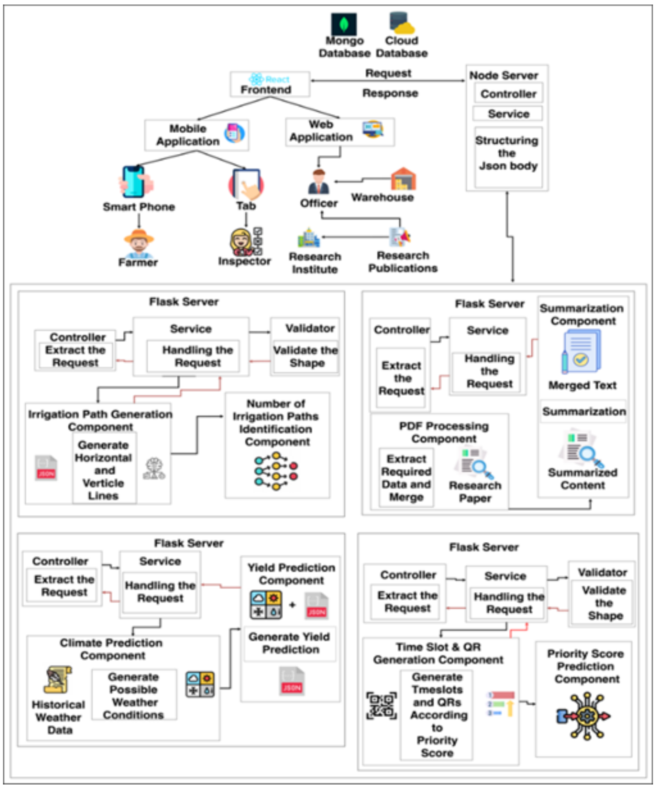


**Figure 1** Overall system diagram

### 3.1. Furrow Irrigation

Paddy cultivation involves various stages, including land preparation, seed selection, transplantation, weeding, and harvesting. One critical aspect of successful rice farming is irrigation, which ensures a consistent water supply for optimal growth [11]. In Sri Lanka, traditional irrigation methods are complemented by specific technologies like diversion structures, field levelling, and furrow construction, contributing to effective water management and supporting successful rice cultivation.

For the implementation, district-wise data on soil composition and types were obtained from the meteorological department. An initial dataset of irrigation path estimations, based on climatic and demographic features, was acquired from the RRDI, and later expanded with the help of research instructors from RRDI. Data preprocessing techniques, including data cleaning, transformation, and feature selection, were applied to ensure data quality, reliability, and suitability for analysis and modelling.

For retrieving the site plan of a paddy field, a map component utilizing GeoJson [24] was implemented. A path creation algorithm was used to generate the optimal furrow irrigation pattern for the given site plan. The resulting design, including the furrow irrigation details, would be converted into a 2D plan and delivered to the user through the map component.

The path creation algorithm takes into account various inputs such as the spacing between irrigation paths, the number of irrigation paths, soil type, paddy type, and water requirement. Additionally, the algorithm considers the elevation (slope) of the land and the water flow through the site when determining the placement of irrigation paths within the paddy field. To calculate the appropriate number of irrigation paths, a supervised machine learning model was developed using multiple algorithms like Support Vector Machine, Linear Regression, Decision Tree, and Random Forest. The model's performance was evaluated using various accuracy metrics. Once the dimensions and shape of the paddy field, as well as its slope, are provided, the trained model will predict the suitable number of irrigation paths for optimal furrow irrigation.

### 3.2. Priority-based Fertilizer Subsidy Distribution

Nutrients like Nitrogen, Phosphorus, Potassium, and micronutrients are vital for paddy crops, and farmers rely on fertilizers for their supply.[12] Proper fertilizer management is crucial for maximizing yields and preventing environmental contamination [12]. To assist farmers in obtaining these nutrients, the Sri Lankan government has implemented a subsidy program. Farmers receive fertilizers containing nitrogen, phosphorus, potassium, and micronutrients either for free or at a discounted rate [6]. While the subsidy program has contributed to agricultural growth and food security, concerns regarding leakages, corruption, and unsustainable fertilizer use have prompted the need for reforms [13].

This component has been initiated as a pilot project for the Kalutara district. Data from farmers who had registered for the fertilizer subsidy program in two Agrarian divisions, namely Kalutara and Agalawatta, were collected through Agricultural Instructors (AIs) assigned to each division. The collected data encompassed a total population of 800 farmers who had registered for the subsidy program. A scientific survey was conducted with the participation of Agricultural Instructors to identify potential factors that could be considered when prioritizing the distribution of fertilizer subsidies among farmers. A priority score was allocated to each farmer based on the survey results.

Data preprocessing techniques (data cleaning, data transformation and feature selection) were employed to ensure the quality, reliability, and suitability of the data for analysis and modelling. Imbalances in the initial dataset were addressed using oversampling, under-sampling, and class weights techniques to enhance the accuracy of predictions.

In addition to the survey results, a correlation analysis was conducted to identify the factors contributing to the priority score. Multiple supervised learning models, including Decision Trees, Random Forest, Gradient Boosting, and Support Vector Machine, were implemented and trained to predict the priority score when provided with a farmer's details. [25] The performance of each model was evaluated using various accuracy measures such as precision, recall, F1 score, and overall accuracy. The best-performing model was selected to generate a priority distribution list. Time slots would be generated based on this priority distribution list, and QR codes would be created for each time slot. These QR codes would then be distributed to the registered farmers through a notification component.

### 3.3. Research Dissemination

Research and development have played a pivotal role in the advancement of paddy cultivation in Sri Lanka. Institutions like the Rice Research and Development Institute (RRDI) have been instrumental in creating and transferring knowledge and technologies to enhance productivity, efficiency, and sustainability in rice farming [14]. The development of high-yielding rice varieties, improved cultivation practices, and the adoption of environmentally friendly farming methods are among the key research findings that have significantly impacted rice production [9]. These advancements have resulted in increased rice productivity, reduced costs, and minimized environmental impact, contributing to the overall economic development and food security of Sri Lanka.

Research publications related to paddy cultivation were obtained from the Rice Research and Development Institute (RRDI) to fulfil the data requirements of this component. The selected research papers have been categorized into three main research categories: pests and diseases, weed science, and yield optimization. To extract the text from the research papers, a Python library called PyPDF2 was used, which reads the pages of the PDF file from a specified URL and concatenates the text.

To identify and extract entities from the text, utilized the advanced Named Entity Recognition (NER) model called Spacy. This NER model processed the text and extracted the abstract and conclusion sections from each research paper, which contained the summary of the research. A sentence similarity comparison model was then utilized to identify unique sentences within the text and assemble them into a coherent paragraph. This paragraph was then forwarded to a Natural Language Processing (NLP) model, which generated a concise and meaningful summary of the research paper, tailored to suit a non-technical audience's comprehension.

To achieve text summarization, a diverse set of established Natural Language Processing (NLP) models, including Facebook/BART-large-CNN, T5 Transformers, and Kullback-Leibler Sum, were employed.[26] These models were carefully selected based on their demonstrated proficiency in generating accurate and coherent summaries. The quality of the generated summaries was evaluated using metrics such as ROUGE, while also considering the readability of the summaries for a non-technical audience. Through this evaluation process, the most effective summarization model was identified, ensuring optimal performance in producing high-quality summaries.

### 3.4. Yield Prediction

Despite the integration of modern scientific techniques, traditional prediction methods for paddy cultivation continue to be practiced in Sri Lanka, alongside advancements in meteorological data, satellite imagery, and other tools. Traditional techniques, based on observations of celestial movements, seasonal changes, and animal behavior, have been complemented by scientific knowledge, although accessibility and funding limitations continue to sustain their use, especially in rural areas [15]. While traditional methods have advantages, such as cultural significance and local knowledge, they also have limitations in terms of accuracy, flexibility, reliance on superstitions, and labor-intensive processes [15].

To acquire the data requirements of this component, existing datasets of district-wise climatic data and yield statistics data were obtained from the meteorological department and the department of census and statistics. These data were collected after analyzing the gaps and suggestions provided by existing studies. Data preprocessing techniques, including data cleaning, feature selection, feature engineering, and data augmentation, were applied to ensure the quality, reliability, and suitability of the data for analysis and modelling. Several supervised learning models, namely Decision Tree, Gradient Boosting, Elastic Net, and Random Forest were implemented and trained using the acquired datasets to predict the expected yield.

In addition to these models, an ensembling model was also trained. When making predictions, both climatic variables (such as rainfall, minimum and maximum humidity, and minimum and maximum temperature) and soil variables (including texture, structure, pH value, and nitrogen value) were taken into consideration. The model takes season, coordination of the paddy field, paddy type, soil type, cultivation method and cultivation month as input parameters and predicts the yield and wastage. Multiple accuracy measurements were used to evaluate the performance of the models, and cross-validation techniques were employed to further assess their effectiveness. The best-performing model, determined through these evaluations, was selected to carry out the yield predictions.

## 4. Results and Discussions

### 4.1. Furrow Irrigation

The path creation algorithm demonstrated an execution time of 71.721077 milliseconds. The algorithms employed in the study included Support Vector Machine (SVM), Linear Regression, Decision Tree (DT), and Random Forest (RF). Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-Squared were utilized as accuracy metrics to evaluate the performance. The accuracies of each model be viewed in the table. [Tab. 1]. The residual plot of the model [Fig. 1] and the graph containing predicted values vs the actual values [Fig. 2] further support this justification.

**Table 1** Furrow irrigation accuracy table

| Regression Modal | MSE | MAE | RMSE | R-Squared |
|---|---|---|---|---|
| SVM(SVR) | 17.7043 | 3.0214 | 4.2076 | 0.6428 |
| Linear Regression | 341.6629 | 13.5857 | 18.4841 | 0.8059 |
| Decision Tree | 3.5929 | 0.9529 | 1.8955 | 0.9880 |
| Random Forest | 2.3100 | 1.0214 | 1.5199 | 0.9987 |

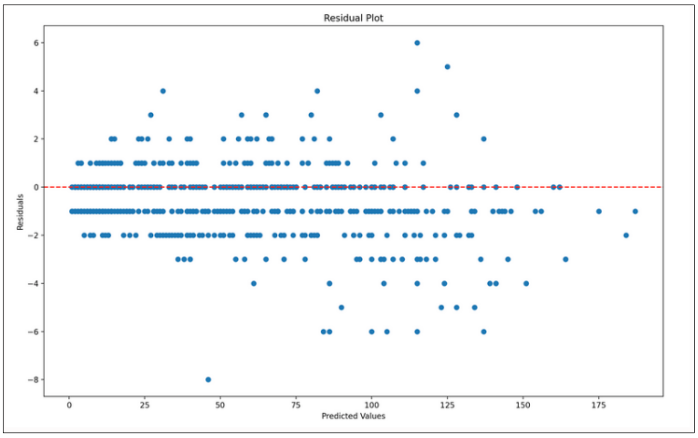


**Figure 2** Residual plot of the Random Forest model

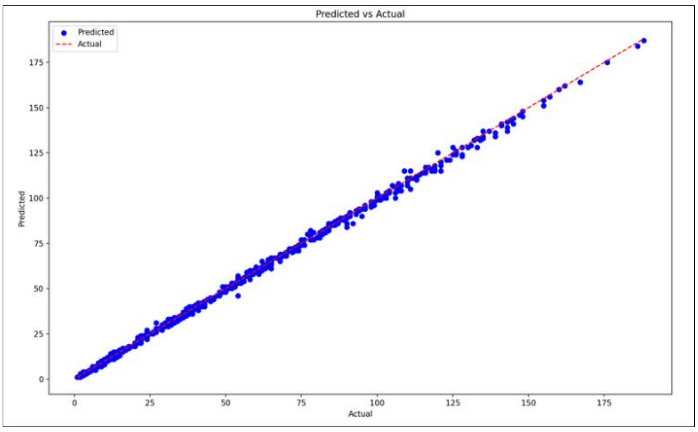


**Figure 3** Predicted vs actual values of the Random Forest model

Among the models tested, the Random Forest (RF) algorithm emerged as the best model due to its superior performance across all accuracy metrics. The RF model achieved an MSE of 2.3100, an MAE of 1.0214, an RMSE of 1.5199, and an R-Squared value of 0.9987. These results surpassed those of the other algorithms, justifying the selection of RF as the optimal model based on the accuracy metrics. In comparison to previous studies[16] [17], it is important to note that this research surpasses conceptual frameworks by providing a practical implementation for determining irrigation paths in paddy fields. Earlier studies lacked a practical approach. However, the current implementation is limited to accurately providing results only for rectangular-shaped fields. The model faces challenges in identifying irregular objects, resulting in decreased accuracy for irregularly shaped fields. The significance of this research lies in its application for small-scale paddy farmers. By providing their site plans and necessary details, farmers can obtain an optimal irrigation site plan tailored to their specific fields, enabling them to enhance irrigation practices and improve efficiency. Future work should focus on increasing the accuracy of the model to deliver precise results for irregularly shaped land. Additionally, implementing a system to assess the optimality of existing irrigation plans in paddy fields would be a valuable area for further exploration.

### 4.2. Priority-based Fertilizer Subsidy Distribution

In this research component, four different algorithms were utilized: Decision Tree (DT), Random Forest (RF), Support Vector Machine, and Gradient Boosting. The evaluation of these algorithms was based on accuracy metrics, specifically Train Accuracy and Test Accuracy. The accuracies of each can be viewed in the below table. [Tab. 2].

Among the models tested, the Gradient Boosting algorithm emerged as the best model, exhibiting superior performance across both Train Accuracy and Test Accuracy metrics. The Gradient Boosting model achieved a Train Accuracy of 0.98 and a Test Accuracy of 0.9625, surpassing the performance of the other algorithms. Comparison between test and train accuracies [Fig. 5] and comparison of the predicted labels against the actual labels of the Gradient Boosting model is depicted in the form of a confusion matrix. [Fig. 4].

**Table 2** Priority-based fertilizer subsidy distribution accuracy table

| Classifier Modal | Train Accuracy | Test Accuracy |
|---|---|---|
| Decision Tree | 0.95 | 0.9375 |
| Support Vector Machine | 0.895 | 0.925 |
| Random Forest | 0.991 | 0.95 |
| Gradient Boosting | 0.98 | 0.9625 |

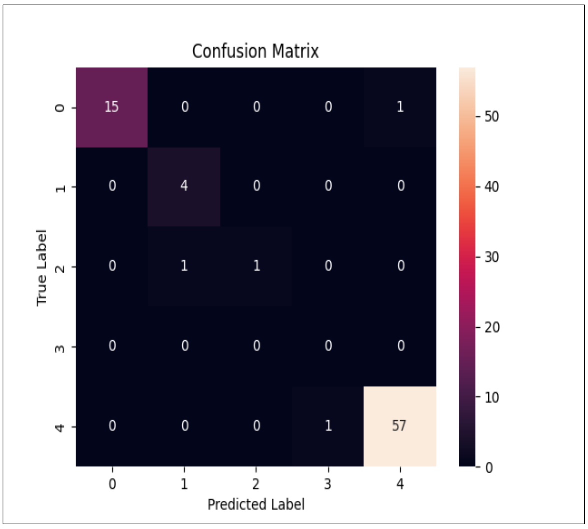


**Figure 4** Confusion matrix of the Gradient Boosting model

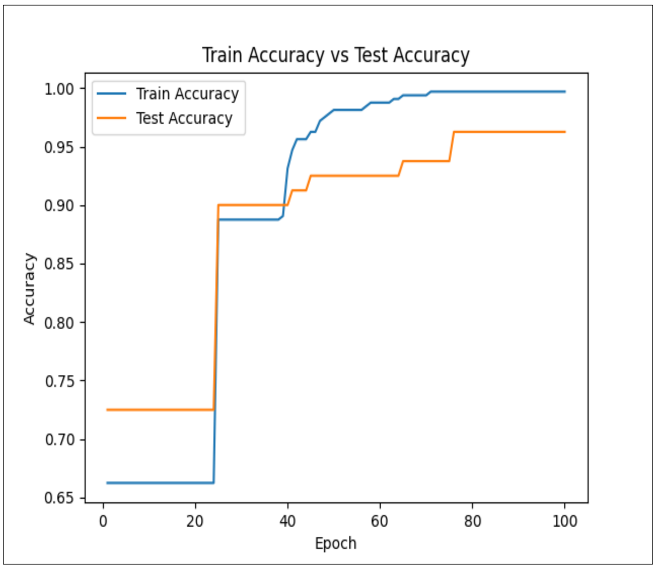


**Figure 5** Test accuracy vs train accuracy of Gradient Boosting model

When comparing this research with previous studies, it is evident that although previous studies [5][6][13] have identified issues in the fertilizer subsidy program, none of them has proposed an implementation to effectively distribute the subsidy among farmers, taking into account the urgency of their needs. This research addresses this gap and provides a method to conveniently distribute the fertilizer subsidy to farmers, ensuring the timely application of fertilizers to their fields. However, there are limitations to consider. The initial implementation of the system was based on a dataset derived from only two agrarian divisions in Kalutara. To provide island-wide coverage and expand the system beyond the pilot-level implementation, a larger dataset needs to be acquired. The significance of this research lies in its ability to provide farmers with a convenient method of receiving the fertilizer subsidy, ensuring that fertilizers are applied to their fields in a timely manner. This approach can greatly benefit farmers by optimizing their agricultural practices. Moving forward, future work should focus on acquiring a larger dataset that can accommodate island-wide prediction capability. By expanding the dataset, the model can be further improved and enhanced to provide accurate predictions and recommendations for farmers across a broader geographic area.

### 4.3. Research Dissemination

In this research component, four different algorithms were employed: Kullback-Leibler Sum Model, Text Summarization Using spaCy, T5 Transformer Model, and BART Model. The evaluation of these algorithms was based on two accuracy metrics: ROUGE F1 Score and Human Readability. The accuracies of each can be viewed in the below table. [Tab. 3]

**Table 3** Research disseminating accuracy table

| NLP Modal | ROUGE F1 Scorer | Human Readability |
|---|---|---|
| Kullback-Leibler Sum Model | 0.55 | Very low |
| Text Summarization Using spaCy | 0.55 | Very low |
| T5 Transformer Model | 0.229 | Moderate |
| BART Model | 0.451 | High |

Among the models tested, the BART Model emerged as the best choice. Although all models exhibited low human readability, the BART Model demonstrated the highest human readability and achieved a moderate ROUGE F1 Score. Comparatively, the Kullback-Leibler Sum Model, Text Summarization Using spaCy, and T5 Transformer Model had lower ROUGE F1 Scores. ROUGE results of the BART model for multiple research paper summarizations are depicted in the figure. [Fig .6].

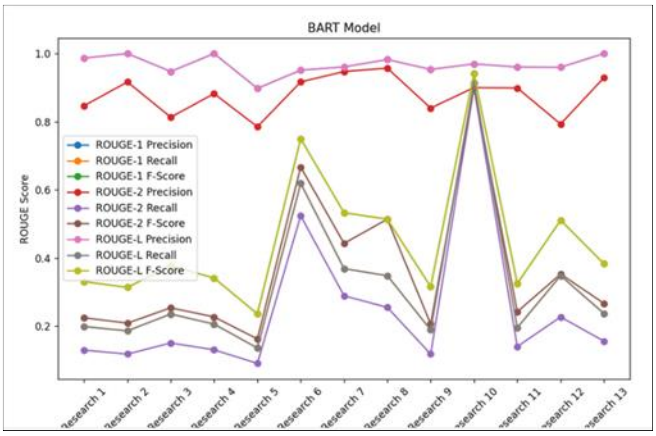


**Figure 6** ROUGH results of the BART Model for research paper summarization

When comparing this research with previous studies [9][10][18][20], it is evident that no effective solution had been provided for disseminating research findings to small-scale farmers. This research aims to address that gap and enable farmers to effectively apply research findings to enhance their productivity. However, there are limitations to consider. The recommendations derived from the research papers are based on an abstract level, which may lack specific implementation details that are crucial for farmers. Future work should focus on automating the process and establishing a direct bridge with the Rice Research and Development Institute (RRDI) to streamline the submission and dissemination of research papers, enabling direct access to comprehensive and practical findings. The significance of this research lies in its potential to empower farmers by providing them with accessible and applicable research findings. By improving the readability and relevance of research papers, farmers can leverage their knowledge to enhance their agricultural practices and productivity. Moving forward, future work will involve transitioning from the current system of submitting research papers through agrarian division levels to an automated process that directly connects with RRDI. This transformation aims to streamline the dissemination of research findings and ensure efficient and timely access for farmers.

### 4.4. Yield Prediction

In this research component, four algorithms were utilized: Decision Tree (DT), Gradient Boosting (GB), Elastic Net, and Random Forest (RF). The evaluation of these algorithms was based on accuracy metrics such as Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-Squared. The accuracies of each can be viewed in the below table. [Tab. 4].

**Table 4** Yield prediction accuracy table

| Regression Modal | Train RMSE | MAE | RMSE | R-Squared |
|---|---|---|---|---|
| Decision Tree | 1.4382 | 4.2140 | 18.2520 | 0.9367 |
| Gradient Boosting | 0.0131 | 2.3979 | 12.6725 | 0.9697 |
| Elastic Net | 2.360 | 1.7084 | 3.694 | 0.9974 |
| Random Forest | 0.3544 | 0.2779 | 3.1173 | 0.9982 |

Among the models tested, the Random Forest (RF) algorithm emerged as the best model based on the values of the accuracy metrics. The RF model exhibited superior performance compared to the other algorithms, achieving an MSE of 2.3100, an MAE of 1.0214, an RMSE of 1.5199, and an R-Squared value of 0.9987. The residual plot of the model [Fig. 7] and the graph containing predicted values vs the actual values [Fig. 8] further depicts the Random Forest model's superior performance. Further analysis through cross-validation techniques was conducted to assess the model's generalization ability. The cross-validation results [Fig .9] showed consistent performance, with a mean cross-validation RMSE score of 2.2469 and a mean cross-validation MAE score of 0.8519. These scores, along with the relatively small differences between the training and cross-validation scores, indicate that the model is not overfitting and is capable of generalizing well to unseen data.

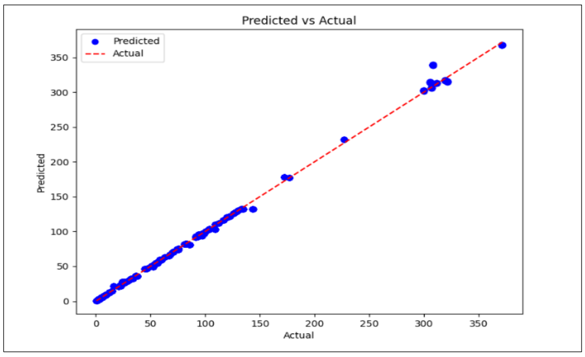


**Figure 7** Predicted vs actual values of the Random Forest model

Comparing this research with previous studies [21][22][23], it is evident that previous studies focused on either climatic variables or soil qualities alone, without considering both factors together in yield predictions. In contrast, this research took a more comprehensive approach by incorporating both climatic variables and soil qualities, leading to more holistic yield predictions in agriculture. However, limitations should be acknowledged. The dataset used in this research was derived from small-scale farmers, limiting the accuracy when predicting yields for larger farming areas. To address this, future work should focus on acquiring a larger dataset that covers a wider geographical region, enabling the model to provide predictions for farming areas across the entire region. Additionally, incorporating feedback from actual customers who use the application will contribute to continuous improvements and enhance the model's predictive capabilities. The significance of this research lies in its ability to provide farmers with acceptable yield predictions, allowing them to plan their future work accordingly. By considering both climatic variables and soil qualities, this research offers a more holistic and practical approach to yield prediction in agriculture. Moving forward, efforts should be directed toward acquiring a larger dataset for wider coverage, improving the model based on customer feedback, and ensuring its effectiveness and reliability in aiding farmers in their decision-making processes.

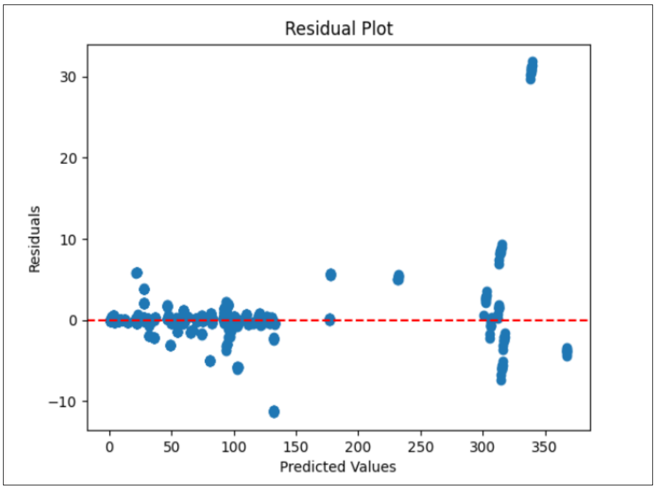


**Figure 8** Residual plot of the Random Forest model

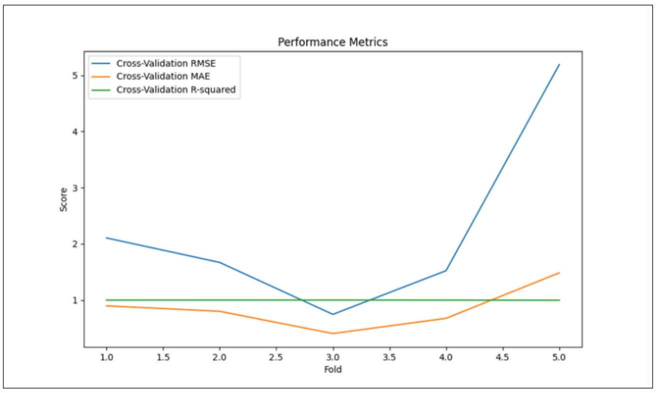


**Figure 9** Cross-validation results of the Random Forest model

## 5. Conclusion

Paddy cultivation has been an important practice in Sri Lanka for over 2,500 years, contributing significantly to food security and economic prosperity. Challenges in rice farming include inadequate and uneven distribution of water due to a lack of suitable bed irrigation mechanisms, ineffective and disorganized fertilizer subsidy programs, and a lack of yield forecasting models for smallholder farmers. Small-scale farmers, who play an important role in rice production in Sri Lanka, often find it difficult to access the latest research results due to the communication gap between farmers and research institutes. To address these challenges, a comprehensive research project was undertaken, focusing on four components: furrow irrigation, distribution of fertilizer subsidies according to priorities, research dissemination and production forecasting. The implementation of the furrow generation algorithm allows the optimal irrigation lines to be determined for rectangular fields. The study also proposes a plan to distribute fertilizer subsidies based on priority, ensuring timely fertilization of fields based on farmers' needs. To improve research dissemination, a BART model was used to improve the readability and relevance of research papers, providing farmers with accessible and applicable results. By leveraging technology and data-driven approaches, farmers can improve irrigation practices, optimize fertilizer use, and make informed decisions for successful rice farming. Furthermore, empowering farmers with accessible research results will promote knowledge transfer and facilitate sustainable farming practices. For example, the accuracy of the furrow irrigation algorithm is limited to rectangular fields, while irregular fields present challenges. In addition, the initial implementation of the fertilizer subsidy distribution system was based on a limited data set, requiring expansion for broader coverage and improved forecasting. Overall, this study contributes significantly to the improvement of rice farming practices in Sri Lanka. By addressing key challenges and leveraging technology and data-driven approaches, it is paving the way for greater efficiency, productivity and sustainability in the country's agricultural sector. Future work should focus on improving the accuracy of irregular fields in the row irrigation component and expanding the dataset to achieve island-wide coverage and efficiency for the fertilizer subsidy model. Furthermore, promoting direct contact with research institutes will ensure quick access to valuable knowledge by farmers, further supporting the development of the rice industry in Sri Lanka.

## Compliance with ethical standards

*Disclosure of conflict of interest*

No conflict of interest to be disclosed.

## References


[1] T. W. Wikramanayake, "THE BEGINNINGS OF RICE CULTIVATION IN SRI LANKA," Journal of the Royal Asiatic Society of Sri Lanka, vol. 53, p. 203–218, 2007.

[2] Dilini D. Perera, A.K.A. Dissanayake, U.D. Raveena Udari, W.A.R. Wickramasinghe, "Policies that Matter or Are Forgettable The Evaluation of Paddy Policies in Sri Lanka:1998 to 2018," Hector Kobbekaduwa Agrarian Research and Training Institute, Colombo, 2021.

[3] Blain Hanson, Irrigation Specialist, U.C. Davis, "A Simple Way to Estimate the Distribution Uniformity of Furrow Irrigation Systems," University of California, California.

[4] "WHAT IS AGRICULTURAL RUNOFF?," 10 12 2022. [Online]. Available:https://www.buschsystems.com/resource-center/knowledgeBase/glossary/what-is-agricultural-runoff#:~:text=Defintion,bodies%20of%20waters%20or%20evaporate.

[5] H. Ekanayake, "The Impact of Fertilizer Subsidy on Paddy Cultivation in Sri Lanka," Staff Studies, vol. 36, 2009.

[6] Weerahewa, Jeevika; Kodithuwakku, Sarath S.; Ariyawardana, Anoma, "The Fertilizer Subsidy Program in Sri Lanka," Food Policy for Developing Countries: Case Studies, 2010.

[7] Prameya R Hegde, Ashok Kumar A R, "Crop Yield and Price Prediction System for Agriculture Application," INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT), vol. 11, no. 07, 2022.

[8] A. A. S. B. J. B. Ranjani Dhanapal, "Crop price prediction using supervised machine learning algorithms," Journal of Physics, 2021.

[9] Dilrukshi, W.P.T, Muddannayake, "A study on agriculture information generation and dissemination at the Department of Agriculture (DOA) in Sri Lanka," Journal of the University Librarians Association of Sri Lanka, vol. 18, no. 1, p. 15, 2014.

[10] M.B.N.Abeysinghe, "Decentralized Agricultural Extension and Social Networks: The Potential for Agricultural Information and Technology Dissemination in Sri Lanka," SRI LANKAN JOURNAL OF AGRICULTURE AND ECOSYSTEMS, vol. 3, no. 2, p. 30, 2021.

[11] R. Shand, "Irrigation and Agriculture in Sri Lanka," Institute of Policy Studies, Sri Lanka, 2002.

[12] "Nutrient Management," Rice Knowledge Bank, 04 04 2023. [Online]. Available: http://www.knowledgebank.irri.org/training/fact-sheets/nutrient-management.

[13] Jeevika Weerahewa, Shashika D. Rathnayaka, Naduni Nayanathara, Devesh Roy, "Decomposition of Productivity Growth in Sri Lanka's Paddy Sector: Roles of Area Expansion and Chemical Fertilizer Use," Sri Lanka Statistical Review, vol. 1, no. 1, 2022.

[14] S.R. Sarathchandra, K.P.S.D. Hennayake, M.P.H.K. Jayaweera, L. Nugaliyadde, K.S. Hemachandra, A.D.N.T. Kumara,, "Assessment of Grain Preference Among Selected Rice Varieties by Field Rat, Bandicota bengalensis and House Rat, Rattus rattus in Sri Lanka," Tropical Agricultural Research, 2022.

[15] D. Hirimuthugodage, "Sri Lankan farmers' traditional knowledge and climate change predictions," 2023. [Online].

[16] Issaka, R. Z., Ibrahim. H., Issah, M.H, "Performance Evaluation Of Furrow Lengths And Field," INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY, vol. 4, no. 10, 2015.

[17] Sisira S. Withanachchi, Sören Köpke, Chandana R. Withanachchi, Ruwan Pathiranage, "Water Resource Management in Dry Zonal Paddy Cultivation in Mahaweli River Basin, Sri Lanka: An Analysis of Spatial and Temporal Climate Change Impacts and Traditional Knowledge," MDPI Climate, vol. 2, no. 4, p. 329–354, 2014.

[18] Sisira S. Withanachchi, Sören Köpke, Chandana R. Withanachchi, Ruwan Pathiranage, "Water Resource Management in Dry Zonal Paddy Cultivation in Mahaweli River Basin, Sri Lanka: An Analysis of Spatial and Temporal Climate Change Impacts and Traditional Knowledge," MDPI Climate, vol. 2, no. 4, p. 329–354, 2014.

[19] C.A.K. Dissanayake, W. Jayathilake, H.V.A. Wickramasuriya, U. Dissanayake, W.M.C.B. Wasala, "A Review on Factors Affecting Technology Adoption in Agricultural Sector," The Journal of Agricultural Sciences - Sri Lanka, vol. 17, 2022.

[20] C.A.K. Dissanayake, W. Jayathilake, H.V.A. Wickramasuriya, U. Dissanayake, W.M.C.B. Wasala, "A Review on Factors Affecting Technology Adoption in Agricultural Sector," The Journal of Agricultural Sciences - Sri Lanka, vol. 17, 2022.

[21] Piyal Ekanayake, Windhya Rankothge, Rukmal Weliwatta, Jeevani W. Jayasinghe, "Machine Learning Modelling of the Relationship between Weather and Paddy Yield in Sri Lanka," Hindawi, 2021.

[22] Vinushi Amaratunga, Lasini Wickramasinghe, Anushka Perera, Jeevani Jayasinghe, Upaka Rathnayake, "Artificial Neural Network to Estimate the Paddy Yield Prediction Using Climatic Data," Hindawi, 2020.

[23] Anusha Indika Walisadeera, Athula Ginige, Gihan Nilendra Wikramanayake, "User centered ontology for Sri Lankan farmers," ScienceDirect, vol. 26, pp. 140-150, 2015.

[24] Morandini, L. and Sinnott, Richard and Tomko, Martin, "Getting the best performance for GeoJSON map visualizations: PostGIS vs CouchDB backend," in *Free and Open Source Software for Geospatial Conference*, Nottingham, UK, 2013.

[25] Er-raji, Naoufal and Benabbou, Faouzia and Danubianu, Mirela and Zaouch, Amal, "Supervised Machine Learning Algorithms for Priority Task Classification," 2018.

[26] F. St-Amant, "Fine-Tuning the BART Large Model for Text Summarization," [Online]. Available: https://towardsdatascience.com/fine-tuning-the-bart-large-model-for-text-summarization-3c69e4c04582.