Research Proposal

Research Project – ICT 4608

Bachelor of Information and Communication Technology (Honors)

Department of Information and Communication Technology
Faculty of Technology
Rajarata University of Sri Lanka

Details of the Research Project

Research Title	:	A Comparative Study of Machine Learning Models for Rice Price Prediction in Sri Lanka
Group Number	:	02
Group Name	:	Code Zen
Submission Date	:	2025/01/31

Details of the Group Members

Name	Registration ID	Index No.	Signature
B.M.N Gayathri	ITT/2020/030	1306	
C.M.V.S Upananda	ITT/2020/112	1376	
Y.M.G.D.L Yapabandara	ITT/2020/119	1381	
A. Sanjayan	ITT/2020/097	1363	
A.R.D. Nawarathna	ITT/2020/065	1336	

Details of Supervisor(s)

Name	:	Mr. Dhanushka Jayasinghe
Designation	:	Lecturer (Temporary)
Department/ Unit/ Institute	:	Information Communication Technology
Contact Details	:	0713380704

Name	:	Mr. Wiraj Wickramaarachchi
Designation	:	Senior Lecturer
Department/ Unit/ Institute	:	Information Communication Technology
Contact Details	:	0776329957

1. Introduction

1.1. Title of the research project

A Comparative Study of Machine Learning Models for Rice Price Prediction in Sri Lanka

1.2. Purpose and significance of the research study

The purpose of this research is to develop an accurate and data-driven model to predict rice retail prices in Sri Lanka, by integrating weather patterns (rainfall, temperature, radiation), rice yield data, and LKR-USD exchange rates. Rice prices fluctuate due to unpredictable climate conditions, varying production outputs, and economic factors, making it difficult for farmers and traders to plan effectively. By leveraging machine learning, this study aims to provide a reliable forecasting tool that helps stakeholders make informed decisions regarding pricing, storage, and trade.

The significance of this study lies in its potential to stabilize the rice market by reducing uncertainty and improving price forecasting accuracy. Accurate predictions can help farmers optimize their harvest planning, reducing losses from sudden price drops while ensuring fair pricing. Additionally, traders and policymakers can use the insights to manage supply chains efficiently and implement policies that safeguard both farmers and consumers against extreme price volatility.

Furthermore, incorporating exchange rate fluctuations into the prediction model adds an economic dimension, as international trade and import policies impact local rice prices. Understanding these relationships can guide better financial planning for farmers and help authorities develop strategies to mitigate the impact of external market shifts. Ultimately, this research aims to enhance food security, economic stability, and agricultural sustainability in Sri Lanka.

1.3. Problem statement

Rice price fluctuations in Sri Lanka are influenced by multiple factors, including climate conditions, crop yield variations, and foreign exchange rate fluctuations. Farmers and market stakeholders often struggle to predict these price changes, leading to financial instability, inefficient resource allocation, and difficulties in planning harvests and sales. Traditional forecasting methods are insufficient in addressing the complex interplay of these variables, necessitating a data-driven approach to improve price predictions and decision-making.

Lack of Standardized Pricing Mechanisms in Local Markets

Prices vary significantly between rural and urban markets, as well as between wholesale and retail transactions.

Different pricing policies across cooperatives, government-regulated markets, and private traders create inconsistencies in price prediction models.

Limited Availability and Reliability of Historical Data

Weather, yield, and economic data may have gaps, inconsistencies, or inaccuracies due to manual data collection methods.

Many small-scale farmers do not maintain systematic records of past yields and market trends, making model training difficult.

Seasonal and Festival-Based Demand Fluctuations

Rice demand surges during certain periods (e.g., Sinhala and Tamil New Year, Vesak, and other festivals), causing short-term price spikes.

Capturing these periodic variations in a predictive model requires specialized time-series forecasting techniques.

Unpredictable Shocks (Natural Disasters and Pandemics)

Floods, droughts, and extreme weather events can lead to sudden supply shocks, distorting price trends.

Global crises like COVID-19 have demonstrated how unexpected disruptions can impact agricultural supply chains and consumer purchasing behavior.

1.4. Research question(s)

- How do key weather factors, such as rainfall, temperature, and radiation, influence retail rice prices in Sri Lanka?
- What is the relationship between rice yield and retail price fluctuations in Sri Lanka?
- How do exchange rate fluctuations impact retail rice prices in Sri Lanka over time?
- What are the most effective machine learning approaches to enhance the accuracy of retail rice price prediction models?

1.5. Aims and objectives

Aim

To develop a machine learning-based predictive model for forecasting retail price of the rice in Sri Lanka, by integrating rainfall, temperature, radiation, rice yield trends, and LKR-USD exchange rate fluctuations.

Objective

To Analyze the impact of weather on retails price of the rice by examining how rainfall, temperature, and radiation affect production and market pricing in Sri Lanka, identifying seasonal patterns and extreme weather effects.

To Evaluate the relationship between rice yield and price fluctuations by assessing historical production data and market trends to understand supply-demand dynamics.

To Assess the influence of exchange rate variations on retails price of the rice by studying LKR-USD fluctuations and their impact on import costs, exports, and agricultural inputs.

To Optimize the prediction model’s accuracy by fine-tuning hyperparameters, validating with historical data, and evaluating forecasting techniques for reliable decision-making.

1.6. Research methodology, research design, techniques and tools to be adopted

1.6.1 Research Methodology

This study adopts a quantitative, data-driven research methodology, utilizing artificial intelligence (AI) and machine learning (ML) algorithms to predict retails price of the rice in Sri Lanka. The research integrates weather data, rice yield, LKR-USD exchange rate fluctuations, and paddy buying prices from local rice mills. By analyzing these variables, the study aims to identify the most accurate machine learning models for forecasting rice prices.

The methodology consists of the following phases:

1. Data Collection: Collecting historical datasets on weather conditions, rice yield, LKRUSD exchange rates, and paddy buying prices.
2. Data Preprocessing: Cleaning, transforming, and normalizing the data to ensure consistency and accuracy.
3. Feature Engineering: Identifying the most influential features (independent variables) that affect rice prices.
4. Model Development: Comparing and implementing various AI and ML algorithms to forecast rice prices.
5. Model Evaluation & Optimization: Evaluating model performance and optimizing parameters to ensure high prediction accuracy.

1.6.2 Research Design

Phases of Research:

1. Data Collection:
    - Weather Data: Gather temperature, rainfall, and solar radiation data from the Sri Lanka Meteorological Department and Rice Research and Development Institute (RRDI), Bathalagoda.
    - Rice Yield Data: Obtain historical production data from the Department of Agriculture and Sri Lanka Census and Statistics Department, along with data from the Paddy Marketing Board.
    - Exchange Rate Data: Collect monthly LKR-USD exchange rates from the Central Bank of Sri Lanka and financial platforms like Yahoo Finance.
    - Paddy Buying Price Data: Obtain historical paddy buying prices from rice mills such as Rathna Rice Pvt Ltd and the Paddy Marketing Board.
2. Data Preprocessing:
    - Handling Missing Data: Impute missing values using mean imputation or interpolation for time-series data.
    - Outlier Detection: Detect and remove outliers using the Interquartile Range (IQR) method.
    - Data Normalization: Normalize datasets (e.g., using Min-Max Scaling or Zscore Normalization) to ensure all variables are on a comparable scale.
    - Time-Series Transformation: Aggregate data into monthly, seasonal, or yearly time-series formats for better analysis.
3. Feature Engineering:
    - Identify key features like seasonal weather patterns, annual rice yields, and exchange rate fluctuations.
    - Derive additional variables reflecting the relationships between these features and their impact on rice prices.

4. Model Selection:
    - Implement multiple AI and ML models for rice price forecasting:
        - Random Forest Regression (RF)
        - XGBoost (Extreme Gradient Boosting)
        - ARIMA (AutoRegressive Integrated Moving Average)
        - Long Short-Term Memory Networks (LSTM)
5. Evaluation & Optimization:
    - Use evaluation metrics like Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R^2 Score to assess the models' prediction accuracy.
    - Cross-validation and hyperparameter tuning will be applied to optimize the models and prevent overfitting.

1.6.3 Techniques to be Adopted

1. Data Preprocessing Techniques:
    - Handling Missing Data: Impute missing values using mean imputation or interpolation for monthly or seasonal time-series data.
    - Outlier Detection: Use the Interquartile Range (IQR) method to remove extreme values from the data.
    - Feature Scaling & Normalization: Normalize data using techniques like MinMax Scaling or Z-score normalization to bring all variables to a comparable scale.
    - Time-Series Transformation: Convert the data into a time-series format, aggregating by month or season to capture trends and cyclic patterns.
2. Machine Learning Algorithms:
    - Random Forest Regression (RF):
        - A powerful ensemble learning method that builds multiple decision trees and combines their predictions.
        - Effective for capturing nonlinear relationships between features and target variables, such as weather, yield, and price.
        - It can handle feature importance analysis and is robust to overfitting.
    - XGBoost (Extreme Gradient Boosting):
        - A popular gradient boosting algorithm that uses decision trees to make predictions.
        - It works by iteratively correcting the errors of previous models, making it highly accurate and efficient.
        - Suitable for datasets with multiple features and nonlinear relationships, like weather data and exchange rates.

- ARIMA (AutoRegressive Integrated Moving Average):
    - A classical time-series forecasting model that analyzes historical price trends and seasonal patterns to make future predictions.
    - It works well when data exhibits linear patterns with little noise, such as monthly rice prices.
    - The ARIMA model uses lagged values to predict future prices, making it useful for simple price trend forecasting.
- Long Short-Term Memory Networks (LSTM):
    - A type of recurrent neural network (RNN) designed for sequential data analysis, which excels in handling time-series data with long-term dependencies.
    - LSTMs are capable of learning complex, nonlinear relationships from sequences of data, such as seasonal weather variations and their longterm effect on rice prices.
    - LSTM is suitable for this study as it can handle long-term dependencies in sequential data, such as past yield and exchange rate trends affecting future rice prices.

3. Evaluation Techniques:
    - Root Mean Squared Error (RMSE): Measures the average magnitude of errors in predictions. It penalizes larger errors more significantly than smaller ones.
    - Mean Absolute Error (MAE): A simpler evaluation metric that measures the average magnitude of errors without squaring them.
    - R^2 Score: Measures how well the model explains the variability in the target variable (rice price). A higher R^2 indicates better model fit.
    - Cross-validation: Ensures that the model generalizes well to unseen data by splitting the dataset into multiple subsets and training the model on different subsets.
4. Optimization Techniques:
    - Hyperparameter Tuning: Use GridSearchCV or RandomizedSearchCV to find the optimal hyperparameters for the machine learning models.
    - Cross-validation: This will help ensure that the model is robust and does not overfit to any particular subset of the data.

1.6.4 Tools for Implementation

- Programming & Data Processing:
    - Python Libraries:
        - NumPy & Pandas: For data manipulation and handling.
        - Scikit-learn: For implementing machine learning models and evaluation metrics.
        - TensorFlow/Keras: For developing deep learning models like LSTM.
        - Statsmodels: For traditional statistical models like ARIMA.
    - Matplotlib & Seaborn: For data visualization and result interpretation.
- Data Collection & Management:
    - Sri Lanka Meteorological Department and RRDI-Bathalagoda for weather data.
    - Department of Agriculture and Paddy Marketing Board for rice yield and paddy buying price data.
    - Central Bank of Sri Lanka and Yahoo Finance for exchange rate data.

1.6.5 Implementation Workflow

1. Preprocessing the Data:
    - Clean and preprocess the raw data, handling missing values and removing outliers.
    - Normalize the data for input into machine learning models.
2. Model Selection and Training:

    o Test multiple machine learning algorithms (RF, XGBoost, ARIMA, LSTM) to determine which produces the most accurate predictions. o Train models on historical data and evaluate their performance using standard metrics.
3. Evaluating and Fine-Tuning:
    - Evaluate model performance using RMSE, MAE, and R^2 Score.
    - Fine-tune models by optimizing hyperparameters using GridSearchCV or RandomizedSearchCV.

1.7. Expected research results and/or innovations

1. Weather-Impact Analysis:

The study will provide insights into how key weather parameters (rainfall, temperature, and solar radiation) influence rice production and market prices in Sri Lanka. This knowledge will enable farmers and traders to anticipate price fluctuations due to seasonal and extreme weather conditions.

2. Yield-Price Dynamics Insight:

The study is expected to uncover strong correlations between rice yield variations and price fluctuations. The results will offer valuable data on how yield-based supply-demand imbalances affect local rice prices, helping farmers make informed decisions on production strategies and storage.

3. Understanding of Exchange Rate Influence:

By analyzing the LKR-USD fluctuations, the research will demonstrate how currency devaluation and appreciation impact local rice prices through changes in import costs and agricultural input prices. This will give policymakers and farmers a better understanding of external economic factors affecting domestic rice pricing.

4. Innovative Decision Support Tool:

The outcome of this study will be the development of a web-based decision support tool that integrates the predictive model. This platform will provide stakeholders with real-time price forecasts, allowing them to make informed decisions regarding the harvesting, storing, and selling of rice. The tool will be user-friendly, accessible online, and designed to streamline decision-making processes.

5. Enhanced Forecasting Techniques:

Through experimentation with various machine learning models and optimization techniques, the research will introduce innovative ways to fine-tune prediction models. By validating with historical data and optimizing hyperparameters, the study will enhance the application of machine learning in agricultural retail price prediction.

1.8. Research schedule / Work Plan

Research Task	Intended Starting Date	Intended Ending Date
Finalize Research Topic	2-December-2024	31-December-2024
Literature Review	1-January-2025	17-January-2025
Research Proposal Writing and Submission	18-January-2025	31-January-2025
Proposal Presentation	5-February-2025	7-January-2025
1st Chapter Submission (Introduction)	6-February-2025	8-February-2025
Writing Literature Review	18-January-2025	13-February-2025
Working on Methodology	14-February-2025	21-June-2025
2nd Chapter submission (Literature Review)	11-February-2025	14-February-2025
Literature Review Presentation	19-February-2025	21-February-2025
Initial Implementation for Experiments	22-February-2025	21-June-2025
Writing Methodology Chapter	1-March-2025	7-March-2025
3rd Chapter Submission (Methodology)	4-March-2025	7-March-2025
Methodology Presentation	9-April-2025	11-April-2025
Progress Presentation 1	18-June-2025	20-June-2025
Working on Thesis	21-June-2025	12-September-2025
Analysis if Results	1-August-2025	12-August-2025
Progress Presentation 2	13-August-2025	15-August-2025
Writing Draft Thesis and Research Paper	16-August-2025	12-September-2025
Final Presentation Demonstrations	1-October-2025	3-October-2025
Final Thesis Submission	4-October-2025	17-October-2025

2. Report of preliminary review of literatures

Accurate yield prediction and profitability optimization are essential for sustainable agriculture. Several studies have explored the application of artificial intelligence (AI) and machine learning (ML) in improving agricultural forecasting, crop selection, and decision-making processes.

The study "Rice Price Forecasting Using the ARIMA Model" examines the factors influencing rice price fluctuations in Eastern Pangasinan, Philippines, using timeseries data from the Philippine Statistics Authority and the ARIMA model for forecasting. The literature reviewed highlights key determinants such as agricultural inputs (fertilizers, fuel, and machinery), climate conditions, supply chain disruptions, market competition, and government policies. Studies indicate that rice prices are projected to rise steadily from 2024 to 2028, with regular well-milled rice increasing by 5.68% annually and special rice by 1.80% to 1.71% over four years. The findings suggest that rising production costs and supply instability contribute to price surges, outpacing the estimated 2% annual increase in consumer salaries, which may strain household economies. Further research comparing ARIMA with alternative forecasting models is recommended to enhance prediction accuracy and inform policy interventions.[1]

Another study explores the impact of the COVID-19 lockdown on rice prices in India using machine learning-based time series forecasting models. The researchers applied ARIMA, ANN, and ELM models to predict price fluctuations caused by supply chain disruptions. Their findings indicate that the ARIMA intervention model estimated a price increase of INR 0.92/kg during the lockdown, while the ELM model outperformed other approaches due to its ability to capture nonlinear patterns in price movements. The study highlights the effectiveness of machine learning models, particularly ELM, in agricultural price forecasting under crisis scenarios. This research underscores the need for adaptive forecasting models that can assist policymakers in mitigating the effects of unforeseen market disruptions.[2]

Another investigation analyzed the use of Artificial Neural Networks (ANNs) for forecasting paddy yield in Sri Lanka. This study focused on climate factors such as temperature, rainfall, and humidity to predict seasonal rice production. The research compared three gradient descent optimization techniques—Batch Gradient Descent (BGD), Stochastic Gradient Descent (SGD), and Mini-batch Gradient Descent (MBGD)—in training ANN models. The findings revealed that BGD outperformed the other methods in terms of accuracy, proving its efficiency for paddy yield forecasting. The study emphasized the limitations of traditional time-series models, such as ARIMA, in handling nonlinear relationships in agricultural data.[6]

Similarly, another study explored the integration of feature-engineered machine learning models to enhance rice yield prediction. This research compared Linear Regression (LR), Support Vector Machine (SVM), k-Nearest Neighbors (KNN), and Random Forest (RF) models using climate data, including rainfall, temperature, and solar radiation. The results demonstrated that Random Forest Regression, with a reduced number of features, achieved similar predictive accuracy to models using a complete feature set, indicating the effectiveness of feature selection in optimizing model performance. This study contributed to food sustainability efforts by improving predictive accuracy, aligning with the United Nations Sustainable Development Goals (SDGs).[7]

In contrast to yield prediction, other research has focused on optimizing crop profitability using ML techniques. A proposed decision-support system for farmers aimed to select the most profitable crops based on historical market trends, production costs, and climatic conditions. By utilizing deep learning approaches such as Long Short-Term Memory (LSTM) networks and Time Series Analysis, the study identified patterns in market demand and supply fluctuations. A key finding was that integrating economic factors into crop selection could help stabilize market prices and maximize farmers' revenue. This study highlighted the role of ML in strategic agricultural planning and market-driven decision-making.[8]

Other research highlights critical findings related to the impact of economic indicators on rice production in Sri Lanka, focusing on machine learning applications. Key economic factors such as GDP, inflation rate, population growth, and imports have shown significant correlations with rice production outcomes. Models like Support Vector Machines (SVM), Linear Regression, and Gaussian Process Regression (GPR) have demonstrated improved prediction accuracy when compared to traditional methods. Further studies stress the importance of including environmental factors, such as climate change, in forecasting models. However, a gap exists in the research regarding the integration of fuzzy logic or hybrid models like Adaptive Neuro-Fuzzy Inference Systems (ANFIS), which may more effectively handle complex, non-linear relationships between economic and agricultural variables. This gap is relevant to research on rice price prediction in Anuradhapura, where machine learning models are also being applied to forecast rice prices using weather data, yield trends, and exchange rate fluctuations. By considering these additional factors, such research could contribute to refining predictive models and providing more accurate forecasts, benefiting stakeholders like farmers, traders, and policymakers.[9]

Agricultural decision support systems play a crucial role in improving forecasting accuracy and aiding stakeholders in making informed decisions. One such system introduced an approach to predicting rice prices, rice production, and minimum rice requirements using a centralized data warehouse. The system utilized time series analysis, particularly the ARIMA model, to analyze historical trends and predict future market behavior. Similar studies have highlighted the influence of economic factors such as GDP, population growth, and inflation on rice price fluctuations, emphasizing the need for machine learning techniques to improve prediction models. While traditional statistical methods like ARIMA provide valuable insights, machine learning models such as Support Vector Machines (SVM), Random Forest, and neural networks have demonstrated superior predictive accuracy in handling complex, non-linear relationships in agricultural data. Moreover, integrating climatic variables like rainfall, temperature, and solar radiation into predictive models has been found to enhance forecasting reliability. Despite these advancements, a significant research gap exists in the adoption of hybrid approaches, such as Adaptive Neuro-Fuzzy Inference Systems (ANFIS), which could further refine predictive accuracy by capturing intricate interactions between economic, environmental, and production variables. Addressing this gap, recent studies have focused on applying machine learning models for rice price prediction in the Anuradhapura District, incorporating economic, climatic, and yield data to develop a robust forecasting framework. By leveraging advanced predictive techniques, this research aims to support farmers, traders, and policymakers in making data-driven decisions to stabilize rice markets and ensure food security.[10]

Several studies have explored rice price forecasting using econometric and statistical models, highlighting the influence of economic and environmental factors on price volatility. Some research has demonstrated that international rice prices, crude oil prices, and exchange rate fluctuations significantly impact domestic rice prices in Sri Lanka, using regression models and time series analysis. Other studies have emphasized the role of climatic factors such as rainfall and temperature variations in affecting rice yields and subsequent market prices. However, traditional forecasting methods, such as ARIMA and basic regression models, have limitations in capturing complex, nonlinear relationships in price prediction. Recent advancements in machine learning, particularly in models like Random Forest, XGBoost, and Long Short-Term Memory (LSTM) networks, offer improved accuracy by handling high-dimensional and time-series data more effectively. Despite their success in global agricultural markets, these advanced AI-driven techniques remain underexplored in the Sri Lankan context, particularly in localized settings such as Anuradhapura. Addressing this research gap, new studies integrate weather patterns, rice yield trends, and exchange rate fluctuations into machine learning models to develop a more precise and dynamic rice price forecasting system, providing valuable insights for farmers, traders, and policymakers.[11]

This paper highlights the growing application of machine learning (ML) models in predicting crop prices, a crucial aspect of agricultural sustainability. Various studies emphasize the effectiveness of regression-based models, such as Decision Tree Regression, Linear Regression, and Random Forest Regression, in forecasting price fluctuations. Research demonstrates the high accuracy of decision trees and XGBoost in price prediction, while Ridge and Lasso regression have been used to address multicollinearity issues in datasets. Additionally, time-series models like ARIMA and LSTM have been explored for capturing trends and seasonality in price data. Advanced neural networks, including General Regression Neural Networks (GRNN), have also shown promise in crop yield prediction. However, challenges such as data availability, seasonal variations, and external market influences remain significant barriers to accuracy. The literature suggests that integrating multimodal data sources, including satellite imagery and weather trends, could further enhance predictive performance. Overall, ML-based price forecasting models contribute to informed decision-making for farmers, policymakers, and traders, supporting agricultural market stability and economic growth.[15]

The study on "Machine Learning Modelling of the Relationship between Weather and Paddy Yield in Sri Lanka" focuses on the application of machine learning techniques to predict paddy yield based on climatic factors. Prior research has established that weather conditions, including temperature, humidity, and rainfall, significantly influence agricultural productivity. Various statistical and machine learning models have been employed in different regions to analyze these relationships, with Random Forest (RF) emerging as a particularly effective tool due to its robustness in handling nonlinear interactions. Several studies have applied regression models, including Multiple Linear Regression (MLR) and Power Regression (PR), to estimate crop yields, often incorporating climate variables such as temperature, humidity, and precipitation. The literature highlights that while traditional regression techniques provide useful insights, machine learning models, especially RF, offer superior predictive accuracy. The study builds on this foundation by developing a machine learning-based crop-weather model for Sri Lanka, emphasizing the importance of weather indices in predicting paddy yield and comparing different modeling approaches for optimal performance.[13]

Research on statistical models for identifying climate contributions to crop yields highlights three primary approaches: time-series models, cross-section models, and panel models. These models have been extensively used to analyze the relationship between climate variables and crop yields at different spatial scales, from site-specific to global levels. The review discusses the strengths and limitations of statistical models compared to process-based crop models, emphasizing their effectiveness in capturing climate impacts despite challenges like collinearity among climate variables and non-climatic trend removal. Additionally, the study underscores the need for improved data accuracy, better integration of adaptation strategies, and the combination of statistical and crop models to enhance predictive capabilities in agricultural climate studies.[14]

The research addresses the significant gap in applying machine learning models for rice price prediction. Despite extensive studies on rice production forecasting, most existing research has primarily focused on univariate time-series forecasting using ARIMA and Artificial Neural Networks (ANN). While ARIMA has proven effective for short-term predictions, there is a lack of comparative analyses involving advanced machine learning models such as Random Forest, XGBoost, and LSTM. Furthermore, much of the existing research emphasizes rice production trends rather than direct price forecasting, limiting its practical application in real-time economic decision-making. Another critical shortcoming is the absence of integrated models that incorporate multiple influencing factors, including weather conditions, rice yield variations, and exchange rate fluctuations. To address this gap, recent studies introduce a multi-factor, machine-learning-based approach to predict rice prices, leveraging diverse data sources to enhance forecasting accuracy. This research aims to provide a comprehensive and data-driven predictive framework, supporting farmers, traders, and policymakers in making informed market decisions.[12]

A comprehensive review of machine learning applications in agricultural price prediction focuses on algorithm selection, research paradigms, and performance metrics. The study reveals that neural networks, especially ANN and SVM, are the most widely used models for predicting agricultural product prices due to their superior accuracy in capturing complex price patterns. The research also emphasizes that most studies adopt a positivist paradigm with quantitative and longitudinal approaches. Additionally, evaluation metrics such as RMSE and MAPE are commonly used to assess model performance. This study provides valuable insights into the evolution of machine learning techniques in agricultural economics and supports the use of AI-based models for enhancing price prediction accuracy.[3]

Other studies employed the ARIMA model to analyze historical price data in Eastern Pangasinan, Philippines, identifying key determinants such as fuel, fertilizer, machinery, and climate conditions. The findings projected a steady rise in rice prices from 2024 to 2028, surpassing the estimated annual salary increase, indicating economic strain on consumers. Similarly, research comparing Linear Regression and Random Forest algorithms to predict premium rice prices in West Java, Indonesia, revealed that Random Forest outperformed Linear Regression, achieving higher accuracy (98.69%), and emphasized factors such as production levels, government policies, and market competition in price variations. These studies underscore the importance of data-driven forecasting for policymakers and stakeholders to mitigate market volatility and enhance food security.[4]

Recent studies on rice price forecasting have employed various statistical and machine learning techniques to improve prediction accuracy and support market stability. Research using the ARIMA model to analyze historical price trends in the Philippines highlights economic and environmental factors influencing rice prices, such as agricultural inputs, climate, and government policies. Comparisons between Linear Regression and Random Forest models for premium rice price prediction in West Java, Indonesia, found that Random Forest achieved superior accuracy (98.69%) in forecasting. Meanwhile, a Multivariate Fuzzy Time Series (MFTS) approach using Fuzzy Vector Autoregressive (FVAR) models has been introduced to predict rice prices and production in Malaysia, demonstrating that Trapezoidal Fuzzy Numbers (TrFNs) improved predictive accuracy over conventional VAR models. These studies underscore the importance of advanced forecasting models in mitigating rice market volatility, enhancing food security, and guiding policymaking. [5]

2.1. References

[1] B. Mark *et al.*, "Rice Price Forecasting Using the Arima Model," *IRE Journals* |, vol. 7, 2024.

[2] S. Rathod, G. Chitikela, N. Bandumula, G. Ondrasek, S. Ravichandran, and R. M. Sundaram, "Modeling and Forecasting of Rice Prices in India during the COVID-19 Lockdown Using Machine Learning Approaches," *Agronomy*, vol. 12, no. 9, p. 2133, Sep. 2022.

[3] "(PDF) Machine Learning for Price Prediction for Agricultural Products," *ResearchGate*.

[4] Irfan Rasyid Muchtar and Afiyati Afiyati, "Comparison of Linear Regression and Random Forest Algorithms for Premium Rice Price Prediction (Case Study: West Java)," *Jurnal Indonesia Sosial Teknologi*, vol. 5, no. 7, pp. 3122–3132, Jul. 2024.

[5] M. Bilal, M. A. Alrasheedi, Muhammad Aamir, S. Abdullah, Siti Mariam Norrulashikin, and Reza Rezaiy, "Enhanced forecasting of rice price and production in Malaysia using novel multivariate fuzzy time series models," *Scientific Reports*, vol. 14, no. 1, Dec. 2024.

[6] E. J. K. P. Nandani and T. T. S. Vidanapathirana, "Forecasting Paddy Yield in Sri Lanka Using Back-propagation Learning in Artificial Neural Network Model," *Journal of the University of Ruhuna*, vol. 12, no. 2, pp. 110–120, Dec. 2024.

[7] Aminda Amarasinghe *et al.*, "Advancing food sustainability: a case study on improving rice yield prediction in Sri Lanka using weather-based, feature-engineered machine learning models," *Deleted Journal*, vol. 6, no. 11, Nov. 2024.

[8] A. Alagalla and L. Weerasinghe, "Best Profitable Crops Prediction with Profit, Cost, and Farmland Optimization using Machine Learning," *International Journal of Computer Information Systems and Industrial Management Applications*, vol. 15, pp. 550–565, 2023, Accessed: Jan. 31, 2025. [Online].

[9] Sherin Kularathne, Namal Rathnayake, M. Herath, Upaka Rathnayake, and Y. Hoshino, "Impact of economic indicators on rice production: A machine learning approach in Sri Lanka," *PLoS ONE*, vol. 19, no. 6, pp. e0303883–e0303883, Jun. 2024.

[10] W. D. R. Somawardhana, G. A. Deegala, Fernando, G. C. Dassanayake, and A. Perera, "Vee-Bissa: The agricultural decision support system," *2013 IEEE 8th International Conference on Industrial and Information Systems (ICIIS)*, pp. 449–453, Dec. 2013.

[11] C. Hathurusingha, N. Abdelhamid, and D. Airehrour, "Forecasting Models Based on Data Analytics for Predicting Rice Price Volatility: A Case Study of the Sri Lankan Rice Market," *Journal of Information & Knowledge Management*, vol. 18, no. 01, p. 1950006, Mar. 2019.

[12] A. Sultana and M. Khanam, "Forecasting Rice Production of Bangladesh Using ARIMA and Artificial Neural Network Models," *Dhaka University Journal of Science*, vol. 68, no. 2, pp. 143–147, Oct. 2020.

[13] P. Ekanayake, W. Rankothge, R. Weliwatta, and J. W. Jayasinghe, "Machine Learning Modelling of the Relationship between Weather and Paddy Yield in Sri Lanka," *Journal of Mathematics*, vol. 2021, p. e9941899, May 2021.

[14] W. Shi, F. Tao, and Z. Zhang, "A review on statistical models for identifying climate contributions to crop yields," *Journal of Geographical Sciences*, vol. 23, no. 3, pp. 567–576, Apr. 2013.

[15] I. Mahmud, P. R. Das, M. H. Rahman, A. R. Hasan, K. I. Shahin, and D. M. Farid, "Predicting Crop Prices using Machine Learning Algorithms for Sustainable Agriculture," *2017 IEEE Region 10 Symposium (TENSYMP)*, pp. 1–6, Sep. 2024.

3. **Recommendation of supervisor(s) on the research problem and research proposal** ***(This section should be filled by the supervisor(s). Supervisor(s) may consider the adequacy and scope of the research problem, quality and adequacy of the reviewed literature, methodology proposed, and the schedule).***

   Comments (if any):

   I certify that, the student engaged continuously with me in developing the proposal and, I am confident that he is adequately competent to defend this proposal.

   Signature(s) of Supervisor(s):

   Date: 01/02/2025

4. **Research proposal defense assessment team** ***(this section should be filled by the department)***

Date of research proposal

defense:

Panel members	Name	Department / Institute
Chair		
Member		
Member		
Member		
Member		

5. **Comments of the assessment team on the research proposal** ***(This should be filled by the chair of the assessment panel. <u>In case of revision or fail, needed revision in the proposal or reasons to fail the proposal should be mentioned here</u>)***

Result of the research proposal	**Excellent / Good / Pass with revisions / Fail**
Score	
Signature of the panel chair	
Date	