e-ISSN: 2582-7464

Computational Intelligence and Machine Learning
Volume 6, Issue 1, April 2025

The Analysis of Crop Production Data for Predicting Future Trends of Highland Vegetables in Benguet

Anna Rhodora M. Quitaleg[1*], Dr. Menchita F. Dumlao[2]

[1] Student, Angeles University Foundation, Angeles City, Philippines
[2] Professor, Angeles University Foundation, Angeles City, Philippines
[*]Corresponding Author Email: quitaleg.annarhodora@auf.edu.ph

Abstract

This study employs machine learning techniques to enhance agricultural prediction to secure future food supplies and support sustainable practices. It aims to forecast trends in highland vegetable crops in Benguet providing invaluable insights for enhancing yields and contributing to vital food security initiatives. Notably, the researchers concentrated on crop production data due to inconsistencies and incompleteness in historical datasets collated. Furthermore, it proposes a mobile platform for data collection from farmers, capturing their production and farming practices. This method intends to streamline the data collection process, enabling LGUs to effectively collate and utilize the information. The methodology employed a structured approach, encompassing data collection, preparation, analysis, modeling, and evaluation using Python programming and Power BI tools. In the analytical phase, a variety of machine learning models were explored, including Linear Regression, Lasso Regression, Ridge Regression, Decision Trees, SVM, and Random Forest. Model evaluation was based on the accuracy, MAE, MSE, and RMSE. The Random Forest model emerged as the most suitable choice, boasting the best metrics for production purposes with an accuracy of 98.94%. The outcomes of this study hold significant potential, not only in reshaping agricultural practices and decision-making but also in fostering sustainable approaches to highland vegetable cultivation.

Keywords

Crop Production, Data Mining, Machine Learning, Predictive Analytics, Sustainable Agriculture.

INTRODUCTION

Agriculture serves as the backbone of numerous economies around the globe, and it is especially pivotal in regions where highland vegetables are a significant source of income and food security[1][2][3]. Benguet, often referred to as the "Salad Bowl of the Philippines," is one such region where the cultivation of highland vegetables plays an essential role in both the local and national economy[1][4]. Understanding the dynamics of crop production in Benguet is crucial not only for planning and decision-making at the farmer level but also for policy formulation aimed at sustainable agriculture.

Predictive analytics offers a powerful tool [5] for understanding future trends in crop production, thereby facilitating better planning and more effective resource allocation. Machine learning algorithms, in particular, have shown great promise in capturing the complexities and nuances of agricultural data [6]. However, accurate prediction requires the incorporation of various factors such as climate, soil conditions, and farming practices, which are often region-specific.

Predictive analytics is a term mainly used in statistical and analytics techniques. This term is drawn from statistics, machine learning, database techniques, and optimization techniques. It has roots in classical statistics. Predictive analytics predicts the future by analyzing current and historical data. Future events and behavior of variables can be predicted using the models of predictive analytics [7].

In the paper "Predictive Analytics in Business Analytics: Decision Tree" [8], the authors conducted a systematic literature review to explore the role of predictive analytics and decision trees in business decision-making. The study, emphasizes the growing importance of predictive analytics for forecasting future trends and enhancing corporate performance. The decision tree methodology is highlighted as a practical tool in this domain. The findings offer valuable insights for both researchers and practitioners in business analytics.

In recent years, data mining and machine learning have emerged as powerful techniques for gaining insights from large, complex datasets. Data mining involves the extraction of patterns, information, and knowledge from large volumes of data. It serves as the foundation upon which machine learning algorithms operate. Machine learning, a subset of artificial intelligence (AI), enables systems to learn from data, identify patterns, and make decisions with minimal human intervention[9][10]. Together, these technologies have been successfully applied in various fields, including healthcare, finance, and now, increasingly, in agriculture [6].

Benguet, the leading producer of highland vegetables in the Philippines, plays a vital role in national food security and the local economy. However, inconsistencies in historical agricultural data and reliance on manual, fragmented data collection methods hinder accurate crop production forecasting. This lack of reliable, data-driven insights affects farmers, policymakers, and LGUs, leading to inefficient resource allocation, unpredictable yields, and economic losses. Without a structured forecasting system, farmers face uncertainties in crop planning, increasing the risks of overproduction and shortages.

To address this issue, this study aims to develop a machine learning-based predictive model for highland vegetable production forecasting, utilizing historical agricultural data from LGUs, the Department of Agriculture (DA), and farmers' records. Additionally, the study proposes a mobile-based data collection platform to streamline and standardize the recording of production data, ensuring more accurate and accessible agricultural information. By applying data-driven forecasting techniques, this research seeks to enhance agricultural planning, reduce production uncertainties, and support the sustainable growth of Benguet's highland farming sector.

In addition, in the alignment with Republic Act No. 11293 [11], also known as the Philippine Innovation Act , this study underscores the necessity for innovation within the agricultural sector. The Act, which has also been cited in the paper entitled 'A Decision Support System for Benguet's Upland Vegetable Crop Prediction using Machine Learning Techniques' [1], aims to foster a culture of innovation throughout the Philippines. It focuses on the development and implementation of sustainable and inclusive solutions across diverse domains.

By employing machine learning algorithms for predictive analytics, this research directly supports the objectives set forth in the Act, facilitating innovative methodologies with the potential to transform traditional farming practices in Benguet. The study thus represents a proactive step towards making agricultural systems not only more efficient but also more resilient, echoing the legislation's emphasis on sustainable and inclusive innovation.

In the agricultural context, machine learning can facilitate the automatic detection of trends and anomalies, provide real-time advice to farmers, and even predict future events based on existing data. These capabilities are particularly relevant for areas like Benguet, where agriculture is a critical part of the local economy but is subject to various external pressures like market demand, climate change, and societal needs.

This paper aims to employ machine learning algorithms to analyze crop production data in Benguet. By integrating factors like Municipality, Farm Type, Month, and Crop Type, the study endeavors to make robust predictions about future crop yields. Through a meticulous study of past and present data, this research contributes to creating a Decision Support System (DSS) that could aid farmers, agricultural organizations, and policymakers in making informed choices.

The research takes a human-centered approach, guided by the principle of empathy towards the farmers in Benguet, especially considering the significant challenges they faced during the pandemic. Many farmers had to discard their produce, turning what should have been valuable food sources into agricultural waste. This study aims not only to predict future trends but also to provide actionable insights that could alleviate such issues. By developing data-driven predictive models, this research aims to assist farmers and other stakeholders in making informed decisions to mitigate such losses and adapt to market trends.

Furthermore, the study aligns with the United Nations' Sustainable Development Goals (SDGs), specifically targeting Goals 9 (Industry, Innovation, and Infrastructure), 11 (Sustainable Cities and Communities), and 12 (Responsible Consumption and Production). By focusing on these areas, the study aims to contribute to a more resilient and sustainable agricultural system that empowers local communities while addressing global challenges.

RESEARCH METHODOLOGY

Research Design

This study employs a quantitative research design, integrating descriptive and predictive analytics to analyze historical crop production data and forecast future trends in Benguet's highland vegetable industry. Using machine learning techniques, it identifies key patterns affecting vegetable yields through a longitudinal approach, capturing seasonal variations over multiple years. Data is sourced from local government units (LGUs), the Department of Agriculture (DA), and farmers to ensure accuracy and comprehensiveness. A mobile platform will also be proposed to enhance real-time data collection and accessibility

Research Methods

This study employs data mining techniques to analyze historical crop production data, ensuring accurate forecasting of highland vegetable yields in Benguet. The process follows key data mining stages, including data collection from various sources (LGUs, DA, and farmers), data analysis through preprocessing techniques like cleaning and normalization, and data modeling using machine learning algorithms such as Random Forest, Decision Trees, and Regression models to identify key patterns and trends. Model evaluation is conducted to assess predictive performance and ensure accuracy. Additionally, Power BI is utilized for data visualization, converting raw data into interactive dashboards for informed decision-making. These procedures enable a comprehensive, data-driven approach to agricultural forecasting.

The Study Area

The study area includes the various topography of the Philippine province of Benguet, which is part of the Cordillera Administrative Region. There are thirteen (13) municipalities in this area, each with distinctive topographical, climatic, and agricultural characteristics. The study's core is comprised of these municipalities, which provide essential data sources for analyzing crop trends and forecasting future agricultural results using machine-learning techniques.

Fig. 1 displays the study area's map, highlighting the geographical distribution of the 13 municipalities in Benguet: Atok, Bakun, Bokod, Buguias, Itogon, Kabayan, Kapangan, Kibungan, La Trinidad, Mankayan, Sablan, Tuba, and Tublay.

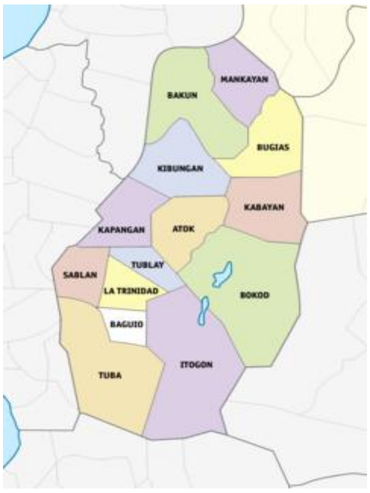

Figure 1. Benguet - The study area

The Dataset

The dataset used in this study comprises essential variables that play a crucial role in predicting crop yields in Benguet. These variables have been carefully curated and collected to capture key aspects of agricultural practices and climatic conditions. Table 1 shows a brief description of the main variables in the dataset, highlighting their names, data types, and units of measurement.

Table 1: Crop Production Dataset Features

Feature	Data Type	Description
Municipality	Categorical	Name of the municipality within the province of Benguet
Farm Type	Categorical	Type of farm, categorizing agricultural practices.
Year	Numerical	The year when the crop production data was recorded.
Month	Categorical	The month when the crop production data was recorded.
Crop	Categorical	Name of the high-value crop being cultivated.
Area Planted	Numerical	The total area of land where the crop was planted. Hectares (ha)
Area Harvested	Numerical	The total area of land where the crop was harvested. Hectares (ha)
Production	Numerical	The total crop production quantity. Metric Tons (mt)
Productivity	Numerical	The crop productivity, indicating the yield per unit area of land Yield per unit area (mt/ha)

Data Mining Procedures

In the context of this study, data mining serves as the foundation for unraveling the complexities of crop trends and yields in the highland vegetable farms of Benguet. The processes [12]taken to collect, prepare, analyze, and model the agricultural datasets are covered in this section, which goes deeply into the exhaustive data mining process.

Data Collection: The research employed a blend of in-person and virtual interviews for gathering data, partnering mainly with the Office of the Provincial Agriculture (OPAG) and the Department of Agriculture (DA). Initial consolidated datasets had limitations due to the few available instances for analysis. To improve predictive power, data was specifically requested on a monthly basis, enhancing granularity and aligning with crop seasonal trends. While the monthly data capture improved the study's scope, issues like data inconsistency and crop variety posed challenges. Nevertheless, stringent data verification measures were implemented to ensure accuracy and analysis quality.

Data Preparation: Data preparation involved a series of steps for data cleaning and processing, and correlation analysis to ensure reliable insights for analysis and modeling.

Data Cleaning and Preprocessing: Data for the study was collated from multiple workbooks and worksheets, representing different years, municipalities, and farm types within Benguet. A focused approach was adopted, narrowing the commodities to the top 10 high-value vegetables in the region. Various preprocessing steps were conducted to address data inconsistencies, missing values, and outliers. Techniques like normalization and feature scaling ensured uniformity across attributes. The dataset was also categorized by farm types—"irrigated" and "rain-fed"—to gain deeper insights into how different agricultural practices influence crop trends and yields in Benguet. This refined dataset served as the foundation for subsequent analyses and insights.

Correlation Analysis: Feature selection was crucial in this study to streamline model inputs and remove noise. A correlation analysis was conducted to understand the interplay among dataset features[13] [14].

The heatmap provides a visual representation of the Pearson correlation coefficients between the various features in the dataset. Our aim is to pinpoint features that demonstrate substantial interrelationships. In the Pearson correlation scale, coefficients can range from 0 to 1: a coefficient of 0 suggests no correlation, while a coefficient of 1 signifies a perfect positive correlation between the features[15].

The dataset revealed a strong correlation of 0.97 between "Area Harvested" and "Production," indicating that higher harvested areas typically yield greater production. Features selected for the study, including "Area Harvested," "Area Planted, " and "Production," each showed a minimum correlation of 0.5, making them valuable for analysis and model building.

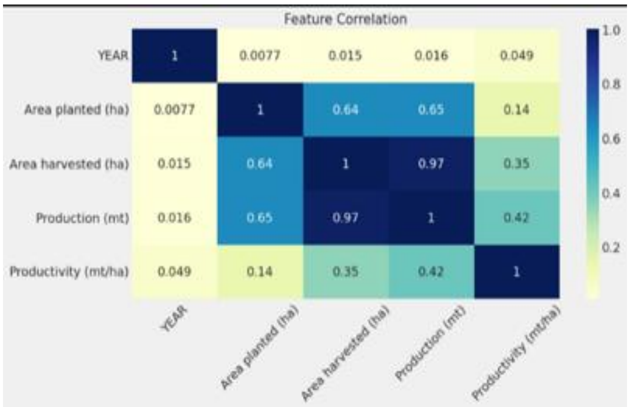

Figure 2. Heatmap

Data Analysis: In this study, key features like Municipality, Farm Type, Month, and Crop Type were scrutinized for inclusion in machine learning models. To ensure algorithmic accuracy and eliminate bias, these categorical variables were encoded into binary format as shown in fig. 3. The code used transformed all specified features into binary identifiers for unbiased model training. Key variables such as farm location, crop types, and harvest timing were provided for predictive modeling. The model aimed to predict 'Production,' representing the estimated crop yield.

	ATOK	BAKUN	BOKOD	BUGUIAS	ITOGON	KABAYAN	KAPANGAN	KIBUNGAN	LA TRINIDAD	MANKAYAN	SABLAN	TUBA	TUBLAY	IRRIGATED	RAINFED
0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0
1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0
2	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0
3	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0
4	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
24723	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
24724	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
24725	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
24726	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
24727	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1

BROCCOLI	CABBAGE	CARROTS	CAULIFLOWER	CHINESE CABBAGE	GARDEN PEAS	LETTUCE	SNAP BEANS	SWEET PEPPER	WHITE POTATO	Area planted (ha)	Area harvested (ha)
0	1	0	0	0	0	0	0	0	0	98.0	120.00
0	1	0	0	0	0	0	0	0	0	115.0	76.00
0	1	0	0	0	0	0	0	0	0	82.0	51.19
0	1	0	0	0	0	0	0	0	0	0.0	98.00
0	1	0	0	0	0	0	0	0	0	54.0	115.00
...	...	...	...	...	...	...	...	...	...	...	...
0	0	1	0	0	0	0	0	0	0	1.0	0.00
0	0	1	0	0	0	0	0	0	0	2.0	0.50
0	0	1	0	0	0	0	0	0	0	0.5	1.00
0	0	1	0	0	0	0	0	0	0	0.0	1.00
0	0	1	0	0	0	0	0	0	0	0.0	1.50

Figure 3. Crop Production Dataset Features in Binary Format

Data Modeling: After analyzing the dataset, the researchers proceeded to data modeling. The dataset was divided into training and testing sets using an 80-20 split, a commonly recommended ratio in both literature and practice. Figure 4 illustrates popular split ratios, emphasizing the efficacy of the 80-20 division for both training and model evaluation.

Model Evaluation: For model evaluation, the dataset was fed into the computer model and preprocessed to avoid overfitting. Pipelines for various models were employed, facilitating standard scaling to improve performance. Post-training, the model's accuracy was assessed using standard metrics such as R-squared and various error measures.

Figure 4. Data Splitting

Machine Learning Algorithms Employed

This study leveraged multiple machine learning algorithms. Machine learning algorithms [16] [17]play a pivotal role in this study by enabling the development of predictive models for crop trends in Benguet's highland vegetable production. These algorithms are employed to harness the power of historical crop production data and environmental factors, allowing the researcher to uncover intricate patterns and relationships within the dataset. This include Linear Regression, Decision Trees, Random Forest, Support Vector Machines, Lasso, and Ridge Regression [6] —for robust crop yield predictions. These algorithms were chosen for their unique capabilities in modeling complex relationships and were carefully tuned. They ranged from simple models like Linear Regression, useful for its interpretability but limited by its linear assumptions, to more complex ones like Random Forest, known for high accuracy but computational intensity. The aim was to construct a versatile Decision Support System for agriculture.

RESULTS AND DISCUSSION

Crop Trends and Patterns: In this section, the researchers delve into the observed trends and patterns of crop production across 13 municipalities in Benguet. This analysis aims to provide a nuanced understanding of the region's agricultural landscape, offering critical insights that can inform future decision-making for farmers, policymakers, and other stakeholders. These visualizations (see Fig. 5) enable the depiction of each municipality and how different factors come together to paint the big picture of farming in these areas.

The researchers observed periods of low to zero production in regions that historically had consistent production levels. This is particularly notable given that Benguet is a major producer of these highland vegetable crops in the Philippines.

The plot highlights a significant drop in crop production in the year 2020 as shown in fig. 6. This decline suggests that the pandemic has had a profound impact on farm productivity in Benguet.

Upon examining the patterns and trends, it was observed that the months with consistently high yields were from June to September. These months align with the typical harvest season, coinciding with the summer and fall seasons.

The graph offers a detailed view of crop production across an eight-year period, from 2015 to 2022. A prominent trend observed was a significant increase in production beginning in May and sustaining through September, indicating a seasonal peak in farming activities during these months.

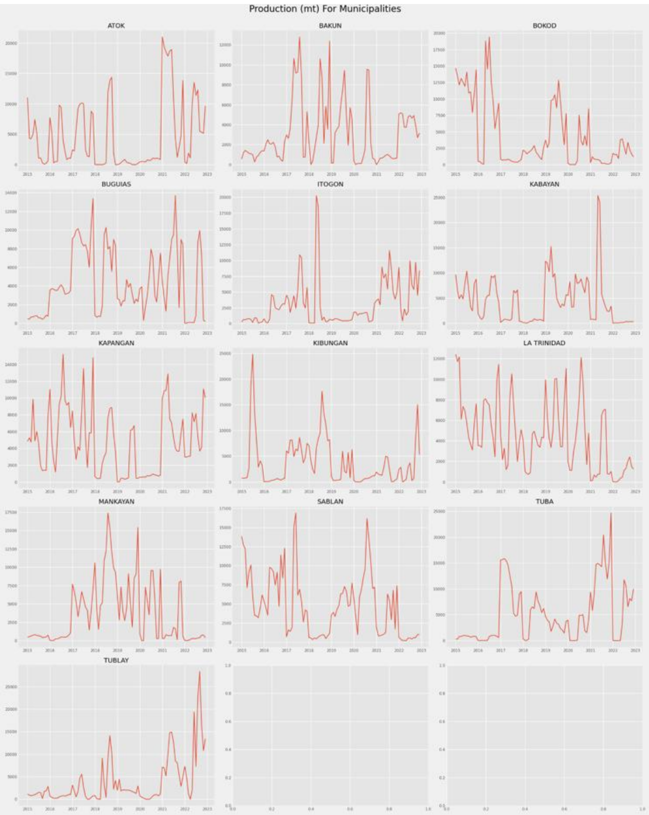

Figure 5. Crop Trends and Patterns in Benguet

Crop Production Predictions: In this section, the researchers delve into the predictive models generated by various machine learning algorithms to forecast future trends in crop production. These models aim to serve as a robust decision-making tool for stakeholders and farmers of Benguet.

The figures show the predictions of the model using the six different algorithms presented in a scatterplot where in the actual values are represented by the red broken line that runs diagonally through the middle while the predictions are the various blue points located near or along the red line.

Most predictions from the models are closely fitted to the line, indicating that the models did a good job at predicting production values. Notice that for larger values, the models are predicting values which are very far from the line which are the actual values. But not all the predictions lie close to the actual values, and these are what are known as outliers. A few outliers can be seen where the actual value is higher which means that the models that were trained were not particularly good at predicting large production values. This is because we have very little data on what affects the predicted value for production. The model relies its predictions mostly on the two features, Area planted, and Area harvested.

The scatter plots as show in fig. 6a using Linear Regression, 6b using Ridge Regression, 6c using Lasso Regression, 6d using Decision Tree Regression, 6e using SVM Regression, and 7f using Random Forest Regression, illustrates both actual and predicted values, with the red dashed line representing the actual values and the blue points denoting the predictions. The trained models strive to align the blue points as closely as possible with the red dashed line. While the models demonstrate a fairly accurate performance in predicting lower values, they are not flawless; some data points deviate significantly from the actual line.

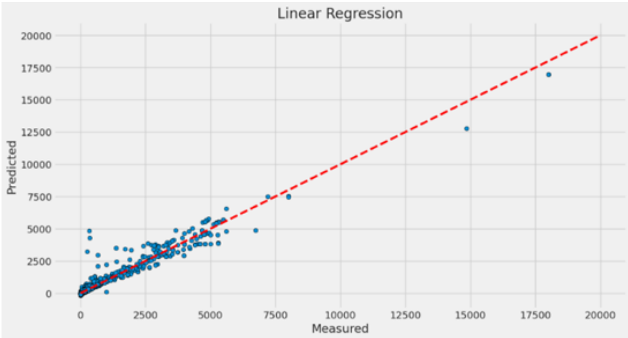

Figure 6a. Prediction using Linear Regression

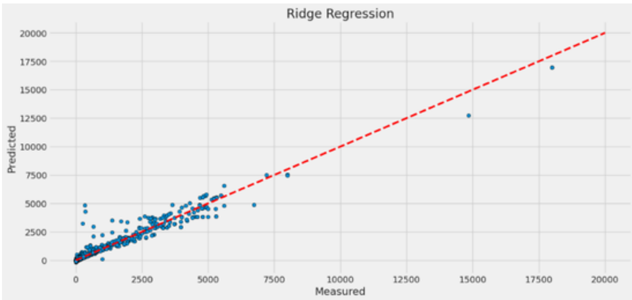

Figure 6b. Prediction using Ridge Regression

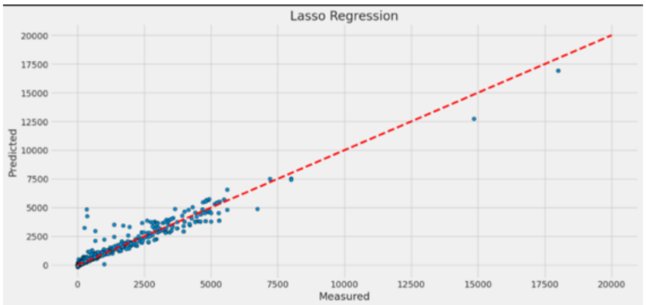

Figure 6c. Prediction using Lasso Regression

Figure 6d. Prediction using Decision Tree Regression

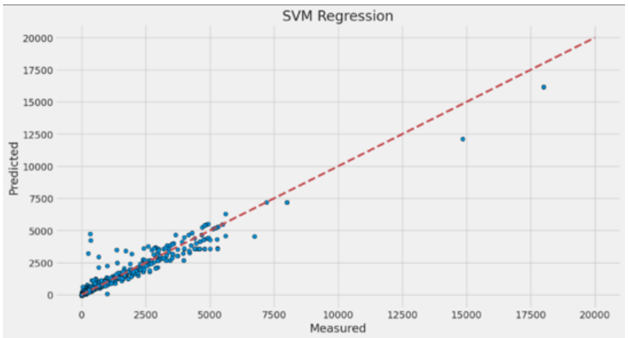

Figure 6e. Prediction using SVM Regression

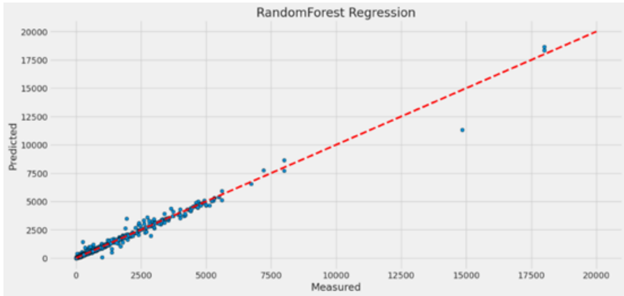

Figure 6f. Prediction using Random Forest Regression

Model Evaluation: After training the models, the researchers compared each using evaluation metrics commonly used for regression problems, such as the R-squared Score (Accuracy), Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE). These metrics allowed for a comprehensive assessment of the models' performance, enabling the researchers to identify the most effective model for predicting the target variable based on the input features. By rigorously evaluating the models using these metrics as shown in the table, the researchers gained valuable insights into the models' predictive capabilities and made informed decisions on their suitability for solving the regression problem at hand.

Table 2: Machine Learning Algorithms Evaluation Metrics

Models	**Accuracy**	**Mean Absolute Error (MAE)**	**Mean Squared Error (MSE)**	**Root Mean Squared Error (RMSE)**
Linear	0.963459	66.827391	24828.216634	157.569720
Lasso	0.963452	66.830537	24832.607325	157.583652
Ridge	0.963399	65.397911	24868.990188	157.699049
Decision Tree	0.981314	16.153847	12696.670742	112.679505
SVR	0.951814	51.633876	32740.116238	180.942301
Random Forest	0.989094	13.511579	7410.065983	86.081740

In the subsequent table, it's evident that the Random Forest algorithm outperforms the other models, boasting an accuracy rate of 98.91%. This is notably higher than the 95-98% range achieved by alternative algorithms. Additionally, Random Forest has the lowest Root Mean Squared Error (RMSE) value, registering at 86.08. This suggests that when employing Random Forest, the potential prediction error is approximately limited to 86 metric tons—significantly lower than the error margins of over 100 metric tons associated with the other algorithms.

The Random Forest model stands out due to its superior performance across metrics, making it ideal for precise future crop trend predictions. Utilizing regression analysis and appropriate model selection enhances our understanding of crop trends and supports strategic agricultural decision-making.

CONCLUSION AND FUTURE SCOPE

This research aims to utilize machine learning algorithms for predicting future crop yields in Benguet, based on eight years of historical production data, to enhance food sustainability. A variety of machine learning algorithms were deployed, among which the Random Forest algorithm proved most effective in minimizing predictive error, as evaluated through standard performance metrics. The study's findings were visualized in easily interpretable graphs and incorporated into a dynamic dashboard for broader dissemination. Although the research objectives were met, the study identifies room for improvement in data consistency and completeness.

This study lays the groundwork for improving agricultural forecasting in Benguet’s highland vegetable industry. Future research can enhance data integration, predictive accuracy, and real-time applications for policymakers and farmers.

Incorporating climate data, market trends, and AI-driven decision support systems can improve resilience and profitability. Expanding machine learning models and leveraging IoT for real-time monitoring will enhance precision farming.

Beyond vegetables, applying this framework to other highland crops and optimizing farm-to-market supply chains can further support sustainable agriculture. These advancements will drive smarter farming, climate resilience, and long-term food security.

ACKNOWLEDGEMENT

The researcher sincerely thanks the Department of Agriculture staff for their invaluable support and expertise, which greatly contributed to this study. Special appreciation is extended to the farmers for sharing their firsthand experiences, shaping both the research and the proposed system. A heartfelt gratitude goes to my family for their unwavering support and to my son and his friend for their contributions to the analytic portion of the research. Above all, heartfelt thanks to the Almighty Father for His wisdom, strength, and blessings throughout this endeavor.

REFERENCES

[1] A. R. Quitaleg and M. Dumlao, "A Decision Support System for Benguet's Upland Vegetable Crop Prediction using Machine Learning Techniques," in *European Chemical Bulletin*, 2023.

[2] S. Lopez-Ridaura, L. Barba-Escoto, C. Reyna, J. Hellin, B. Gerard, and M. van Wijk, "Food security and agriculture in the Western Highlands of Guatemala," *Food Secur.*, vol. 11, no. 4, pp. 817–833, Aug. 2019, doi: 10.1007/s12571-019-00940-z.

[3] "The State of Food Security and Nutrition in the World 2021," *State Food Secur. Nutr. World 2021*, Jul. 2021, doi: 10.4060/CB4474EN.

[4] "Averting an agricultural and ecological crisis in the Philippines' salad bowl." https://news.mongabay.com/2020/03/averting-an-agricultural-and-ecological-crisis-in-the-philippines-salad-bowl/ (accessed Sep. 07, 2023).

[5] S. Kant Gupta and S. Malik, "Technoarete Transactions on Intelligent Data Mining and Knowledge Discovery Application of Predictive Analytics in Agriculture".

[6] K. G. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, "Machine Learning in Agriculture: A Review," *Sensors 2018, Vol. 18, Page 2674*, vol. 18, no. 8, p. 2674, Aug. 2018, doi: 10.3390/S18082674.

[7] V. Kumar and M. L., "Predictive Analytics: A Review of Trends and Techniques," *Int. J. Comput. Appl.*, vol. 182, no. 1, pp. 31–37, Jul. 2018, doi: 10.5120/IJCA2018917434.

[8] C. S. Lee, P. Yeng, S. Cheang, and M. Moslehpour, "Predictive Analytics in Business Analytics: Decision Tree," 2022.

[9] L. O. Colombo-Mendoza, M. A. Paredes-Valverde, M. D. P. Salas-Zárate, and R. Valencia-García, "Internet of Things-Driven Data Mining for Smart Crop Production Prediction in the Peasant Farming Domain," *Appl. Sci. 2022, Vol. 12, Page 1940*, vol. 12, no. 4, p. 1940, Feb. 2022, doi: 10.3390/APP12041940.

[10] J. Gladju, B. S. Kamalam, and A. Kanagaraj, "Applications of data mining and machine learning framework in aquaculture and fisheries: A review," *Smart Agric. Technol.*, vol. 2, p. 100061, Dec. 2022, doi: 10.1016/J.ATECH.2022.100061.

[11] "Republic Act No. 11293." https://lawphil.net/statutes/repacts/ra2019/ra_11293_2019.html?fbclid=IwAR2qEgsgBg80KPyKQue_9-MUXnLp1FQaahhYhGmbBJuLgC7rQghmlGD8QP0 (accessed. May 10, 2020).

[12] "Data Mining Techniques - GeeksforGeeks." https://www.geeksforgeeks.org/data-mining-techniques/ (accessed May 20, 2023).

[13] "How to create a seaborn correlation heatmap in Python? - GeeksforGeeks." https://www.geeksforgeeks.org/how-to-create-a-seaborn-correlation-heatmap-in-python/ (accessed July 15, 2023).

[14] P. Roberto De Melo Barros Junior *et al.*, "Multi-fractal detrended cross-correlation heatmaps for time series analysis," *Sci. Reports 2022 121*, vol. 12, no. 1, pp. 1–14, Dec. 2022, doi: 10.1038/s41598-022-26207-w.

[15] S. Lopez-Ridaura, L. Barba-Escoto, C. Reyna, J. Hellin, B. Gerard, and M. van Wijk, "Food security and agriculture in the Western Highlands of Guatemala," *Food Secur.*, vol. 11, no. 4, pp. 817–833, Aug. 2019, doi: 10.1007/S12571-019-00940-Z/FIGURES/7.

[16] S. Ray, "A Quick Review of Machine Learning Algorithms," *Proc. Int. Conf. Mach. Learn. Big Data, Cloud Parallel Comput. Trends, Prespectives Prospect. Com. 2019*, pp. 35–39, Feb. 2019, doi: 10.1109/COMITCON.2019.8862451.

[17] M. Batta, "Machine Learning Algorithms - A Review," *Int. J. Sci. Res.*, vol. 18, no. 8, pp. 381–386, 2018, doi: 10.21275/ART20203995.