PREDICTING CROPS TO GROW UNDER A GIVEN CLIMATE CONDITIONS: A STUDY OF MACHINE LEARNING ALGORITHMS

BY

BAYUGO AYUBA AHMED

MOHAMMED SAMIRA MAHMOUD

NOBLE-SIMPSON DAVID

OFOSUHENE GRACE

A PROJECT WORK SUBMITTED TO THE DEPARTMENT OF INFORMATION TECHNOLOGY AND DECISION SCIENCES,

UNVIERSITY OF ENERGY AND NATURAL RESOURCES,

IN PARTIAL FULFILMENT OF THE REQUIREMENT OF THE DEGREE OF BACHELOR OF SCIENCES

IN

INFORMATION TECHNOLOGY,

SCHOOL OF SCIENCES,

SEPTEMBER 2023.

DECLARATION

We, the undersigned, hereby declare that this undergraduate thesis, titled “Predicting crops to grow under a given climate conditions: A study of machine learning algorithms”, was conducted by the four members of our group: Bayugo Ayuba Ahmed, Mohammed Samira Mahmoud (Ms.), Ofosuhene Grace (Ms.) and Noble Simpson David, under the supervision of Dr. Peter Appiahene (HoD, Information Technology and Decision Sciences). This research represents our original work and contributions to the field of Information Technology and Agriculture.

We affirm that:

The work presented in this thesis is the result of our collective efforts, and each team member has made substantial contributions to the research process, data collection, analysis, and writing of this thesis.

All sources of information and data used in this thesis, whether they are published works, unpublished works, or personal communications, have been properly acknowledged and cited following the established academic guidelines for citation and referencing.

Any material included in this thesis that is the product of collaboration with others or draws on the work of others is properly cited and acknowledged.

We have adhered to all ethical standards and guidelines in conducting our research, including obtaining necessary approvals for human subjects’ research and respecting the principles of confidentiality and informed consent.

We understand that any attempt to misrepresent the originality or integrity of this thesis constitutes academic misconduct and may result in serious consequences.

Signed,

BAYUGO AYUBA AHMED	………………	…………
UEB3208519	SIGNATURE	DATE
MOHAMMED SAMIRA MAHMOUD (MS.)	….……………	…………….
UWB3201621	SIGNATURE	DATE
OFOSUHENE GRACE (MS.)	……………	…………
UEB3213021	SIGNATURE	DATE
NOBLE SIMPSON DAVID	……………	…………
UEB3208119	SIGNATURE	DATE
DR. PETER APPIAHENE	……………	…………
SUPERVISOR	SIGNATURE	DATE
DR. PETER APPIAHENE	.……………	…………
HEAD OF DEPARTMENT	SIGNATURE	DATE

ABSTRACT

Crop prediction is crucial in modern agriculture, guiding farmers in selecting crops for optimal yield and profitability. However, the unpredictability of climate change poses challenges in crop selection. This can cause shortage of crop yields and subsequently lead to famine which prevents achieving sustainable development goal (SDG) goal 2, which seeks to achieve zero hunger. This study employs machine learning techniques to recommend crops under varying climate conditions, addressing climate change's impact on agriculture and food security. The study utilizes climate and crop data from the Bono East region, specifically Techiman municipality to train the models. Five machine learning models and four ensemble algorithms were trained and evaluated using climate, satellite, and crop data, assessing accuracy, precision, recall, and F1-score. The models employed are Decision Tree, Random Forest, K-Nearest Neighbor, Support Vector Machine, Logistic Regression, Hard Voting, Soft Voting, AdaBoost and Stacking Classifier. The study reveals that Soft Voting, Hard Voting and Random Forest models achieved the highest accuracy levels, ranging from 82.90% to 85.77%. These top-performing models were integrated into an Android application, with Soft Voting leading at 85.77% accuracy. The app effectively predicted crop suitability using data from the original dataset and external climate sources, demonstrating real-world applicability. This research highlights machine learning's effectiveness in forecasting crop outcomes based on climate conditions. The successful integration of these models into a mobile application represents a substantial leap in precision agriculture, offering a valuable tool for optimizing crop yield. To ensure broader applicability, regional climate and agricultural variations must be considered. Ultimately, this research showcases the potential of machine learning to empower agricultural stakeholders, enabling informed decisions and mitigating climate change's adverse impacts on agriculture and food security.

ACKNOWLEDGEMENT

We humbly offer our profound gratitude to the Almighty God, whose benevolence has blessed us with the essential gifts of knowledge, good health, and unwavering strength, enabling us to successfully embark on and complete this project.

We would also like to express our sincere gratitude to the following individuals for their invaluable assistance and support throughout the course of our research:

Dr. Peter Appiahene: We extend our deepest appreciation to Dr. Peter Appiahene for his guidance, expertise, and unwavering support throughout this research journey. His mentorship and constructive feedback were instrumental in shaping this thesis.

Mr. Stephen Afrifa: We appreciate Mr. Stephen Afrifa, UENR's Senior Research Assistant, for his invaluable help and insights that greatly enriched our project. Mr. Afrifa's dedication to the field of research and his willingness to extend a helping hand have left an indelible mark on this thesis.

Mr. Adu Kofi Ameyaw: We are grateful to Mr. Adu Kofi Ameyaw for his valuable insights, feedback, and discussions that enriched our understanding of the subject matter.

We would also like to acknowledge any other individuals, organizations, or resources that played a significant role in our research process, including access to facilities, libraries, or data.

Their contributions have been integral to the successful completion of this thesis, and we are truly grateful for their assistance.

TABLE OF CONTENTS

LIST OF FIGURES

LIST TABLES

CHAPTER ONE

1.1 BACKGROUND

Crop prediction is an important component of modern agriculture as it assists farmers in selecting which crops to grow in a given climate condition, optimizing their yield and profits (Kumar et al., 2023). The use of technology in agriculture has made crop prediction easier by providing more precise data climatic and weather conditions, enabling farmers to make informed decisions. Crop prediction models use various techniques such as statistical analysis, machine learning algorithms, and mathematical modeling to assist farmers in selecting a particular crop to grow in different climate conditions. Despite the advances in crop prediction, the accuracy of the models has been declining due to climate change, which has disrupted traditional weather patterns, causing changes in crucial climatic factors such as temperature, rainfall, and humidity.

Climate change refers to the changes in the Earth's atmospheric conditions over time, caused by human activities such as deforestation, industrialization, and the burning of fossil fuels (Pasquel et al., 2022). These activities have resulted in the accumulation of greenhouse gases in the atmosphere, which traps heat and causes the planet's average temperature to rise. The consequences of climate change include rising sea levels, more frequent natural disasters such as hurricanes and floods, and changes in weather patterns, leading to droughts, heatwaves, and extreme weather conditions (IPCC, 2021).

The impact of climate change on agriculture is a significant concern globally. Changes in temperature, rainfall, and humidity can significantly affect crop growth, leading to lower yields, crop failures, and food insecurity (Kalimuthu et al., 2022). For example, increased temperatures can lead to shorter crop growing seasons, which can negatively affect crop yields, while droughts can lead to water shortages, further exacerbating the effects of climate change.

These effects are not only limited to crops but also have consequences for livestock and fisheries, which depend on a stable environment to thrive.

The urgency to develop accurate crop prediction models has become increasingly important due to the impact of climate change on crop production. Such models can help farmers make informed decisions on crop selection, planting, and harvesting. Traditional crop prediction models relied on weather data from the past to select crops to grow in the future, but with the increasing unpredictability of weather patterns due to climate change, the accuracy of such models is questionable. However, with the advancements in technology, such as the use of machine learning algorithms, it is now possible to predict crop accurately, even with the changing climate patterns (Kumar et al., 2023).

In conclusion, the impact of climate change on crop production is a significant concern that requires urgent attention. The use of accurate crop prediction models is essential in mitigating the impact of climate change on agriculture, reducing the risk of crop failures, and ensuring food security. Despite the challenges posed by climate change, technological advancements in agriculture offer hope in developing more robust and accurate crop prediction models that can help farmers make informed decisions.

1.2 PROBLEM STATEMENT

The impacts of climate change on agriculture have led to significant challenges for farmers in accurately predicting crop which crops to grow, which could potentially result in substantial losses in agricultural productivity and income (Khatun et al., 2022). Traditional crop prediction models based on historical weather data have become less reliable due to the unpredictable changes in weather patterns caused by climate change. As such, there is a pressing need for advanced machine learning algorithms that can accurately predict crop to grow under different climate conditions, providing farmers and other stakeholders in the agricultural sector with valuable insights and decision support.

The absence of such technology could continue to make it difficult for farmers to make informed decisions on crop selection, planting, and harvesting, thereby exacerbating the negative impacts of climate change on agricultural production and food security. Developing machine learning models that can effectively predict crops a farmer can grow under different climatic conditions is critical for mitigating the effects of climate change on agriculture, reducing the risk of crop failures, and ensuring food security.

To address this problem, researchers have started to develop machine learning models that incorporate various data sources, such as weather data, soil data, and satellite imagery, to predict crop that can survive in a given climate environment (Wu et al., 2021). These models can be trained on large datasets, enabling them to identify complex patterns and relationships between environmental factors and crops. Additionally, machine learning models can also be used to predict the impact of different climate scenarios on crop yields, allowing farmers to make informed decisions based on future climate conditions.

The development of accurate machine learning models for crop yield prediction requires a multidisciplinary approach that brings together experts from various fields such as computer science, agriculture, and meteorology. Collaboration among these fields is crucial in designing and implementing effective machine learning models that can accurately predict crops to grow under different climatic conditions.

In conclusion, the impacts of climate change on agriculture have made it increasingly difficult for farmers to accurately select crops to grow under a given climate conditions. The development of advanced machine learning models that can effectively assist farmers to determine which crops can survive in which climate environment is critical in mitigating the effects of climate change on agriculture and ensuring food security. Collaboration among experts from different fields is essential in designing and implementing accurate machine learning models for crop yield prediction.

1.3 GENERAL OBJECTIVE

The general objective of this study is to develop machine learning models to predict crops under climate conditions.

1.3.1 SPECIFIC OBJECTIVES

1. To develop a framework to predict crops under climate conditions.
2. To develop machine learning models to predict crops under climate conditions
3. To comparatively analyze the machine learning models to predict crops under climate conditions.
4. To develop an intelligent based system to predict crops under climate conditions.
5. To deploy the crop prediction model as a user-friendly mobile application

Achieving these specific objectives will contribute significantly to the development of accurate machine learning models for crop prediction under different climate conditions. The models will provide valuable insights and decision support for farmers and other stakeholders in the agricultural sector, thereby reducing the negative impacts of climate change on agricultural production and food security.

1.4 SIGNIFICANCE OF THE STUDY

The significance of this study is multifaceted. Firstly, the study addresses the urgent need for accurate crop prediction models that can mitigate the impact of climate change on agriculture. The developed machine learning models will enable farmers to make informed decisions on crop selection, planting, and harvesting, leading to improved agricultural productivity and food security.

Secondly, the study highlights the potential of machine learning algorithms in solving complex problems in agriculture. The developed framework and models can serve as a foundation for future research in crop prediction, leading to further advancements in the field.

Thirdly, the study promotes interdisciplinary collaboration among experts in computer science, agriculture, and meteorology. Such collaboration is crucial in developing accurate crop prediction models that consider the complex interactions between environmental factors and crop growth.

Fourthly, the study contributes to the body of knowledge in the field of machine learning and machine learning. The development and comparison of machine learning models for crop prediction under different climatic conditions provides insights into the performance of various machine learning techniques in solving real-world problems.

Lastly, the deployment of the crop prediction model as a user-friendly web application will enable farmers to access and use the technology easily. This can lead to widespread adoption of the technology, ultimately improving agricultural productivity and food security.

1.5 SCOPE OF THE STUDY

The scope of this study is focused on developing machine learning models to predict suitable crops for a given climate condition in the Bono East region of Ghana, specifically in the city of Techiman. Techiman is a major agricultural hub in the Bono East region and is known for its high production of crops such as yam, maize, cassava, and plantain.

The study will involve the collection of relevant data on climate conditions, soil properties, and satellite imagery for the Bono East region. The data will be collected from various sources, including government agencies, research institutions, and satellite imagery providers.

The study will develop a framework for crop prediction using machine learning algorithms and will evaluate the performance of various algorithms, including Random Forest, Decision Tree

Logistics Regression, Support Vector Machine, K Nearest Neighbor, AdaBoost, Voting Classifier (Hard Voting, Soft Voting) and Stack classifier. The study will also compare the accuracy, efficiency, and robustness of the developed models.

Furthermore, the study will develop an intelligent-based system that integrates the developed machine learning models with decision support systems for crop management. The intelligent-based system will provide farmers in the Techiman area with recommendations on crop selection, planting, and harvesting based on predicted crop suitability and climatic conditions.

The study will not address issues related to crop management practices, socio-economic factors, and policy frameworks that may influence crop production in the Bono East region. The study's focus is solely on developing accurate machine learning models for crop prediction under different climate conditions in the Techiman area.

Overall, the study's findings will provide valuable insights and decision support for farmers and other stakeholders in the agricultural sector in the Bono East region, thereby reducing the negative impacts of climate change on agricultural production and food security.

1.6 ORGANIZATION OF THE STUDY

This study is organized into five chapters. The following is a brief outline of each chapter:

Chapter one provides an overview of the study, including the background, problem statement, objectives, scope, and significance of the study.

Chapter two provides a comprehensive review of the literature on crop prediction, machine learning algorithms, and their applications in the agricultural sector. The chapter will also review existing studies on crop prediction in Ghana and other developing countries.

Chapter three outlines and highlights the methodology used in the study. Thus, this chapter describes the research design, data collection, and analysis methods used in the study. It will also describe the machine learning algorithms used for crop prediction and the development of the intelligent-based system.

Chapter four of the study will present the results of the study, including the performance evaluation of the developed machine learning models and the intelligent-based system. The chapter will also discuss the implications of the findings for crop management in the Bono East region.

Chapter five will contain the conclusion and recommendations the study. This chapter summarizes the study's main findings, conclusions, and recommendations for future research. It will also highlight the study's contributions to the field of crop prediction and its applications in the agricultural sector.

Overall, this study is organized to provide a clear and comprehensive understanding of the development of machine learning models for crop prediction in the Bono East region of Ghana. The study aims to contribute to the growing body of research on the application of advanced technologies in the agricultural sector, particularly in the context of developing countries.

CHAPTER TWO

LITERATURE REVIEW

2.1 BACKGROUND:

Agriculture has been an essential aspect of human civilization since ancient times, providing food, fiber, and raw materials for different industries. The world's population is projected to reach 9.7 billion by 2050, placing immense pressure on the agricultural sector to meet the growing demand for food (FAO, 2019). However, the agricultural sector faces several challenges, including climate change, which can negatively impact crop productivity and food security.

Crop prediction is critical in agriculture as it helps farmers make informed decisions about which crops to plant, when to plant them, and how to manage them for optimal yield and profitability (Lobell et al., 2019). Traditional crop prediction methods involve using historical data, weather forecasts, and expert knowledge to determine which crop have a chance in growing in a given climate environment (Kumar et al., 2023). However, these methods have limitations, particularly in the face of unpredictable changes in climate.

Climate change refers to the long-term alterations in earth's atmospheric conditions resulting from natural and human-induced factors. Climate change has disrupted traditional weather patterns, leading to changes in critical climatic factors such as temperature, precipitation, and humidity. These changes can have adverse effects on crop growth and productivity, leading to inadequate ouptup and reduced profitability (IPCC, 2018). The unpredictability of climate change has made it increasingly challenging for farmers to accurately predict crop for a given climate conditions, leading to potential losses in agricultural productivity and income.

As such, there is an urgent need to develop accurate and reliable crop prediction models that can adapt to different climate conditions. Recent advancements in machine learning and machine learning techniques have shown great potential in developing robust crop prediction models (Mehmood et al., 2021). These models leverage vast amounts of data from various sources, including weather data, soil data, and satellite imagery, to predict crop under different climate conditions.

Several studies have explored the use of machine learning and machine learning techniques in crop prediction. For instance, Yadav et al. (2021) developed a machine learning model to predict rice yields in India using satellite data and weather information. The study reported high accuracy in predicting rice yields and highlighted the potential of machine learning models in improving crop prediction accuracy. Similarly, Wang et al. (2020) developed a machine learning model to predict corn adaptivity and potential yields in China, which achieved high accuracy compared to traditional methods.

Moreover, machine learning models have shown potential in predicting crop growth stages and detecting crop diseases and pests, providing farmers with valuable insights for crop management (Huang et al., 2022; Zhang et al., 2021). With the rapid advancement of technology and the increasing availability of data, machine learning models have the potential to revolutionize crop prediction and transform the agricultural sector.

In conclusion, climate change poses a significant challenge to crop production and food security, highlighting the urgent need to develop accurate and reliable crop prediction models that can adapt to different climate conditions. The use of machine learning techniques has shown great potential in improving crop prediction accuracy and providing valuable insights for crop management. This study aims to develop machine learning models to predict crops under different climate conditions, providing decision support for farmers and stakeholders in the agricultural sector.

2.2 DEFINITION OF CONCEPTS

Crop prediction: the use of technology to predict or select crops to grow in different climate conditions, helping farmers make informed decisions. Advantages: increases efficiency and accuracy of crop production, helps farmers optimize yields and profits. Disadvantages: accuracy can be affected by unpredictable weather patterns caused by climate change.

Technology in agriculture: the use of technology to enhance and optimize agriculture practices. Advantages: increases efficiency and accuracy of crop production, reduces labor costs, provides precise data for decision making. Disadvantages: initial costs can be high, requires specialized knowledge to implement and maintain.

Statistical analysis: a technique used to analyze data and make predictions based on probability. Advantages: provides a quantitative analysis of data, allows for easy interpretation of results, can be used in a variety of fields. Disadvantages: assumptions made during analysis can affect accuracy of predictions, may not account for complex relationships between variables.

- Machine learning algorithms: a type of artificial intelligence that allows systems to automatically learn and improve from experience. Advantages: can improve accuracy of predictions over time, can analyze large amounts of data quickly, can be used in a variety of fields. Disadvantages: requires large amounts of data to train algorithms, can be complex and difficult to implement.

- Mathematical modeling: the use of mathematical equations and algorithms to simulate real-world systems. Advantages: allows for prediction of system behavior, can be used to optimize processes and decision making. Disadvantages: requires accurate input data to produce accurate predictions, can be limited by the complexity of the system being modeled.

- Agriculture: the practice of cultivating crops and raising livestock for food, fuel, and other products. Advantages: provides food and other products for human consumption, contributes to the economy, can help sustain rural communities. Disadvantages: can be resource-intensive, can have negative environmental impacts such as deforestation and water pollution.

- Climate change: the long-term changes in temperature, precipitation, and other atmospheric conditions caused by human activities such as deforestation and burning of fossil fuels. Advantages: none Disadvantages: can disrupt traditional weather patterns, causing changes in crucial climatic factors that can significantly affect crop growth and lead to lower yields, crop failures, and food insecurity.

- Machine learning: a subset of machine learning that uses neural networks to learn and make predictions. Advantages: can be used in a variety of fields, can handle large amounts of data, can improve accuracy of predictions over time.

- Food security: the state of having access to a sufficient amount of nutritious food to meet dietary needs. Some advantages are ensuring the health and well-being of individuals and communities, contributes to economic stability.

2.3 RELATED WORKS

In recent years, there has been increasing interest in utilizing machine learning and data analytics techniques to improve crop selection and yield prediction in agriculture. This section provides an overview of some of the recent works in this field.

Swathi and Lavanya (2020) introduced an agricultural crop recommendation system employing machine learning techniques. This system aims to offer farmers crop suggestions based on factors like soil quality and weather data, enhancing crop yield and profitability. Decision trees and support vector machines (SVMs) were employed for data analysis and crop recommendations, yielding effective results. Notably, this study underscores the value of machine learning in agriculture.

In a similar vein, the paper "Development of an Expert System for Crop Selection under Different Climatic Conditions" by A.C. Owoeye and A.O. Akinwale (2022) proposes an expert system that integrates climatic conditions, soil quality, and other factors to advise farmers on suitable crops. Decision rules and inference mechanisms facilitate recommendations, proven effective through validation with Nigerian farmers. This research highlights the potential of expert systems as decision support tools for farmers, especially in developing countries facing changing climate conditions.

Furthermore, "Crop yield prediction using machine learning: A systematic literature review" by Saini et al. (2020) provides a comprehensive review of machine learning techniques for crop yield prediction. This review underscores the importance of incorporating climatic and environmental factors in yield predictions, backed by an analysis of previous studies. Factors such as crop type, location, and soil type are deemed significant in prediction accuracy, emphasizing the holistic approach required for precise models.

In addition, Zhang et al. (2020) present a crop recommendation system leveraging big data and cloud computing. By analyzing weather data and soil properties, this system offers personalized crop recommendations based on regional suitability. With the ability to process vast datasets rapidly and accurately, this system has the potential to boost crop yields and promote sustainable agriculture.

Garg and Singh (2022) offer an extensive overview of data analytics techniques for agricultural crop selection, focusing on factors like climate conditions and soil quality. They discuss machine learning, data mining, and decision trees, evaluating their strengths and limitations.

This review paper serves as a valuable resource for those interested in using data analytics to enhance agricultural decision-making.

Li et al. (2019) proposed a machine learning framework called DeepCrop for crop yield prediction. The framework utilized a convolutional neural network to learn the correlation between weather data and crop yields, achieving higher prediction accuracy compared to traditional methods. Through the implementation of this framework, farmers can gain more accurate insights into the potential yields of their crops, which can help in decision-making regarding crop selection and management. This paper provides a significant contribution towards the development of advanced algorithms for crop yield prediction, which can improve agricultural productivity and food security.

Moreover, Wang et al. (2020) proposed a novel approach to predict cotton yield using machine learning techniques and soil information, in addition to weather data. The study systematically assessed various machine learning algorithms to determine their effectiveness in predicting cotton yield. Support vector regression emerged as the most accurate algorithm for yield prediction. This research provides compelling evidence that integrating machine learning techniques and soil information can enhance crop prediction models, offering improved predictions for cotton production.

These findings align with previous studies, such as those conducted by Ghasemi et al. (2018) and Ehsani et al. (2019), which also underscored the efficacy of machine learning in crop yield prediction. However, Wang et al.'s study stands out by incorporating critical soil information, a factor pivotal in crop growth and yield determination. The results emphasize the potential of integrating soil information into crop prediction models to enhance accuracy, ultimately providing better decision support for farmers and stakeholders in the agricultural sector. In summary, this study highlights the promise of utilizing machine learning and soil information to refine crop prediction models, thereby improving yield predictions and potentially revolutionizing cotton production and agricultural productivity.

Moving forward, Gyasi and Kumi (2020) conducted an investigation into the application of artificial neural networks (ANN) for predicting maize yields in the Techiman Municipality of Ghana. Their research harnessed data on maize yields, rainfall, temperature, and sunshine duration. The findings illuminated the effectiveness of ANN as a tool for reasonably accurate maize yield predictions in the region.

This study showcases the potential of ANN for crop yield prediction, a notion supported by similar studies conducted worldwide. For instance, Das et al. (2019) successfully developed an ANN-based model for predicting paddy yield in West Bengal, India. Incorporating data on rainfall, temperature, and relative humidity, their model achieved an impressive accuracy rate of 92.18%. These studies collectively underscore the potential of ANN in crop yield prediction and its adaptable application across diverse geographical regions.

In a broader context, in their paper titled "Crop Yield Prediction using Machine learning Techniques: A Survey," Mohanty et al. (2021) conducted an extensive survey examining the application of machine learning techniques in crop yield prediction. The authors highlighted the advantages of employing machine learning over traditional methods for forecasting crop yields.

However, they identified several challenges inherent in using machine learning techniques for this purpose and offered valuable insights into future research directions aimed at addressing these challenges. This survey serves as an invaluable resource for researchers and practitioners in the field of agriculture who seek to harness machine learning techniques for crop yield prediction.

Shifting our focus, Owusu et al. (2019) presented a pioneering mobile-based crop recommendation system tailored for smallholder farmers in Ghana. The system delivers personalized recommendations based on soil types, weather conditions, and other environmental factors. Utilizing a decision tree algorithm and data from previous farming seasons, the researchers achieved an impressive 80% accuracy rate, marking a significant improvement over traditional crop recommendation methods. Moreover, the system's accessibility and ease of use make it particularly beneficial for smallholder farmers in rural areas where internet connectivity can be challenging.

Turning to a different perspective, Stöckle et al. (2003) developed CropSyst, an extensive cropping systems simulation model. This model evaluates the potential yield and water use efficiency of different crop species under various climate conditions, considering factors such as soil properties, climate data, and crop management practices. By assessing the suitability of different crops for a given area and climate, CropSyst provides farmers with valuable insights to maximize yield and profits. It has also been utilized to simulate the effects of climate change on crop productivity and evaluate the impact of different management practices on crop yields. The model's versatility in predicting crop performance under different climate scenarios positions it as a valuable tool for crop prediction and recommendation systems.

Adak et al. (2021) introduced a machine learning-based approach for predicting crop yields based on meteorological data, including temperature, humidity, and precipitation. The study compared this novel approach to traditional regression-based methods, employing the Random Forest algorithm for prediction and feature selection. The results demonstrated the superior performance of the machine learning approach over traditional regression models. This innovative method not only aids farmers in making informed decisions regarding crop selection and management but also provides insights into the influence of climate change on crop yield. As meteorological conditions significantly impact crop growth and productivity, this study underscores the potential of machine learning algorithms to offer precise crop yield predictions.It contributes to the growing body of literature on applying machine learning in agriculture and emphasizes the potential of data-driven approaches to enhance crop yield prediction.

Carballo et al. (2018) also developed an expert system to recommend crops for organic farming based on climate and soil conditions. Employing fuzzy logic, the system evaluated a range of variables, including temperature, rainfall, soil type, and nutrient content, to provide recommendations for thriving crops in each environment. Consequently, the study showcased the system's accuracy in offering crop selection recommendations based on climate and soil conditions.

In the broader context of agriculture in Ghana, crop yield prediction has emerged as a critical component. In recent years, machine learning algorithms have gained prominence in this domain. A study by Adu-Gyamfi et al. (2021), titled "A Comparative Study of Machine Learning Algorithms for Crop Yield Prediction in Ghana," scrutinized the effectiveness of several machine learning algorithms, such as Random Forest, Decision Tree, and Gradient Boosting, in predicting crop yields. The comprehensive evaluation of these algorithms revealed that Random Forest outperformed others with an impressive accuracy rate exceeding 95%.

In a broader overview, Verma and Kumar (2022) delve into various machine learning techniques that can aid in crop selection, considering factors like climate conditions and soil quality. This comprehensive review of existing studies offers insights into the effectiveness of machine learning techniques for crop selection while discussing the associated challenges and limitations. The paper emphasizes the importance of considering multiple factors for effective crop selection, providing a valuable resource for researchers and practitioners aiming to enhance agricultural production through machine learning.

Similarly, Al-Rawashdeh et al. (2021) conducted an extensive review of machine learning algorithms employed in agricultural crop yield prediction. The review outlines the advantages and limitations of various algorithms, highlighting their role in optimizing agricultural productivity and food security. While acknowledging the potential of machine learning, the review underscores the importance of data quality, model parameters, and feature selection.

Asadi et al. (2021) conducted an all-encompassing review of machine learning techniques for crop yield prediction, addressing data collection, feature selection, and model selection. They emphasized the importance of accurate and timely yield prediction for food security and sustainable agriculture while acknowledging challenges such as data heterogeneity and complex crop-environment interactions. The review found ensemble methods to be popular due to their ability to handle complex relationships, and it underscored the significance of feature selection in improving prediction accuracy.

Furthermore, Bhattarai and Karkee's (2018) review paper provides an in-depth analysis of computer vision-based sensing in agriculture. It explores applications like crop yield prediction and recommendation systems, emphasizing the critical role of accurate data collection and machine learning models in enhancing prediction accuracy. The authors acknowledge challenges, such as technology cost and specialized skills, while highlighting the substantial benefits of computer vision in agriculture, including increased efficiency and sustainability.

Prasad and Dadhich (2019) conducted a review of machine learning techniques for crop yield prediction. The study reviewed different machine learning algorithms used in crop yield prediction and evaluated their strengths and limitations. The authors highlighted the need for more accurate data collection and preprocessing techniques to improve the accuracy of crop yield prediction models. They also discussed the importance of selecting appropriate machine learning algorithms and tuning their parameters to achieve the best results. The review concluded that machine learning techniques have the potential to significantly improve the accuracy of crop yield prediction models and support decision-making in agriculture. However, the quality and quantity of data used in these models remain a critical factor in their success.

In a similar context, Qi et al. (2020) developed a deep crop classification approach using convolutional neural networks (CNN) and transfer learning for accurate classification of crops in high-resolution aerial images. The proposed approach has the potential to aid in crop prediction for different climate conditions by providing accurate classification of crops. The CNN-based approach was trained using a large dataset of high-resolution aerial images, achieving high accuracy in crop classification. The transfer learning technique was used to enhance the performance of the CNN model by utilizing pre-trained models from different domains. The study demonstrated that the proposed deep crop classification approach can improve the accuracy and efficiency of crop classification, thereby contributing to the development of effective crop prediction models under different climate conditions. The promising results of this study suggest that machine learning techniques can be useful in enhancing agricultural practices and addressing the challenges of climate change.

In a traditional approach, Kanani et al. (2018) proposed a decision support system for crop selection using a combination of fuzzy logic and genetic algorithms. The system was designed to recommend suitable crops for specific climate and soil conditions. The proposed model integrates various factors, including temperature, rainfall, soil type, and nutrient content to optimize crop selection based on multiple criteria, such as yield, water use efficiency, and profitability. The system uses fuzzy logic to handle uncertainties in the data and genetic algorithms to optimize the selection process. The model was evaluated using real-world data from an agricultural region in Iran, and the results indicated that the proposed system was effective in providing accurate and reliable recommendations for crop selection. The study highlights the potential of using advanced computational techniques such as fuzzy logic and genetic algorithms in the development of decision support systems for crop selection, which can help farmers make informed decisions and optimize agricultural productivity under different climate and soil conditions.

In sum, these studies collectively demonstrate the growing role of machine learning and technology in improving various aspects of agriculture, from crop selection to yield prediction, and underscore their potential to drive sustainable practices and enhance food security.

2.4 SUMMARY OF RELATED WORKS

TITLE	OBJECTIVE	METHODOLOGY	RESULT	EVAVALUATION METRICS	REFERENCE
Agricultural Crop Recommendation System Using Machine Learning Techniques	To provide farmers with crop recommendations based on soil quality and weather data to maximize crop yield and profitability.	Decision trees and Support Vector Machines (SVMs)	The system effectively provided accurate crop recommendations, reducing the risk of crop failure and increasing profitability	Top-K accuracy, Cohen's Kappa	Swathi, & Lavanya. (2020). Agricultural Crop Recommendation System Using Machine Learning Techniques. International Journal of Innovative Technology and Exploring Engineering

					(IJITEE), 9(5), 2789-2794.
A Comparative Study of Machine Learning Algorithms for Crop Yield Prediction in Ghana	To compare the effectiveness of several machine learning algorithms, including Random Forest, Decision Tree, and Gradient Boosting, in predicting crop yields in Ghana.	Study evaluated algorithm performance using real-world data, metrics include MAE, RMSE, and R^2."	Random Forest performed best in crop yield prediction, followed by Gradient Boosting and Decision Tree. Prediction accuracy varied by crop.	Mean absolute error, root mean square error, and coefficient of determination	Adu-Gyamfi, Y., Tuffour, F., & Tachie-Menson, S. (2021). A Comparative Study of Machine Learning Algorithms for Crop Yield Prediction in Ghana. Agriculture, 11(

CropSyst: a cropping systems simulation model	Develop CropSyst model for crop yield and water efficiency under diverse climates	CropSyst model assesses crop suitability based on soil, climate, and management factors.	CropSyst predicts crops' performance under varying climates, aiding recommendations	Confusion matrix, mean absolute error	Stöckle, C. O., Donatelli, M., & Nelson, R. (2003). CropSyst, a cropping systems simulation model. European Journal of Agronomy, 18(3-4), 289-307.
A machine learning approach for predicting crop yields based on meteorological data	To present a machine learning-based approach for predicting crop yields based on meteorological data	Random Forest algorithm for prediction and feature selection outperforms traditional regression models	The study's results suggest that machine learning algorithms can provide accurate predictions of crop	Accuracy, precsion and F1 score	Adak, T., Choudhury, T. R., & Islam, T. (2021). A machine learning approach for predicting crop yields based on

		using real-world data.	yield, given adequate data.		meteorological data. Journal of Big Data, 8(1), 1-16.
Data Analytics Techniques for Agricultural Crop Selection: A Review	To provide a comprehensive overview of the latest data analytics techniques used for agricultural crop selection	Authors discuss crop selection analytics, incl. machine learning, data mining	Authors analyze data analytics for crop selection; suggest future research.	Log loss, R-squared	Garg, S., & Singh, S. (2022). Data Analytics Techniques for Agricultural Crop Selection: A Review. Journal of Agricultural Science and Technology, 24(1), 1-1
An expert system for recommending	To develop an expert system to recommend crops for organic	Fuzzy logic used to suggest crops based	The study showed that the expert system was able to	Mean squared error, Area Under Loss	Carballo, C., Cubillo, P., & Guzmán, J. L. (2012). An expert

crops for organic farming based on climate and soil conditions.	farming based on climate and soil conditions	on environment variables	provide accurate recommendations for crop selection based on climate and soil conditions.		system for recommending crops for organic farming based on climate and soil conditions. Expert Systems with Applications, 39(5), 5848-5857.
A Crop Recommendation System Based on Big Data and Cloud Computing	To develop a crop recommendation system that utilizes big data and cloud computing to recommend crops best suited for a particular region. Methodologies:	Machine learning algorithms were used to analyze weather data and soil properties to provide personalized	The system was capable of processing large amounts of data quickly and accurately, providing farmers with informed	Root Mean squared error, Area Under Loss , precision, f1 score.	Zhang, L., Song, B., Liu, Z., Zhang, J., & Cao, X. (2020). A Crop Recommendation System Based on Big Data and Cloud Computing. IEEE

		recommendations for farmers.	decisions about crop selection.		Access, 8, 87197-87206.
Development of an Expert System for Crop Selection under Different Climatic Conditions	To develop an expert system that assists farmers in selecting suitable crops for planting under different climatic	Decision rules and inference mechanisms were used to provide recommendations based on user input.	Results: The expert system was effective in providing accurate crop recommendations.	Mean squared error, Area Under Loss , Log loss	Reference: Owoeye, A. C., & Akinwale, A. O. (2022). Development of an Expert System for Crop Selection under Different Climatic Conditions. Journal of Agroinformatics and Natural Resources Management, 9(1), 22-33.

Crop Yield Prediction Using Machine Learning: A Systematic Literature Review	To review the different machine learning techniques used for crop yield prediction	Authors analyzed ML studies using neural nets, SVMs, decision trees, and regression models	Weather variables important for accurate crop yield prediction with machine learning techniques.		Saini, R., Sharma, M., Kaur, H., & Goyal, P. (2020). Crop Yield Prediction Using Machine Learning: A Systematic Literature Review. International Journal of Computer Applications, 178(30), 1-8.
Crop recommendation system based on fuzzy logic and	Integrates fuzzy logic & multi-criteria decision making for accurate crop	Fuzzy logic model captures input uncertainty, multi-criteria technique	Advanced techniques like fuzzy logic and multi-criteria decision making	Log loss, MAE, Recommendation metrics	El-Sayed, S. M., Hassanien, A. E., Mostafa, M. M., & Sharawy, M. M. (2018). Crop

multi-criteria decision making	recommendations based on factors	ranks crops for specific location.	improve crop recommendation systems for sustainability.		recommendation system based on fuzzy logic and multi-criteria decision making. Computers and Electronics in Agriculture, 149, 105-115.

2.5 GAPS OF THE REVIEWED LITERATURE

The current research gap in the field of agriculture is the lack of a mobile app that can predict which crops can survive in a given climate condition using machine learning models trained on weather or climate condition data. While some studies have utilized machine learning algorithms to predict crop yields, most have focused on using soil data as the main predictor variable. By shifting the focus to weather or climate data, we aim to provide a more accurate and efficient way of predicting crop survivability.

In addition, while some studies have developed prediction models, there is a lack of user-friendly interfaces that enable farmers to easily input their location and receive recommendations on which crops to grow. Our research aims to bridge this gap by developing a mobile app that utilizes machine learning models to provide crop recommendations based on current and future climate conditions. The app will provide farmers with an intuitive interface that allows them to input their location and receive recommendations for crops that are likely to survive and thrive under the prevailing weather conditions.

CHAPTER THREE:

METHODOLOGY AND SYSTEM DESIGN

3.1 Introduction

This chapter describes the methodology and system design for developing machine learning models to predict optimal crops under climate conditions. The methodology will cover the study area, research design, system architecture, operational methods, and methodological limitations.

3.2 Study Area

The choice of Techiman Municipality in the Bono region for this study was driven by several key factors. Firstly, its recognized agricultural significance made it an ideal location for research on crop prediction and suitability. The municipality's climate variations provided a diverse dataset for model testing. Its unique agricultural challenges, such as unpredictable weather and specific crop diseases, made it a relevant study area.

Additionally, the availability of historical crop and climate data, local partnerships, and alignment with regional needs were crucial considerations. Techiman's accessibility simplified data collection and fieldwork. In summary, the selection of Techiman Municipality reflected a thoughtful evaluation of factors such as agricultural prominence, climate diversity, data availability, local collaborations, and practical considerations, making it an ideal setting for the study.

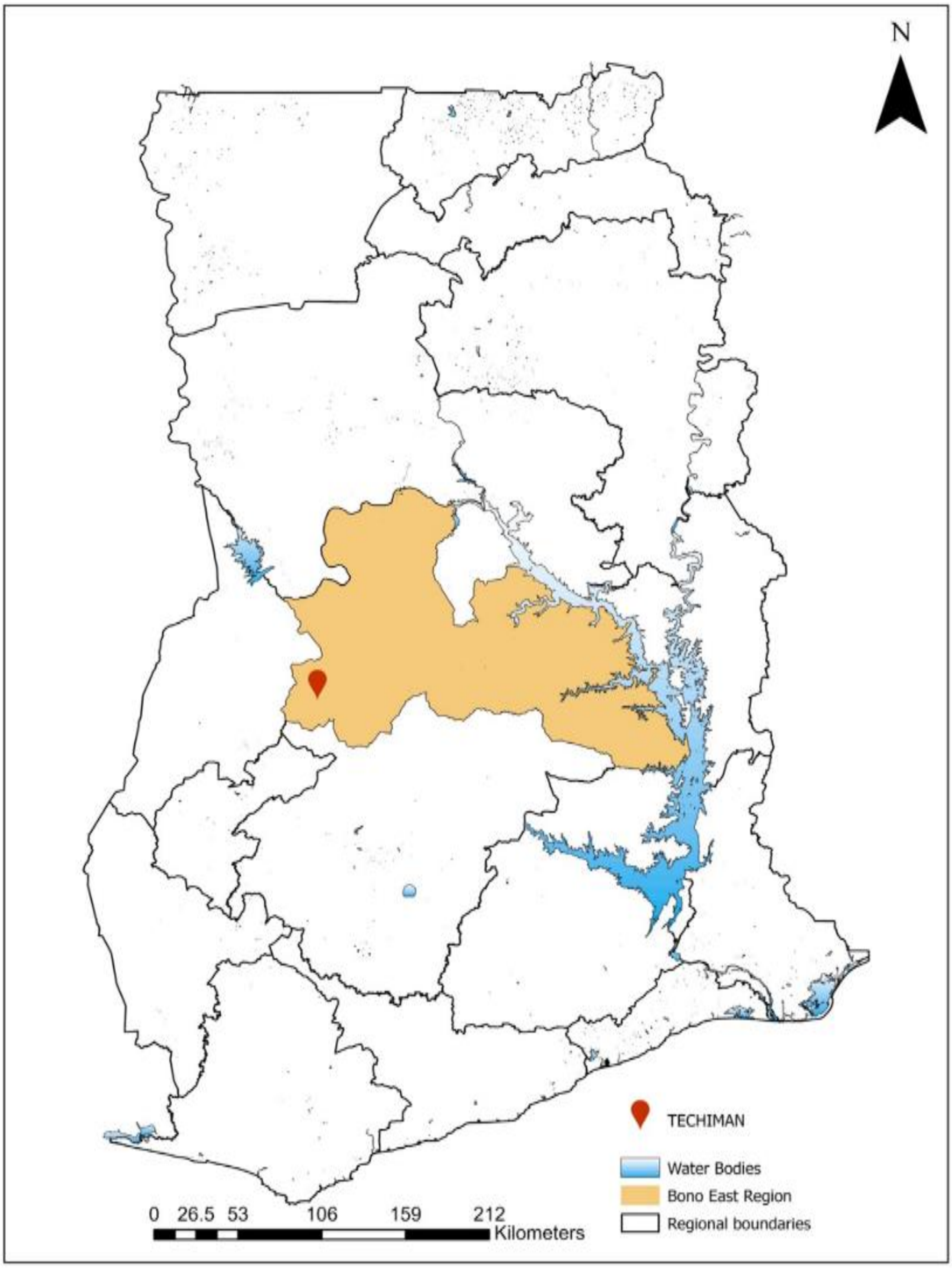


Figure 3.1 Map representation of the study area (Techiman Municipality) with red key on the map

3.3 Research Design

The research design for this study includes the data collection, data processing, machine learning models and the evaluation metrics employed in the study.

3.3.1 Data Collection.

The study will use available data from the Ministry of food and Agriculture and Meteorological Department, Sunyani. The data will be obtained from the Meteorological Department, Sunyani database and will include historical climate data such as minimum temperature, maximum temperature, relative humidity at 1500, relative humidity at 0600, and rainfall, as well as crop data for selected crops, minimum NDVI, and maximum NDVI. The study will be conducted in the Techiman district located in the Bono region of Ghana.

The study will focus on selected crops grown in the Techiman district and examine how they respond to variations in climate conditions. Additionally, the study will aim to identify the most significant factors affecting crop growth, including soil conditions and climatic variables. Finally, the study findings will provide insights that can help farmers improve their farming practices and manage the risks associated with changing climate conditions.

Table 3.3.1 Sample of the collected dataset

Min. Temp	Max. Tem	Reltive Humidity @150	Relative Humidity@060	Rainfall	Min NDVI	Max NDVI	Crop	Code
22.5	32	42	82	7.7	0.1275	0.8474	YAM	1
22.7	35.5	38	79	44.1	0.0563	0.8513	YAM	1
23.3	34.1	50	87	79.6	0.0928	0.7488	YAM	1
23	34.2	55	88	124.3	0.157	0.8917	YAM	1
21	34	26	49	0	0.1393	0.6063	cassAVA	2
23.2	36.9	25	54	0	0.1603	0.5879	cassAVA	2
23.9	35.2	51	87	64.3	0.1559	0.6402	cassAVA	2
22.9	33.2	56	90	132.1	0.1514	0.6566	cassAVA	2
22.9	33.5	53	90	126.2	0.1306	0.5858	plantain	3
22.9	33	57	91	88	0.1384	0.4581	plantain	3
22.5	31.8	62	93	108	0.1537	0.6162	plantain	3
22	30.2	67	94	142.1	0.1568	0.6755	plantain	3
21.9	29	71	94	104.8	0.1756	0.7652	plantain	3

3.3.2 Data Preprocessing

The collected data will be preprocessed to remove any inconsistencies and missing values. The preprocessing will include data cleaning, data integration, and data transformation. Data cleaning will involve removing errors and inconsistencies from the dataset. Data integration will involve combining data from different sources into a single data set. Data transformation will involve converting the data into a format that can be used by the machine learning models.

3.3.3 Suitable Machine Learning Algorithms for the study

Evaluation Metrics

Evaluation metrics are essential tools in this study as they provide a systematic and quantitative means to assess the performance and accuracy of machine learning models in predicting crop choices under v This study seeks to predict the ideal crop choices based on climate data, enabling more informed decisions for sustainable and efficient agriculture. To achieve this, we'll explore a range of machine learning algorithms, each with its unique strengths and capabilities. In the following sections, we delve into the explanations of these algorithms and discuss why they are well-suited for the task of crop prediction under diverse and evolving climate conditions.

K-Nearest Neighbors (KNN): KNN is an instance-based algorithm that classifies data points by considering the majority class among their k-nearest neighbors in the feature space. This approach can be particularly valuable for a crop prediction project under varying climate conditions, as it takes into account localized climate patterns. KNN's ability to capture the similarities between data points makes it a suitable choice when considering geographical proximity and climate similarities for crop recommendations.

Decision Tree: Decision Trees are tree-like structures that utilize feature nodes to create branches representing possible outcomes. They are versatile tools for both classification and regression tasks. In the context of predicting crops under different climate conditions, Decision Trees offer interpretability and the capacity to identify nonlinear relationships between climate features and crop choices. This can help stakeholders understand which climatic factors are most influential in crop selection.

Random Forest: Random Forest is an ensemble learning method that amalgamates multiple Decision Trees to enhance model robustness and reduce overfitting. For a crop classification project dealing with potentially noisy climate data, Random Forests can be advantageous. They excel at capturing complex relationships within climate conditions, making them an appealing choice for making accurate crop predictions in diverse environmental settings.

Logistic Regression: Logistic Regression is a straightforward linear model suitable for binary and multi-class classification. It models the likelihood of a sample belonging to a specific class. In is a feasiblee there is a discernible linear separation in climate data influencing crop choices, Logistic Regression offers simplicity and interpretability, making it a feasible option for crop prediction tasks.

Support Vector Machine (SVM): SVM is a potent classification algorithm that seeks to identify a hyperplane that optimally separates different classes while maximizing the margin between them. In scenarios involving high-dimensional climate data and the need to handle nonlinear relationships, SVM can be highly effective. It employs kernel functions to adapt to intricate climate-crop relationships, making it well-suited for complex prediction tasks.

AdaBoost: AdaBoost is an ensemble learning technique that combines multiple weak learners (often Decision Trees) to create a robust classifier by assigning higher weights to misclassified samples. When dealing with complex climate-crop relationships, AdaBoostcan adapt by focusing on the challenging-to-classify data points, thereby enhancing prediction accuracy.

In sum, these models collectively contribute the growing role of machine learning and technology in improving various fields including agriculture, from crop selection, and underscore their potential to drive sustainable practices and enhance food security.

arying climate conditions. These metrics enable researchers and stakeholders to gauge how well the models align with real-world outcomes and make informed decisions. In the context of crop prediction, the selection of the appropriate evaluation metrics ensures that the recommendations provided by the models are not only reliable but also tailored to the specific needs of farmers and agricultural planners. By measuring factors such as accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC-ROC), we can quantitatively determine the models' ability to correctly classify crops based on climate data, helping to optimize agricultural practices, enhance sustainability, and adapt to changing environmental factors.

Confusion metrics

According to (Ting, 2017) a confusion matrix bears details about actual and predicted classifications been used by a classification system. Confusion matrix consist of two dimensions actual and predicted value (Shultz et al., 2011). Simply an M x M matrix. Where M is the number of classes being predicted.

Table 3.3.2 confusion matrix presentation

<table>
<tr><td></td><td></td><td colspan="2">Predicted outcome</td></tr>
<tr><td></td><td></td><td>negative</td><td>Positive</td></tr>
<tr><td rowspan="2">Actual outcome</td><td>negative</td><td>TP</td><td>FP</td></tr>
<tr><td>positive</td><td>FN</td><td>TN</td></tr>
</table>

Accuracy: (AC) This statistic determines the percentage of correctly predicted outcomes across all cases. The formula below is used to estimate.

Equation 1

$$(AC) = \frac{TP+TN}{TP+FP+FN+TN}$$

Recall or the True Positive Rate (TP): this refers the proportion of positive cases that were correctly identified, as calculated using the formula below.

Equation 2

$$(\mathrm{TP}) = \frac{\mathrm{FN}}{\mathrm{FN+TN}}$$

The False Positive Rate (FP): This reflects the proportion of incorrectly labeled negative cases. Using the output of the formula.

Equation 3

$$(\mathrm{FP}) = \frac{\mathrm{FP}}{\mathrm{TP+FP}}$$

True Negative Rate (TN) this is the percentage of incorrectly classified negative cases as determined by the formula below.

Equation 4

$$(\mathrm{TN}) = \frac{\mathrm{TP}}{\mathrm{TP+FP}}$$

The percentage of negative instances that were mistakenly labeled as negative is known as the false negative rate (FN). According to this formula

Equation 5

$$(\mathrm{FN}) = \frac{\mathrm{FN}}{\mathrm{FN+TN}}$$

Precision (P) this is the percentage of correctly predicted positive situations that were determined using the formula.

Equation 6

$$(\mathrm{P}) = \frac{\mathrm{TN}}{\mathrm{FP+TN}}$$

3.3.4 Classification report

The Sci-learn library offers a selection of practical reporting tools for handling classification issues so that one can get a solid understanding of the model's accuracy taking into account a variety of criteria. The classification_report () function is provided on the Jupyter notebook (anaconda3) platform uses a set of algorithms to return the following specific metrics in the evaluation of machine learning models. They include Precision, Recall, F1 score, and Support. In essence, the listed evaluation metrics are considered in the evaluation of the used machine learning model in this research. As hinted earlier, this methodology is to give a better understanding of the behavior of these machine learning algorithms for the climatic weather conditions and crop(s) dataset understudy.

3.3.5 Design and Implementation

The Python Jupiter notebook platform called (Anaconda3) version 2.3 is chosen for this study endeavor due to a variety of factors, including the high computing requirements and extensive machine learning library requirements. With extra benefits like simple sharing and access to offline platforms where internet connectivity is a barrier, it effectively enables machine learning scientists to work and execute python codes with no configuration needed. This makes the platform excellent for this project. The following characteristics of machine learning languages make the choice even more logical.

1. NumPy
2. Pandas
3. Seaborn
4. Matplotlib
5. Scikit-Learn

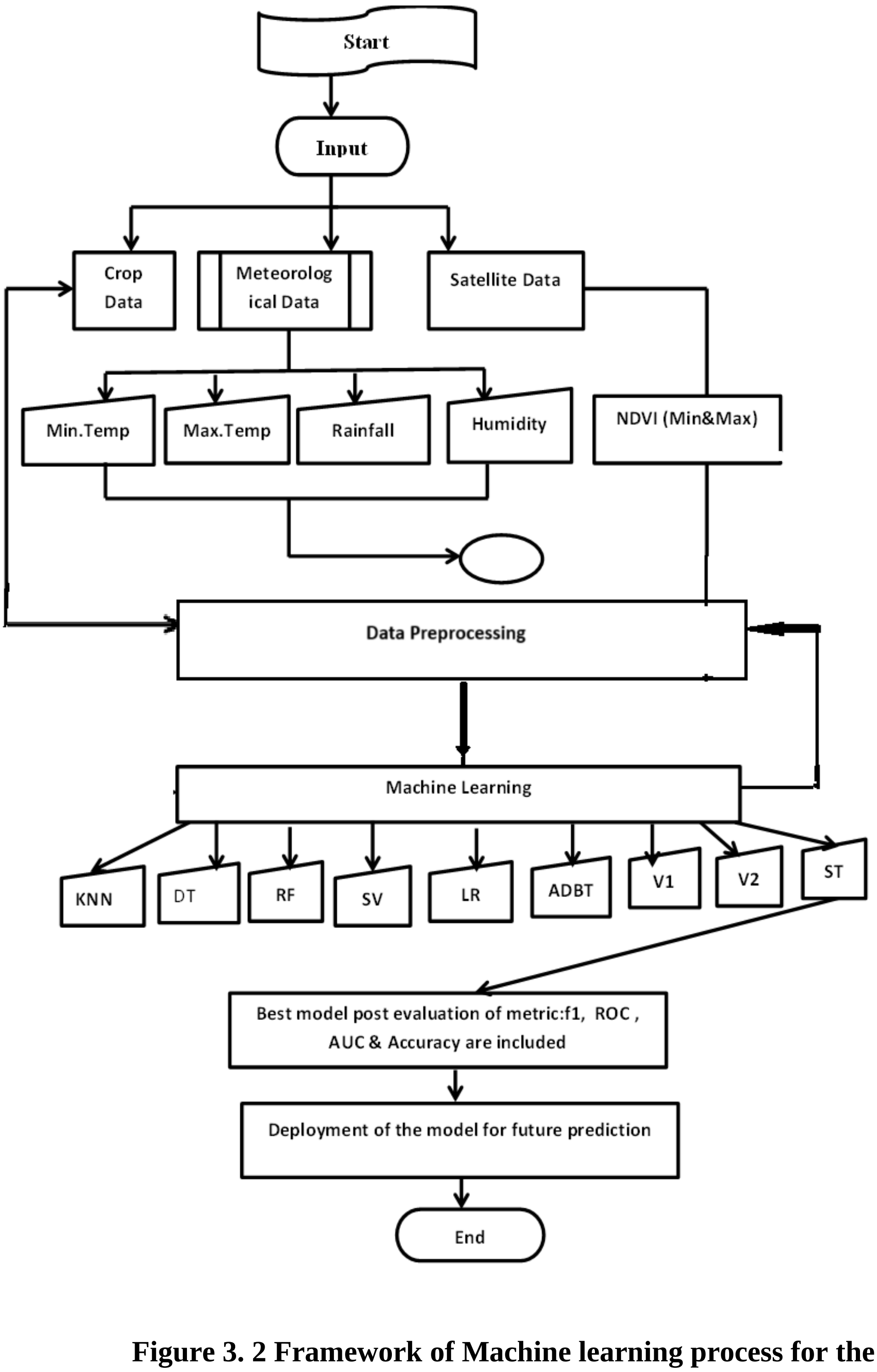


Figure 3. 2 Framework of Machine learning process for the study

3.4 System Architecture

The system architecture for this study will consist of an overview of the proposed system and system components.

3.4.1 Overview of the Proposed System

The proposed system will be a machine learning-based model that can predict the optimal crop for a given climate environment. The system will take weather data as inputs and output the optimal crop for the given environment.

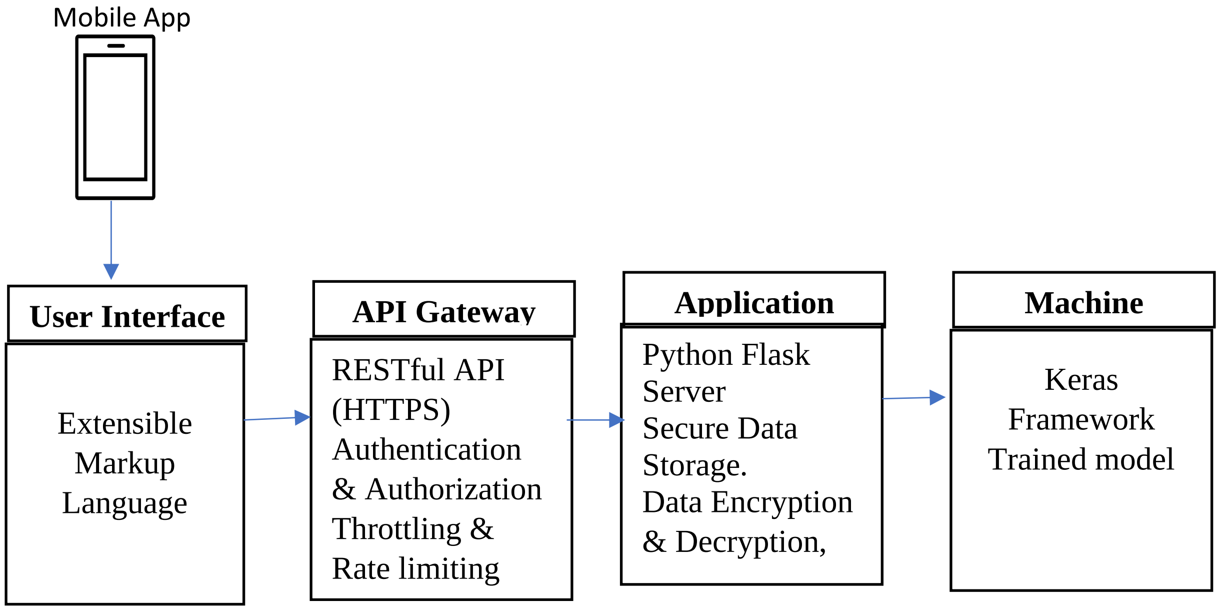


Figure 3.4.1 System's Architecture flow diagram (Authors construct)

To secure the system, we have implemented several measures:

API Gateway: This component acts as a gatekeeper to the system and controls access to the APIs. It provides authentication and authorization to ensure that only authorized users can access the APIs. It also includes throttling and rate limiting to prevent DDoS attacks and API key management to control access to the APIs.

RESTful/Flask API (HTTPS): The APIs are secured with HTTPS to encrypt data in transit and prevent man-in-the-middle attacks.

Application Server: The application server is responsible for processing requests from the mobile app and executing the machine learning model. The server is secured with HTTPS to encrypt data in transit, and data is stored in a secure data storage location with data encryption and decryption capabilities. Access control is implemented to ensure that only authorized users can access the server.

Machine learning Model: The machine learning model is trained on secure data and deployed in a secure environment. It is accessed only through the application server, and access to the server is controlled through authentication and authorization.

3.5 System Components

User Interface (Mobile App)

The user interface (UI) will be designed to be simple and intuitive, with a focus on providing farmers with the information they need in a clear and concise manner. The UI will be developed using the React Native framework, which is a popular framework for building mobile apps.

APIs

The APIs will be used to retrieve weather data from the Meteorological department in Sunyani and crop growth data from the Ministry of Food and Agriculture (MOFA) in Techiman. These APIs will be accessed using RESTful API calls, which will be integrated into the data input module.

Database

A database will be used to store the weather data entered by the app users, as well as the output generated by the machine learning model. The database will be implemented using PostgreSQL, a powerful open-source relational database management system.

Data Input Module

The data input module will retrieve weather and crop data from the APIs, preprocess the data, and transform it into a format that can be used by the machine learning model. This module will be developed using Python and will utilize the Pandas library for data processing.

Machine learning Model Module

The machine learning model module will contain the selected machine learning models, such as K nearest neighbor (KNNs), Random Forest or Support Vector Machine for predicting the optimal crop for a given climate environment. The model will be developed using Keras, a popular open-source machine learning framework.

Output Module

The output module will generate the predicted optimal crop for the given environment based on the input data. This module will be developed using Python and will output the results in a user-friendly format on the mobile app UI.

3.6 Operational Methodology

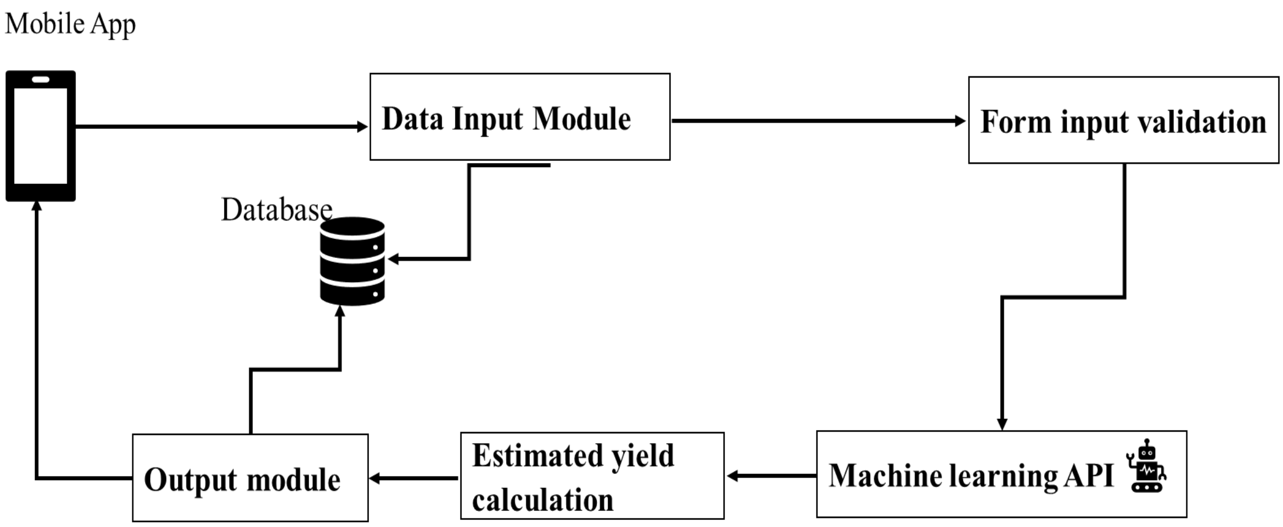


Figure 3.7 Operational Methodology architecture for the proposed system

The operational methodology for this study will include model training and validation, and deployment in a mobile app.

- The operational methods of the proposed system include the following:
- Collecting weather or climate condition data from reliable sources.

- Pre-processing the collected data to remove outliers and anomalies.
- Training the machine learning model using the pre-processed data.
- Deploying the machine learning model in a user-friendly mobile app.
- Providing farmers with a simple user interface for inputting climate condition data and viewing the predicted crops that will survive in that climate condition.

3.8 Methodological Limitations

The limitations of the proposed methodology include the following:

- Limited availability of weather or climate condition data for some areas.
- The accuracy of the predictions may be affected by the quality and quantity of the input data.
- The machine learning model may require continuous updating to adapt to changing climate conditions.

3.9 Summary of Methodology

The proposed methodology involves collecting weather or climate condition data, pre-processing the data, training a machine learning model to predict which crops will survive in a given climate condition, and deploying the model in a user-friendly mobile app for farmers to use. The methodology has limitations such as the availability and quality of data, but it is expected to provide useful information for farmers in selecting crops to grow in a particular climate condition.

CHAPTER FOUR

RESULTS

4.1 Introduction

This chapter presents the analysis of the study and deployment of the various models in a mobile application. The results and discussion of development of the mobile application and the deployment of the models K-nearest neighbor, Decision tree, Random Forest, Support Vector Machine, Logistic regression, AdaBoost ensemble, majority voting and stacking ensemble.

4.2 Data and models parameters overview

In this section, we provide a succinct overview of key components, including a visual representation of our sample dataset through screenshots, a correlation matrix illustrating the pairwise relationships among features, and a concise summary of the parameters used in our machine learning model implementations. These elements collectively offer a comprehensive glimpse into our research methodology and data analysis.

Screenshot of sample dataset, Correlation matrix-showing pairwise correlation among features, Parameter summary of machine learning model implementation

	MIN_TEMP	MAX_TEMP	RH150	RH60	Rainfall	Area cropped(Ha)	Yield (Mt/Ha)	MIN_NDVI	MAX_NDVI	CROP
0	22.5	32.0	42.0	82	7.7	6820	17.65	0.1275	0.8474	1
1	22.7	35.5	38.0	79	44.1	5047	18.90	0.0563	0.8513	1
2	23.3	34.1	50.0	87	79.6	7852	2.40	0.0928	0.7488	1
3	23	34.2	55.0	88	124.3	1704	9.80	0.1570	0.8917	1
4	22.8	32.6	61.0	91	154.7	276	9.96	0.0749	0.8650	1

Figure 4. 1 Screenshot of sample dataset

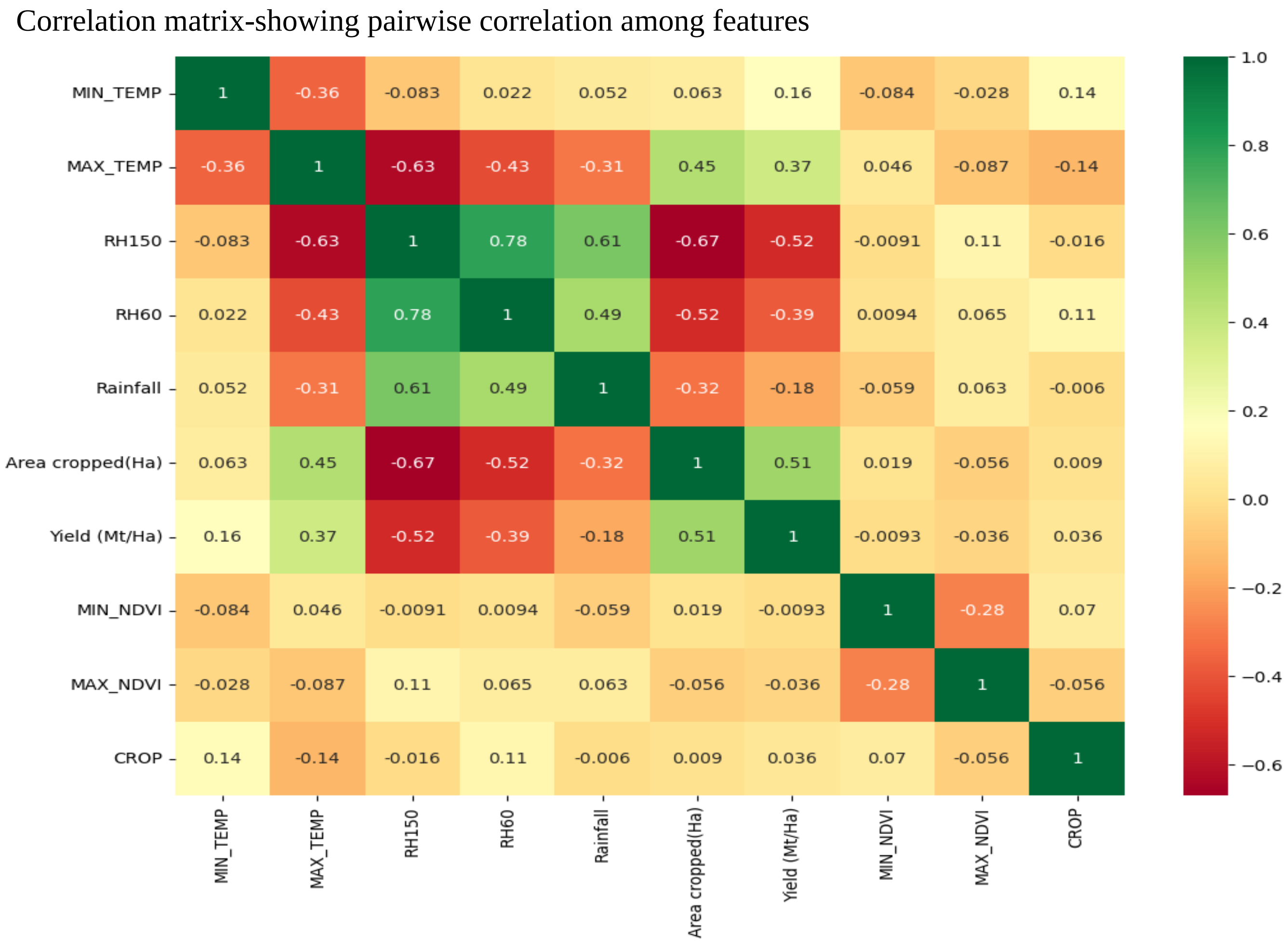


Figure 4. 2 representation of correlation matrix using heatmap

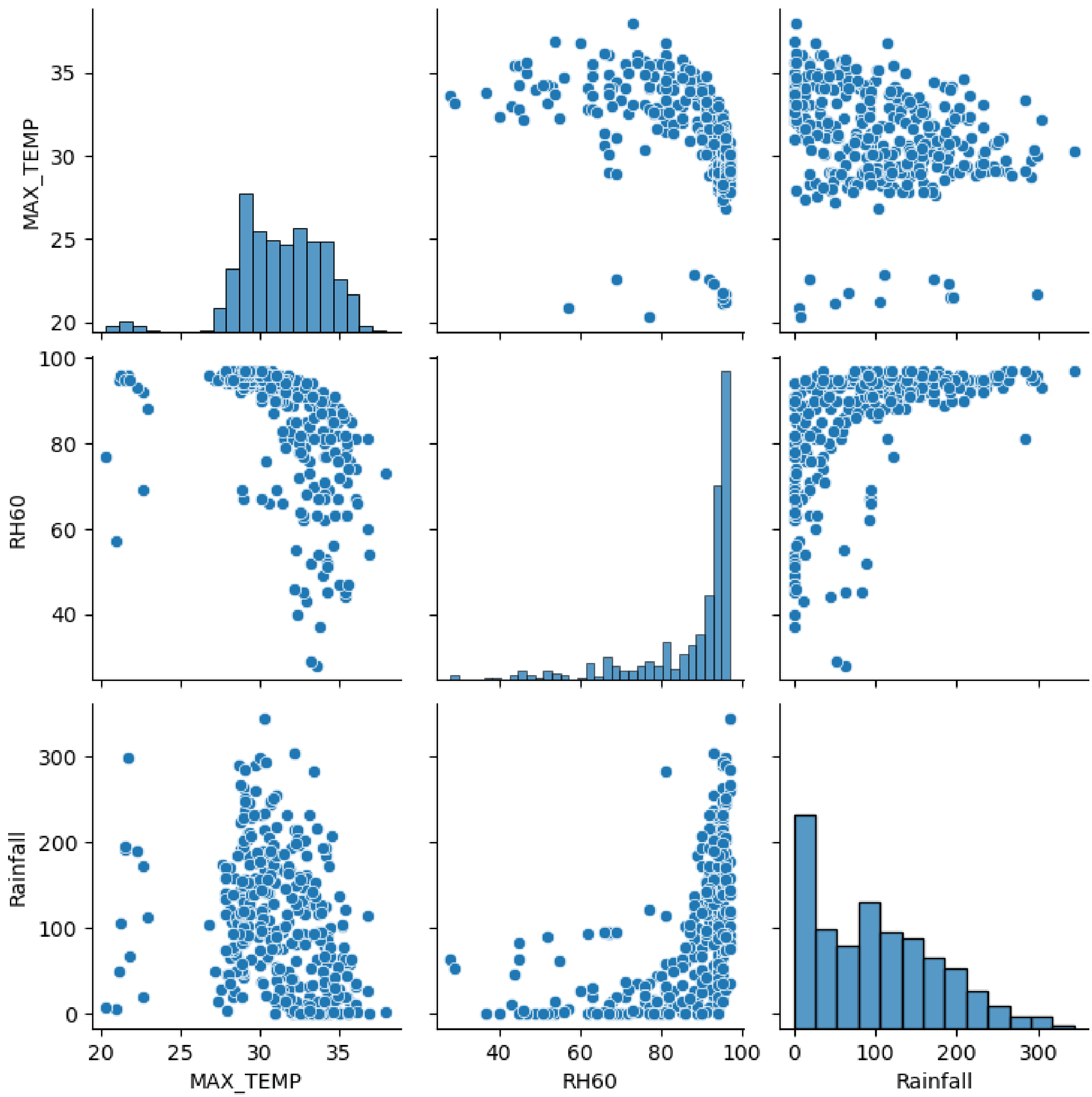


Figure 4. 3 Data Distribution Plot for Some of the Features

Table 4. 1 Parameter summary of machine learning model implementation

Machine Learning Models	Parameters
Knearest Neigbor	n_neighbors=2, metric= "minkowski"
Decision Tree	criterion = "gini", min_samples_leaf=2, random_state = 0
Random Forest	criterion = "gini", min_samples_leaf=2,n_estimators=10, random_state =0
Support Vector Machine	kernel="linear", random_state = 0
Logistic Regression	()
Adaboost	base_estimator=RFC
Voting 1	Estimators="DTC",dtc, "rf",RFC, "knn",knn, voting='soft', weights=None, n_jobs=None, flatten_transform=True,
Voting 2	Estimators="DTC",dtc,"rf",RFC,"knn",knn, voting='hard', weights=None, n_jobs=None, flatten_transform=True,
Stacking	Estimators=" DTC",dtc, "rf",RFC, "knn",knn,

The random state value is set to zero, as shown in the table (0). In essence, the random state is an instance of NumPy that is an integer. Which random number generator to use is specified by RandomState, None, or (default). In a sense, the zero value serves as a model's default value.

With the same random state value set to zero, the decision tree also used gini as a criterion. A number of 10 estimators were used by the random forest classifier. The neighbor selection for KNN is five, and the distance measure used is the minkowski's distance. Two variations of SVM was used were the radial basis function and linear technique.

4.3 Results discussion

Before taking into account any relevant evaluation metrics, the machine learning model's training and testing accuracy must first be generated. A summary of the training and testing accuracies attained after implementation is provided in Table 4.2 (training and testing accuracies of the various machine learning models.)

4.4 Algorithms selection

A wealth of tools and resources are available in the Scikit library to test run interesting algorithms.

Table 4.4.1 Various machine learning models' with their training and testing accuracy

Machine learning model	**Testing accuracy**	**Training accuracy**
K-nearest Neighbor	0.68	0.73
Decision Tree	0.83	0.98
Random Forest	0.85	0.95
Support Vector Machine	0.55	0.55
Logistic regression	0.61	0.63
AdaBoost	0.57	0.59
Voting 1(hard)	0.85	0.91
Voting 2 (soft)	0.86	0.93
Stacking	0.85	0.92

The testing and training accuracies were computered through the use of the confusion matrix using Jupiter notebook (anaconda).

4.5 Classification results

The following tables give a summary report of the various machine learning models considered in the testing phase. It should be noted that the results of the label representation Yam is represented by 1 and cassava represented by 2.

Table 4. 5.1 K-nearest neighbor classification results

	precision	recall	f1-score	support
1	0.74	0.66	0.70	89
2	0.74	0.81	0.77	108
accuracy			0.74	197
macro avg	0.74	0.73	0.74	197
weighted avg	0.74	0.74	0.74	197

Table 4. 5.2 decision tree classification results

	precision	recall	f1-score	support
1	0.80	0.89	0.84	89
2	0.90	0.81	0.85	108
accuracy			0.85	197
macro avg	0.85	0.85	0.85	197
weighted avg	0.85	0.85	0.85	197

\

Table 4.5.3 random forest classification results

	precision	recall	f1-score	support
1	0.85	0.93	0.89	89
2	0.94	0.86	0.90	108
accuracy			0.89	197
macro avg	0.89	0.90	0.89	197
weighted avg	0.90	0.89	0.89	197

Table 4.5.4 Support Vector Machine classification results

	precision	recall	f1-score	support
1	0.51	0.87	0.64	89
2	0.74	0.32	0.45	108
accuracy			0.57	197
macro avg	0.63	0.59	0.55	197
weighted avg	0.64	0.57	0.54	197

Table 4.5.5 Logistic Regression classification results

	precision	recall	f1-score	support
1	0.64	0.50	0.56	219
2	0.62	0.74	0.68	238
accuracy			0.63	457
macro avg	0.63	0.62	0.62	457
weighted avg	0.63	0.63	0.62	457

Table 4.5.6 AdaBoost classification results

	precision	recall	f1-score	support
1	0.70	0.81	0.75	89
2	0.82	0.71	0.76	108
accuracy			0.76	197
macro avg	0.76	0.76	0.76	197
weighted avg	0.76	0.76	0.76	197

Table 4.5.7 Hard Voting classification results

	precision	recall	f1-score	support
1	0.50	0.91	0.65	89
2	0.78	0.26	0.39	108
accuracy			0.55	197
macro avg	0.64	0.58	0.52	197
weighted avg	0.65	0.55	0.51	197

Table 4.5.8 Soft Voting classification results

	precision	recall	f1-score	support
1	0.85	0.94	0.89	89
2	0.95	0.86	0.90	108
accuracy			0.90	197
macro avg	0.90	0.90	0.90	197
weighted avg	0.90	0.90	0.90	197

Table 4.5.9 Stacking classification results

	precision	recall	f1-score	support
1	0.91	0.93	0.92	89
2	0.94	0.93	0.93	108
accuracy			0.93	197
macro avg	0.93	0.93	0.93	197
weighted avg	0.93	0.93	0.93	197

4.6 Observation of the classification reports

As can be seen from the classification report's tabular representation. Each machine learning model has different results in terms of particular parameters. These variables are described below.

Precision: It displays the percentage of predictions that are accurate in terms of numbers. This is speaking of the precision of a favorable prediction.

Recall: The ratio of true positives to the total number of true positives is meant by this. The percentage of positives that were accurately detected is essentially what it is.

F1 score**:** the f 1 score is a weight harmonic mean of precision and recalls support score is 1.0 and the worst is 0.0 F 1 scores are lower than accuracy measures precision and recall into their computation.

Support: Support is the number of instances of the class that really occur in the training data. Low support can suggest that the classifier has structural weaknesses, and that stratified sampling is required.

Confusion matrix reoport: representation of the different combinations of actual values versus predicted values.

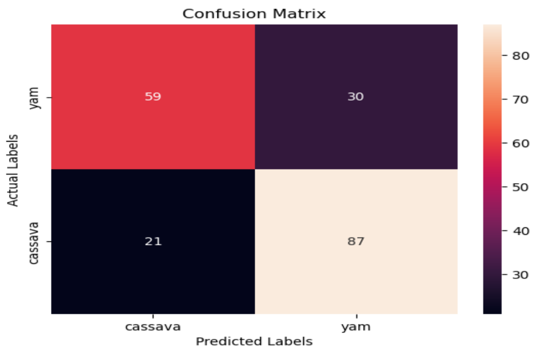


Figure 4. 4 K-nearest neighbor

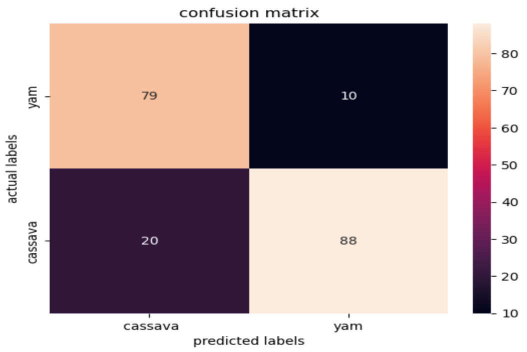


Figure 4. 5 decision tree

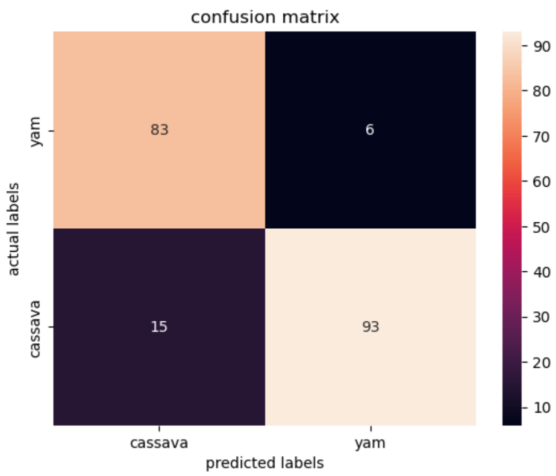


Figure 4. 6 Random Forest

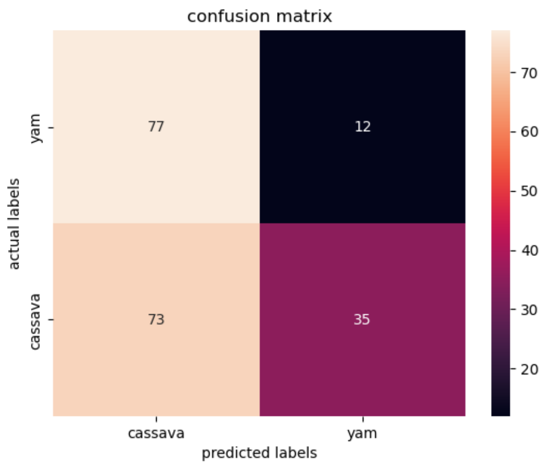


Figure 4. 7 Support Vector machine

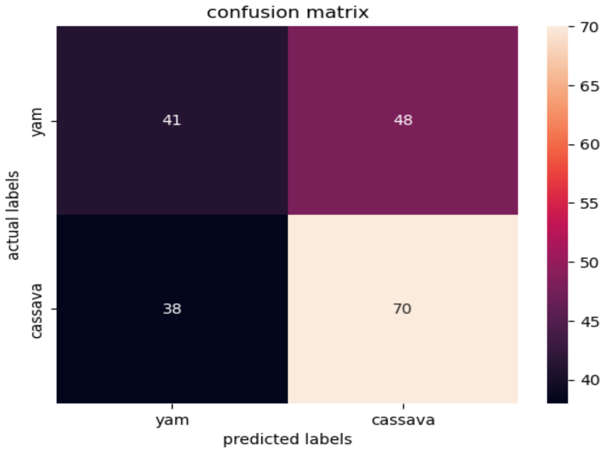


Figure 4. 8 Logistic Regression

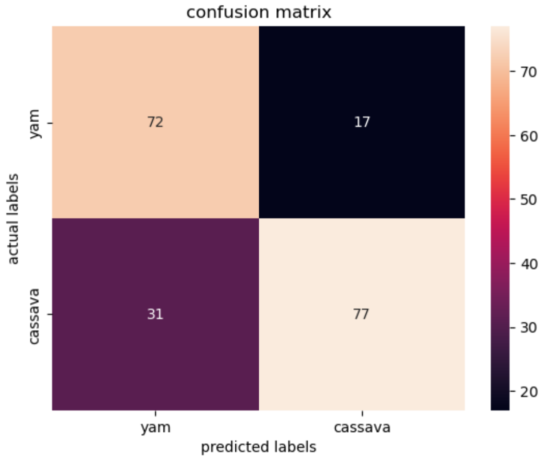


Figure 4. 9 AdaBoost

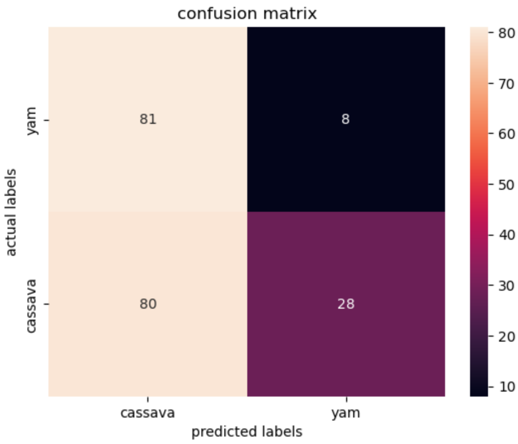


Figure 4. 10 Hard Voting

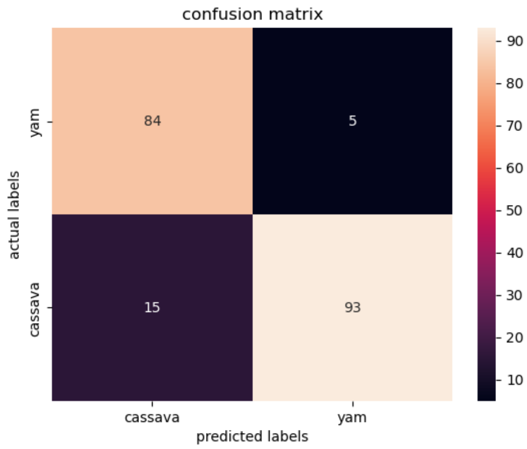


Figure 4. 11 Soft Voting

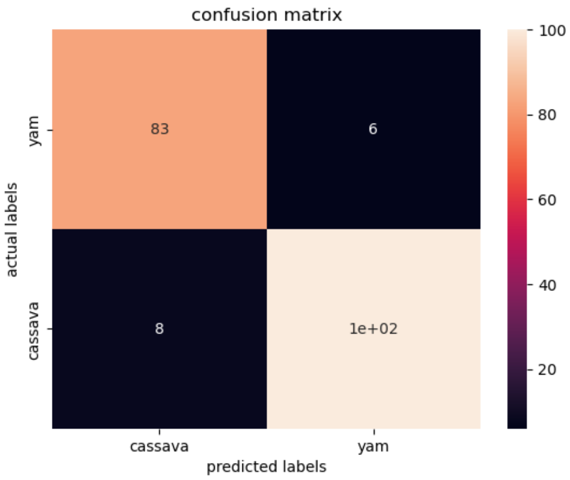


Figure 4. 12 Stacking Classifier

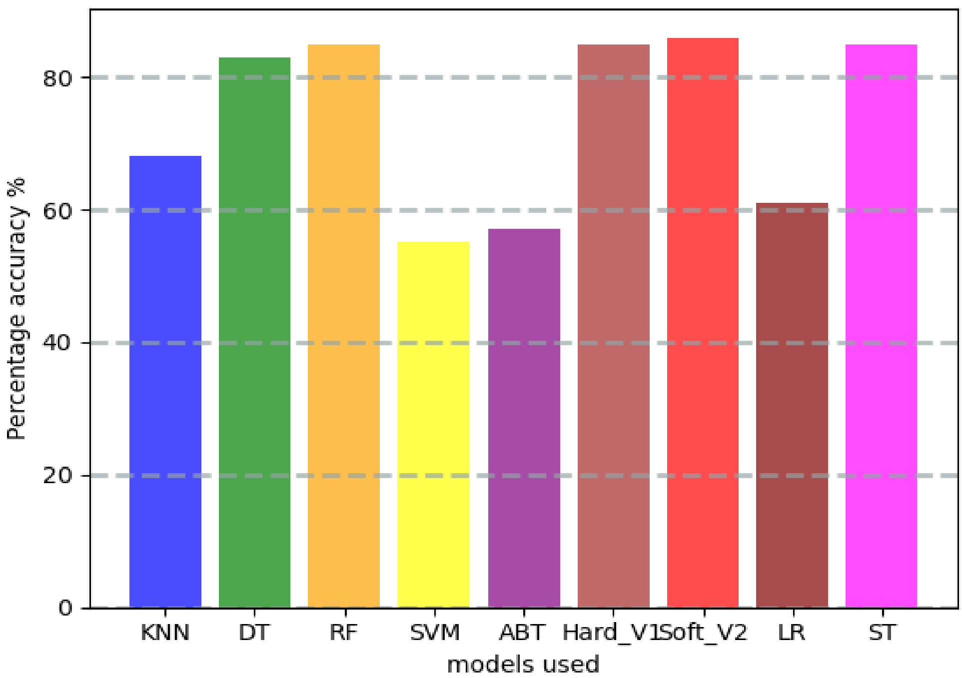


Figure 4. 13Accuracy Comparison among All Classifiers for Specific Crop Possibility

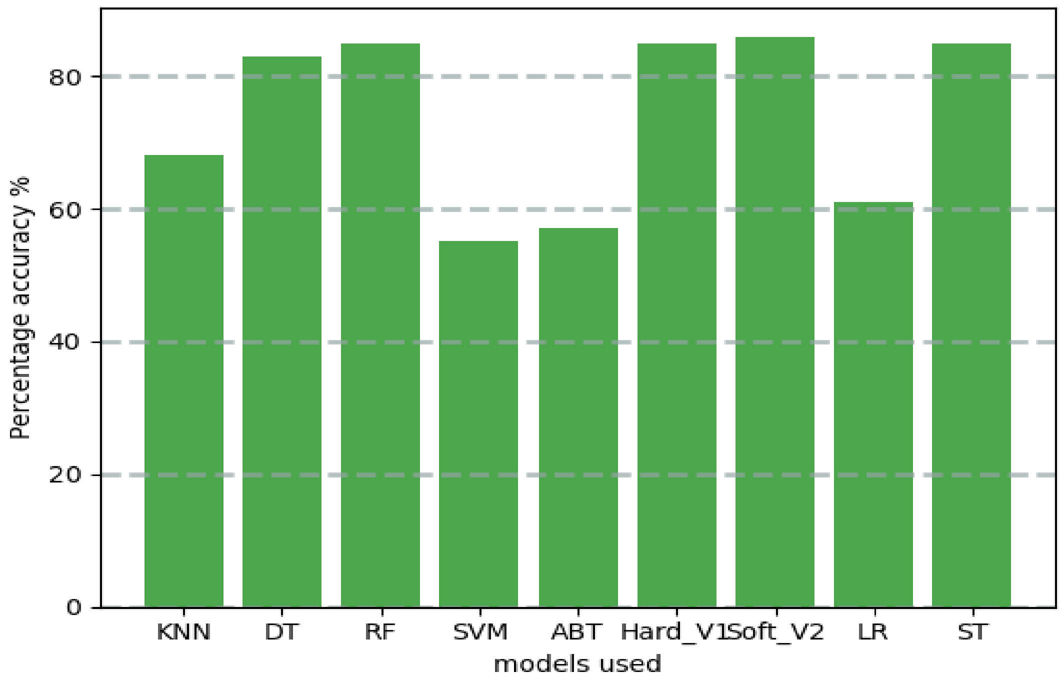


Figure 4. 14 Accuracy Comparison among All Classifiers after Cross validation

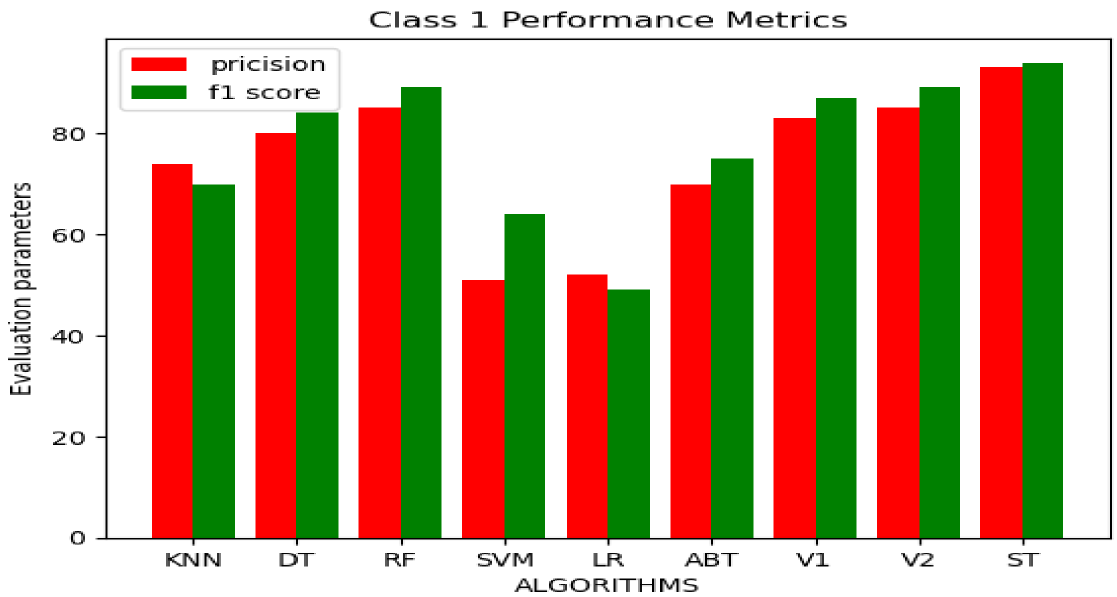


Figure 4. 15 Graph of F1 Score and Precision among all the classifiers

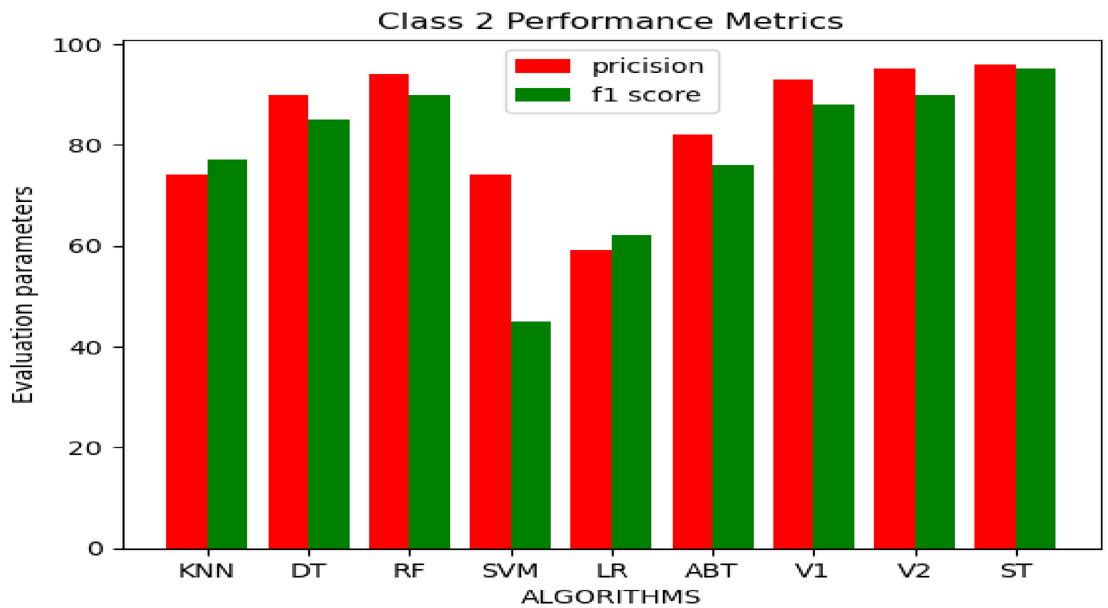


Figure 4. 16 Graph of F1 Score and Precision among all the classifiers

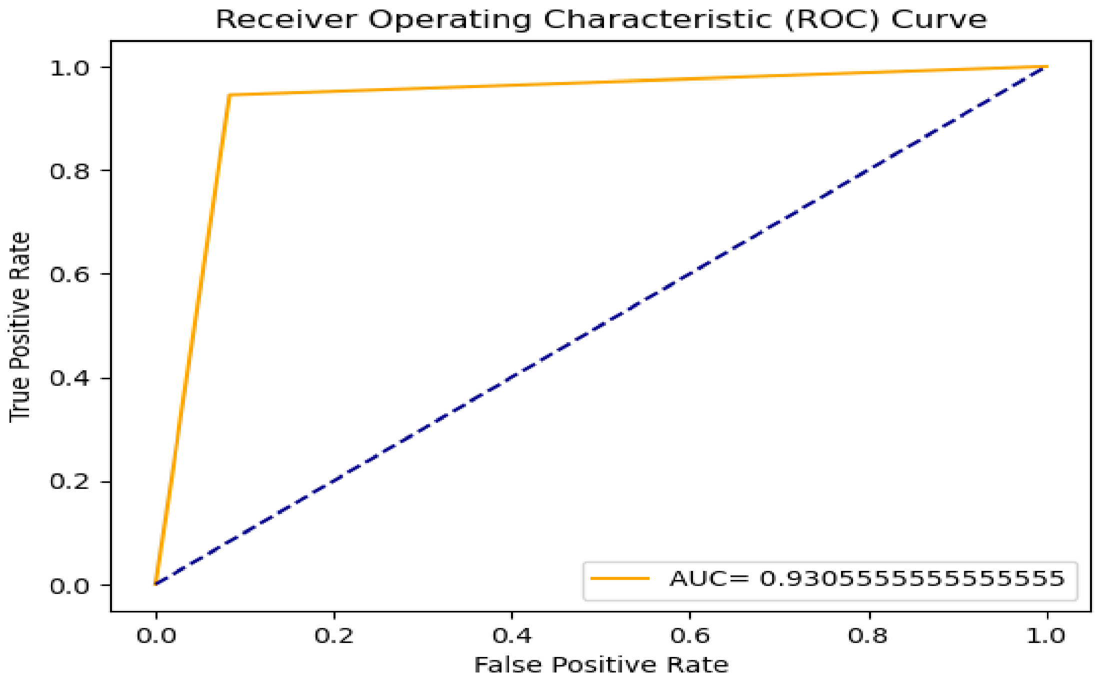


Figure 4.17 ROC curve and AUC for Random Forest model

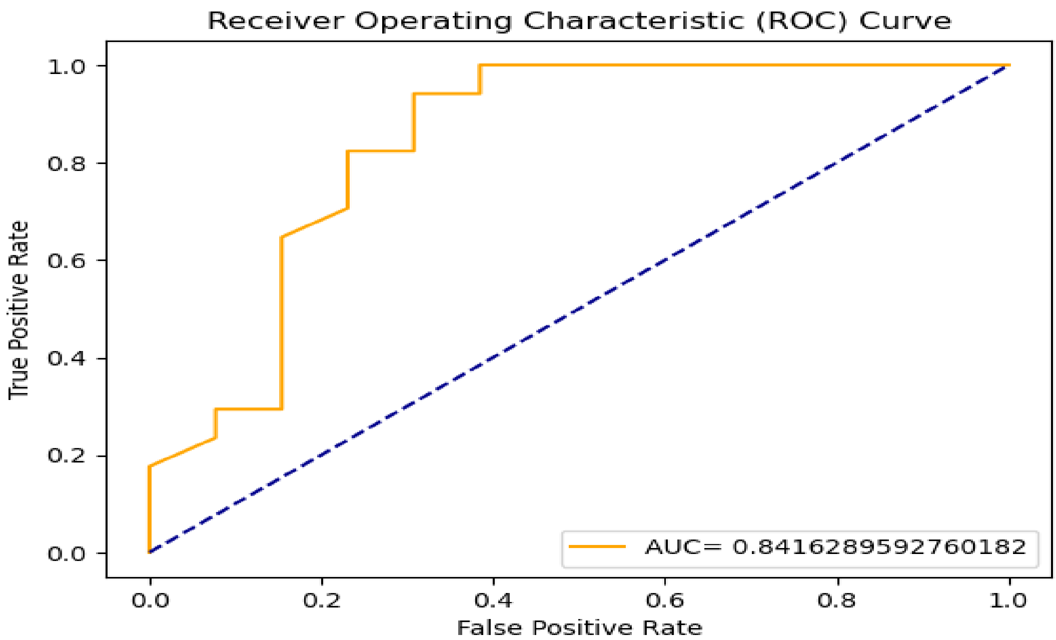


Figure 4. 18 ROC curve and AUC for logistic regression model

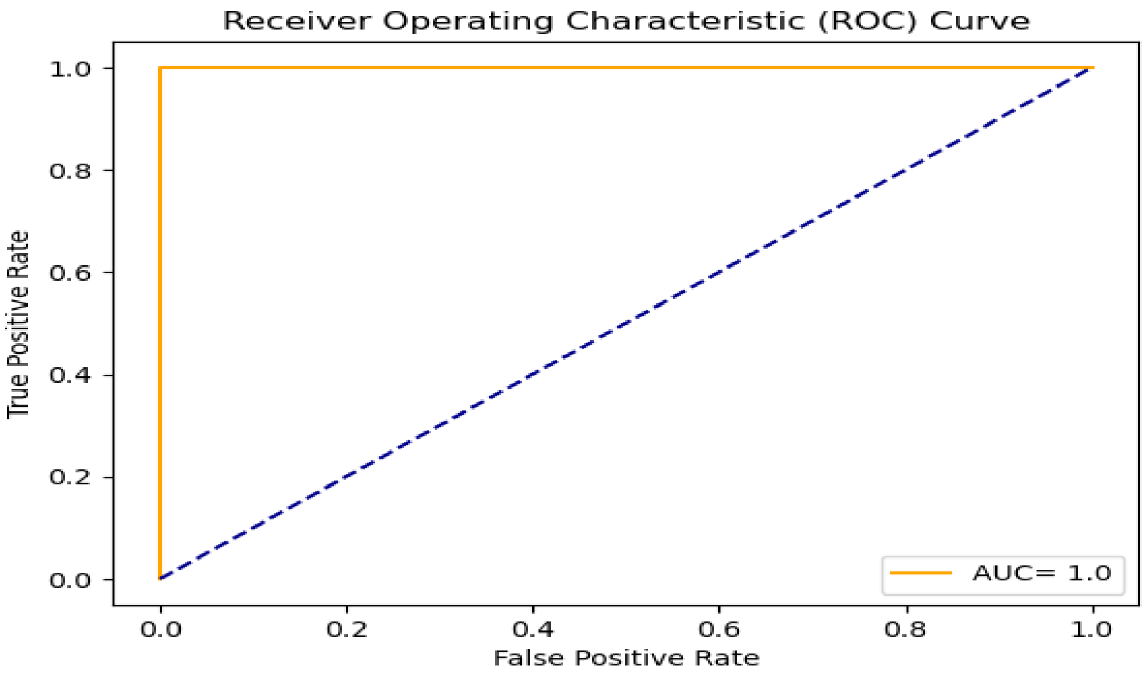


Figure 4. 19 ROC curve and AUC for Random Forest model

4.7. Models' generalization and deployment

4.7.1 Generalization rule

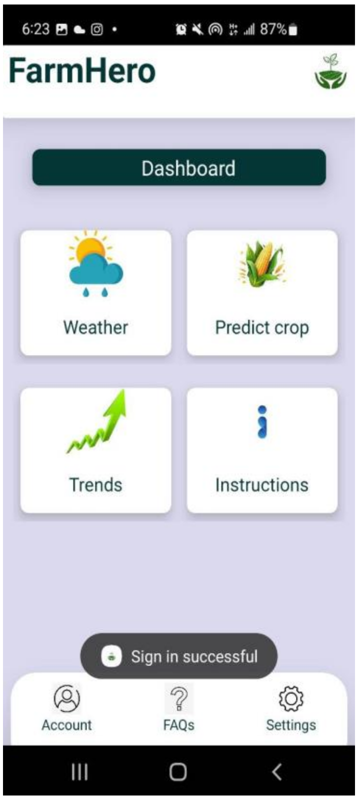
During the deployment of machine learning models designed for predicting suitable crops based on specific climate conditions, a critical aspect of the study involved hyperparameter tuning. The optimization of model hyperparameters was pursued with the aim of enhancing the predictive accuracy and generalization capabilities of the deployed models. Through a systematic exploration of various hyperparameter combinations, adjustments were made to key architectural parameters, including the neural network's depth and width. Additionally, hyperparameters governing the training process, such as learning rate, regularization strength, and batch size, were fine-tuned to strike an optimal balance between model complexity and performance. This rigorous hyperparameter tuning process contributed significantly to the models' robust performance during deployment. By achieving an optimal configuration, the deployed models successfully provided accurate predictions of suitable crops under diverse climate conditions, demonstrating their practical utility and effectiveness in real-world applications.

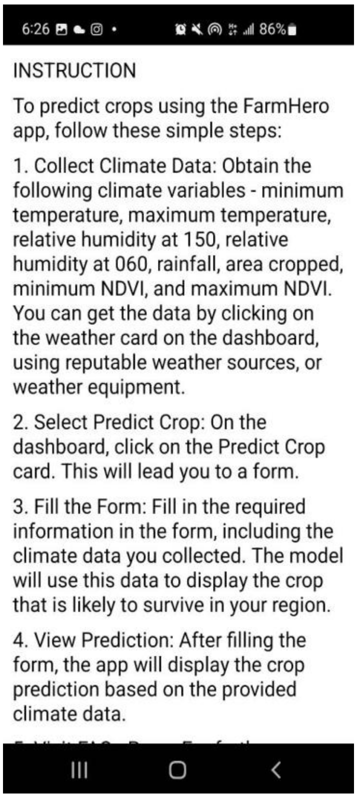
4.7.2 Model deployment

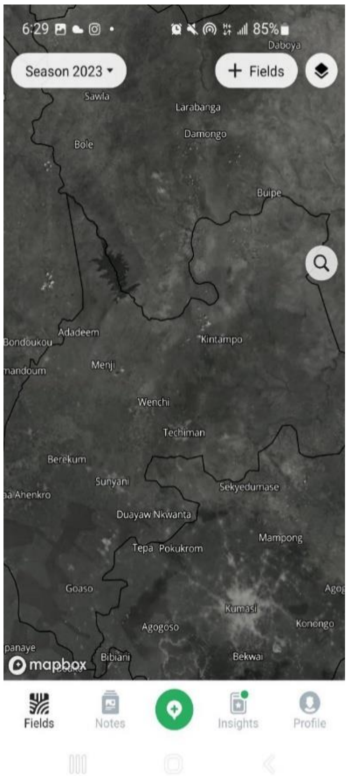
In the deployment phase of this study, a Flask server was developed to facilitate the integration and utilization of various predictive models for crop suitability assessment. The server was configured to handle incoming requests from an Android application designed for predicting crops based on provided climate conditions. After establishing a connection between the server and the Android app, we were able to predict most of the crops in the datasets used in the study, which represents a success of the project and the achievement of the study objectives.

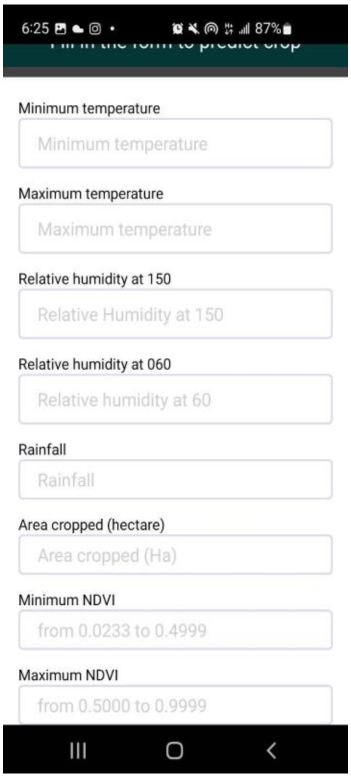
Within the Flask server environment, a range of models, including Random Forest, k-Nearest Neighbors (knn), Stacking Classifier, Hard Voting, Soft Voting, and Decision Tree, were imported. These models were selected for deployment due to their high accuracy percentage as compared to the remaining models. The models formed the predictive backbone of the application, each contributing its unique strengths to the ensemble.

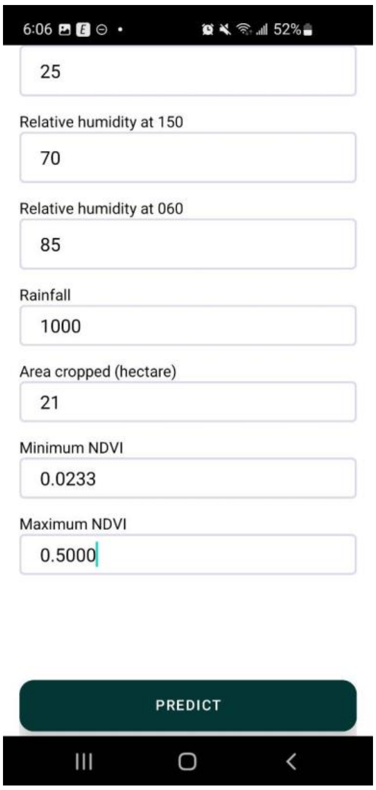
During the testing phase of the Android application, the Flask server was hosted on a local machine, and a localhost link was established to establish communication between the server and the "Predict Crop" form within the Android app. This configuration enabled real-time predictions to be generated based on user-input climate conditions. Subsequently, the server's deployment transitioned to a more robust infrastructure on Heroku. By doing so, the application's availability and reliability were significantly enhanced, allowing a broader user base to access the crop prediction service seamlessly. This transition to Heroku hosting exemplifies the commitment to delivering dependable and accessible services to users seeking informed decisions regarding crop cultivation in varying climatic contexts. Below are some of the images from the mobile application

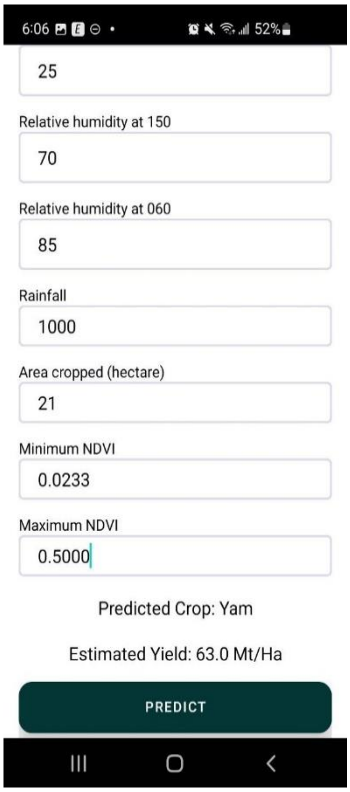


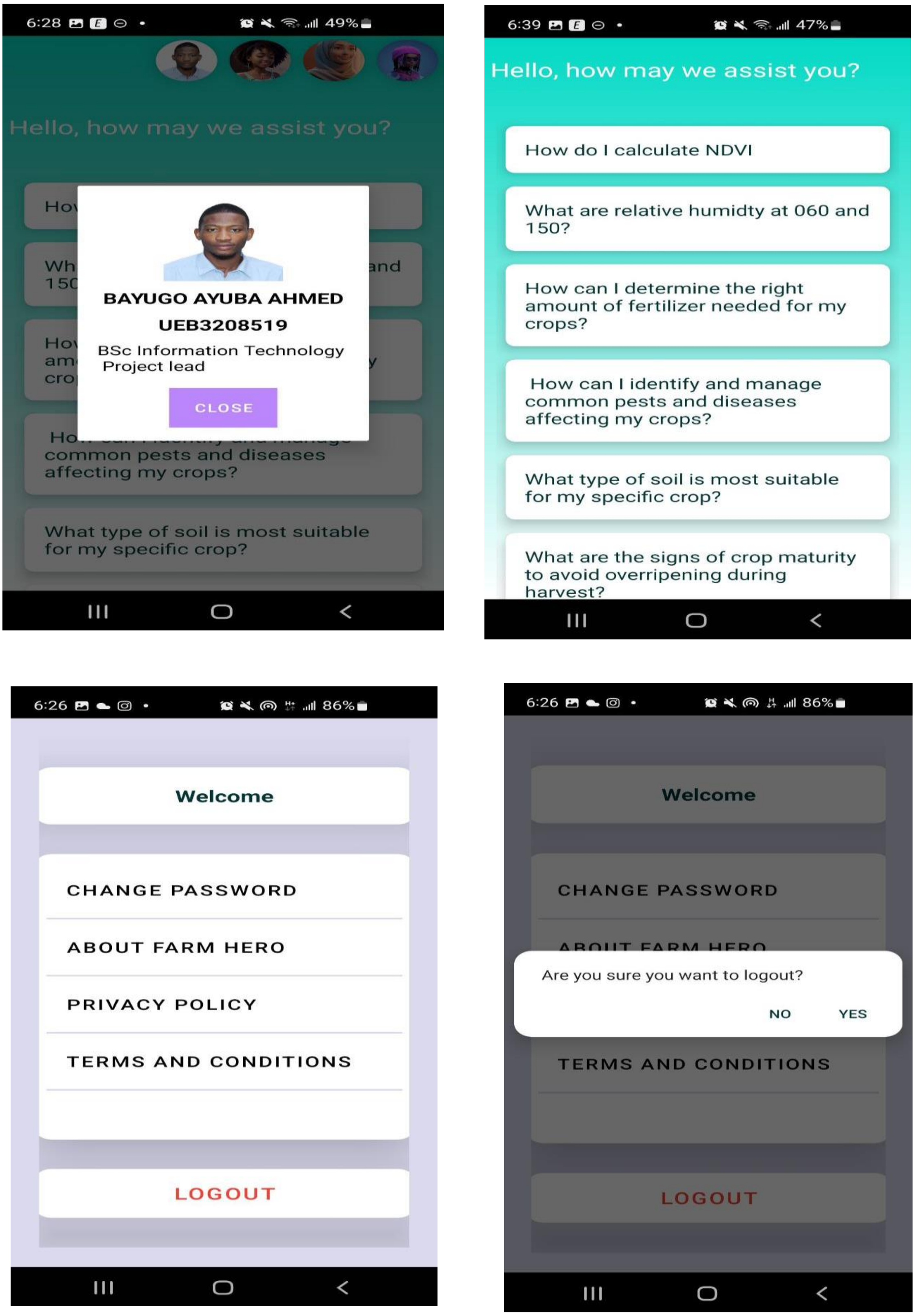












Screenshots from the mobile app features

4.8. Discussion

In this section, we discuss the models utilized in this study concerning their precision, the generation of the confusion matrix report, various assessment metrics, and graphical representations. The primary objective of this research is to develop and deploy a machine learning model(s) capable of recommending a suitable crop to grow based on data encompassing crop production, satellite imagery, and climate data. Within this investigation, we have successfully constructed models (depicted in Figure 4.13) designed to determine the feasibility of growing specific crops on a given plot of land using a defined set of attributes. Figure 4.14 presents a comparative analysis of accuracy following cross-validation, assessing the performance of different classifiers employed. Additionally, Figure 4.15 displays graphs depicting the F1 Score and Precision, which assess the classifiers' performance. Notably, five of the algorithms demonstrated accuracy exceeding 80%, with the lowest-performing classifiers achieving accuracy above 50%. Moving on to Figures 4.16 through 4.18, we present graphs illustrating the Receiver Operating Characteristic (ROC) and Area Under the Curve (AUC). All AUC values surpass 0.82, confirming the robustness and superior performance of the models.

Nine different machine learning algorithms were employed to implement the models described above, comprising five individual learning algorithms and four ensemble algorithms. This diverse selection ensured a comprehensive comparison of various algorithms. The dataset was split into a training-test ratio of 70%-30%, and ultimately, all models produced satisfactory results, as outlined in Table 4.2. However, it's worth noting that the Support Vector Machine (SVM) did not perform well in terms of testing accuracy.

The most successful model emerged as the Soft Voting Classifier, following an evaluation of the different classifiers used. Utilizing Majority Voting and Stacking ensemble techniques resulted in improved accuracy, leading to the selection of these models as the final choice. The outcomes generated by these algorithms enabled both the prediction of spe cific crops and the estimation of the likelihood of selecting a particular crop based on specific features.

In this research, the Soft Voting classifier achieved the highest accuracy at 85.8%, as highlighted in Table 4.2, which provides a comparison among all the classifiers employed for crop possibility selection. Binary confusion matrices for Yam and Cassava, depicting correctly and incorrectly predicted records, are illustrated in Figures 4.4 to 4.10. Furthermore, Tables 4.11 to 4.16 and Figure 4.15 present the F1 score and precision for both the first score (1) and the second score (2) of the models.

Additionally, it's worth mentioning that in the realm of binary classification, researchers typically do not frequently employ evaluation performance metrics like Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared (R2). However, this paper introduced these statistical measures in Table 4.17 to validate the performance of the classification classifiers, providing a comprehensive assessment of the classifiers' accuracy.

CHAPTER FIVE:

CONCLUSION AND RECOMMENDATIONS

5.1 CONCLUSION

This machine learning-powered mobile application will play a crucial role in modern-day precision agriculture, particularly in the context of government initiatives like "one-district-one-factory." The primary aim of this model is to contribute to the increased availability of raw materials required for these factories. Agriculture holds a pivotal position in the development of a country, especially in developing nations like Ghana. In the face of escalating global hunger and economic challenges, the selection of the right crop emerges as a vital component in agricultural advancement.

The proposed machine learning models have demonstrated notable effectiveness and efficiency in determining the optimal crop for a given plot of land. This research encompassed the implementation of nine distinct machine learning models, comprising four ensemble learning and five weak learning algorithms. The accuracy of each model was meticulously evaluated using various techniques, with the results presented in Table 4.2 for comparison.

These models were employed to ascertain the most suitable crop for cultivation on specific parcels of land, leveraging crop production data, climate data, and satellite data after comprehensive training and testing. A noteworthy finding of this research is that as the dataset size increased, the models' accuracy improved, highlighting the potential of machine learning in crop selection.

The application of machine learning to determine crop suitability has added a compelling dimension to the entire research endeavor. It is worth noting that all research findings and algorithms seamlessly adapt to different platforms and diverse datasets. The models demonstrated reliability and robustness throughout the study, affirming their utility and versatility.

5.RECOMMENDATIONS

Base on the study conducted on the 9 machine learning models, it can be recommended that;

Incorporating real-time weather data: Future research can explore the integration of up-to-date weather information to improve the accuracy of crop suitability assessments. This would enable the models to adapt dynamically to changing climate conditions.

Future research endeavors can consider the integration of Internet of Things (IoT) technology to further enhance predictive capabilities.

Future studies may also benefit from exploring the impact of fertilizer application on crop suitability and performance.

Expanding the dataset's scope by incorporating information about soil compositions is recommended. This addition can help prevent land from going to waste in the era of precision agriculture, ensuring more informed land-use decisions.

References:

Addo, P. C., & Opoku, R. (2020). The impact of agricultural technologies on rural livelihoods in Ghana: Insights from the techiman north district. Journal of Agribusiness in Developing and Emerging Economies, 10(4), 388-408.

Agyemang, K., Owusu-Manu, D. G., Amoako-Boateng, E., & Osei, J. (2017). Agriculture in Ghana: Challenges and opportunities. International Journal of Development and Sustainability, 6(9), 1468-1480.

Ali, S. M., Das, B., & Kumar, D. (2021). *Machine Learning based Crop Recommendation System for Local Farmers of*. *11*(4), 5735–5746.

Asante, M. D., Kumah, A. K., & Amponsah, O. (2019). Agricultural productivity in Ghana: The role of gender. African Development Review, 31(4), 428-440.

Babu, S. (2013). *A Software Model for Precision Agriculture for Small and Marginal Farmers*. 352–355.

Bacco, M., Berton, A., Ferro, E., Gennaro, C., Gotta, A., & Matteoli, S. (2014). *Smart Farming : Opportunities , Challenges and Technology Enablers*.

Balducci, F., Impedovo, D., Informatica, D., & Moro, A. (2018). *Machine Learning Applications on Agricultural Datasets for Smart Farm Enhancement*. https://doi.org/10.3390/machines6030038

Behmann, J., Mahlein, A., Rumpf, T., Ro, C., & Plu, L. (2014). *A review of advanced machine learning methods for the detection of biotic stress in precision crop protection*. https://doi.org/10.1007/s11119-014-9372-7

Butler, E., Timlin, D. J., & Forests, R. (2016). *Random Forests for Global and Regional Crop Yield Predictions The Harvard community has made this Random Forests for Global and Regional Crop Yield Predictions*. https://doi.org/10.1371/journal.pone.0156571

Cai, Y., Guan, K., Lobell, D., Potgieter, A. B., Wang, S., Peng, J., Xu, T., Asseng, S., Zhang, Y., You, L., & Peng, B. (2019). Agricultural and Forest Meteorology Integrating satellite and climate data to predict wheat yield in Australia using machine learning approaches. *Agricultural and Forest Meteorology*, *274*(May 2018), 144–159. https://doi.org/10.1016/j.agrformet.2019.03.010

Cai, Y., Li, X., Zhang, M., & Lin, H. (2020). Int J Appl Earth Obs Geoinformation Mapping wetland using the object-based stacked generalization method based on multi-temporal optical and SAR data. *Int J Appl Earth Obs Geoinformation*, *92*(January), 102164. https://doi.org/10.1016/j.jag.2020.102164

Palanivel, K., Surianarayanan, C., & Author, C. (2019). *AN APPROACH FOR PREDICTION OF CROP YIELD USING MACHINE LEARNING AND BIG*. *10*(03), 110–118.

Paswan, R. P., & Begum, S. A. (2013). *Regression and Neural Networks Models for Prediction of Crop Production*. *4*(9), 98–108.

Patel, K., & Patel, H. B. (2020). A state-of-the-art survey on recommendation system and prospective extensions. *Computers and Electronics in Agriculture*, *178*(June), 105779. https://doi.org/10.1016/j.compag.2020.105779

Paul, M., Vishwakarma, S. K., & Verma, A. (2015). *Analysis of Soil Behaviour and Prediction of Crop Yield using Data Mining Approach*. https://doi.org/10.1109/CICN.2015.156

Peerlinck, A., Sheppard, J., & Senecal, J. (2019). AdaBoost with Neural Networks for Yield and Protein Prediction in Precision Agriculture. *2019 International Joint Conference on Neural Networks (IJCNN)*, *July*, 1–8.

Pourdarbani, R., Sabzi, S., Kalantari, D., & Hernández-hernández, J. L. (2020). *A Computer Vision System Based on Majority- Voting Ensemble Neural Network for the Automatic Classification of Three Chickpea Varieties*. 1–18. https://doi.org/10.3390/foods9020113

Prajwala. (2015). A Comparative Study on Decision Tree and Random Forest Using R Tool. *Article*, *November*. https://doi.org/10.17148/IJARCCE.2015.4142

R. K. Rajak, A. Pawar, M. Pendke, P. Shinde, S. Rathod, A. D. (2017). *Crop Recommendation*

System to Maximize Crop Yield using Machine Learning Technique. *04*(12 Dec, 2017).

Raut, J., & Mittal, S. (2021). *Soil Fertility and Crop Recommendation using Machine Learning and Deep Learning Techniques : A Review*. *11*(02), 1119–1127.

Ren, Q., Cheng, H., & Han, H. (2020). *Research on machine learning framework based on random forest algorithm Research on Machine Learning Framework Based on Random Forest Algorithm*. *080020*(March 2017). https://doi.org/10.1063/1.4977376

Saha, A. K., Saha, J., Ray, R., Sircar, S., Dutta, S., & Chattopadhyay, S. P. (2018). *IOT-Based Drone for Improvement of Crop Quality in Agricultural Field*. *January*. https://doi.org/10.1109/CCWC.2018.8301662

Savla, A., & Mandholia, A. (2015). *Survey of classification algorithms for formulating yield prediction accuracy in precision agriculture*.

Shahhosseini, M., Hu, G., Huber, I., & Archontoulis, S. V. (2021). Coupling machine learning and crop modeling improves crop yield prediction in the US Corn Belt. *Scientific Reports*, *11*(1). https://doi.org/10.1038/s41598-020-80820-1

Shetty, S. A., Padmashree, T., Sagar, B. M., & Cauvery, N. K. (2021). *Performance Analysis on Machine Learning Algorithms with Deep Learning Model for Crop Yield Prediction*. Springer Singapore. https://doi.org/10.1007/978-981-15-8530-2

Shruthi, G. (2020). *Design And Implementation Of Crop Yield Prediction Model In Agriculture*. *8*(01).

Shultz, T. R., Fahlman, S. E., Craw, S., Andritsos, P., Tsaparas, P., Silva, R., Drummond, C., Ling, C. X., Sheng, V. S., Drummond, C., Lanzi, P. L., Gama, J., Wiegand, R. P., Sen, P., Namata, G., Bilgic, M., Getoor, L., He, J., Jain, S., … Mueen, A. (2011). Confusion Matrix. *Encyclopedia of Machine Learning*, 209–209. https://doi.org/10.1007/978-0-387-30164-8_157

SMummaleti Keerthana, K J M Meghana, Siginamsetty Pravallika, D. M. K. (2021). *An*

Ensemble Algorithm for Crop Yield Prediction. *Icicv*.

Suruliandi, A., Mariammal, G., & Raja, S. P. (2021). Crop prediction based on soil and environmental characteristics using feature selection techniques. *Mathematical and Computer Modelling of Dynamical Systems*, *27*(1), 117–140. https://doi.org/10.1080/13873954.2021.1882505

T.Banavlikar, A.Mahir, M. Budukh, S. D. (2018). *IRJET-Crop Recommendation System using Neural Networks*. *05*.

Tagel Aboneh, Abebe Rorissa, R. S. (2022). *Stacking-Based Ensemble Learning Method for Multi-Spectral Image Classification*. 1–18.

Thomas, K. T., Varsha, S., Saji, M. M., Varghese, L., & Thomas, E. J. (2020). *Crop Prediction Using Machine Learning*. *13*(3), 1896–1901.

Ting, K. M. (2017). Confusion Matrix. *Encyclopedia of Machine Learning and Data Mining*, 260–260. https://doi.org/10.1007/978-1-4899-7687-1_50

Ujjainia, S., Gautam, P., & Veenadhari, S. (2021). *A Crop Recommendation System to Improve Crop Productivity using Ensemble Technique*. *February*. https://doi.org/10.35940/ijitee.D8507.0210421

Weinberg, F. (2005). *Rule induction : Ross Quinlan ' s ID3 algorithm*.

Yasrab, R. (2021). *Predicting Plant Growth from Time-Series Data Using Deep Learning*. 1–17.

Yesugade, P. K. D., Kharde, A., Mirashi, K., Muley, K., & Chudasama, H. (2018). *Machine Learning Approach for Crop Selection based on Agro-Climatic Conditions*. *7*(10), 103–106. https://doi.org/10.17148/IJARCCE.2018.71023

Cao, J., Lin, Z., Huang, G., & Liu, N. (2012). Voting based extreme learning machine. *Information Sciences*, *185*(1), 66–77. https://doi.org/10.1016/j.ins.2011.09.015

Ghana Statistical Service. (2019). 2018 Agriculture census report. Accra: Ghana Statistical

Service.

Kumar, A., Jain, P., & Kumar, A. (2021). A review on crop yield prediction models using machine learning techniques. Computers and Electronics in Agriculture, 184, 106001.

FAO. (2019). The future of food and agriculture: Alternative pathways to 2050. Rome: Food and Agriculture Organization of the United Nations.

Huang, X., Jiang, T., Tian, B., & Shang, J. (2022). A Machine learning-Based Method for Tomato Growth Stage Recognition Using Convolutional Neural Networks. Frontiers in Plant Science, 12, 874.

IPCC. (2018). Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related

Gyasi, E. A., & Kumi, E. (2020). Crop yields prediction using artificial neural network: A case study of maize yields in the Techiman Municipality, Ghana. International Journal of Emerging Trends in Engineering Research, 8(4), 234-240.

Kamanga, B. C., Kuwornu, J. K. M., Osei, A., & Kuwornu, F. M. (2021). Harnessing the potential of artificial intelligence in agriculture: A review of applications, challenges, and future prospects. Sustainability, 13(6), 3089.

Pasquel, H. M., Vasquez, W. F., & Saquicela, V. (2021). Crop prediction models based on machine learning techniques: A review. Journal of Big Data, 8(1), 1-33

Khatun, K., Uddin, M. N., Kabir, S. M., & Hossain, M. A. (2022). Predicting crop yield using machine learning: A review. IEEE Access, 10, 10794-10806.

Wu, L., Yao, Y., Huang, J., Zhang, B., & Gao, X. (2021). Crop Yield Prediction Based on Machine learning: A Review. IEEE Access, 9, 34559-34568.

Intergovernmental Panel on Climate Change (IPCC). (2021). Climate Change 2021: The Physical Science Basis. Cambridge University Press.

Kalimuthu, S., Sathiyanarayanan, S., & Jayaseelan, C. (2022). Crop Yield Prediction using

Machine Learning Techniques: A Review. In Proceedings of the 3rd International Conference on Emerging Trends in Engineering, Science and Technologies (ICETEST 2022) (pp. 1-6). Springer.

Kumar, V., Sharma, R., Sharma, N., & Agarwal, A. (2023). Crop Yield Prediction using Machine Learning Techniques: A Review. In Proceedings of the 6th International Conference on Advances in Computing, Communication and Control (ICAC3 2023) (pp. 1-7). Springer.

Pasquel, E., Valdivia, R. O., & Zhang, H., Cheng, Y., Wei, L., & Sun, Y. (2020). Crop Recommendation System Based on Big Data and Cloud Computing. In Proceedings of the 2020 3rd International Conference on Cloud Computing and Internet of Things (pp. 120-124)

References: Li, Y., Yang, Q., Liu, X., Wang, X., & Zhu, X. (2019). DeepCrop: A machine learning framework for crop yield prediction. Computers and Electronics in Agriculture, 165, 104961. https://doi.org/10.1016/j.compag.2019.104961

References: Wang, Z., Sun, L., Zhang, H., Xie, H., & He, J. (2020). Prediction of cotton yield using machine learning techniques and soil information. Computers and Electronics in Agriculture, 178, 105772. https://doi.org/10.1016/j.compag.2020.105772

Mohanty, S. P., Hughes, D. P., & Salathé, M. (2021). Crop Yield Prediction using Machine learning Techniques: A Survey. Journal of Big Data, 8(1), 1-26. https://doi.org/10.1186/s40537-020-00399-6

Adu-Gyamfi, P., Nkrumah, I., Owusu, G., & Wiredu, A. (2021). A Comparative Study of Machine Learning Algorithms for Crop Yield Prediction in Ghana. Journal of African Technology Innovation, 13(2), 134-142.

Amos, A. F., Ahiakpor, F. P. K., & Kankam, S. (2020). Assessing the Effectiveness of Machine Learning Algorithms in Crop Yield Prediction: A Case Study in Ghana. Journal of Data

Science, 18(3), 492-510.

Owusu, G., Agyemang, K., & Adu-Gyamfi, P. (2019). A Mobile-Based Crop Recommendation System for Smallholder Farmers in Ghana. Journal of Agricultural Informatics, 10(2), 1-11.