5-2024

Recommendation System using machine learning for fertilizer prediction

Durga Rajesh Bommireddy

RECOMMENDATION SYSTEM USING MACHINE LEARNING

FOR FERTILIZER PREDICTION

A Project

Presented to the

Faculty of

California State University,

San Bernardino

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

in

Computer Science

---

by

Rajesh Bommireddy

May 2024

RECOMMENDATION SYSTEM USING MACHINE LEARNING

FOR FERTILIZER PREDICTION

A Project

Presented to the

Faculty of

California State University,

San Bernardino

by

Rajesh Bommireddy

May 2024

Approved by:

Dr. Qingquan Sun, Advisor, School of Computer Science

Dr. Bilal Khan, Committee Member

Dr. Yan Zhang, Committee Member

ABSTRACT

This project presents the development of a sophisticated machine-learning model aimed at enhancing agricultural productivity by predicting the optimal fertilizer suited to specific crop requirements. Leveraging a diverse set of features including soil color, pH levels, rainfall, temperature, and crop type, our model offers tailored recommendations to farmers. Three powerful algorithms, Support vector machines (SVM), Artificial neural networks (ANN), and XG-Boost, were implemented to facilitate the prediction process. Through comprehensive experimentation and evaluation, we assessed the performance of each algorithm in accurately predicting the best fertilizer for maximizing crop yield. The project not only contributes to the advancement of machine learning techniques in agriculture but also holds significant implications for sustainable farming practices and food security.

ACKNOWLEDGEMENTS

Firstly, I wish to express my gratitude to my esteemed advisor, Dr. Qingquan Sun, and respected committee members, Dr. Bilal Khan, and Dr. Yan Zhang. They have provided guidance, knowledge, and wonderful support. Their belief in me makes me strive to achieve my goal in the timeline. Their valuable input, mentorship, and guidance greatly influenced the course and excellence of this endeavor.

I express my gratitude to the teachers and staff at California State University, San Bernardino, whose availability of crucial resources and a supportive academic atmosphere were vital in ensuring the successful completion of this project. Their unwavering commitment to academic achievement has served as a consistent source of motivation.

DEDICATION

I would like to dedicate my master's research to my advisor, Dr. Qingquan Sun, with the deepest thanks and respect. His guidance and motivation helped me to complete my project successfully. Every time I met him, he always motivated me to show my work should look unique to others. I truly believe I gained immense knowledge under his guidance and support.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER ONE:

INTRODUCTION

Any country's capacity to maintain economic stability is greatly dependent on its agriculture industry. Farmers frequently rely on peer feedback and personal testimonials to enhance their crop-cultivation methods; however, this approach may not always produce the desired outcomes and may even cause financial disasters. We must give up conventional farming methods to ensure the security of the world's food supply. These methods often overlook elements that affect crop yield potential, soil composition, crop water requirements, and land suitability when creating plans for fertilizer and crop rotation. Technological breakthroughs may be compared to groundbreaking developments in agricultural science. Numerous technologies, such as big data, deep learning, machine learning, and digital image processing, have a significant positive impact on agriculture. Contemporary methodologies, such as data mining algorithms, recommendation systems, and digital image processing, are used to gather and scrutinize information. Recommendation system technologies are used in several sectors, including online commerce, medical treatment, and movie suggestions. Machine learning can greatly improve recommendation systems in terms of their usefulness and accuracy. Various agricultural data, such as photos of land and crops, may be obtained, copied, and examined using digital image processing methods. Later, this data may be used for purposes such as identifying different crop diseases, measuring soil minerals, and assessing the total amount of crop growth. Understanding the significance of precise agriculture and the role that artificial intelligence plays in optimizing the selection of crops and the use of pesticides. These technologies have the potential to enhance agricultural development by facilitating the recommendation of appropriate crops, fertilizers, and pesticides.

Problem Statement

The main issue we identified is the challenges faced by farmers in making decisions on crop selection and pesticide use, considering factors such as soil conditions, weather patterns, and insect infestations. Emphasizing the need for personalized, evidence-based guidance for enhancing agricultural techniques.

The main problem in the general agriculture area is that farmers often face cropping issues, which are based on certain values they need to know. These values also require them to calculate using historical data, a process that can be challenging due to the reliance on variables such as soil, rainfall, temperature, and humidity. Subsequently, the farmer has relied on some recommendation systems to forecast the optimal crop. Furthermore, Machine learning does not recommend any pesticides. Additionally, I have identified flaws in the current method like water scarcity, soil degradation, pests and diseases, and access and resources. So, to solve the above problems, I came up with a machine learning approach. In that machine learning approach, we implemented algorithms to calculate the historical data that we provide. My machine learning algorithms will archive the above problem statement based on the data.

CHAPTER TWO:

LITERATURE SURVEY

In countries, like India that are still developing agriculture plays a role in driving the economy. To ensure the quality and yield of crops remain high it is essential to consider factors such as temperature, water availability, and profitability when selecting crops, fertilizers, and pesticides. Researchers are utilizing technology to recommend crops based on market trends, water requirements, and soil conditions. They are also guiding insecticides diagnosing plant diseases and suggesting fertilizer options. This research presents Agri Rec, a method using machine learning to suggest crops and fertilizers. Furthermore, it suggests a technique that evaluates the qualities of crops, soils, and fertilizers to identify the fertilizer mix, for specific crop soil pairings. The effectiveness of this method was tested using 5,000 soil samples from Gujarat covering twenty-four types of crops. The results showed a 92.11% accuracy rate in fertilizer recommendations and 95.85% accuracy, in crop suggestions. Surpassing existing recommendation methods by fourfold [1].

This study has its advantages and disadvantages. Ignoring challenges and concentrating solely on the benefits could skew the research outcomes. Restrict the advantages for customers. SSM is compared to production methods that aim to eliminate or resist environmental variations rather than adjust to them.

SSM plays a role, in developing information hubs setting up risk mitigation plans improving biotechnology, and protecting the environment. To create a recommendation system for contemporary agriculture Big Data is an advanced technique in computer engineering that has progressed significantly. It has applications in numerous fields such as agriculture, medicine, and retail. Utilizing big data enables the discovery of a greater quantity of historical data, which may be used to construct a new system capable of functioning as a predictive system. Agriculture is a crucial component of India's economy. However, many of the agricultural methods currently employed are outdated [2].

We can improve agriculture and make it more productive and cost-effective by using cutting-edge technology like IoT, data mining, and big data. Accurate projections of agricultural productivity are essential to the progress of agriculture. We have developed a machine learning technique called Random Forests that takes weather and environmental variations into account when predicting crop productivity. We have gathered an extensive data set on agricultural yields from several sources [3].

The Random Forest model had significant predictive capabilities in forecasting crop yields and consistently outperformed the MLR model across all performance metrics that were evaluated. Numerous research studies have indicated that the Random Forest (RF) algorithm is highly effective, in forecasting crop results in scenarios and demonstrates an accuracy level, in analyzing data [4].

Agriculture stands out as a sector where analytics plays a role. By understanding past data and uncovering patterns accurate future forecasting becomes possible. With 55% of Indians engaged in agriculture utilizing analytics and predictions could offer insights to them. It's important to take into account factors that can have an effect, on crop production. The integration of data analytics, into the industry has seen growth, in technology adoption in recent years. This study examines government crop production data to forecast crop yield based on characteristics [5].

In agriculture, farmers utilize machine learning to receive recommendations, on insecticides, fertilizers, and other necessary products customized to their needs with accuracy. Various machine learning methods are applied across domains on a basis. This research delves into the relationship, between independent variables affecting the forecast offering insights into predicting the optimal type of fertilizer using techniques such, as multiple linear regression and lasso regression [6].

In today's society where people lead lives agriculture plays a vital role. Enhancing crop yield involves selecting the types of crops to plant. Different methods of machine learning can be applied to create recommendation systems, for the field of agriculture. These systems use parameters to suggest crops for cultivation. By leveraging data science farmers can access recommendation systems to assist them in making decisions about crop planting. The use of the Light GBM Machine Learning Algorithm addresses inefficiencies, in systems and enhances accuracy and reliability. Analyzing data sets based on soil levels and environmental factors enables the generation of crop recommendations. Light GBM, a gradient-boosting framework that employs decision trees is employed to optimize model efficiency and reduce memory usage. The recommendation system built using the LightGBM algorithm is robust, precise, and dependable [7].

Agriculture plays a role in the lives of 60% to 70% of India's population serving both as a source of livelihood and, as a driver of the economy. Factors such as soil quality, weather conditions, rainfall patterns, droughts, pest infestations, and the use of fertilizers and pesticides pose challenges for farmers affecting crop yields and income levels while leading to food shortages. By utilizing the Random Forest Algorithm to suggest the crops based on soil characteristics like pH levels, temperature, humidity rainfall amounts, and nutrient content (phosphorous. P, nitrogen. N, potassium. K) it becomes possible to address these issues effectively. Through leveraging the Kaggle dataset, for model training purposes the main objective is to develop a model that can recommend crops for specific types of soil while demonstrating the superior performance of the Random Forest Algorithm compared to other methods [8].

CHAPTER THREE:

SYSTEM REQUIREMENTS

Frameworks

Python

In this project, we used the Python version of 3.6.2. Python is a powerful coding language widely used in machine learning, Artificial Intelligence, etc., due to its simplicity and robustness.

Flask

In this project, we utilized Flask, a Python web framework to create web applications, with the assistance of the Jinja 2 engine in our project. Flask was instrumental in crafting the user interface, for our development work.

Pandas

Pandas is a widely used Python library which is used for data analysis. In this project, we rely on pandas for analyzing the data. Data handling will be efficient by using pandas which helps us to improve the performance of the model.

NumPy

NumPy is also one of the Python libraries used for mathematical computations. NumPy is very helpful in machine learning and developing machine learning models where complex array multiplications are required. In our project, we have used NumPy for array multiplications.

Scikit-learn.

Scikit-learn is one of the libraries used for machine learning projects for identifying and classification tasks. It is used for building machine learning models.

Matplotlib

Matplotlib is Python plotting used for creating high-quality visualization graphs to analyze the data and frequency. In our project, we have used Matplotlib to visualize the model parameters, etc.

Hardware requirements

System: Intel core I5.

Storage: 1 TB.

Ram: 8GB Memory

Software requirements

Operating system: Windows 10

Programming Language: Python 3.6.2.

Development environment: Visual Studio Code

Tool: Flask for building user interface

Dataset: Crop and fertilizer

CHAPTER FOUR:

SYSTEM DESIGN

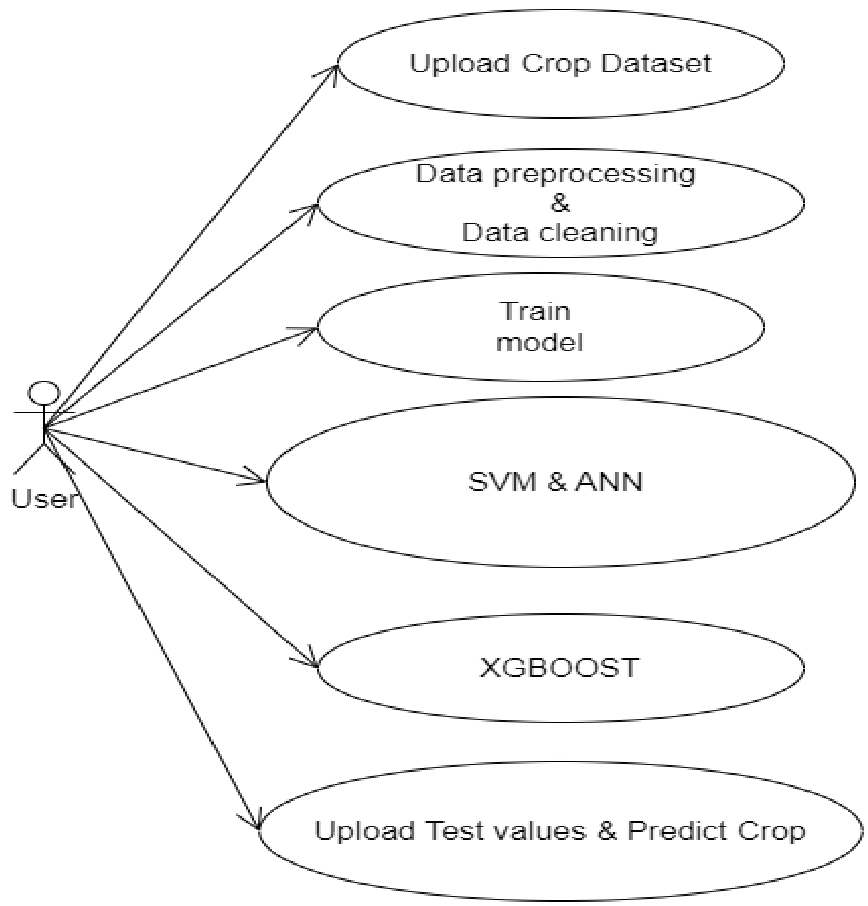
UML Diagrams

UML diagrams give a visual description of the parts and relationships in intricate systems of machine learning. They are very priceless. while addressing several components and procedures. UML diagrams aid in the planning and organization of the system during the creation of machine learning initiatives. It facilitates efficient data preprocessing, deployment componentry, and machine learning model organization.

Goals

The main aims in the design of the UML are as follows:

- Giving users an available and expressive visual modeling Language that makes easily understandable models.
- It opens the gate for specializing and advancements in the concepts.
- Be unique in developing the project and coding languages.
- It gives the idea for making the modeling language.
- Increases the chance of object-oriented tools in the market.
- Usage of modularizing concepts such as collaborations, frameworks, patterns, and components.
- Finally combining best practices.



.

Figure 1. Use Case Diagram of Fertilizer Recommendation System

Figure 1 describes the interaction between the Fertilizer recommendation system and the user. The user requests crop details from the system and input features in the first use case. The recommendation system then uses the inputs to train the model and updates it based on user preferences. The user is then presented with the anticipated output together with the input values provided by the user. These outcomes, along with the user-entered values and the recommended fertilizer as projected by the recommendation system, are shown on the prediction page.

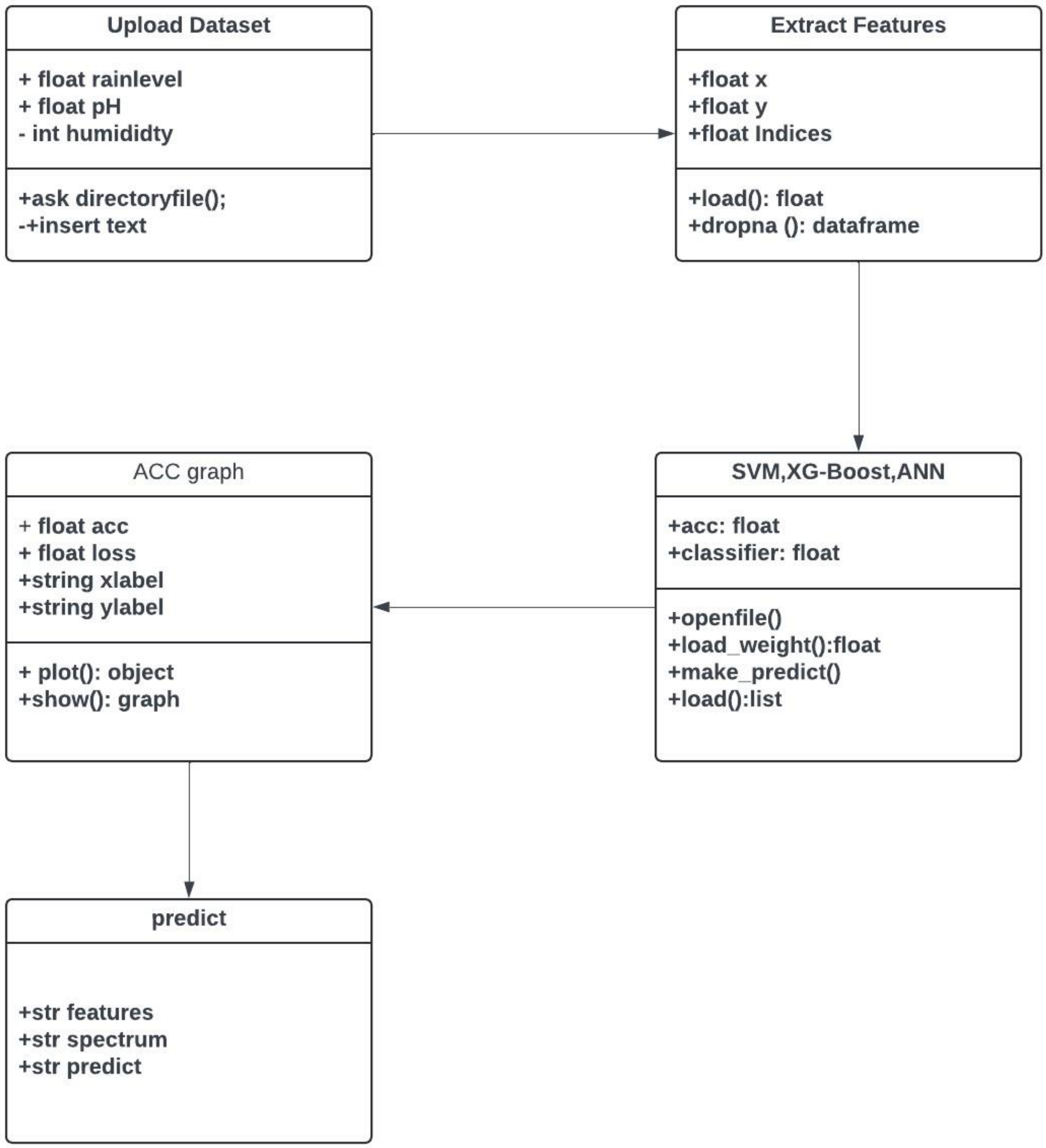


Figure 2. Class Diagram of Fertilizer Recommendation System

Figure 2 describes the class diagram that illustrates the relationship between the classes. Class diagram mainly defines the attributes and relationships in developing the system. In our System, we have five main classes: Upload Dataset, Extract features, Listed models, Accuracy Graph, and Predict. The dataset class is used for importing the data from the file and makes it ready for the feature extraction class. After feature extraction, it will be used to train the model to get the predicted output. The connections between the classes are represented by denoting arrows which gives the flow of the classes to implement the system.

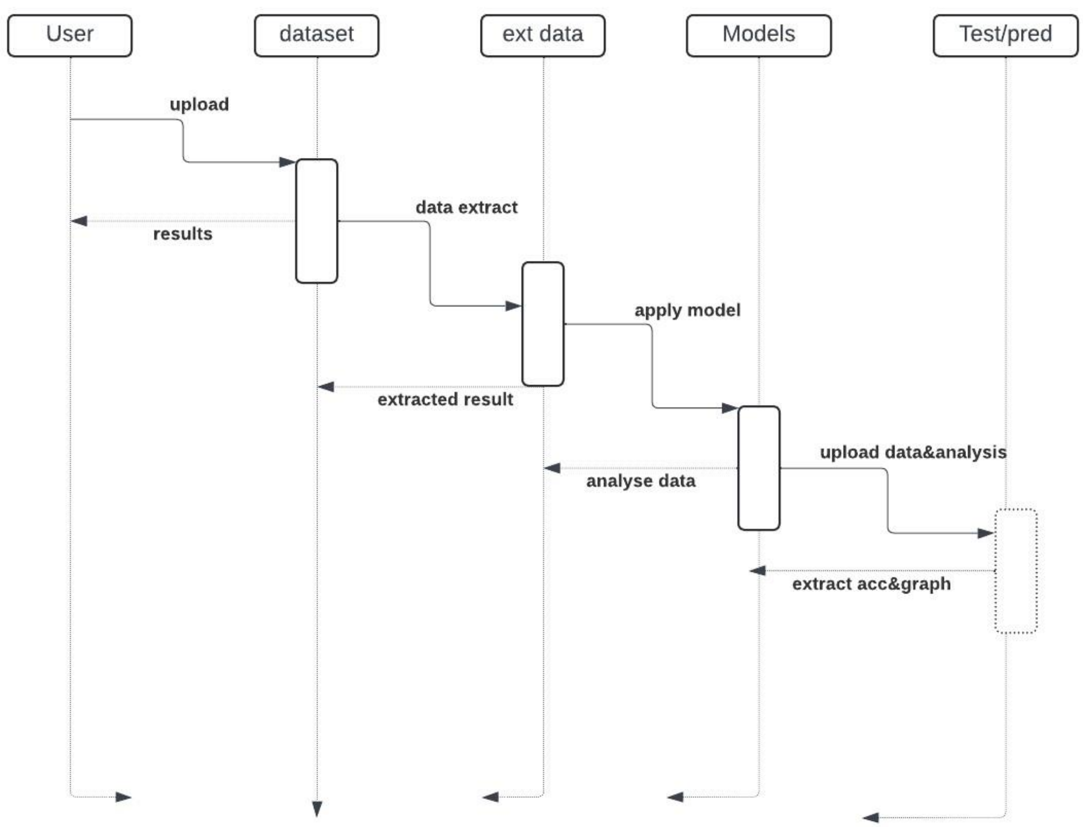


Figure 3. Sequence Diagram of Fertilizer Recommendation System.

Figure 3 describes the user interactions with the objects and components. Sequence diagrams are mainly used to give information about how the signals are transmitted between the objects and classes. It mainly sends messages to trigger the events which need to be processed. In our case, we have 5 sequences that show how a user will interact with the data being loaded and make it ready for extraction. Once the data is split apply those data for models to train the data. Finally, test data will be given the predicted sequence to display the fertilizer.

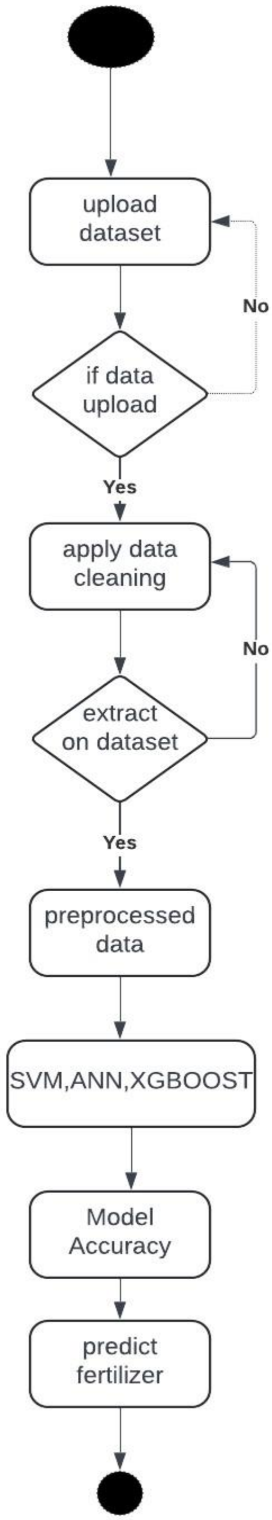

upload dataset
No
if data upload
Yes
apply data cleaning
No
extract on dataset
Yes
preprocessed data
SVM,ANN,XGBOOST
Model Accuracy
predict fertilizer

Figure 4. Activity Diagram of Fertilizer Recommendation System

Figure 4 describes the flow of activities and how the system should follow the process. Initially, the user will open the web page to enter the input values. After it will be ready for preprocessing and make it ready to train the model. The user can check the test data to predict the best fertilizer that needs to be used for that recommended crop. It clearly shows how the process behaves visually.

CHAPTER FIVE:

IMPLEMENTATION

In developing the project, I have taken the platform of Python with version 3.6.2 version by using the inbuilt function of Python for developing web applications. I have used Flask here for the user interface. In this project, we have written logical programming code for implementing the machine learning models instead of using inbuilt machine learning models. I have used libraries like NumPy, pandas, matplotlib, and Scikit-learn for developing web applications.

Modules

- Upload Dataset of crop and fertilizer data
- Extract Features from the dataset.
- Train Machine learning Models
- Model Accuracy & Loss Graph
- Dynamic input to Test & Predict Fertilizer.

Data collection

At the start of the machine learning process, the first task is to collect the data. This phase focuses on locating and addressing all data-related challenges. It is essential to pinpoint the sources of data during this phase since information is gathered from corresponding sources. This stage holds importance in the process. The precision of predictions improves as more data is gathered.

I have constructed an extensive dataset that includes historical crop data, soil properties, weather data, nuisance data, and chemical effectiveness information. Utilize algorithmic learning to provide recommendations for suitable crops and appropriate herbicides. Develop a user-friendly interface to facilitate farmers' access and use of these proposals effortlessly.

Data preparation

Once we have collected the data our next task is to prepare it for the stages. This involves organizing the data into a format and adjusting it for our training procedures. During this phase, we group the data. Then randomly determine its orientation. This stage follows two processes:

- Data investigation:

  This process aims to grasp how the data functions. It investigates and recognizes the characteristics and arrangement of the data. A thorough investigation of the data leads to results. Through this, we can detect any data points and anomalies.

- Data preprocessing:

  Performed data cleansing and processing on the acquired information, ensuring that outliers and missing values are properly addressed. To ensure trustworthiness, it is important to both normalize and standardize the data. I analyze data correlations to understand the relationships between different variables.

Data analysis

This is one of the important steps in building a model to analyze the data characteristics. It makes the dataset suitable for evaluating the models which we are implementing in our project. We can use different techniques like cluster analysis, normalization, and other major techniques for preparing the data that fits into the selected models.

Instructing the Model

In this phase, we work on enhancing our models' performance by training them. We utilize our curated dataset to train the model, with existing data ensuring it comprehends attributes and traits effectively.

Examining the Model

After we've trained our models using a dataset we move on to the phase where we test them. This involves obtaining the accuracy of our model, with test data that is already available. By evaluating the models using our testing data we can determine their accuracy levels.

Algorithm Implementation

In my project for fertilizer recommendations, I used machine learning techniques such as gradient boosting, neural networks with regression techniques, and support vector machines (SVM). I implemented the methods using Python's functional programming approach, without relying on libraries wherever possible.

Considered using classification techniques such as logistic regression or deep learning when providing pesticide recommendations. To evaluate the method, the information should be divided into several sets for testing and training. Enhancing the modeling of hyperparameters may be achieved using techniques such as grid scanning or optimization methods.

SVM

My study utilizes a supervised learning approach called support vector machine (SVM) in the field of machine learning to carry out regression and classification tasks. During my API development, I have found that Support Vector Machines (SVMs) are very efficient in dealing with binary classification problems. These problems involve dividing the components of a dataset into two distinct groups. These are agricultural products, namely crops and fertilizer. The objective of support vector machine algorithms is to identify the most effective decision boundaries that can separate data points into distinct groups. Within the context of feature spaces, multidimensional spaces are sometimes referred to as hyperplanes. The goal is to increase the space, between the data points of each group and the dividing line to improve the differentiation, among sets of data. Support Vector Machines (SVMs) are important tools for analyzing complex data sets that cannot be successfully divided by a linear boundary. Nonlinear support vector machines (SVMs), also known as SVMs, use a mathematical method to transform data into a higher-dimensional space, facilitating the identification of a boundary.

The following kernel functions are often selected for Support Vector Machines (SVMs):

The linear kernel. In this process, the information is transformed into a space using a fundamental kernel function and then separated by a straight boundary.

In support vector machine and machine learning, a particular type of kernel function called a polynomial kernel is employed. The degree of resemblance between data items in each context is measured using a mathematical formula. More than three dimensions in space. This kernel function is more efficient than the linear kernel. It can transform the data into a higher-dimensional space, which makes it possible to distinguish non-linear data.

The RBF kernels. The Gaussian kernel function is commonly acknowledged as the predominant and effective option for Support Vector Machines (SVMs) in many classification tasks.

The sigmoid kernel. This kernel function bears similarity to the RBF kernel; however, it has a unique structure that may be beneficial for some classification problems.

In the API development, we used the linear kernel that makes some functional implementations those are hinge loss, fit, and prediction to generate accuracy and model implementation. Here, we used the fit function to train the SVM model using an optimization technique called stochastic gradient descent. Once we updated the loss using the loss function at each epoch then we used the predict function to make predictions on the trained SVM model.

<u>ANN</u>

An Artificial neural network, commonly referred to as an ANN is an algorithm inspired by the workings of the brain and is utilized to tackle issues. This approach utilizes learning techniques inspired by the networks found in human biology. The development of networks was driven by a desire to replicate the intricate processes of the human brain. While artificial neural networks (ANNs) function similarly to networks they are not exact replicas. ANNs exclusively process structured input data.

In this algorithm, I implemented a regression model. The rationale for utilizing artificial neural networks instead of linear regression for regression analysis is based on the limitation of linear regression in identifying only linear relationships between the independent and dependent variables, thus being unable to capture intricate non-linear connections. To properly comprehend the intricate non-linear correlation between the characteristics and the desired outcome, we need additional techniques. One such technique to examine is the use of artificial neural networks. Artificial neural networks provide the ability to acquire an understanding of the intricate connection between qualities and objectives by integrating activation functions at every layer. Next, we will examine the operation and meaning of artificial neural networks. The code was performed entirely in the Flask framework. We will be using the crop and fertilizer dataset information to arrive at predictions on the median fertilizer recommendation. The directory with datasetname.csv in the flask application contains the data.

The datasets method is used for data scaling. Rescaling the data would lead to quicker convergence towards the global optimum for optimizing loss functions. In this context, I used API implementation the standard level scaling algorithm class, which carries out z-score equalization. Z-score equalization involves subtracting each data point from its mean and dividing it by the standard deviation of the data. We will be doing training using the increased dataset.

To create an output, it multiplies the number of input values by matching weight, adds up the weighted inputs, and then runs the total through a sigmoid activation function. To minimize a loss function that gauges the difference between expected and actual outputs, then the network modifies the weights of connections between neurons during training by utilizing an optimization technique such as gradient descent.

Both forward and backpropagation are used in the training process:

Forward propagation: Predictions are formed as input data is passed through the network in a forward manner.

Backpropagation: The network is trained with the estimated errors between the expected and actual outputs. Gradient descent is then used to propagate these errors backward through the network and modify the weights. Calculated the output using sigmoid activation.

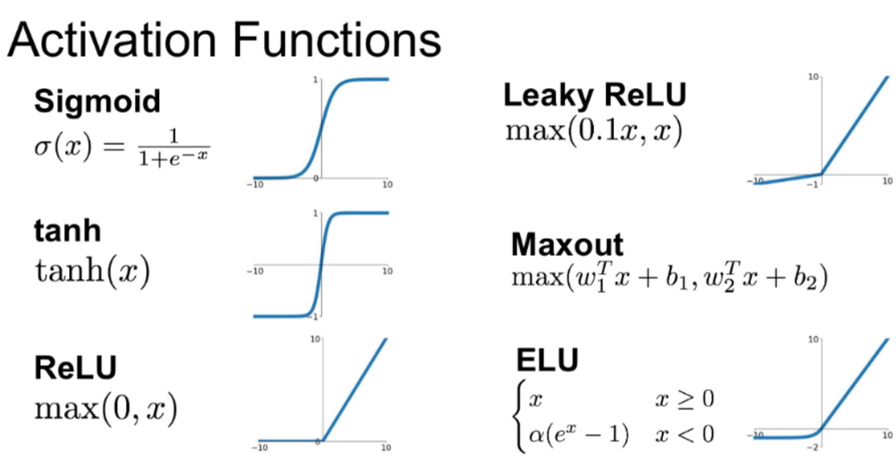


Figure 5. List of Activation Functions

<u>XG-Boost</u>

The XG-Boost Classifier has shown superior effectiveness, simplicity, and resilience. XG-Boost is my preferred way of implementing a machine learning technique that falls under the category of ensemble learning, namely the gradient boosting framework. The system utilizes decision trees as the fundamental basis for learning and integrates regularization techniques to increase the performance of the model. XG-Boost is a widely acclaimed program known for its fast computation, analysis of feature significance, and handling of missing data. It is often used for applications like regression, classification, and ranking.

Therefore, it is not surprising that CERN acknowledged it as the most effective method for categorizing signals from the Large Hadron Collider. The problem presented by CERN necessitated a scalable system capable of handling data produced at a pace of 3 petabytes per year, while successfully differentiating an exceptionally uncommon signal from background sounds in an intricate physical process. The XG-Boost Classifier is the most effective, simple, and resilient option.

- Bagging

  Implementing an ensemble technique in which multiple models are trained, and their predictions are aggregated to provide a final prediction. Despite being regarded as models that are easily understood, decision trees exhibit substantial variety in their behavior. Let us analyze a training dataset that we split into two halves using a random procedure. Now, let's use each component to train a decision tree to get two models. When using any of these models, they will provide unique results. Due to this attribute, decision trees are recognized for their substantial diversity. Bagging and boosting are powerful aggregation strategies that effectively reduce the variance of any learner. The bagging strategy utilizes many decision trees as its fundamental learners, which are produced concurrently. We educate the learners by using replacement-sampled data. The final forecast is derived by calculating the average of the outputs from all learners. Ensemble learning is a technique in which many models are trained and their predictions are aggregated to get a final prediction.

- Enhancing

  Boosting is the iterative construction of trees, where each subsequent tree is specifically intended to reduce the errors generated by the previous tree. Each tree in the system assimilates information from its predecessors and adapts to rectify any lingering flaws. Thus, the next tree in the series will gain information from an improved version of the residuals.

  The learners utilized in boosting are weak learners that have a large bias and just marginally better predictive ability than random guessing. Each of these feeble learners offers vital information for prediction, enabling the boosting technique to construct a potent learner by adeptly blending these feeble learners. The ideal robust learner minimizes both inherent bias and unpredictability.

In our project, we used the extreme gradient boosting model where we constructed the tree structure based on a depth-first approach and trained the weak learners based on updating the previous iteration. After that taking the best split based on calculating Gini Impurity.

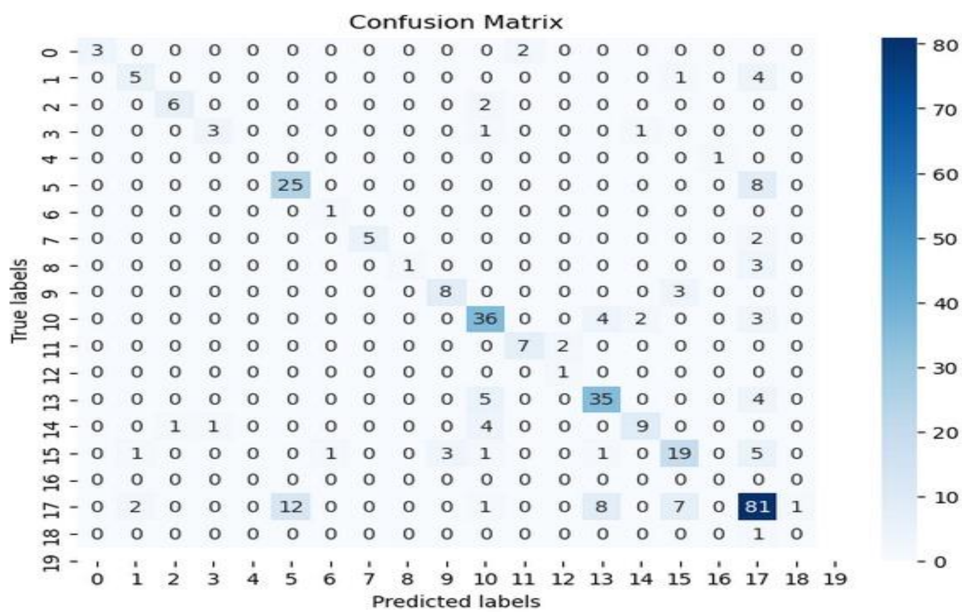


Figure 6. Confusion matrix

The above figure represents the confusion matrix. This representation is used to obtain the performance of the model. It shows the model predictions compared to the actual predictions. We can observe a few labels have better predictions and some are very low due to imbalance in the dataset.

Conclusion

Used machine learning to highlight the successes and efforts of the Crop along with the Pesticide recommendation System. Assessed the possible benefits to sustainability, pesticide consumption reduction, and agricultural output. Mention upcoming improvements, like integrating IoT devices or images from satellites for ongoing monitoring. With this idea, implemented a machine learning system that helps farmers to suggest the best fertilizer that gives a better yield.

CHAPTER SIX:

RESULTS

Figure 7. User Home page

The above figure illustrates users are accessing the web application where trained machine learning models are running in the backend to check the prediction of fertilizer. This page is filled with a lot of buttons like choose file, upload, preprocess, machine learning models, and prediction. These are used to access the information after uploading their dataset to trigger the classification process to visualize the result.

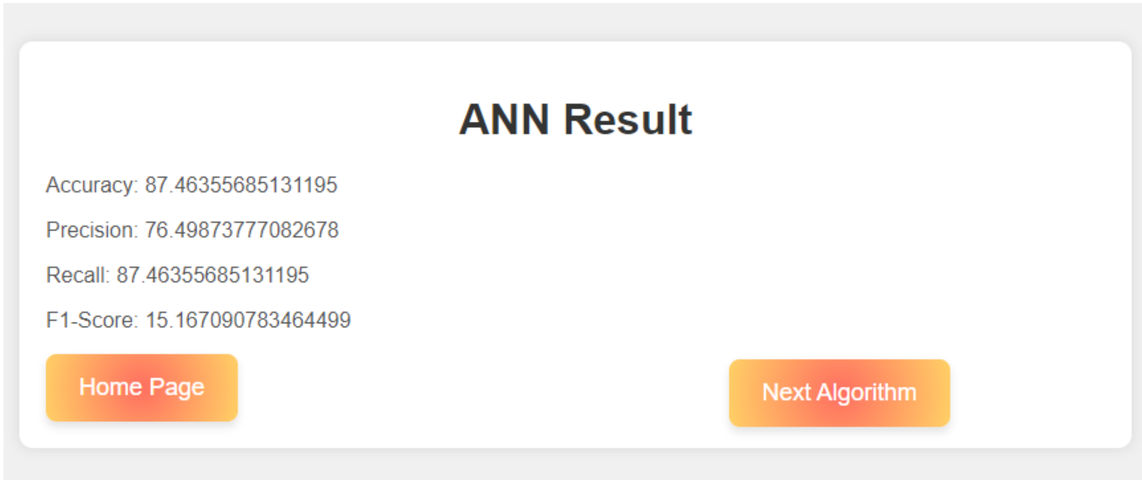


Figure 8. Performance of the ANN model

The above figure describes the performance of the ANN model after sending the preprocessed trained data to the trained ANN model. This model achieved an accuracy of 87.45% and it also gives information on performance metrics which are calculated during the process. It achieves a good precision score but registered a low F1 score due to the imbalance in the data.

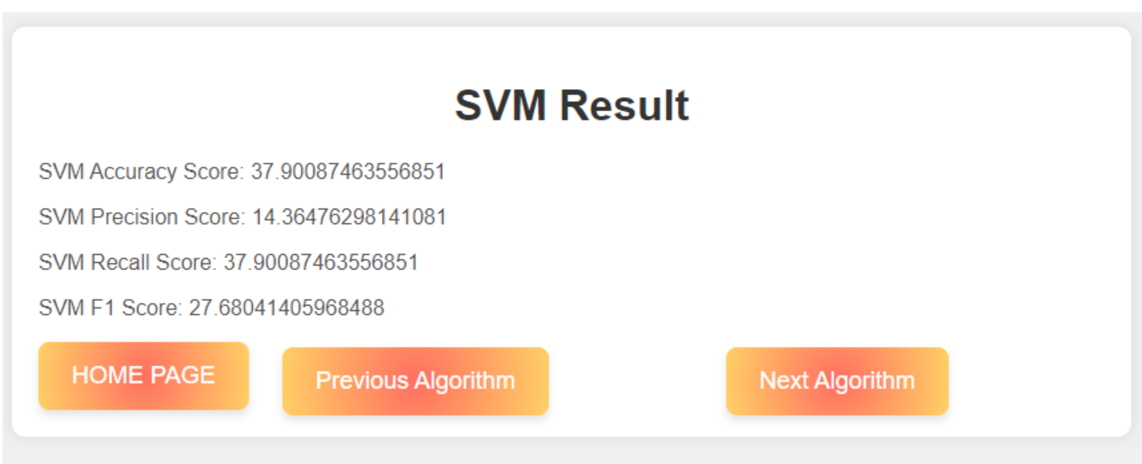


Figure 9. Performance of the SVM model

After receiving trained data, the SVM model's performance is obtained in the above image, where it attained a 37.9% accuracy rate. Because of the imbalance in the dataset, this model performs poorly when compared to other models. The figure above also includes a list of the performance metrics, which are determined during model implementation.

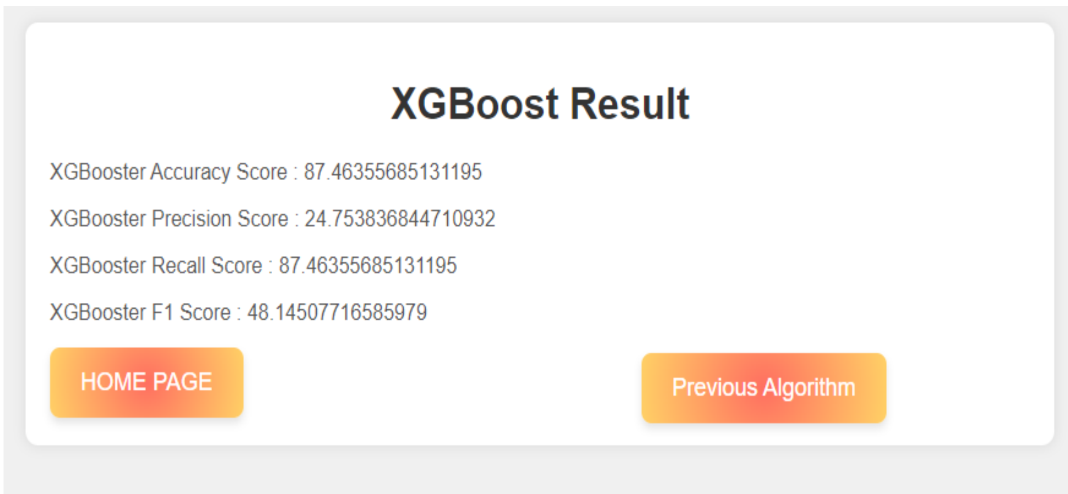


Figure 10. Performance of the XG-Boost model

After applying the trained XG-Boost model, the model's performance is shown in the above figure. Due to the XG-Boost models' excellent handling of imbalance and tabular data, this model performs well on all the other three models with an accuracy of 87.465%. It also achieved better performance metrics results with a very good recall score.

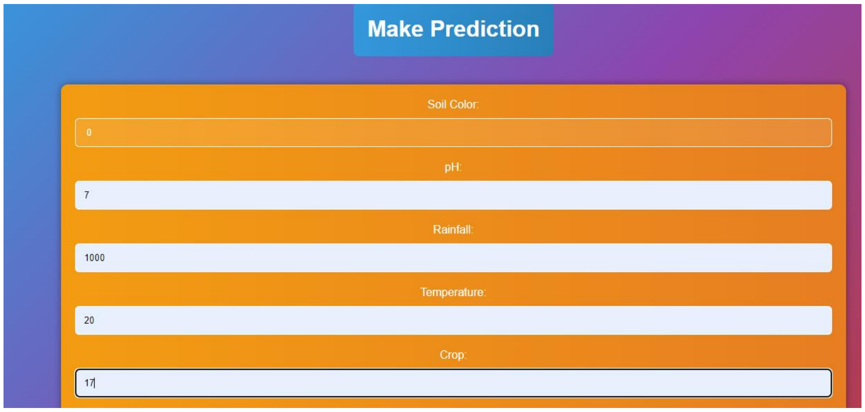


Figure 11. A section used to make a prediction

The above figure gives information on fertilizer names after entering the values of soil characteristics like soil color, pH, temperature, rainfall, and crop in the corresponding fields. Once the values are entered then the system predicts the fertilizer name on the below page.

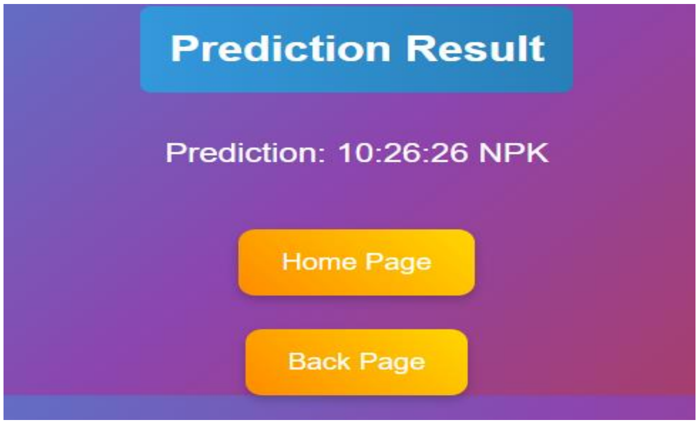


Figure 12. Fertilizer prediction page

The above figure displays the suggested fertilizer name to the user which can be used in his field to achieve better results. This entire recommendation system is developed user user-friendly because farmers can enter known features like soil color, temperature, rainfall, etc.

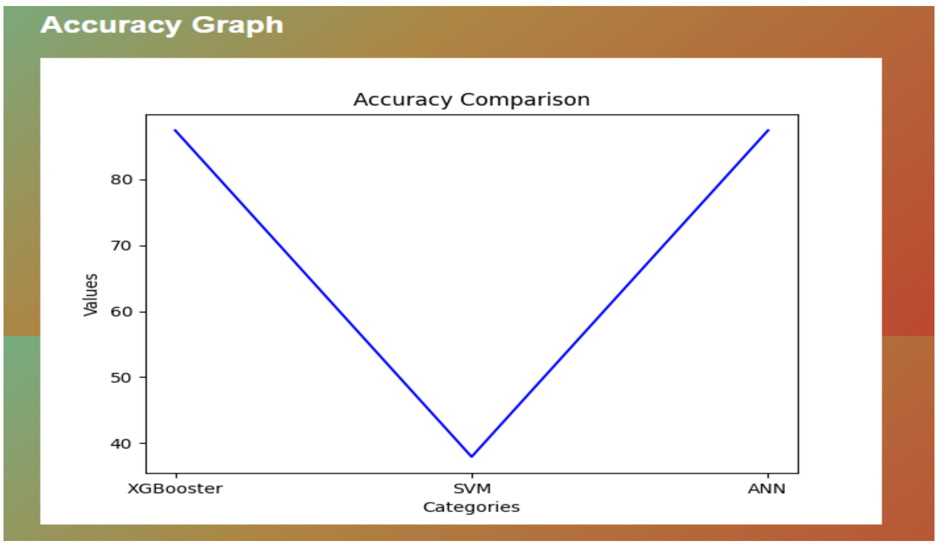


Figure 13. Accuracy comparison Graph

The above plot shows how accuracy varies among the models. The XG-Boost model achieves better accuracy than other models and the SVM model has less accuracy with 37.4%. The accuracy of is ANN model was like XG-Boost with a very slight difference.

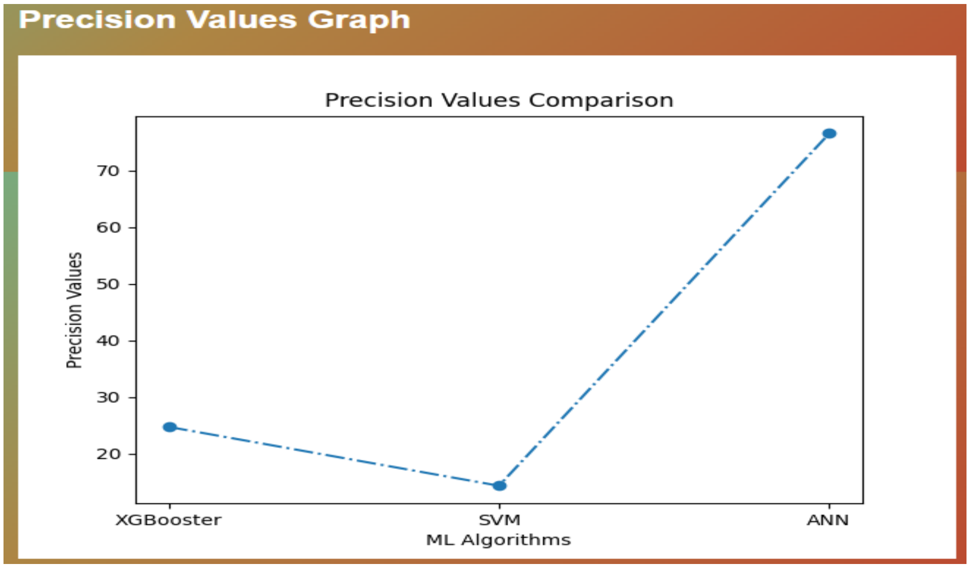


Figure 14. Precision Value Comparison Graph

The above figure represents precision metric comparison which is obtained by positive predictions attained during the model implementation. In all three models, ANN got a better precision score compared to other models. SVM achieves a lower precision rate due to less performance.

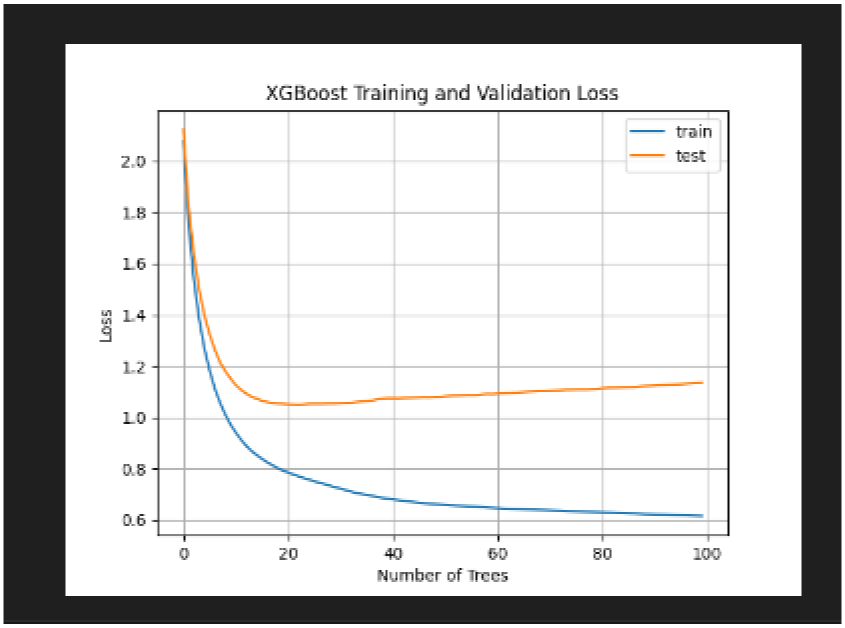


Figure 15. XG-Boost validation loss curve.

This figure gives the information about XG-Boost model validation loss curve with the number of trees on the vertical plane and loss on the horizontal. The training loss displays the error in training data, while the validation loss represents the error in validation data. Initially, both losses are high as the model starts with weak learners. We can see how the loss varied over the sequential number of trees that were added.

REFERENCES

[1] Nowak, P. (1998, June). Agriculture and change: the promises and pitfalls of precision. Communications in soil science and plant analysis, *29*(11–14), 1537–1541.doi.org/10.1080/00103629809370047

[2] Patel*, K., & B. Patel, H. (2023, May 10). Multi-criteria Agriculture Recommendation System using Machine Learning for Crop and fertilizer prediction. Current Agriculture Research Journal, *11*(1), 137–149.doi.org/10.12944/carj.11.1.12

[3] N, L. (2018, May 31). Crop Recommendation System for Precision Agriculture. International Journal for Research in Applied Science and Technology, *6*(5), 1132–1136.doi.org/10.22214/ijraset.2018.5183

[4] Karthikeyan, R., Gowthami, M., Abhishek, A., & Karthikeyan, P. (2018, September 1). Implementation of Effective Crop Selection by Using the Random Forest Algorithm. International Journal of Engineering & Technology, *7*(3.34), 287.doi.org/10.14419/ijet.v7i3.34.19209

[5] Jin, Z., Prasad, R., Shriver, J., & Zhuang, Q. (2016, December 8). Crop model- and satellite imagery-based recommendation tool for variable rate N fertilizer application for the US Corn system. Precision Agriculture, *18*(5), 779–800 doi.org/10.1007/s11119-016-9488-z

[6] Arora, D. Y., & Gambhir, M. A. (2020, September). Design and analysis of a prediction model for crop yield production in agriculture. International Journal of

Innovative Research in Computer Science & Technology, *8*(5). doi.org/10.21276/ijircst.2020.8.5.5.

[7] K. Monika, B. Ramprakash, S. Muthuramalingam and K. Mirdula, "Crop Fertilizer Prediction using Regression analysis and Machine Learning algorithms," 2022 5th International Conference on Contemporary Computing and Informatics (IC3I), Uttar Pradesh, India, 2022, pp. 1261-1266, doi: 10.1109/IC3I56241.2022.10072846.

[8] M. B. Begum, G. Sivakannu, J. Eindhumathy, J. S. Priya, M. Mahendran and R. R. Kumar, "Enhancing Agricultural Productivity with Data-Driven Crop Recommendations," 2023 Second International Conference on Augmented Intelligence and Sustainable Systems (ICAISS), Trichy, India, 2023, pp. 1055-1063, doi: 10.1109/ICAISS58487.2023.10250657.

[9] S. S. Sistla, K. Ketineni, N. Uppalapati and N. Lanka, "Crop Recommender Framework Utilizing ML Approach," 2023 IEEE 8th International Conference for Convergence in Technology (I2CT), Lonavla, India, 2023, pp. 1-6, doi: 10.1109/I2CT57861.2023.10126287.

[10] S. M. G, S. Paudel, R. Nakarmi, P. Giri and S. B. Karki, "Prediction of Crop Yield Based-on Soil Moisture using Machine Learning Algorithms," 2022 2nd International Conference on Technological Advancements in Computational Sciences (ICTACS), Tashkent, Uzbekistan, 2022, pp. 491-495, doi: 10.1109/ICTACS56270.2022.9988186.

[11] P. Kumar, K. Bhagat, K. Lata, and S. Jhingran, "Crop recommendation using machine learning algorithms," 2023 International Conference on Disruptive Technologies (ICDT), Greater Noida, India, 2023, pp. 100-103, doi: 10.1109/ICDT57929.2023.10151325.

[12] S. Gawade, G. Rout, P. Kochar, V. Ahire and T. Namboodiri, "AGROFERDURE: Intelligent Crop Recommendation System For Agriculture Crop Productivity Using Machine Learning Algorithm," 2023 International Conference on Computer, Electronics & Electrical Engineering & their Applications (IC2E3), Srinagar Garhwal, India, 2023, pp. 1-9, doi: 10.1109/IC2E357697.2023.10262476.