International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
DOI: https://doi.org/10.17762/ijritcc.v11i3.6338
***Article Received**: 30 December 2022 **Revised**: 31 January 2023 **Accepted**: 10 February 2023*

Crop Yield Prediction Using Gradient Boosting Neural Network Regression Model

[1]Purushotam Naidu k, [2]Dr. V. Lakshmana Rao, [3]Chaitanya Sree Gunturu, [4]Akkina Niharika, [5]Ch. Rohitha Anupama, [6]G. Srivalli

[1]Assistant Professor, Department of Computer Science and Engineering, Gayatri Vidya Parishad College of Engineering for Women, Visakhapatnam,Andhra Pradesh, India
purushotam.k30@gmail.com

[2]Assistant Professor, Department of Computer Science and Engineering, Gayatri Vidya Parishad College of Engineering for Women, Visakhapatnam,Andhra Pradesh, India
lakshman@gvpcew.ac.in

[3]Research Scholar, Department of Computer Science and Engineering, Gayatri Vidya Parishad College of Engineering for Women, Visakhapatanam, Andhra Pradesh, India
chaitanyasreegunturu@gmail.com

[4]Research Scholar, Department of Computer Science and Engineering, Gayatri Vidya Parishad College of Engineering for Women, Visakhapatanam, Andhra Pradesh, India
akkinaniharika111@gmail.com

[5]Research Scholar, Department of Computer Science and Engineering, Gayatri Vidya Parishad College of Engineering for Women, Visakhapatanam, Andhra Pradesh, India
chitikinarohitha795@gmail.com

[6]Research Scholar, Department of Computer Science and Engineering, Gayatri Vidya Parishad College of Engineering for Women, Visakhapatanam, Andhra Pradesh, India
srivalligovada@gmail.com

Abstract— The finest utility sector is agriculture, especially in emerging nations like India. Utilizing historical data in agriculture can change the context of decision-making and increase farmer productivity. Approximately a part of India's population is employed in agriculture, however this sector contributes just 14% of the country's GDP. This can be explained in part by farmers not making sufficient decisions on yield forecast. By examining numerous climatic elements, such as rainfall, and land characteristics, such as soil type and ground water salinity, as well as historical records of crops cultivated, the suggested machine learning technique tries to estimate the agricultural yield for a certain location. Finally, we anticipate that our proposed Machine Learning Gradient Boosting Neural Network Regression (Grow Net) model was predicting the accurate yield. Finally our system is expected to predict the yield based on dataset we have taken. We were compared our proposed algorithm with various Machine Learning algorithms such as Random Forest, Support Vector Machine, KNN, Multi-layer Perceptron Regressor, Gradient Boosting Regressor and results shows that proposed was given best RMSE ,MAE and R^2 value.

Keywords-Random Forest, Support Vector Machine, KNN, Multi-layer Perceptron Regressor, Gradient Boosting Regressor and Grow Net.

I. INTRODUCTION

In India, agriculture is thought to be the fundamental and most common kind of culture. Due to the fact that ancient people farm their own land, their requirements have been fulfilled. Modern people are unfamiliar with the significance of planting crops at the proper moment and places. The foundations for agricultural production are the natural resources and inputs needed, such as nitrogen, phosphorus, potassium, temperature, pH value, rainfall, and humidity. Predicting crop yields is a significant agricultural issue. To predict agricultural productivity, the current approach builds K nearest neighbours (KNN). Applying models like Random Forest, Support Vector Machine, K nearest neighbor(KNN), Multi-layer Perceptron Regressor(MLP),Gradient Boosting Regressor, and GrowNet to the dataset that comprises the aforementioned resources improves the suggested model.

Machine learning considered to be a part of artificial intelligence. Machine learning algorithms build a mathematical model using sample data, also referred to as "training data" in order to provide predictions or decisions without being explicitly instructed to do so. Machine learning uses a number of methods to build mathematical models and predict outcomes based on previously collected information or data. With the use of machine learning, a machine may anticipate outcomes without being explicitly programmed and automatically learn from data. When a machine learning system gets new data, it forecasts the outcome using the prediction models it has built

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
DOI: https://doi.org/10.17762/ijritcc.v11i3.6338
***Article Received**: 30 December 2022 **Revised**: 31 January 2023 **Accepted**: 10 February 2023*

from prior data. The volume of data used to construct a better model that predicts results determines how accurately the output will be.

The use of data is crucial to machine learning. We've used the crop dataset from Kaggle for our project. Data is collected from several cities in the state of Rajasthan in order to create and implement agricultural production prediction systems. State, district, crop, production, area, yield, rainfall, soil, and temperature are included in the statistics. By employing data preparation, we have eliminated all the null values, noise, and duplicate data from the data we have collected. There are 48 columns and 3649 rows in the dataset. We choose the top 15 characteristics using mutual information feature selection. We have used ML models to execute those features.

II. RELATED WORK

One of the oldest nations in the world, India has a strong agricultural industry. However, recent changes in agricultural trends are mostly attributable to globalization. There are several factors that have impacted the status of Indian agriculture aspects. Farmers all around the world can know the crop output depending on their land and weather conditions thanks to crop yield predictions made using more precise models. As part of our study, we employed the gradient boosting approach for improved accuracy and used both machine learning and deep learning models. The present method, which includes machine learning and neural network models, helps farmers comprehend agricultural production and also boosts the country's GDP.

Raja et al. [1] Proposed in agriculture crop prediction is crucial and heavily reliant on soil and environmental factors including temperature, humidity, and rainfall. Classification Techniques: Naive Bayes (NB), Decision Trees (DT), Support Vector Machine (SVM), K Nearest Neighbor (KNN), Random Forest (RF).

Haufler et al. [2] Proposed a Field Measurement Tested and Simulated Canopies using Pilot Study. The machine learning algorithm's performance evaluations when applied to the gathered field data show that, the estimated crop yield was less than 8% in 81% of test situations and the forecast error was less than 1.3%, which is crucial.

Suresh et al. [3] Proposed that in India's economy, agriculture is one sector that is essential. Inputs used in agriculture, including chemicals, fertilizers, insecticides, and soil condition.

Rashid et al. [4] Proposed Image identification of agricultural disease must be heavily utilized in intelligent agriculture. The data that were retrieved show that ANN and RF are the two most used algorithms for predicting crop production. LR, SVR, CNN, LASSO and SVM, as well as other well-known algorithms, were each applied in the research.

Khaki et al. [5] proposed deep learning framework employs proposed convolution neural networks (CNNs) and recurrent neural networks (RNNs) to estimate agricultural production based on environmental data and management strategies.

Nishant et al. [6] Proposed The yield of almost all crops cultivated in India is predicted in this study.The application error when the models were applied separately was approximately 4%, for Lasso it was around 2%, The final result after stacking was less than 1% for Kernel Ridge.

Qiao et al. [7] Proposed Presented Accurate and on-time crop output forecasts based on remote sensing data are essential for ensuring food security. A 3-D CNN is originally utilised to extract spatial-spectral data for the evaluation of agricultural production prediction. Some of the techniques used include 3-D convolution neural networks, Multi kernel GP, and Gaussian Process(f(x) gp(m(x), k (x, x)).

Latha Jothi et al. [8] Proposed Climate change and other environmental changes are becoming a serious danger to agriculture. KNN model-based categorization of ground water levels. Later, the proposed method was compared to the C&R tree algorithm, and it topped it by achieving greater accuracy of 90%.

Bhanumathi et al. [9] proposed the CRY algorithm for crop production, which forecasts outcomes for datasets linked to agriculture using beehive clustering methods. It was done using the k-means, apriori, and bayes algorithms. For classification, linear regression, ANN, and KNN are used.

Bhosale et al. [10] Proposed that the average contribution of AR, AUC, and FPI on crop yield is shown by this R2 value to be 70%. The association between a group of factors, including AR, AUC, and FPI, and their impacts on rice crop yield is established using regression analysis.

Pantazi et al.[11] Proposed In comparison to the high yield category, which achieved 83%, the medium yield category achieved 70%. The SKN model has an average overall performance of 81.65%, greater than the CP-ANN model's 78.3% and the XY-F model's 80.92%.

Khaki et al. [12] Proposed for the production of food on a worldwide scale, a crop yield prediction is crucial. To predict the phenotype based on genotype data, this method uses the

ISSN: 2321-8169 Volume: 11 Issue: 3
DOI: https://doi.org/10.17762/ijritcc.v11i3.6338
***Article Received**: 30 December 2022 **Revised**: 31 January 2023 **Accepted**: 10 February 2023*

DNN model, which can capture both the linear and nonlinear impacts of genetic markers.

III. METHODOLOGY

First, The Rajasthan crop dataset is the one that we use, and Kaggle is where we got it. We pre-processed the dataset to get rid of duplicates, NULL values, and missing values, which produced the final pre-processing data. Once the data has been cleaned of all noise, certain ML models are used to train and test the data. We utilized some performance metrics to assess the models' performance after doing training and testing on the data shown in Fig 1.

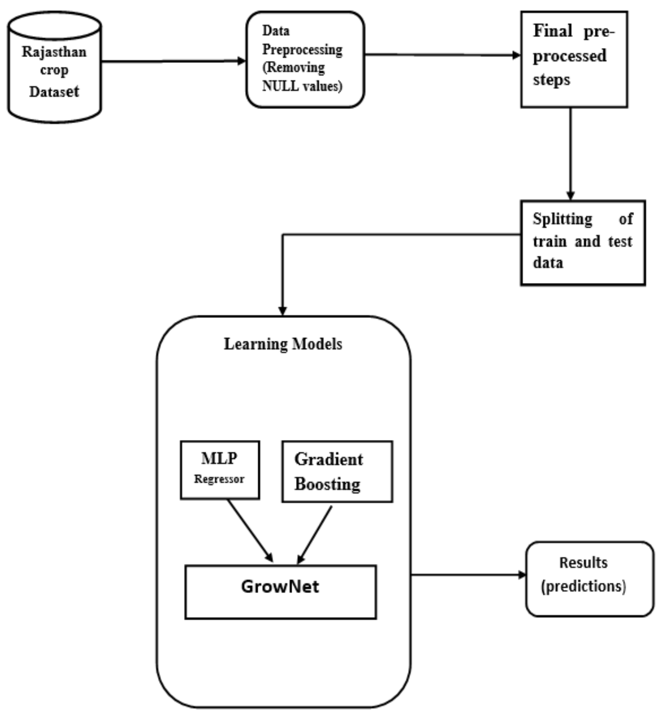

Fig 1: Architecture Diagram

A. Data collection

The first module, which focuses on data collection, is this one. We take into account the Kaggle dataset (Rajasthan Dataset for yield prediction Kaggle). The state government and farmers in different areas of Rajasthan provided the initial sources of the data. The data includes features such as state, district, crop, production, area, yield, rainfall (monthly rainfall and annual rainfall), soil (predominantly reddish medium texture Sandalized desert soils, Lithosoils in foothills and alluvials in plains, brown soils, clay loam, coarse sand in texture, deep soils in valleys, eastern portion alluvial, high soluble salts & The data obtained contains null, duplicate, and missing values.

B. Data Preprocessing

Preparing unprocessed data for learning models is referred to as data preparation. The data used in this project has already undergone some preprocessing. 46 columns and 3649 rows make up the dataset. The It describes the process of converting unprocessed raw data into numerical characteristics that may be processed while keeping the original dataset's contents intact. The two stages of the data pre-processing are feature selection and data normalisation. By scaling the data in the range of 0 to 1, the data may be normalised. Prior to normalization, categorical characteristics are converted to numeric format. For feature extraction, methods can be utilised such the Pearson coefficient, chi-square test, information gain, etc.

C. Model Building

In this stage, a machine learning and deep learning model are created by learning from training data and generalising it, then using that knowledge to predict the future and achieve its goal. The model is constructed using a variety of methods, including as Random Forest, Support Vector Machine, K Nearest Neighbors, and others. The Gradient Boosting approach is then used to each model to reduce error rates. Calculating the models performance metrics comes after model creation and is an important phase. The performance of the models is evaluated using recall, accuracy, and precision. The top model's approach is recognised as the best machine learning algorithm for predicting crop yields when these metrics are used to compare the models.

1) Random Forest

A forest is created by growing and combining various decision trees using the supervised machine learning method Random Forest. Both R and Python support its usage for classification and regression issues. Random Forest generates several decision trees that are then blended to produce a more accurate forecast. The Random Forest model is built on the premise that different uncorrelated models (decision trees) perform noticeably better when combined than when used independently.

Each tree provides a classification or a "vote" when using Random Forest for categorization. The categorization with the most "votes" is chosen by the forest. When performing regression with Random Forest, the forest selects the mean of all tree outputs. Random forest is a popular machine-learning algorithm that makes use of the supervised learning technique. It can be used to solve classification and parameter estimation ML problems.

Random Forest model averages the results from multiple decision trees applied to various subgroups of the input dataset. Higher accuracy and overfitting are prevented by the larger number of trees in the forest. The greater number of trees within the forest prevents higher accuracy and overfitting. It operates well and generates accurate predictions of the outcome even with the massive dataset. When a significant amount of data is absent, accuracy can still be maintained. The random forest

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
DOI: https://doi.org/10.17762/ijritcc.v11i3.6338
***Article Received**: 30 December 2022 **Revised**: 31 January 2023 **Accepted**: 10 February 2023*

takes the dataset as the input and it will divide the data in to the samples of the data it will divide the data into row samples and column samples. And it will divide the data into different decision trees and the output of all decision trees will be average and predict the result shown in Fig 2.

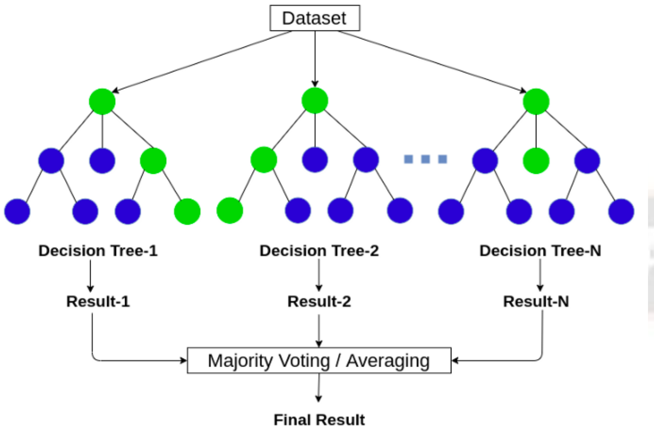

Fig 2: Random Forest

2) Support Vector Machine

SVM is a supervised learning technique used to classify data points into classes, using extreme vectors and points to create a hyperplane. The regression algorithm known as Support Vector Regression, as its name suggests, enables both linear and non-linear regressions. The Support Vector Machine is the basis for how this approach operates. A further configurable parameter for SVR is ε(epsilon). The diameter of the tube surrounding the estimated function is determined by the value of epsilon (hyperplane). The algorithm does not penalized points that fall within of this tube because they are regarded as accurate predictions. Imagine a tube with an estimated function (hyperplane) in the centre and bounds on either side determined by ε. This is a straightforward way to think about SVR. Imagine a tube with an estimated function (hyperplane) in the centre and bounds on either side determined by ε. This is a straightforward way to think about SVR. Shown in figure3.

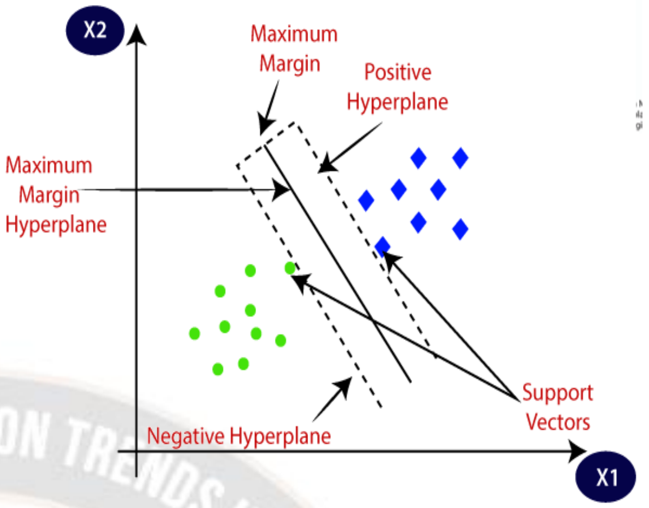

Fig 3: Support vector machine

3) K Nearest Neighbor

One of the most powerful techniques for predicting unknown parameters is the use of algorithmic machine learning.

A crop forecast model that utilizes machine learning is the goal of this project. The application's objective is to predict crop yield to help farmers choose the best seeds for plantations. Numerous machine learning (ML) approaches, such as regression model, support vector machines, neural networks, and K-Nearest Neighbor, can be used (K-NN). K-NN is explored in this work. Problems involving classification and regression can be resolved using the supervised machine learning method known as k-nearest neighbor (K-NN). Although it is easy to use and pick up, a serious factor with it is that as data usage rises, it's becoming considerably slower. Our objective is to use a model where information is concentrated in a small number of groups to predict how a following instance will be classified. Based on the smallest distance between both the query point and the training examples, K-NN determines the k-nearest neighbors. The predictive search object is then selected from the assembled k-nearest neighbors by a simple majority. It can also be used for regression-output, which was previously mentioned as the item's reward. The most popular metric used in the K-NN technique for calculating distance is Euclidean distance.

4) Multi-layer Perceptron Regressor

Among feedforward artificial neural networks, multilayer perceptron Multi-layer perceptron is a subclass. When referring to networks made up of multiple layers of perceptron, the term "MLP" is sometimes used ambiguously, sometimes broadly to apply to any feedforward ANN. An MLP has different levels of input layers and a directed graph joining the output and input layers.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
DOI: https://doi.org/10.17762/ijritcc.v11i3.6338
***Article Received**: 30 December 2022 **Revised**: 31 January 2023 **Accepted**: 10 February 2023*

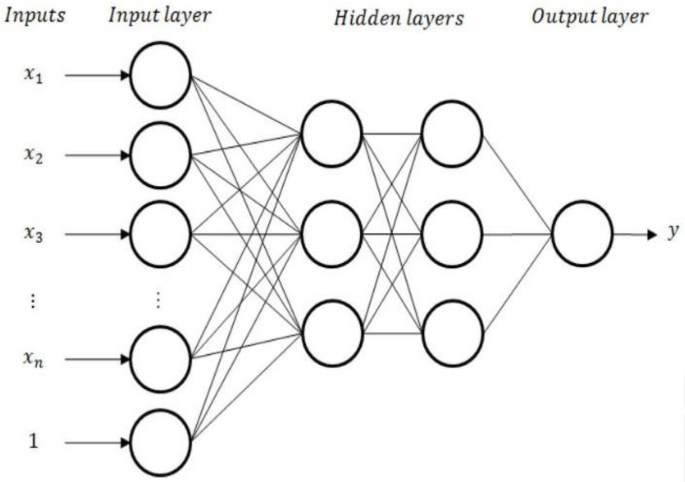

Fig 4: Multi-layer Perceptron Regressor

5) Gradient Boosting Regressor

In machine learning, "boosting" is a method of developing composite models by combining several simple models. The fact that weak learners are inserted one at a time while maintaining the model's current trees unchanged is another reason how some boosting is described as an additive model. The ultimate, full model gets stronger as we combine increasingly more simple models. Gradient boosting gets its name due to the fact that the technique utilizes gradient descent to reduce loss.

In gradient boosting, decision tree algorithm is used as the poor learners. Decision Tree uses a tree description of the data to resolve the machine learning issue. Each node in the tree of the tree visualization reflects a class name, while each internal node speaks for an attribute. The squared error is generally the loss function. A differentiable loss function is needed. Gradient Boosting Regression uses the same residuals techniques as linear regression. The difference between the existing forecast and the established correct target value is obtained using gradient boosting regression.

This variance is called to as residual. After that, a weak classifier that transfers features to that residual is learned using gradient boosting regression. This technique moves the model closer to the desired outcome by adding the excess forecasted by a weak algorithm to the input of the current model.

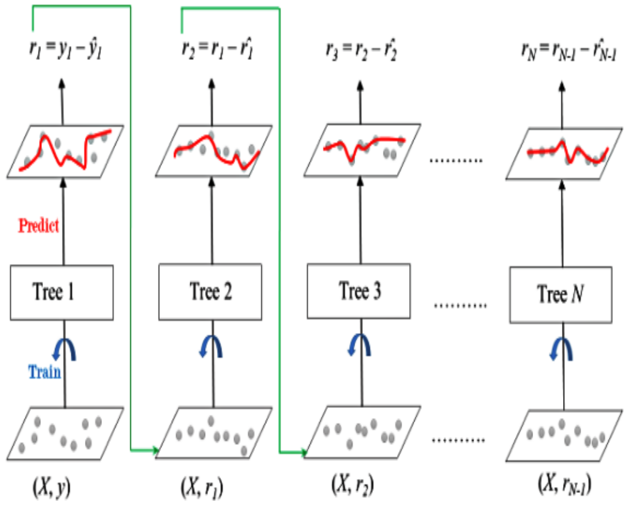

Fig 5: Gradient Boosting Regressor

6) GrowNet

Primitive neural networks are used as weak learners in a unique gradient boosting scheme. This unified framework incorporates general loss functions and presents specific examples for classification, regression, and learning to rank. To address the problem of greedy function optimization in traditional gradient boosting decision trees, an entirely corrective step is added. On various datasets, the suggested model exceeded results obtained using state-of-the-art boosting techniques in all three tasks. To clarify the impact of each modelling framework and model hyperparameter, an ablation study is conducted. Simple is the GrowNet concept. GrowNet uses neural network models as weak learners as opposed to a gradient-boosting decision tree (GBDT). The results of the prior classes are used as inputs by GrowNet.

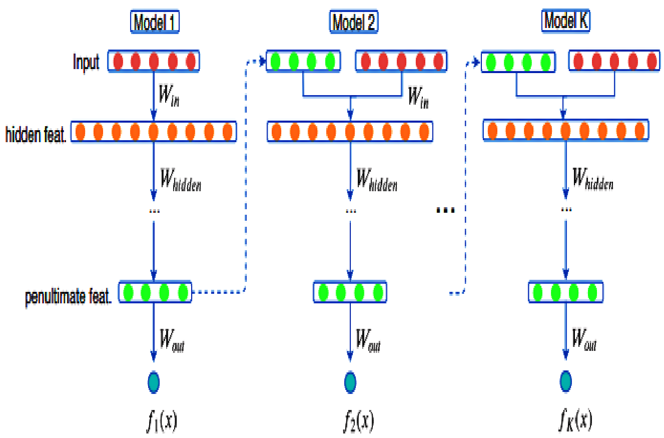

Fig 6: GrowNet

D. Performance Analysis Metrics

To establish the link between dependent and independent variables, methods of machine learning like regression are

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
DOI: https://doi.org/10.17762/ijritcc.v11i3.6338
***Article Received**: 30 December 2022 **Revised**: 31 January 2023 **Accepted**: 10 February 2023*

applied. We will use the placement dataset, which is a simple linear dataset, to illustrate each assessment measure using the sci-kit-learn module.

1. **Mean Absolute Error(MAE)**

A really simple method known as MAE is used to calculate the relative difference between the current and projected values. By summing up all the errors and dividing the result by the total number of observations, the MAE is determined. Furthermore, because this is a loss, obtaining a minimal MAE is our objective.

$$\mathbf{MAE} = \frac{\mathbf{1}}{\mathbf{N}} \Sigma \left| \mathbf{y} - \hat{\mathbf{y}} \right|$$

Where, Σ = Sum of absolute value of residual

y =Actual Output

$\hat{y}$ =Predicted Output

$\frac{1}{N}$ = Divide by total Number of datapoints

2. **Root Mean Squared Error(RMSE)**

The mean squared error's square root is a simple calculation. Since the resulting value you receive utilizes the same unit as the required output variable, loss analysis is straightforward. Because MSE is less tolerant to outliers than MAE, we must use the NumPy. NumPy squared function for computing RMSE.

$$\mathbf{RMSE=\sqrt{MSE}}$$

$$\mathbf{RMSE} = \sqrt{\frac{\mathbf{1}}{\mathbf{n}} \sum_{\mathbf{j=1}}^{\mathbf{n}} \left(\mathbf{y_j} - \hat{\mathbf{y}}_{\mathbf{j}}\right)}$$

3. **R Squared(R^2)**

Instead of measuring performance by the number of wells it did, the R^2 score is a statistic that shows how well your model works. As we've shown, whereas MAE and MSE dependent on context, the R^2score is context-independent. R squared offers a baseline model that neither of the other measures give, allowing us to compare other models against it. In classification problems, we refer to a similar thing as a boundary, which is fixed at 0.5. A regression line's superiority to a mean line is normally estimated by R squared.

$$\mathbf{R^2} = \mathbf{1} - \frac{\mathbf{RSS}}{\mathbf{TSS}}$$

Where, RSS=Sum squared error of regression line

TSS=Squared sum error of mean line

IV. EXPERIMENTATION AND RESULTS

A. Dataset

Data on the State, District, Year, Season, Crop, Area, Production, Yield, Soil, Salinity of Water, and Rainfall are included in the dataset. The remaining 20% of the data is used for testing, and the remaining 80% is used for training.

B. Result

In this study, soil and environmental parameters were incorporated into a dataset from agriculture that was not made available to the general public. For the purpose of this research, information gathered from the agricultural community was therefore used. The metrics of mean absolute error (MAE), root mean squared error (RMSE), and R squared (R2) were used in this study to evaluate the effectiveness of the selecting features and regression approaches.

1) Random Forest

The performance of the Random Forest model is displayed on the graph. The green line in this graph shows the prediction of the chosen features, the blue line shows the prediction of all dataset features, and the yellow line shows the actual yield shown in Figure 7.

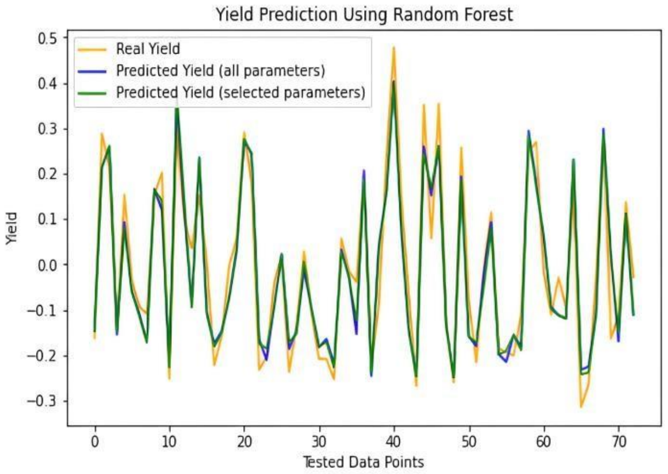

Fig 7: Yield Prediction Using Random Forest

2) Support Vector Regressor

The performance of the Support vector model is displayed on the graph. The green line in this graph shows the prediction of the chosen features, the blue line shows the prediction of all dataset features, and the yellow line shows the actual yield shown in Fig 8.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
DOI: https://doi.org/10.17762/ijritcc.v11i3.6338
***Article Received**: 30 December 2022 **Revised**: 31 January 2023 **Accepted**: 10 February 2023*

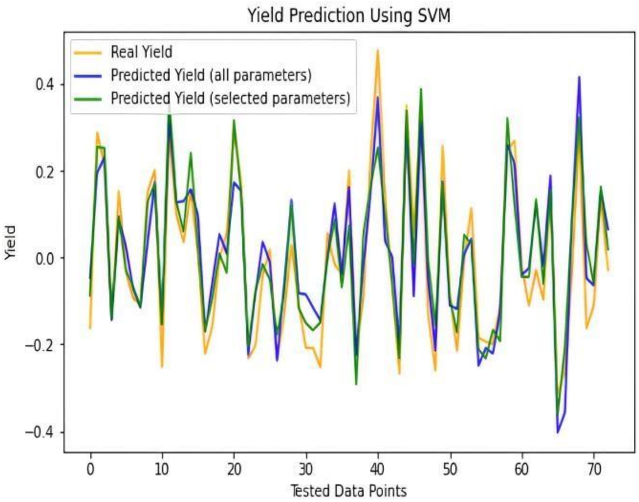

Fig 8: Yield Prediction Using Support Vector Machine

3) K Nearest Neighbor

The performance of the KNN model is displayed on the graph. The green line in this graph shows the prediction of the chosen features, the blue line shows the prediction of all dataset features, and the yellow line shows the actual yield shown in Fig 9.

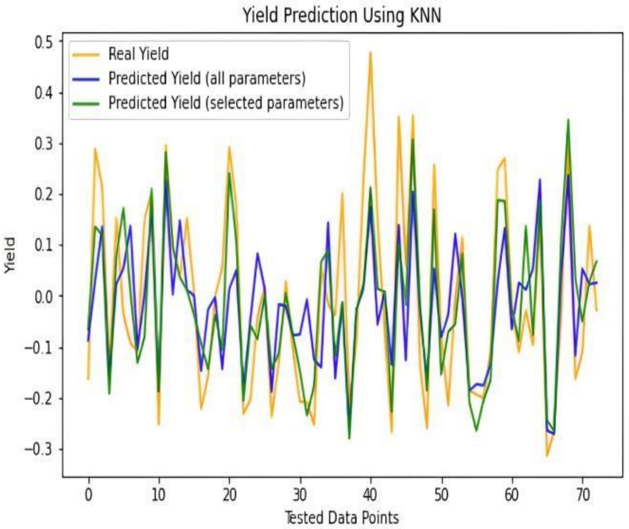

Fig 9: Yield Prediction Using KNN

4) Multi-layer Perceptron Regressor

The performance of the MLP model is displayed on the graph. The green line in this graph shows the prediction of the chosen features, the blue line shows the prediction of all dataset features, and the yellow line shows the actual yield shown in Fig 10.

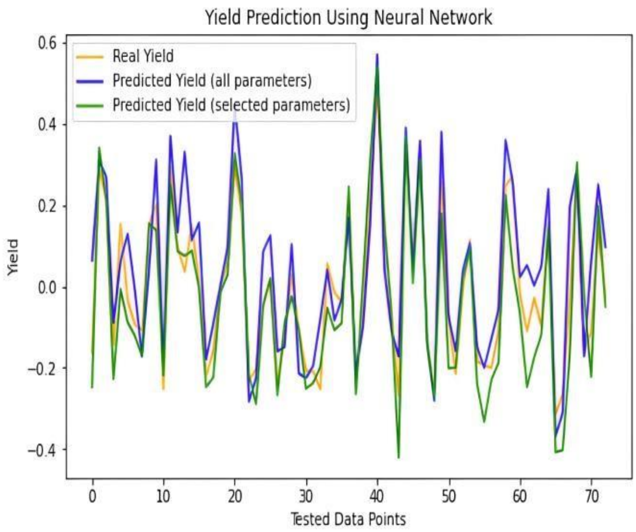

Fig 10: Yield Prediction Using MLP

5) Gradient Boosting Regressor

The performance of the Gradient Boosting Regressor model is displayed on the graph. The green line in this graph shows the prediction of the chosen features, the blue line shows the prediction of all dataset features, and the yellow line shows the actual yield shown in Figure 11.

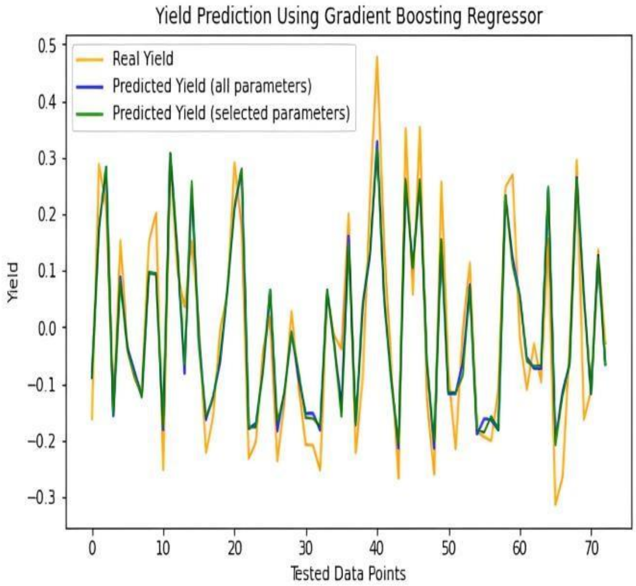

Fig 11: Yield Prediction Using Gradient Boosting Regressor

6) GrowNet

The performance of the GrowNet model is displayed on the graph. The green line in this graph shows the prediction of the chosen features, the blue line shows the prediction of all dataset features, and the yellow line shows the actual yield shown in Fig 12.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
DOI: https://doi.org/10.17762/ijritcc.v11i3.6338
***Article Received**: 30 December 2022 **Revised**: 31 January 2023 **Accepted**: 10 February 2023*

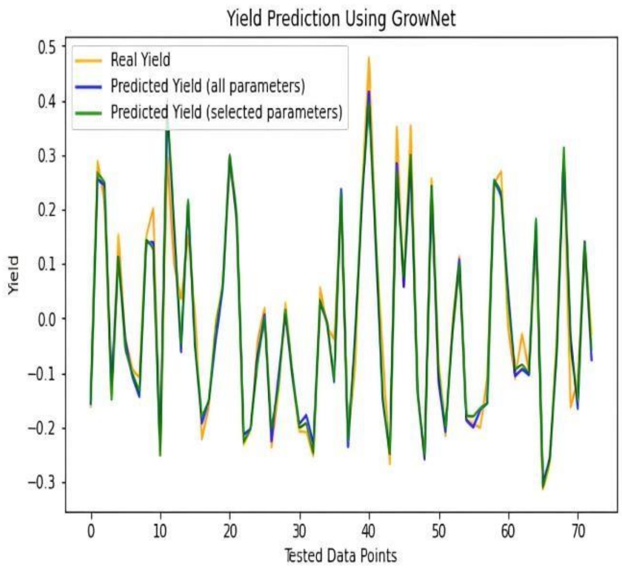

Fig 12: Yield Prediction Using GrowNet

According to Table 1, the K Nearest Neighbor algorithm provides the least accuracy of 70%. Where the Support Vector Machine and Multi-Layer Perceptron Regressor have an accuracy of 83%. Gradient Boosting Regressor provides an accuracy of 87%. And Random Forest provides an accuracy of 88%. The highest accuracy was provided by the GrowNet Model with accuracy of 96%.

TABLE. 1: Performance Analysis Metrics

Models Used	**MAE**	**RMSE**	**R2 score**
Random Forest	0.050	0.063	0.887
Support Vector Machine	0.060	0.075	0.839
K Nearest Neighbour	0.079	0.101	0.705
Multi-layer Perceptron Regressor	0.333	0.416	0.832
Gradient Boosting Regressor	0.268	0.3578	0.876
Grow Net	0.165	0.200	0.961

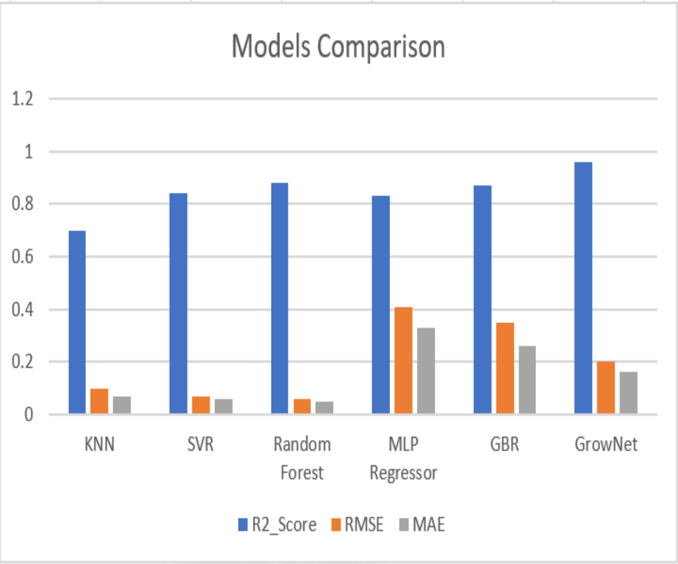

Fig. 13: Models Comparison

V. CONCLUSION

This paper's primary objective is to determine with precision the yield of different crop in Rajasthan. In this situation, we had to build the model using the training data, which was made by splitting the real data set into the training data set and the data set that was sent to different machine learning algorithms, which accounted to 80% of the data set. The qualified model is provided the test data set when the training set is complete, which is then used to compare the model's accuracy to that of other models. The K Nearest Neighbour algorithm, which is the least effective of the three, has an accuracy rate of 70%, according to our research. The MLP and Support Vector Machine techniques provides an accuracy of 83%, which are comparable to adequate when comparison towards the Gradient Boosting Regressor and Random Forest. The best result is generated by GrowNet Regression Model of accuracy 96% which is highest among these algorithms.

REFERENCES

[1] S. P. Raja, B. Sawicka, Z. Stamenkovic and G. Mariammal, "Crop Prediction Based on Characteristics of the Agricultural Environment Using Various Feature Selection Techniques and Classifiers," in IEEE Access

[2] A. F. Haufler, J. H. Booske and S. C. Hagness, "Microwave Sensing for Estimating Cranberry Crop Yield: A Pilot Study Using Simulated Canopies and Field Measurement Testbeds," in IEEE Transactions on Geoscience and Remote Sensing

[3] Suresh, N., Ramesh, N. V. K., Inthiyaz, S., Priya, P. P., Nagasowmika, K., Kumar, K. V. N. H., … Reddy, B. N. K. (2021). Crop Yield Prediction Using Random Forest Algorithm.2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS).

ISSN: 2321-8169 Volume: 11 Issue: 3
DOI: https://doi.org/10.17762/ijritcc.v11i3.6338
***Article Received**: 30 December 2022 **Revised**: 31 January 2023 **Accepted**: 10 February 2023*

[4] Rashid, M., Bari, B. S., Yusup, Y., Kamaruddin, M. A., & Khan, N. (2021). A Comprehensive Review of Crop Yield Prediction Using Machine Learning Approaches With Special Emphasis on Palm Oil Yield Prediction. IEEE Access

[5] Khaki, S., Wang, L., &Archontoulis, S. V. (2020). A CNN-RNN Framework for Crop Yield Prediction. Frontiers in Plant Science

[6] Khaki, S., & Wang, L. (2019). Crop Yield Prediction Using Deep Neural Networks. Frontiers in Plant Science

[7] Nishant, P. S., SaiVenkat, P., Avinash, B. L., & Jabber, B. (2020). Crop Yield Prediction based on Indian Agriculture using Machine Learning. 2020 International Conference for Emerging Technology (INCET).

[8] M. Qiao et al., "Exploiting Hierarchical Features for Crop Yield Prediction Based on 3-D Convolutional Neural Networks and Multikernel Gaussian Process," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing

[9] Dr. V. LathaJothi, Neelambigai A, NithishSabari S, Santhosh K, 2020, Crop Yield Prediction using KNN Model, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) RTICCT – 2020 (Volume 8 – Issue 12).

[10] Bhosale, S. V., Thombare, R. A., Dhemey, P. G., &Chaudhari, A. N. (2018). Crop Yield Prediction Using Data Analytics and Hybrid Approach. 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA).

[11] X. E. Pantazi; D. Moshou; T. Alexandridis; R. L. Whetton;A.M. Mouazen., "Wheat yield prediction using machine learning and advanced sensing techniques" Journal by Elsevier BV. Satorre, E.; Slafer, G. (Eds.) Wheat: Ecology and Physiology of Yield Determination; Haworth Press Technology and Industrial: Binghamton, NY, USA, 1999.

[12] Reynolds, M.; Foulkes, M.; Slafer, G.; Berry, P.; Parry, M.; Snape, J.; Angus, W. Raising yield potential in wheat. J. Exp. Bot. 2009, 60, 1899–1918.

[13] Reynolds, M.; Bonnett, D.; Chapman, S.; Furbank, R.; Manés, Y.; Mather, D.; Parry, M. Raising yield potential of wheat. l. Overview of a consortium approach and breeding strategies. J. Exp. Bot. 2011, 62, 439–452.

[14] Parry, M.; Reynolds, M.; Salvucci, M.; Raines, C.; Andralojc, P.; Zhu, X.; Price, G.; Condon, A.; Furbank, R.Raising yield potential of wheat. ll. Increasing photosynthetic capacity and efficiency. J. Exp. Bot. 2011,62, 453–467.

[15] Foulkes, M.; Slafer, G.; Davies, W.; Berry, P.; SylvesterBradley, R.; Martre, P.; Calderini, D.; Griffiths, S.;Reynolds, M. Raising yield potential of wheat. III. Optimizing partitioning to grain while maintaining lodging resistance. J. Exp. Bot. 2011, 62, 469– 486.

[16] Zhang, W.; Kokubun, M. Historical changes in grain yield and photosynthetic rate of rice cultivars released in the 20th century in Tohoku Region. Plant Prod. Sci. 2004, 7, 36–44.

[17] Grassini, P.; Eskridge, K.; Cassman, K. Distinguishing between yield advances and yield plateaus in historical crop production trends. Nat. Commun. 2013, 4, 2918.

[18] Kumar, P. M., Hong, C. S., Babu, G. C., Selvaraj, J., & Gandhi, U. D. (2021). Cloud-and IoT-based deep learning technique incorporated secured health monitoring system for dead diseases. Soft Computing, 1-16.

[19] Phusanisa Charoen-Ung and Pradit Mittrapiyanuruk. Sugarcane Yield Grade Prediction Using Random Forest with Forward Feature Selection and Hyper-parameter Tuning. Springer International Publishing AG, part of Springer Nature 2019 H. Unger et al. (Eds.): IC2IT 2018, AISC 769, pp. 33–42, 2019.https://doi.org/10.1007/978-3-319-93692-5_4

[20] Ana Paula Marques Ramosa, Lucas Prado Oscob, Danielle Elis Garcia Furuyaa. A random forest ranking approach to predict yield in maize with uav-basedvegetation spectral indices.Computers and Electronics in Agriculture 178 (2020) 105791.https://doi.org/10.1016/j.compag.2020.105791

[21] Safa, M., Samarasinghe, S., 2011. Determination and modelling of energy consumption in wheat production using neural networks: "A case study in Canterbury province, New Zealand. Energy 36 (8), 5140–5147.https://doi.org/10.1011/j.energy.

[22] Pahlavan, R., Omid, M., Akram, A., 2012. Energy input output analysis and application of artificial neural networks for predicting greenhouse basil production. Energy 37,171–176.

[23] Soltanali, H., Nikkhah, A., Rohani, A., 2017. Energy audit of Iranian kiwifruit production P.S. Maya Gopal and R. Bhargavi Computers and Electronics in Agriculture 165 (2019) 104968 using intelligent systems. Energy 139, 646–654.

[24] Taheri-Rad, A., Khojastehpour, M., Rohani, A., Khoramdel, S., Nikkhah, A., 2017. Energy flow modeling and predicting the yield of Iranian paddy cultivars using artificial neural networks. Energy 135, 405–412.

[25] Gonzalez-Sanchez, A., Frausto-Solis, J., Ojeda-Bustamante, W., 2014. Attribute selection impact on linear and nonlinear regression models for crop yield prediction. Sci.World J. 5, 429– 434. https://doi.org/10.1155/2014/509429.

[26] Devi, G. U., Priyan, M. K., & Gokulnath, C. (2018). Wireless camera network with enhanced SIFT algorithm for human tracking mechanism. International Journal of Internet Technology and Secured Transactions, 8(2), 185-194.

[27] Hansen, J.W., Indeje, M., 2004. Linking dynamic seasonal climate forecasts with crop simulation for maize yield prediction in semi-arid Kenya. Agricultural and Forest Meteorology 125 (1–2), 143–157.