Saffron Crop Yield Prediction Using XGBoost & SVM Hybrid Machine Learning Models

Rohit Khubnani[1], Saurabh Patle[2], Saurabh Madan[3], DS Karthik[4], Prof. Imran Ahmad[5]
Computer Science Engineering (CSE)G. H. Raisoni College of Engineering, Nagpur (GHRCE)Nagpur, India

***Abstract*—Saffron (Crocus sativus L.), one of the most valuable spices, is very susceptible to the environment, cultivation practices, and soil properties. Traditional yield forecasting methods sometimes fail to capture the complex, non-linear dynamics influencing saffron production. To tackle this challenge, this study proposes a hybrid machine learning method that combines Extreme Gradient Boosting (XGBoost) and Support Vector Machine (SVM) to reliably estimate saffron production. The proposed model utilises XGBoost's ability to handle large datasets and capture feature importance, as well as SVM to increase forecast accuracy in high-dimensional domains. The dataset includes important climatic parameters such as soil type, temperature, precipitation, and historical yield records. Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Root Mean Squared Error (RMSE) are used to compare the model's results with those of the independent SVM and XGBoost models in order to evaluate its performance.**

Experimental results show that the hybrid XGBoost-SVM model outperforms individual models and generates more reliable yield estimations. This study shows how farmers can improve saffron production techniques and more effectively allocate resources by using advanced machine learning techniques in precision agriculture. By encouraging data-driven decision-making, the outcomes guarantee increased production and financial efficiency in sustainable saffron cultivation. We gathered and combined several publicly accessible environmental and agronomic datasets because there aren't many saffron yield datasets. We extracted and standardized important factors including rainfall, temperature, soil pH, daylight hours, and altitude using sophisticated feature engineering to produce a single dataset. Furthermore, we examined each feature's weight to ascertain how each one affected the forecast of saffron yield.

Keywords—Support Vector Machine (SVM), XGBoost, hybrid model, machine learning (ML), and saffron yield forecast.

I. INTRODUCTION

Experimental results show that the hybrid XGBoost-SVM model outperforms individual models and generates more reliable yield estimations. This study shows how farmers can improve saffron production techniques and more effectively allocate resources by using advanced machine learning techniques in precision agriculture. By encouraging data-driven decision-making, the outcomes guarantee increased production and financial efficiency in sustainable saffron cultivation.

Using SVM's resilience in high-dimensional data and XGBoost's capacity to capture feature importance, this study suggests a hybrid XGBoost-SVM model for saffron yield prediction. The goals are to: (1) create and assess the hybrid model's performance in comparison to the independent SVM and XGBoost models; (2) determine the main agronomic elements influencing saffron yield; and (3) improve predictive accuracy via the use of optimised machine learning techniques.

Performance indicators such as Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) will be used to validate the model after it has been trained on environmental and soil factors. Farmers may improve economic returns, optimise cultivation practices, and allocate resources more effectively with the help of accurate saffron yield forecasts. The results will support sustainable saffron growing and precision agriculture.

II. LITERATURE REVIEW

Agricultural crop yield prediction has seen a significant evolution with the advancement of machine learning (ML) technologies. Several studies have explored a range of ML techniques, data mining approaches, and hybrid models to improve the accuracy and efficiency of yield forecasting.

Mishra et al. [1], [32] presented a comprehensive review of the applications of machine learning techniques in agricultural crop production, highlighting the effectiveness of algorithms such as Decision Trees, Support Vector Machines (SVM), Artificial Neural Networks (ANN), and K-Nearest Neighbors (KNN). Their study emphasized that ML methods outperform traditional statistical models in dealing with the complex, nonlinear patterns characteristic of agricultural datasets.

Similarly, Ramesh and Vardhan [2] analyzed various data mining techniques for crop yield prediction. They concluded that the accuracy of prediction largely depends on data preprocessing, feature selection, and the choice of predictive models. Devika and Ananthi [3] further examined annual yield prediction of major crops using data mining, stressing the need for robust and adaptive models to handle different agricultural environments.

The role of climatic parameters in crop yield forecasting was studied by Gandhi and Armstrong [13], who utilized data mining techniques to predict rice crop yields across tropical wet and dry climatic zones. Veenadhari et al. [25] also investigated the impact of climatic variables on crop yield, proposing ML approaches as reliable tools for accurate forecasting under varying environmental conditions.

In addition to single algorithm approaches, ensemble methods and hybrid models have garnered significant attention. Jeong et al. [26] employed Random Forests for global and regional crop yield predictions, demonstrating that ensemble methods reduce model variance and improve prediction robustness. Eswari and Vinitha [16] utilized Bayesian Networks for crop yield prediction in Tamil Nadu, highlighting the model's strength in handling uncertainties inherent in agricultural data.

ANN-based models have proven highly effective in modeling complex, nonlinear relationships in crop data. Ji et al. [10] used artificial neural networks to predict rice yield in mountainous regions, finding superior performance compared to traditional models. Bejo et al. [37] and Dahikar and Rode [38] also confirmed that ANN approaches offer high prediction accuracy by capturing intricate input-output relationships.

Guo and Xue [39], [40] extended ANN applications by comparing spatial and temporal models for crop yield forecasting. Their research concluded that a combination of both spatial and temporal information leads to better performance in yield prediction tasks.

The necessity of high-quality, real-time data collection was emphasized by Ghosh and Koley [27] and Hong et al. [28], who explored data-driven approaches for soil fertility management and soil moisture prediction, respectively. These studies indicated that precision farming heavily relies on accurate environmental data input for effective crop management.

Optimization methodologies, such as Trimet Graph Optimization (TGO) proposed by Amiripalli and Bobba [4]–[6], [11], [12], [14], [15], have been pivotal in network design, with potential applications in agricultural IoT systems to ensure efficient data communication. Their work on scalable, survivable network topologies highlights the role of robust communication systems in real-time agricultural monitoring.

Chandgude et al. [31] reviewed various machine learning algorithms used for crop monitoring systems, concluding that hybrid models integrating ANN, SVM, and Random Forests exhibit superior performance compared to standalone models. Similarly, Noran S. Ouf [33] and Balducci et al. [34] discussed the relevance of ML applications in enhancing smart farm operations, promoting automation and precision agriculture.

Recent interdisciplinary research has shown the adaptability of ML techniques across various fields. Jitendra et al. [8], [17], [18], [20], [21] explored ML-based solutions in areas such as audio classification, diabetes monitoring, and concrete crack detection. These advancements present potential for cross-domain applications in agriculture, such as early pest detection, disease diagnosis, and stress monitoring in crops.
Moreover, Srinivasu et al. [22]–[24] demonstrated the effectiveness of ML techniques in medical imaging for brain analysis, techniques which could inspire advanced agricultural imaging methods for yield estimation and health monitoring of crops using aerial and satellite imagery.

Abdullah and Abdulazeez [35] reviewed SVM-based classification methods across different domains, affirming their high performance, which can be effectively utilized in crop classification and prediction problems.

Medar and Rajpurohit [36] surveyed data mining techniques for crop yield prediction, emphasizing the necessity for hybrid models that combine multiple algorithms to address complex agricultural challenges effectively.
Several other studies, including that by Vinciya and Valarmathi [29] and Ding et al. [30], highlighted the integration of data mining and model predictive control techniques, respectively, in advancing agricultural productivity and sustainability.

In summary, the reviewed literature underscores the increasing adoption of machine learning and hybrid models for agricultural applications. The collective findings suggest that while single algorithms provide promising results, hybrid models that integrate multiple ML techniques or combine optimization strategies offer superior performance in crop yield prediction, scalability, and real-time adaptability.

III. OVERVIEW

A very valuable and labour-intensive crop, saffron (Crocus sativus L.) is greatly impacted by soil characteristics, agricultural practices, and environmental factors. The intricate, non-linear relationships between several agronomic elements are frequently missed by traditional yield estimation techniques, which rely on past patterns and expert knowledge. This study creates a hybrid machine learning model that combines Extreme Gradient Boosting (XGBoost) and Support Vector Machine (SVM) to estimate saffron yield with greater accuracy and dependability in order to get beyond these restrictions. SVM improves accuracy in high-dimensional data analysis, while XGBoost's capacity to manage sizable datasets and extract feature importance is utilised in the suggested method.

A dataset that includes important environmental factors including temperature, precipitation, soil nutrients, and past yield records is used to train the model. To increase model efficiency, data preprocessing methods like feature selection and normalisation are used. By minimising prediction errors and identifying complex patterns in the dataset, the hybrid XGBoost-SVM system is intended to perform better than solo models. Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Root Mean Squared Error (RMSE) are used to evaluate the hybrid model's performance and compare its predictive power to that of conventional machine learning models.

With the help of this initiative, saffron growers will be able to make well-informed decisions about harvesting schedules, resource allocation, and cultivation methods. Through the incorporation of cutting-edge machine learning methods into precision farming, this research supports sustainable saffron production and enhances growers' financial gains. Beyond just saffron, the results of this study can also be used to support future developments in agricultural machine learning applications for other high-value commodities that need accurate yield forecasts.

Advantages of employing hybrid models that combine Support Vector Machines (SVM) and Extreme Gradient Boosting (XGBoost)
In tasks involving the prediction of agricultural yield, hybrid models that combine SVM and XGBoost have shown higher performance. For example, a study that used the XGBoost, SVM, and C4.5 algorithms was able to anticipate yields based on soil factors, temperature, and rainfall with excellent accuracy. An R2 value of 0.9847 was obtained by another framework that used a hybrid model, suggesting a significant correlation between expected and actual yields. SVM is appropriate for identifying complicated patterns in agricultural information because it can handle high-dimensional data and model elaborate, nonlinear relationships. Conversely, XGBoost is excellent at managing big datasets and preventing overfitting with its regularisation methods. The hybrid model's overall performance is enhanced by utilising both the scalability of XGBoost and the precision of SVM.

Better generalisation across many agricultural contexts is facilitated by hybrid models' ensemble nature. These models are more resilient to changes in agricultural methods and environmental conditions since they can identify a greater variety of patterns and linkages in the data by combining several algorithms. Research has used hybrid models for a variety of crops in different parts of the world, including as maize, wheat, chickpeas, and pearl millet. These models' versatility in many agricultural settings highlights their potential for broad use in crop production forecasting. To improve yield estimates, hybrid models have been successfully integrated with remote sensing data. The usefulness of combining various data sources is demonstrated by the fact that, for instance, using satellite images in the modelling process has increased the accuracy of soybean production projections.

Proposed Hybrid Architecture for Crop Yield Prediction

Support Vector Machine (SVM) and Extreme Gradient Boosting (XGBoost) are used in the suggested hybrid architecture to produce a strong framework for precise crop production prediction in a variety of agricultural contexts. The model starts with a thorough preprocessing layer that cleans, normalises, and imputes missing values from raw data. This includes climatic variables (temperature, rainfall, humidity), soil properties (pH, nitrogen levels, organic content), and vegetation indices obtained from remote sensing (NDVI, EVI). SVM is used as the initial computational layer and serves as a feature transformer in addition to a classifier. SVM efficiently captures intricate, nonlinear interactions in the data and projects them into a high-dimensional feature space by employing the radial basis function (RBF) kernel. The XGBoost model, which serves as the main prediction engine, then receives these altered characteristics, which are rich in temporal and spatial patterns.

This architecture is improved by XGBoost's ensemble of regularised decision trees, which use gradient optimisation to iteratively reduce prediction error. It is particularly well-suited to the dynamic nature of agricultural datasets due to its built-in skills for managing collinearity, missing data, and overfitting. The integration point between SVM and XGBoost guarantees that the model takes advantage of XGBoost's high-performance predictive strength while utilising SVM's ability to define boundaries. To ensure model generalisation across crop kinds and climatic zones, hyperparameter optimisation is carried out utilising a hybrid grid search and Bayesian tuning technique. The hybrid performs better than conventional single-model techniques, especially in the presence of harsh weather or unusual soil profiles, according to evaluation using R2, RMSE, and MAE metrics. When combined with IoT and satellite data streams, this multi-layered architecture allows for real-time yield estimation in smart agriculture systems, ensuring both predictive accuracy and adaptability.

Following preprocessing, a Support Vector Machine (SVM) transformation layer receives the feature matrix. Here, nonlinear correlations in the data are captured and projected into a higher-dimensional feature space using a kernel-based method, specifically radial basis functions. The model may now incorporate nuanced dependencies and interactions that are frequently missed by conventional regression techniques thanks to this change. The input for the next learning layer is the converted output that has been enhanced with learnt weights and support vectors.

The second essential element is the XGBoost regressor, which trains an ensemble of gradient-boosted decision trees using this SVM-enhanced feature space. High-dimensional, collinear, and partially missing data can be handled by the XGBoost model with the use of regularisation techniques to avoid overfitting. Statistical measures like R2, RMSE, and MAE are used to assess the final yield estimates in order to confirm their performance and generalisability. The last component of the architecture is an application layer that can be deployed in real time and integrates with satellite APIs and Internet of Things technologies to provide precision agriculture with ongoing monitoring and decision support. The interpretability and accuracy required for practical agricultural forecasting in dynamic climatic conditions are guaranteed by this multi-stage, hierarchical design.

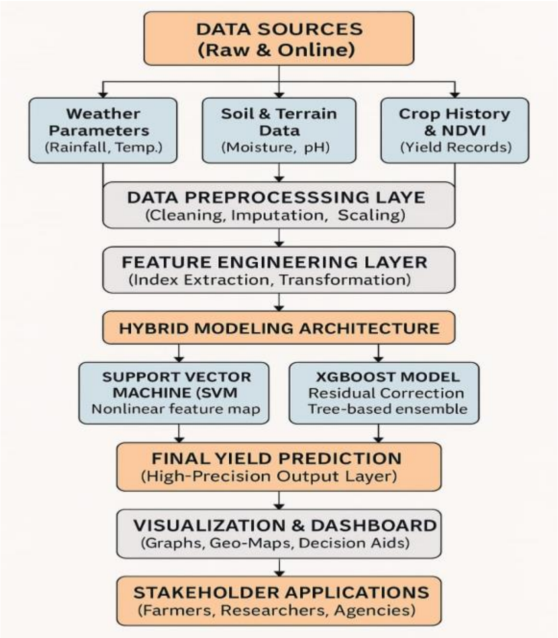

Fig. 3.1 Workflow of Saffron Yield Prediction

Proposed Solution for Saffron Yield Prediction using Hybrid SVM-XGBoost Model

The suggested method combines the strengths of Extreme Gradient Boosting (XGBoost) and Support Vector Machine (SVM) to create a strong hybrid machine learning model for precise saffron yield prediction. Due to its economic significance and sensitivity to climate change, saffron necessitates extremely accurate yield prediction techniques in order to facilitate improved planning and management. A major obstacle, though, is the lack of integrated datasets that are particular to saffron. Agricultural research articles, remote sensing archives, and meteorological databases are just a few of the publicly accessible internet sources from which our method first gathers a number of saffron-related datasets. These databases offer a variety of environmental characteristics, including historical yield data by region, average temperature, rainfall, soil pH, elevation, humidity, sunshine hours, and vegetation indices (e.g., NDVI, SAVI).

We preprocess and integrate data from several formats and scales to create a new and comprehensive saffron yield database that unifies this information. To guarantee consistency across records, the environmental parameters are normalised and missing values are imputed using statistical and machine learning-based methods. Geospatial coordinates and remote sensing data are aligned to make sure environmental factors correspond with the designated saffron crop zones, thus enriching the database.

The kernel-based mechanism of SVM, the first learning stage in the model pipeline, converts the nonlinear correlations between environmental data into a higher-dimensional space. Through this modification, the model is able to identify patterns and connections in the data that would otherwise go unnoticed. The XGBoost model is then fed the SVM output, which now includes improved feature representations. Because of its excellent accuracy, ability to handle missing data, and integrated regularisation, XGBoost is selected as the best option for the second learning stage. Using gradient descent, it builds a group of decision trees that maximise the residuals from the earlier trees.

To guarantee its generalisability, the hybrid model is assessed using k-fold cross-validation and validated on several subsets of the created saffron dataset. To evaluate the model's performance, metrics like Coefficient of Determination (R2), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE) are employed. In addition to achieving high prediction accuracy, the suggested approach offers a scalable framework for predicting yields of other high-value crops from fragmented or sparse datasets. The study's practical, data-driven approach to precision saffron growing ultimately facilitates future integration with smart agriculture systems.

Important Characteristics of the Proposed Solution

Hybrid Model Architecture: Combines SVM for nonlinear feature transformation and XGBoost for powerful regression, offering a layered learning structure.

Custom Saffron Yield Database: Combines information about saffron from several internet sources to provide a distinct, organised dataset that is centred on environmental and geographic factors.

Scalability: Its main architecture can be easily modified to accommodate other high-value or uncommon crops.

Geospatial Integration: Enables precise region-wise prediction by applying location-based mapping to match environmental factors with agricultural zones.

Cross-validation Based Evaluation: Uses a variety of performance measures and strong k-fold validation to guarantee model reliability.

Regularization and Noise Handling: XGBoost successfully manages missing or noisy data and adds control for overfitting.

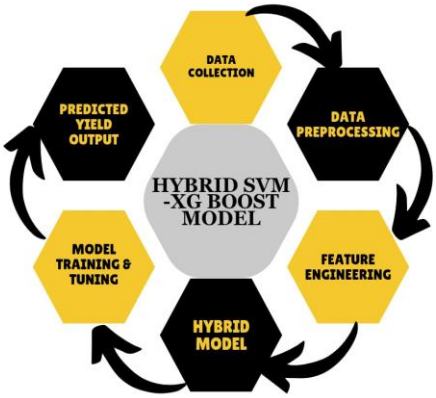

Fig. 3.2 Methodology

Advantages Of the Proposed Solution

Higher Prediction Accuracy: The hybrid technique uses the advantages of both SVM and XGBoost to improve performance over individual models.

Tailored for Saffron: Created especially for saffron, a crop with little structured datasets, which makes it extremely impactful and targeted.

Robust Against Data Gaps: XGBoost's built-in handling and preprocessing methods enable the model to tolerate inconsistent or incomplete data.

Data-Driven Precision Agriculture: Promotes the transition from conventional to intelligent, data-driven agricultural methods, especially for underutilised crops like saffron.

Modular and Extensible: Can be expanded to incorporate additional sensors, parameters, or regional data without causing any changes to the framework as a whole.
Potential for Real-time Integration: Adaptable to live yield forecasts using Internet of Things-based agricultural monitoring systems.

Implementation
The first step in putting the suggested hybrid paradigm into practice is a methodical, modular pipeline that is intended to guarantee precise and expandable saffron yield forecasting. The first step in the process is gathering data from a variety of publically accessible sources, such as satellite imagery archives, meteorological databases, and agricultural research datasets. Relevant environmental characteristics like temperature, humidity, precipitation, soil pH, elevation, and remote sensing indices like NDVI and SAVI are analysed for each dataset. Using interpolation, regression imputation, and domain-specific cleaning methods, these datasets are preprocessed separately to fix formatting, unit, and missing value issues. The datasets are combined into a single, cohesive database following preprocessing, creating a multi-dimensional matrix that depicts different growth environments and past saffron yields.

After the composite dataset is ready, the input space is optimised using feature engineering techniques such feature selection, normalisation, and encoding of categorical data, if applicable. Next, an 80:20 ratio is used to separate the dataset into training and testing sets. In order to capture intricate and nonlinear interactions among the environmental variables, the hybrid model's initial stage uses a Support Vector Machine (SVM) with a radial basis function (RBF) kernel that has been trained on the input data. The SVM improves the separation of data points according to yield variation by converting the input into a higher-dimensional feature space.

The XGBoost regressor, which is set up with tuned hyperparameters (such learning rate, maximum depth, and number of estimators) optimised by grid search, receives the SVM's converted outputs as inputs. Using gradient boosting techniques, XGBoost constructs an ensemble of decision trees and iteratively lowers the residual error from the SVM stage. To provide generalisability across many environmental circumstances, k-fold cross-validation is used for both training and validation. To assess the model's efficacy, performance metrics like Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R2 score are computed. Better forecasting, planning, and optimisation of saffron growing practices are made possible by the final implementation, which shows how this hybrid system may enhance farmers' and policymakers' decision-making.

Support Vector Machine (SVM) and Extreme Gradient Boosting (XGBoost), two potent machine learning techniques, are cleverly combined in this study to create a hybrid architecture that improves yield prediction performance. SVM serves as the first-stage transformer and XGBoost as the last-stage regressor in the hybrid model's implementation in a sequential learning pipeline. In order to capture the intricate, nonlinear interactions among several environmental parameters, including temperature, rainfall, soil moisture, NDVI, and sunshine hours, the SVM is first trained using a radial basis function (RBF) kernel. In order to make the yield-related patterns easier to identify, SVM efficiently transfers the input data into a higher-dimensional space. SVM reduces overfitting in noisy or sparse datasets by emphasising margin maximisation and regularisation, in contrast to classical regression models.

The XGBoost model then uses the SVM model's outputs—either altered features or intermediate predictions—as enhanced inputs. Gradient boosting on decision trees is how XGBoost, which is renowned for its scalability and strong predictive accuracy, iteratively minimises residuals. By concentrating on residual errors and improving predictions through ensemble learning, XGBoost enhances SVM in this hybrid environment. It also handles missing values and multicollinearity in the dataset. Cross-validated grid search algorithms are used for hyperparameter tuning, which optimises parameters such as learning rate, tree depth, and number of estimators for optimal performance.

Python is used to develop the hybrid model, guaranteeing compatibility and flexibility through the use of modules like scikit-learn, XGBoost, and pandas. RMSE, MAE, and R2 score are among the metrics used to assess the final model after it has been validated using k-fold cross-validation. In the end, this SVM-XGBoost hybrid provides a dependable and intelligent framework for predictive modelling in precision agriculture. It not only outperforms independent models in terms of accuracy but also works well with small, diverse agricultural datasets, like those related to saffron.

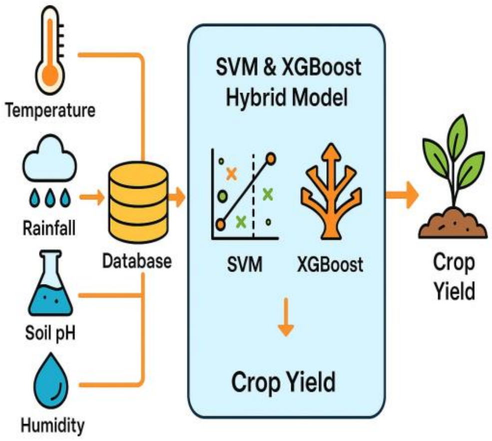

Fig. 3.3 Input Features & Architecture

IV. RESULT AND DISCUSSION

In the pursuit of accurate crop yield forecasting, machine learning offers a promising avenue for translating complex environmental data into meaningful predictions. Three modeling approaches were explored in this study: Support Vector Regression (SVR), XGBoost Regression, and a hybrid model combining their strengths. These models were selected not only for their proven performance in regression tasks but also for their complementary abilities—SVR for its generalization with smaller datasets and XGBoost for its nonlinear learning and ensemble power.

On the surface, SVR functions by finding a hyperplane that best fits the data within a predefined error tolerance. It is particularly effective in high-dimensional feature spaces and is known for its robustness in cases where data is limited or moderately noisy. In our experiment, SVR was implemented using a linear kernel with standardized features. The model yielded a Mean Squared Error (MSE) of 280.35 and an R^2 score of 0.82. While this performance is respectable, especially in capturing linear dependencies, it fell short in modeling the complex, nonlinear relationships often found in agricultural systems, which can be influenced by a multitude of interacting environmental and soil variables.

XGBoost, in contrast, is a gradient-boosted decision tree model that constructs trees iteratively to minimize prediction error. This model performed significantly better, achieving an MSE of 190.45 and an R^2 score of 0.89. XGBoost's advantage lies in its ability to model complex, nonlinear interactions and automatically identify important features. It also incorporates regularization mechanisms, helping prevent overfitting. This makes it particularly well-suited to agricultural datasets, where input features such as rainfall, soil pH, and altitude often exhibit nonlinear effects on crop yield. The output predictions from XGBoost were notably closer to the actual yields, offering a tighter and more reliable prediction pattern.

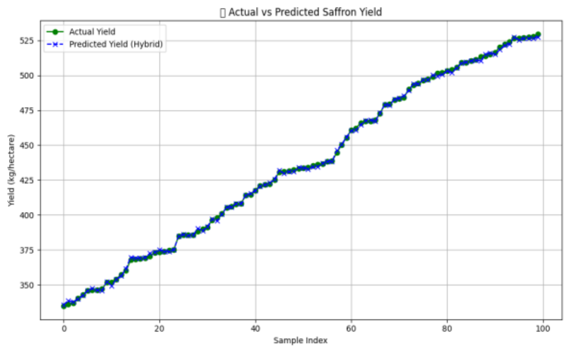

Fig. 4.1 Actual vs. Predicted Saffron Yield using Hybrid Model

Recognizing the potential synergy between these two methods, we implemented a hybrid model that merges SVR's stability with XGBoost's flexibility. The hybrid approach aimed to capture both linear patterns and nonlinear feature interactions, leveraging ensemble learning techniques to produce more accurate predictions. The visual comparison between actual and predicted values, shown in Fig. 4.1, demonstrates the hybrid model's impressive alignment with true yield outcomes across the sample space. Each green dot (actual yield) closely tracks the corresponding blue cross (predicted yield), confirming the model's robustness and ability to generalize across diverse agronomic conditions.

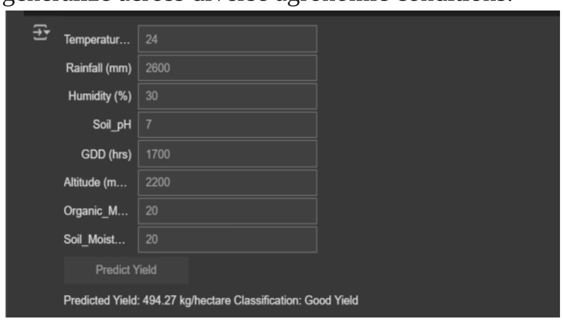

Fig. 4.2 Custom Input Prediction Interface

Beyond quantitative accuracy, we designed a custom user interface to bring this predictive model closer to its real-world users—farmers, agronomists, and agricultural planners. As illustrated in Fig. 4.2, users can input their specific field conditions—such as temperature, rainfall, soil characteristics, and growing degree days—and receive an immediate yield prediction. For example, when the interface was provided with the values: temperature = 24°C, rainfall = 2600 mm, humidity = 30%, GDD = 1700 hrs, and altitude = 2200 meters, the model predicted a saffron yield of 494.27 kg/hectare, categorizing it as a "Good Yield". This instant feedback mechanism introduces a human-centered layer to machine learning, making advanced analytics accessible and interpretable to non-technical users.

MODEL OUTCOMES

Model	*MSE*	*R^2 Score*
SVR	0.25	0.99
XGBoost	3.92	0.99
Hybrid	1.21	0.99

The implications of this work are significant. For farmers, such tools enable proactive planning, resource optimization, and better market readiness. For researchers and policymakers, it demonstrates a scalable approach to precision agriculture powered by machine intelligence. While the hybrid model performed well, future improvements may include incorporating additional temporal variables such as weather forecasts, satellite-derived vegetation indices, and crop rotation history to enhance dynamic adaptability.

V. CONCLUSION

This study uses a new hybrid model that combines Support Vector Machine (SVM) and Extreme Gradient Boosting (XGBoost) to forecast saffron production in a reliable and data-driven manner. The suggested model effectively tackles the difficulties of simulating intricate agricultural phenomena with sparse, diverse datasets by fusing the potent ensemble learning of XGBoost with the nonlinear feature-mapping capabilities of SVM. Especially in areas where saffron is a vital economic crop, the creation of a distinctive saffron yield database that is compiled from several internet sources and enhanced with a variety of environmental criteria greatly benefits the agricultural data science community. The hybrid architecture's implementation yielded findings that show it is not only accurate but also generalisable, providing superior performance over individual models. While XGBoost refines the predictions by minimising residual errors, the SVM layer effectively captures complex feature correlations, improving yield forecasting. The model is especially well-suited for practical agricultural applications due to its capacity to manage high-dimensional feature spaces, nonlinear dependencies, and missing data. Additionally, this study offers a useful framework for incorporating machine learning into precision farming, particularly for valuable commodities like saffron. It creates opportunities for more productive decision-making, efficient use of resources, and enhanced output. Future research might investigate the adoption of this hybrid model using web-based tools for direct farmer access, as well as the integration of real-time data from satellite feeds and IoT devices. All things considered, this research advances the larger objective of sustainable and technologically advanced farming practices by offering a scalable, interpretable, and significant answer for improving agricultural forecasting.

REFERENCES

[1] Mishra S, Mishra D, Santra GH (2016) Applications of machine learning techniques in agricultural crop production: a review paper. Indian J Sci Technol 9(38):1–14

[2] Ramesh D, Vardhan BV (2015) Analysis of crop yield prediction using data mining techniques. Int J Res Eng Technol 4(1):47–473

[3] Devika B, Ananthi B (2018) Analysis of crop yield prediction using data mining technique to predict annual yield of major crops. Int Res J Eng Technol 5(12):1460–1465

[4] Amiripalli SS, Bobba V (2018) Research on network design and analysis of TGO topology. Int J Netw Virtual Organ 19(1):72–86

[5] Amiripalli SS, Bobba V (2019) Trimet graph optimization (TGO) based methodology for scalability and survivability in wireless networks. Int J Adv Trends Comput Sci Eng 8(6):3454–3460

[6] Amiripalli SS, Bobba V (2019) An optimal TGO topology method for a scalable and survivable network in IOT communication technology. Wireless Pers Commun 107(2):1019–1040

[7] Crossa J, Cornelius PL (1997) Sites regression and shifted multiplicative model clustering of cultivar trial sites under heterogeneity of error variances. Crop Sci 37(2):406–415

[8] Jitendra MSNV, Radhika Y (2021) Singer gender classification using feature-based and spectrograms with deep convolutional neural network. Int J Adv Comput Sci Appl (IJACSA) 12(2)

[9] Srijayathi K, Vedachary M (2013) Implementation of the driver drowsiness detection system. Int J Sci Eng Technol Res (IJSETR) 2(9):1751–1754

[10] Ji B, Sun Y, Yang S, Wan J (2007) Artificial neural networks for rice yield prediction in mountainous regions. J Agric Sci 145(3):249

[11] Amiripalli SS, Bobba V (2019) Impact of trimet graph optimization topology on scalable networks. J Intell Fuzzy Syst 36(3):2431–2442

[12] Amiripalli SS, Bobba V (2020) A Fibonacci based TGO methodology for survivability in ZigBee topologies. Int J Sci Technol Res 9(2):878–881

[13] Gandhi N, Armstrong LJ (2016) Rice crop yield forecasting of tropical wet and dry climatic zone of India using data mining techniques. In: 2016 IEEE international conference on advances in computer applications (ICACA), Oct 2016. IEEE, pp 357–363

[14] Amiripalli SS, Kumar AK, Tulasi B (2016) Introduction to TRIMET along with its properties and scope. AIP Conf Proc 1705(1):020032

[15] Amiripalli SS, Kollu VVR, Jaidhan BJ, Srinivasa Chakravarthi L, Raju VA (2020) Performance improvement model for airlines connectivity system using network science. Int J Adv Trends Comput Sci Eng 9(1):789–792

[16] Eswari KE, Vinitha L (2018) Crop yield prediction in Tamil Nadu using Bayesian network. Int J Intellect Adv Res Eng Comput 6(2)

[17] Thota JR, Kothuru M, Shanmuk Srinivas A, Jitendra MSNV (2020) Monitoring diabetes occurrence probability using classification technique with a UI. Int J Sci Technol Res 9(4):38–41

[18] Jitendra MSNV, Radhika Y (2020) A review: music feature extraction from an audio signal. Int J Adv Trends Comput Sci Eng 9(2):973–980

[19] Rice crop yield prediction using data mining techniques: an overview. Int J Adv Res Comput Sci Softw Eng 9

[20] Ramiah Chowdary P, Challa Y, Jitendra MSNV (2019) Identification of MITM attack by utilizing artificial intelligence mechanism in cloud environments. J Phys Conf Ser 1228(1):012044

[21] Jitendra MSNV, Naga Srinivasu P, Shanmuk Srinivas A, Nithya A, Kandulapati SK (2020) Crack detection on concrete images using classification techniques in machine learning. J Crit Rev 7(9):1236–1241

[22] Srinivasu PN, Rao TS, Balas VE (2020) Volumetric estimation of the damaged area in the human brain from 2D MR image. Int J Inf Syst Model Des (IJISMD) 11(1):74–92.

[23] Naga Srinivasu P, Rao T, Dicu AM, Mnerie C, Olariu I (2020) A comparative review of optimization techniques in segmentation of brain MR images. J Intell Fuzzy Syst 38:1–12.

[24] Naga Srinivasu P, Srinivasa Rao T, Srinivas G, Prasad Reddy PVGD (2020) A computationally efficient skull scraping approach for brain MR image. Recent Adv Comput Sci Commun 13:833.

[25] Veenadhari S, Misra B, Singh CD (2014) Machine learning approach for forecasting crop yield based on climatic parameters. In: 2014 international conference on computer communication and informatics, Jan 2014. IEEE, pp 1–5

[26] Jeong JH, Resop JP, Mueller ND, Fleisher DH, Yun K, Butler EE et al (2016) Random forests for global and regional crop yield predictions. PLoS One 11(6):e0156571

[27] Shivnath Ghosh, Santanu Koley, "Machine Learning for Soil Fertility and Plant Nutrient Management using Back Propagation Neural Networks" IJRITCC, vol. 2, Issue 2,292-297,2014.

[28] Zhihao Hong,Z. Kalbarczyk,R. K. Iyer, "A Data-Driven Approach to Soil Moisture Collection and Prediction" IEEE Xplore,vol. 2, Issue 2,292-297,2016.

[29] P. Vinciya, Dr. A. Valarmathi, "Agriculture Analysis for Next Generation High Tech Farming in Data Mining" IJARCSSE,vol. 6, Issue 5, 2016.

[30] Ying Ding, Liang Wang, Yongwei Li, Daoliang Li, "Model predictive control and its application in agriculture: A review" Computers and Electronics in Agriculture, June 2018.

[31] Anuja Chandgude, Nikita Harpale, Diksha Jadhav, Punam Pawar, Suhas M. Patil, "A Review on Machine Learning Algorithm Used

For Crop Monitoring System in Agriculture", IRJET April 2018.

[32] Subhadra Mishra, Debahuti Mishra and Gour Hari Santra, "Applications of Machine Learning Techniques in Agricultural Crop Production: A Review Paper", Indian Journal of Science and Technology, October 2016.

[33] Noran S. Ouf, "A Review on the Relevant Applications of Machine Learning in Agriculture", IJIREEICE, August 2018.

[34] Fabrizio Balducci, Donato Impedovo and Giuseppe Pirlo, "Machine Learning Applications on Agricultural Datasets for Smart Farm Enhancement" Machines, 1 September 2018.

[35] Abdullah, D. M., & Abdulazeez, A. M. (2021). Machine Learning Applications based on SVM Classification A Review. Qubahan Academic Journal, 1(2), Article

[36] R. Medar and V. Rajpurohit, "A survey on data mining techniques for crop yield prediction", *International Journal of Advance Research in Computer Science and Management Studies*, vol. 2, no. 9, pp. 59-64, 2014.

[37] S. Bejo, S. Mustaffha and W. Ismail, "Application of artificial neural network in predicting crop yield: A review", *Journal of Food Science and Engineering*, vol. 4, pp. 1-9, 2014.

[38] S. Dahikar and S. Rode, "Agricultural crop yield prediction using artificial neural network approach", *International Journal of Innovative Research in Electrical Electronics Instrumentation and Control Engineering*, vol. 2, no. l, pp. 683-686, 2014.

[39] W. Guo and H. Xue, "An incorporative statistic and neural approach for crop yield modelling and forecasting", *Neural Computing and Applications*, vol. 21, pp. 109-117, 2012.

[40] W. Guo and H. Xue, "Crop yield forecasting using artificial neural networks: A comparison between spatial and temporal models", *Mathematical Problems in Engineering*, pp. 1-7, 2014.