Grenze International Journal of Engineering and Technology, June Issue

Smart AgroClimate: Leveraging Machine Learning for Integrated Weather & Soil Intelligence to Drive Sustainable, Climate-Resilient Precision Agriculture

Snehal Rathi[1], Shravani Kurkute[2], Yash Baraskar[3], Nikhil Gaikwad[4], Shraddha Bhadane[5] and Prerna Phadnis[6]
[1,2,4-6]Department of Computer Science Engineering Vishwakarma Institute of Information Technology, Pune, India
Email: snehal.rathi@viit.ac.in, shravanikurkute9, gaikwadnikhil193, shraddhabhadane2005, prernaphadnis6}@gmail.com
[3]Department of Electronics and Telecommunication Vishwakarma Institute of Information Technology, Pune, India
Email: yashbaraskar1306

***Abstract*— Accurate soil moisture estimation is essential for farm operations and water resource management. Direct soil moisture models are complex due to factors like soil type, crops, and weather conditions. This paper explores machine learning (ML) techniques to select relevant soil and weather features, based on agricultural and ML studies, and compares them with existing patents to design a cost-effective weather station. Field experiments demonstrate the potential to optimize resource use, maximize crop yields, and support sustainable agriculture. The research highlights ML's role in soil nutrient management, fertilizer guidance, and crop quality assessment. It emphasizes intelligent irrigation scheduling through ML-based control strategies, benefiting economic, social, and environmental aspects. The paper presents a general architecture for smart agriculture, reviews datasets and techniques, addresses challenges with long-term data processing, and suggests future research for scalable, domain-independent irrigation management and resource optimization.**

***Index Terms*— Resilient Agriculture, Intelligent Agricultural Systems, Optimizing Fertilizer, Control Strategies Based on Machine Learning.**

I. INTRODUCTION

Soil moisture significantly influences water distribution in the agricultural hydrological cycle, affecting key farming activities such as crop selection, tilling, planting, fertilizing, and harvesting through processes like infiltration, evaporation, runoff, heat, and gas fluxes [1]. Machine learning algorithms analyze interactions between crop needs, environmental factors, and soil nutrients using advanced models like neural networks, logistic regression, random forests, and LGBM [2]. Contemporary research has advanced agricultural cultivation techniques, field monitoring, and crop yield prediction using ML. Enhancing the dependability and efficiency of automated tasks is essential for sustaining human life. Predicting crop output is challenging due to numerous factors, including humidity, temperature, greenhouse effects, evaporation rate, wind speed, water content, and solar radiation [16]. Studies indicate that accurate crop yield forecasting requires analyzing multiple variables and their interactions, not just educated guesses. There is a growing trend to use machine learning methods to track key crop productivity parameters. Soil quality, particularly temperature and moisture, is crucial for crop growth. However, expensive sensors make comprehensive soil monitoring across large areas difficult.

Grenze ID: 01.GIJET.11.2.135_20

Manipulating soil properties externally in controlled environments like greenhouses is more feasible, but this approach limits widespread agricultural automation [3]. Machine learning methods have gained prominence by surpassing physics-based models' limitations. Ali et al. assert that ML models can understand and predict complex non-linear relationships from sparse data without extensive user knowledge, enabling fusion of vague sources with unknown distributions [4]. Research aims to design sensors and machine learning algorithms to predict agricultural weather and soil conditions, including temperature, humidity, pH, rainfall, and nutrients (Potassium(K), Phosphorus(P), Nitrogen(N)).

The objectives of this study are:

- To analyze association pattern of various weather and soil parameters.
- Create predictive models for suggested crops based on soil and weather conditions.
- Analysis of how well machine learning models perform in predicting with crop along different algorithms.

We aim to establish a machine learning framework in agriculture, boosting crop yield, cutting resource waste, and achieving sustainable farming.

II. LITERATURE SURVEY

Various machine learning techniques—Decision Trees, KNN, Random Forests, SVM, and Neural Networks—are used in agriculture. They are categorized by interpretability, a key feature that allows users to understand and describe each classifier effectively [5]. Soil moisture influences nutrient availability, microorganism movement, molecule transport, soil pH, gas and solvent diffusion (especially oxygen), and metabolism, all essential for soil fertility and productivity. Hurricanes and other weather events significantly impact soil biota and plants, thereby altering global biochemical cycles and ecological services, such as crop productivity, within agricultural settings (van Bergen & Inspect Pollinators Initiative, 2013) [15]. With increasing interest in machine learning, agriscience researchers face new challenges and opportunities. ML models enhance plant disease detection, crop yield prediction, health assessment, and classification, driving efficient agricultural production. Additionally, guidelines for soil, fertilization, water, and weather management improve farming practices [14]. Strothkamper, Tian, and Duman et al. argue that ASCs must increase farming efficiency due to depleting water and fossil fuels, decreasing arable land, and higher consumer demand for sustainable, transparent food chains [6]. Machine learning monitors soil health and detects agricultural diseases. Plantix uses ML to identify plant flaws and diseases from soil patterns, providing farmers with data and advice via smartphones. In Tanzania, deep neural networks detect cassava illnesses and insect damage [7]. ML algorithms predict rainfall and weather using sensors, climate data, and satellite imagery, preventing crop loss for thousands of farmers. Precision farming also manages livestock by monitoring animal well-being, productivity, and reproduction through cameras and sensors, using computer vision to stop disease spread [8]. Chen et al. created a neural network–based rice monitoring system that predicted wet season yields more accurately than official government estimates [9]. Machine learning models can be descriptive or predictive, depending on research goals. Predictive models forecast future events, while descriptive models describe data [10]. Holgren and Thuresson’s forestry work, Engman and Gurney’s hydrology, and Hultquist and Seelan’s agricultural studies have generated vast remote sensing data in geological sciences. As data volumes and diversity [11]. Research on mobile soil sensors indicates that advancements will enhance precision agriculture by providing higher measurement densities at lower costs. Although commercially available options are limited, ongoing prototype development continues. This paper reviews recent approaches for measuring soil’s mechanical, physical, and chemical properties and explores their potential applications [12]. The agri-food sector is increasingly utilizing AI and ML to reduce food waste, enhance machine cleaning, improve production hygiene, and combat illnesses and pests. Automated systems swiftly collect and analyze vast data for single food items, applying AI in supply chain, soil, crop, disease, and pest management [13].

III. METHODOLOGY

A. There are several essential steps in the research

In Fig.1. the machine learning models, of which Gradient Boosting was the most accurate, Showed a lot of potential for crop recommendations. While simpler models like KNN and linear SVC performed decently, more complex models like Random forests and Gradient Boosting showed superior performance. This illustrates the importance of complex models in precision farming decision-making. Important steps in data preprocessing: Heatmap to check for missing values which shows through colors whether there are any missing values in the data set. One of the most important steps that is needed in preparing your data to be Identifying and correcting missing values is crucial for building accurate models, as data can be absent due to reasons like incorrect entry, sensor failures, or measurement issues. Missing data can degrade analysis quality and lead to erroneous conclusions. Therefore, systematically detecting and addressing missing values before analysis or modeling is essential. A heatmap is an effective visual tool for locating missing values, as it displays and color-codes their positions. This makes it easier to identify trends, such as concentrations in specific rows or columns or random distribution across the dataset.

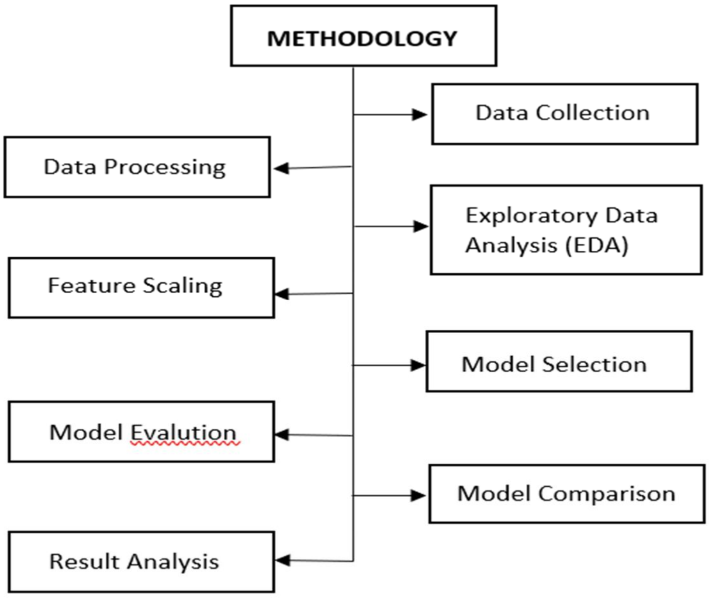

Figure.1. Research methodology steps

B. Selection of Machine Learning Model

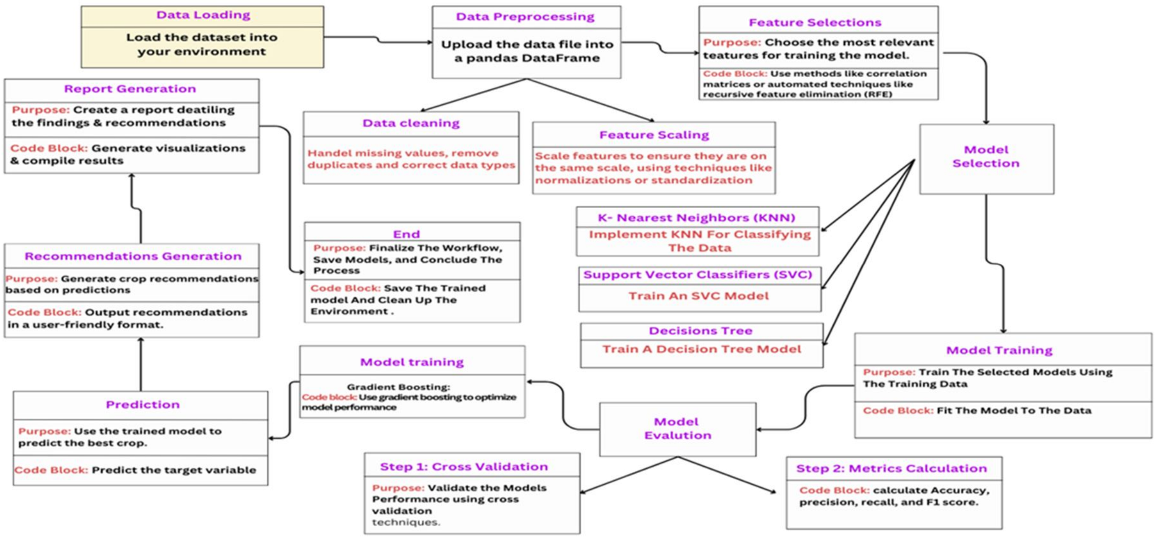

Figure.2. workflow of machine learning model

The Fig 2. A flow diagram for creating a machine learning model begins with preprocessing (data cleaning, handling missing values, scaling). Then, data is loaded into the environment. Feature selection uses algorithms like Recursive Feature Elimination (RFE) to identify top features. Model selection includes methods such as K-Nearest Neighbors (KNN), Support Vector Machines (SVM), and Decision Trees. During training, the data is split, models are tuned and trained, and evaluated using cross-validation metrics like accuracy, precision, and recall. Finally, reports and recommendations are generated based on the trained model.

IV. Result and Discussion

Based on Fig.3. The Bar Graph illustrates machine learning models' effectiveness in forecasting crop recommendations. Gradient Boosting achieved the highest accuracy but is more complex. Polynomial Kernel SVC was second with a higher overfitting risk. Decision Tree and RBF Kernel SVC showed great accuracy; the Decision Tree is easy to interpret but may overfit, while SVC requires more computing power. KNN and Linear Kernel SVC are suitable for smaller datasets and less intensive tasks, providing reliable results.

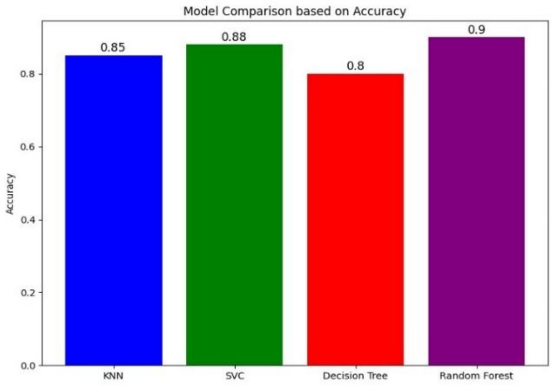

Figure.3. Comparison of different machine learning models

A. K-Nearest Neighbors (KNN):

KNN classifier, an instance-based non-parametric algorithm, achieved high precision on a dataset. It struggles with high dimensionality and noise, resulting in worse performance than other models. Sensitive to k value, slight adjustments enhance functionality. Best for small, distinct-class datasets.

B. Support Vector Classifier(SVC):

SVC surpasses KNN in accuracy across various kernels, especially on datasets with clear class separation. Optimal performance requires hyperparameter tuning via GridSearchCV. The polynomial kernel performs best, while RBF outshines the linear kernel.

C. Decision Tree classifier:

Despite its interpretability, the Decision Tree classifier had the lowest accuracy. It tends to overfit complex patterns and poorly generalizes, especially with significantly differing crop characteristics.

D. The Random Forest Classifier:

Random Forest Classifier, an accurate ensemble method using decision trees, reduces overfitting and enhances robustness. It identifies temperature, N, and K as vital for crop prediction and recommendations.

V. Conclusion

The study evaluated several machine learning methods for crop recommendation, as shown in Table I. Gradient Boosting, by integrating weak learners, achieved the highest accuracy but with increased computational cost and complexity. Polynomial Kernel SVC was the second best for non-linear decision boundaries, though it was more susceptible to overfitting. Decision Tree and RBF Kernel SVC also showed high accuracy; however, Decision Trees were more prone to overfitting despite being easier to interpret. Random Forest demonstrated high accuracy, reliability, adaptability, and performed well across various datasets while avoiding overfitting.

In the meantime, the accuracy of k-NN and Linear Kernel SVC, respectively which qualified them for smaller, less computationally demanding tasks. In conclusion, Random Forest offers a solid mix of performance and efficiency, whereas Gradient Boosting stands out for its exceptional accuracy.

Reference

[1] Acharya, Umesh, Aaron LM Daigh, and Peter G. Oduor. "Machine learning for predicting field soil moisture using soil, crop, and nearby weather station data in the Red River Valley of the North." Soil Systems 5.4 (2021): 57.

[2] Islam, Md Reazul, et al. "Machine learning enabled IoT system for soil nutrients monitoring and crop recommendation." Journal of Agriculture and Food Research 14 (2023): 100880.

[3] Iqbal, Shahid, et al. "Implementing Machine Learning Models–An Analysis of Agricultural Weather and Soil Data." Technical Journal 29.01 (2024): 13-19.

[4] Ali, I.; Greifeneder, F.; Stamenkovic, J.; Neumann, M.; Notarnicola, C. Review of machine learning approaches for biomass and soil moisture retrievals from remote sensing data. Remote Sens. 2015, 7, 16398–16421. [Google Scholar] [CrossRef] [Green Version]

[5] Condran, Sarah, et al. "Machine learning in precision agriculture: a survey on trends, applications and evaluations over two decades." IEEE Access 10 (2022): 73786-73803.

[6] Sharma, Rohit, et al. "A systematic literature review on machine learning applications for sustainable agriculture supply chain performance." Computers & Operations Research 119 (2020): 104926.

[7] Mekonnen, Yemeserach, et al. "Machine learning techniques in wireless sensor network based precision agriculture." Journal of the Electrochemical Society 167.3 (2019): 037522.

[8] Sharma, Abhinav, et al. "Machine learning applications for precision agriculture: A comprehensive review." IEEE Access 9 (2020): 4843-4873.

[9] Chen, Chun, and Heather Mcnairn. "A neural network integrated approach for rice crop monitoring." International Journal of Remote Sensing 27.7 (2006): 1367-1393.

[10] Van Klompenburg Thomas, Ayalew Kassahun, and Cagatay Catal. "Crop yield prediction using machine learning: A systematic literature review." Computers and electronics in agriculture 177 (2020): 105709.

[11] Chlingaryan, Anna, Salah Sukkarieh, and Brett Whelan. "Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review." Computers and electronics in agriculture 151 (2018): 61-69.

[12] Adamchuk, Viacheslav I., et al. "On-the-go soil sensors for precision agriculture." Computers and electronics in agriculture 44.1 (2004): 71-91.

[13] Bhat, Showkat Ahmad and Nen-Fu Huang. "Big data and ai revolution in precision agriculture: Survey and challenges. "leee Access 9 (2021): 110209-110222.

[14] Y.Zhang, M. Wang, D.Zhao, C. liu, and Z. Liu, "Early weed identification based on deep learning: A review, " Smart Agric. Technol., vol. 3, no. October 20222, p.100123, 2023, doi:10.1016/j.atech.2022.100123.

[15] Furtak, Karolina, and Agnieszka Wolińska. "The impact of extreme weather events as a consequence of climate change on the soil moisture and on the quality of the soil environment and agriculture–A review." Catena 231 (2023): 107378.

[16] 1. S. Rathi, S. Wamanacharya, S. Shah and V. Khedkar, "Enhanced Agricultural Decision-Making: Merging Climatic Data with Machine Learning Algorithms," 2024 5th International Conference on Data Intelligence and Cognitive Informatics (ICDICI), Tirunelveli, India, 2024, pp. 859-863, doi: 10.1109/ICDICI62993.2024.10810951.