iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** III **Month of publication:** March 2024

DOI: https://doi.org/10.22214/ijraset.2024.59354

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue III Mar 2024- Available at www.ijraset.com

SmartFarm: A Machine Learning-Based Analytical System for Soil Classification and Crop Recommendation in Precision Agriculture

Prof. Shrikant V. Sonekar[1], Shubham Dnyaneshwar Bhongekar[2]
[1]*Prof, Department of Computer science and Engineering, JD College of Engineering & Management, Nagpur, Maharashtra, India*
[2]*P.G. Student, Department of Computer science and engineering, JD College of Engineering & Management, Nagpur, Maharashtra, India*

Abstract: ***SmartFarm is an innovative machine learning-driven analytical system designed to revolutionize precision agriculture through accurate soil classification and tailored crop recommendations. The project focuses on collecting comprehensive soil and historical crop data, employing rigorous preprocessing techniques, and implementing knowledge-based classification and non-parametric classifiers such as decision trees and neural networks. The system integrates these models, ensuring a cohesive approach to support seamless decision-making for farmers. With a user-friendly interface, SmartFarm enables farmers to input soil data and receive personalized crop recommendations. The project's scalability, adaptability, and commitment to iterative improvement through user feedback make it a promising solution for enhancing agricultural productivity and sustainability. The outcomes contribute to the evolving landscape of precision agriculture, emphasizing the power of machine learning in informed decision support systems. The outcomes underscore the potential of machine learning in fostering informed decision-making and sustainable farming in the evolving landscape of precision agriculture.***

Keyword: ***Machine Learning, Soil Classification, Crop Recommendation, Smart Farming, Agricultural Productivity, Knowledge-based Classification.***

I. INTRODUCTION

Agriculture, as the primary engine of food production, is confronted with the formidable task of meeting the escalating global demand for sustenance. With a burgeoning population, this challenge intensifies, necessitating innovative solutions to enhance agricultural productivity while ensuring the judicious use of resources. Precision agriculture emerges as a beacon of hope, embodying a paradigm shift in farming practices through the integration of advanced technologies. This project introduces SmartFarm, an ambitious and pioneering venture that harnesses the power of machine learning to address two critical facets of precision agriculture: precise soil classification and personalized crop recommendations. The ever-growing demands on agricultural lands necessitate a meticulous understanding of soil health. Soil classification, a cornerstone of agricultural management, is integral to tailoring farming practices based on the unique characteristics of the land. In parallel, the need to optimize crop selection to maximize yields and resource utilization underscores the significance of personalized crop recommendations. SmartFarm aims to bridge these gaps by providing an analytical system that amalgamates cutting-edge machine learning techniques with the intricacies of soil science and agronomy.As the global population surges, agricultural systems must evolve to keep pace with the growing demand for food production. Traditional farming methods, while resilient, often lack the precision required to address the nuanced complexities of varied soil types and environmental conditions. Precision agriculture, as a holistic approach, seeks to overcome these limitations by deploying technology-driven solutions. SmartFarm represents a stride towards this precision, promising to revolutionize how farmers make decisions about soil management and crop cultivation.

II. LITERATURE SURVEY

1) 'Crop recommendation system for precision agriculture' by S. Pudumalar et al. [1] employs the data mining-based technique that uses research data of soil characteristics, soil types, crop yield data and advises farmers on the best crop to plant depending on site-specific factors. This lowers the likelihood of choosing the incorrect crop and raises productivity. In order to recommend a crop for the site-specific parameters with high accuracy and efficiency, this work proposes a recommendation system through an ensemble model with majority voting technique, using Random Tree, CHAID, K-Nearest Neighbour, and Naive Bayes as learners.

©IJRASET: All Rights are Reserved

2) Bhimanpallewar R et al. [2] propose a machine learning technique called "A Machine Learning Approach to Assess Crop Specific Suitability for Small/Marginal Scale Croplands," where the output is the suitability level for the corresponding crop and the input parameters are the existing availability of components in the soil, environmental parameters, and the selected crop. This technique assists in making decisions on how to make the soil more suitable or whether to keep the property undeveloped for a while since it isn't competent.

3) Senagi K et al.'s paper, "Using parallel random forest classifier in predicting land suitability for crop production," [3] makes use of an enhanced machine learning (ML) approach to forecast land suitability for crop production (sorghum) in the presence of soil property data. Parallel Random Forest (PRF), Linear Regression (LR), Linear Discriminant Analysis (LDA), KNN, Gaussian Naïve Bayesian (GNB), and Support Vector Machine (SVM) are the methods used to set up the studies. There aren't many other initiatives that categorise soil by composition. 'Analysis of Agricultural Soils by using Data Mining Techniques' by Ramesh Babu and Rajesh Reddy [4], 'Analysis of Soil Behaviour and Prediction of Crop Yield by Supriya D [5], 'Classification of agricultural soil parameters in India' by Sirsat M et al. [6], 'Crop Recommendation System Using Neural Networks' [7], and 'Using Machine Learning for Land Suitability Classification' [8] are a few of these.

III. PROBLEM DEFINITION

The modern agricultural landscape faces critical challenges in precisely managing soil health and optimizing crop selection. Conventional methods of soil analysis often prove time-consuming and lack the granularity required for efficient decision-making in contemporary agriculture. Additionally, farmers grapple with the absence of personalized crop recommendations tailored to the unique attributes of their soil. The overarching problem tackled by this project is the need for a comprehensive and data-driven approach to soil classification and crop recommendation. SmartFarm, the proposed solution, seeks to leverage machine learning algorithms to address these challenges. It aims to overcome the inefficiencies in traditional soil classification methods, the absence of personalized crop recommendations, and the limited integration of advanced technologies in agriculture. By providing accurate soil classifications and tailored crop suggestions, SmartFarm aspires to revolutionize decision-making in agriculture, enhancing productivity and contributing to the broader goals of sustainability. The project's success hinges on bridging the gap between traditional farming practices and the transformative potential of machine learning, offering a scalable and adaptable solution to meet the diverse needs of agricultural ecosystems globally.

IV. METHODOLOGY

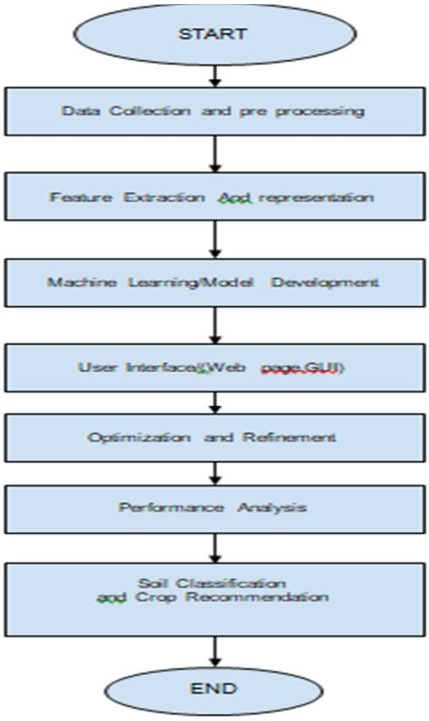


Fig 1: flow chart of model

The technique of the suggested system is made up of numerous blocks, as indicated in fig (1).

1) *Data Collection:* Data collection is the most common approach for gathering and analysing information from various sources. The dataset must have the following qualities to provide an approximate data set for the system. These criteria will be considered for crop recommendation: i) soilPH ii) Humidity iii) NPK levels iv) crop data v) temperature.
2) *Data Pre-Processing:* After collection of data from different sources, the next step is to pre-processed it before the model can be trained. Starting with reading the acquired dataset and going through data purification, data pre-processing can be conducted in numerous ways. When cleansing information, Some dataset characteristics are redundant. are not taken into account while cropping. prediction. As a result, we must remove undesirable properties and datasets that contain some missing data. We must drop or fill these missing values with unwanted nan values with better precision in order to obtain them.Feature Engineering: Using domain expertise, feature engineering extracts features (characteristics, traits, and attributes) from raw data. The idea is to use these extra attributes to improve the quality of ML results.
3) *Training Set:* A training set is a data set that contains data that has been labelled. Both input and output vectors are included. The model is trained using supervised machine learning algorithms using this dataset.
4) *Testing Set:* A testing set is a data set that is devoid of labelled data. It predicts the outcome with the assistance of the training data set. It is unaffected by the training data set.
5) *Machine Learning Algorithm:* Machine learning prediction algorithms [12] [13] necessitate extremely accurate estimation based on previously learned data. Predictive analytics historical information is the use of data, statistical methods, and machine learning approaches to forecast future results. The goal is to go beyond simply understanding what happened to providing the best feasible remedy and a prediction of what will happen next.

KNN, Decision Tree, Nave Bayes, Random Forest, and Gradient Boosting methods are used in this model.

A. *K-Nearest Neighbor Classifier*

KNN is a type of supervised machine learning that can be used to solve a wide range of issues. Classification and regression are examples of challenges that can be addressed. The number of nearest neighbours to a newly forecasted unknown variable is represented by the symbol K.

The distance between the data points is calculated using the Euclidean distance formula.

Euclidean Distance b/w A and B =

$\sqrt{}$ (X2 -X1)2 + (Y2-Y1)2. (1)

B. *Decision Tree*

Decision Trees (DTs) come under supervised learning for classification and regression. A tree representation is used to solve the problem, with each leaf node representing a class label. and the tree's interior node represents attributes.

Entropy:

H(S) = -∑ Pi(S) log2 Pi(S) (2)

Information Gain:

IG(S,A) = H(S) - ∑v€Values(A)(|Sv|/S) H(Sv) (3)

C. *Naive Bayes*

The Bayes theorem is used to create a simple probabilistic classifier called Naive Bayes. Naive Bayes classifiers assume that the value of one feature is independent of the value of any other feature given the class variable.

P(A|B) = (P(B|A) * P(A))/P(B) (4)

D. *Random Forest*

Random Forest is an ensemble learning method that creates a huge number of distinct models to solve classification, regression, and other issues. At training time, decision trees are used. Algorithm of random forest makes decision trees based on various data samples and then

forecasts data from each subset, then votes on it gives the system a better solution RF uses the bagging method for data training, which improves the correctness of the outcome.

Gini Index =1-∑ (Pi)2 i=1 = 1-[(P+)2+ (P-)2] (5)

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

E. Gradient Boosting

Gradient Boosting also comes under supervised machine learning technique for solving classification and regression issues. It's a poor prediction model with an ensemble. As with previous boosting approaches, a gradient-boosted trees model is built stage by stage.

1) *Crop Recommendation:* Based on the N P K, temperature,humidity, and ph, the model will recommend the optimum crop to grow on the given soil.
2) *Performance Analysis:* Performance analysis is a specialised subject that uses systemic objectives to improve performance and decision-making.

ALGORITHM	ACCURACY
Logistic Regression	0.9636363636363636
Naive Bayes with accuracy	0.9954545454545455
Support Vector Machine	0.9681818181818181
K-Nearest Neighbors	0.9590909090909091
Decision Tree	0.9863636363636363
Random Forest	0.990909090909091
Bagging	0.9886363636363636
AdaBoost	0.9818181818181818
Gradient Boosting	: 0.1409090909090909

Table 1: Algorithm tested with accuracy

V. PROPOSED SYSTEM

The proposed system, SmartFarm, is a revolutionary analytical platform that integrates machine learning to address critical challenges in precision agriculture, specifically focusing on accurate soil classification and personalized crop recommendations. Through a systematic methodology, SmartFarm collects comprehensive soil and historical crop data, preprocesses information using advanced techniques, and employs machine learning models, including knowledge-based classification and decision tree classifiers. The integration of these models creates a cohesive system capable of providing farmers with precise soil classifications and tailored crop recommendations, thus facilitating informed decision-making. The user-friendly interface ensures accessibility, while the system's scalability and adaptability promise relevance across diverse agricultural contexts. With rigorous evaluation metrics and a continuous feedback loop for iterative improvement, SmartFarm stands poised to revolutionize decision support in agriculture, bridging the gap between traditional farming practices and the transformative potential of machine learning.

The proposed system plans to implement this in 2 phases:

- Phase 1: mechanism that classifies the soil according to the fertility grades, nutrients etc.
- Phase 2: find the relation between crop nutrient requirements and soil classification groups found in earlier phase; this can be done by clustering of crops with similar soil nutrient, and fertility requirements into classified soil labels.

The proposed system plans to analyse soil through following data:

- The biochemical composition of the soil
- Soil images
- Satellite land imagery and remote sensing data (wherever available and possible)

For mapping the crops to the soil classification groups, we can analyse attributes like:

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

The macro and micro nutrients requirements of the crop

- The pH level of the soil
- Water holding capacity & electrical conductivity of the soil. The proposed system plans to use combination of one or more of the following approaches:
- Classification based on decision trees, deep learning using NN (Neural Networks), SVM(Support Vector Machines) etc.
- Statistical tools like – Bayes distribution, regressions

VI. OBJECTIVES

The SmartFarm project is driven by a comprehensive set of objectives aimed at harnessing the power of machine learning to revolutionize precision agriculture. The primary goals of this ambitious initiative are designed to address critical challenges in soil classification and crop recommendation, while promoting sustainable and data-driven farming practices. At the core of SmartFarm's objectives is the development of accurate soil classification models. Traditional soil analysis methods often prove time-consuming and lack the specificity required for contemporary agriculture. SmartFarm aims to bridge this gap by implementing state-of-the-art machine learning algorithms, including knowledge-based classification, decision tree classifiers, and neural networks. The objective is to create models that can precisely categorize soil types based on a multitude of parameters such as pH levels, nutrient content, and moisture. By doing so, the project seeks to provide farmers with a nuanced understanding of their soil health, laying the foundation for informed decision-making. Complementing the soil classification objective is the aspiration to enable personalized crop recommendations. Agriculture is not a one-size-fits-all endeavor, and optimal crop choices vary based on the unique characteristics of the soil. SmartFarm endeavors to develop a robust machine learning model that synthesizes insights from historical crop data and the intricacies of soil classifications. The aim is to empower farmers with personalized and data-driven recommendations for crop cultivation, ensuring that agricultural practices align with the specific conditions of their land. This objective reflects a commitment to maximizing yields and resource utilization while minimizing environmental impact.

Scalability and adaptability represent key dimensions of SmartFarm's objectives. Agriculture spans a spectrum of geographical regions, soil types, and farming practices. SmartFarm is designed with scalability in mind, ensuring that it can be deployed across diverse agricultural landscapes. The system's adaptability is a crucial consideration, allowing it to accommodate different environmental conditions and the unique needs of farmers globally. This objective reflects a commitment to developing a solution that is not confined to specific regions or farming methods but can flexibly cater to the dynamic nature of agriculture on a global scale.

Rigorous evaluation metrics constitute another cornerstone of SmartFarm's objectives. The project acknowledges the importance of validating the performance of machine learning models to in still confidence in their accuracy. Metrics such as precision, recall,are employed to assess the models' effectiveness in soil classification and crop recommendation. Additionally, validation against ground truth data obtained from field measurements and historical records ensures that SmartFarm's recommendations align with observed agricultural outcomes.

This commitment to robust evaluation mechanisms underscores the project's dedication to delivering reliable and impactful solutions. Establishing a continuous feedback loop is an integral objective for SmartFarm. The agricultural sector is dynamic, and user feedback plays a pivotal role in ensuring the system's relevance and effectiveness. By engaging with farmers and stakeholders, SmartFarm aims to collect insights that inform iterative refinements of machine learning models and the user interface. This ongoing dialogue with end-users is not just a post-implementation formality but a strategic approach to enhance the system based on real-world experiences and evolving agricultural practices. Documentation of the methodology, results, and user guidelines is a final but critical objective for SmartFarm. Transparent and comprehensive documentation serves multiple purposes. It facilitates knowledge transfer, allowing stakeholders to understand the intricacies of the machine learning models and system architecture. User documentation ensures that farmers can effectively utilize the SmartFarm system, maximizing its potential impact on their farming practices.

SmartFarm's objectives collectively chart a course for a transformative approach to precision agriculture. By addressing challenges in soil classification and crop recommendation, fostering usability and accessibility, emphasizing scalability and adaptability, employing rigorous evaluation metrics, and engaging in a continuous feedback loop, the project aspires to contribute significantly to the evolution of sustainable and data-driven farming practices globally. The integration of machine learning into precision agriculture through SmartFarm represents a forward-looking endeavor that seeks to bridge the gap between traditional farming practices and the transformative potential of cutting-edge technologies.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

VII. RESULTS

The accuracy values of each algorithm serve as a critical metric in evaluating their effectiveness in the SmartFarm model. The preference for Random Forest suggests its suitability for handling the complexity and variability inherent in soil conditions, showcasing its potential as a robust tool for precise crop recommendations. The model's reliance on machine learning algorithms underscores its commitment to providing farmers with accurate and data-driven insights, facilitating informed decision-making in agriculture.

The GUI plays a pivotal role in SmartFarm's data acquisition process, fetching essential parameters such as moisture, temperature, humidity, and pH from the IoT database. This data, sourced directly from the field area, is then stored in the IoT database for further analysis. The GUI's database encapsulates the user-entered features and serves as a crucial input for the crop prediction model.

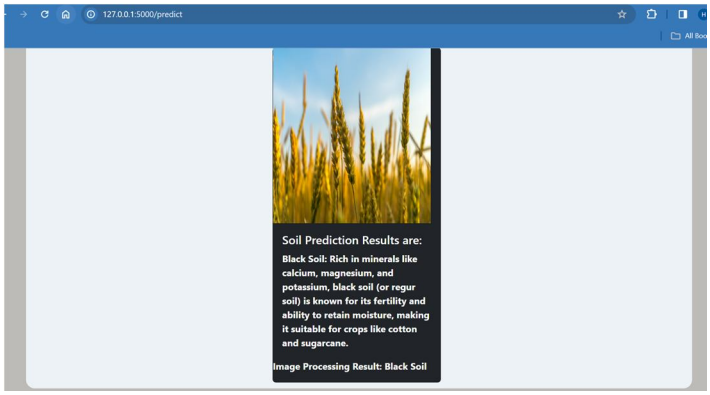


Fig2: GUI of the model

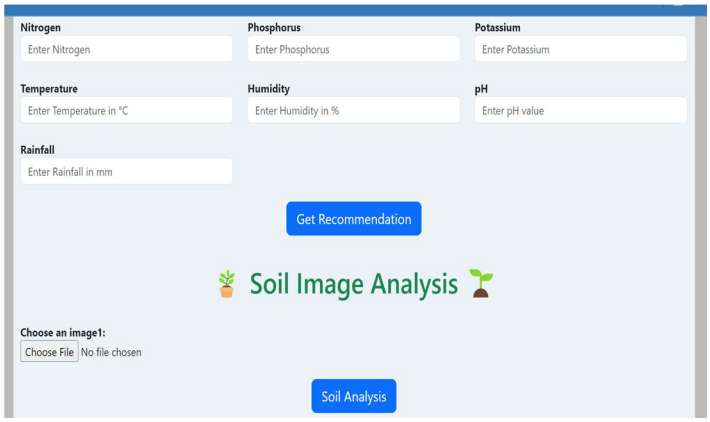


Fig3: Recommendation given the model

The heart of the SmartFarm system lies in its crop prediction model. Fueled by the data collected through the GUI, this model undergoes training to comprehend the intricate relationships between environmental variables and optimal crop choices. The result of this model is then seamlessly integrated into the crop suggestion interface within the GUI.

The crop suggestion interface is a user-friendly component that translates the predictions from the trained model into practical recommendations for farmers. By correlating the predicted outcomes with the features entered by the user, the interface suggests the most suitable crops for cultivation.

This integration of advanced technologies facilitates a precision decision-making process for farmers, enabling them to make informed choices based on data-driven insights.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

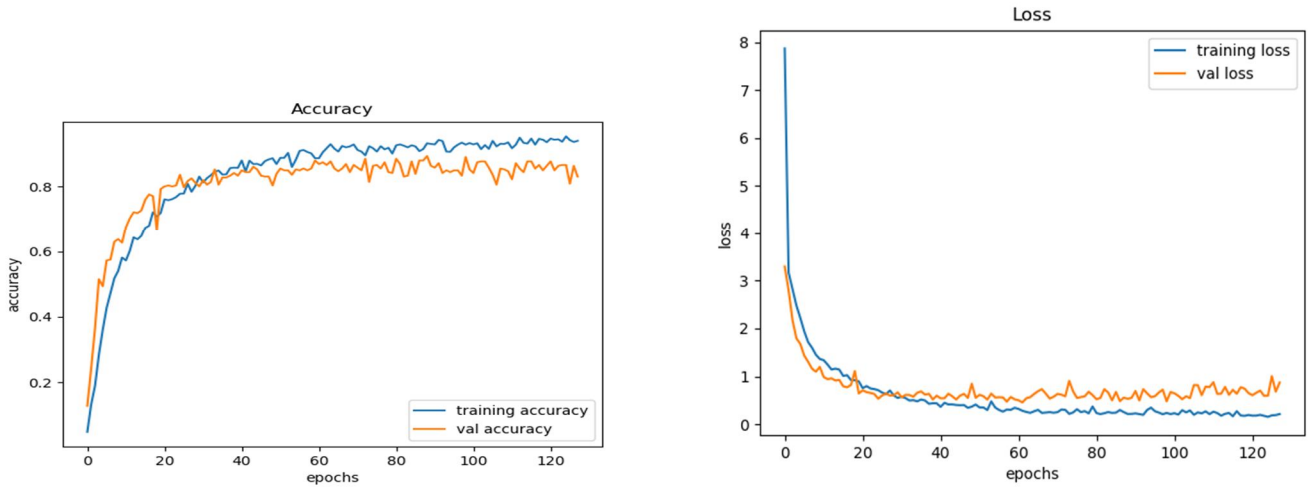


Fig 4:model accuracy vs loss of the Model

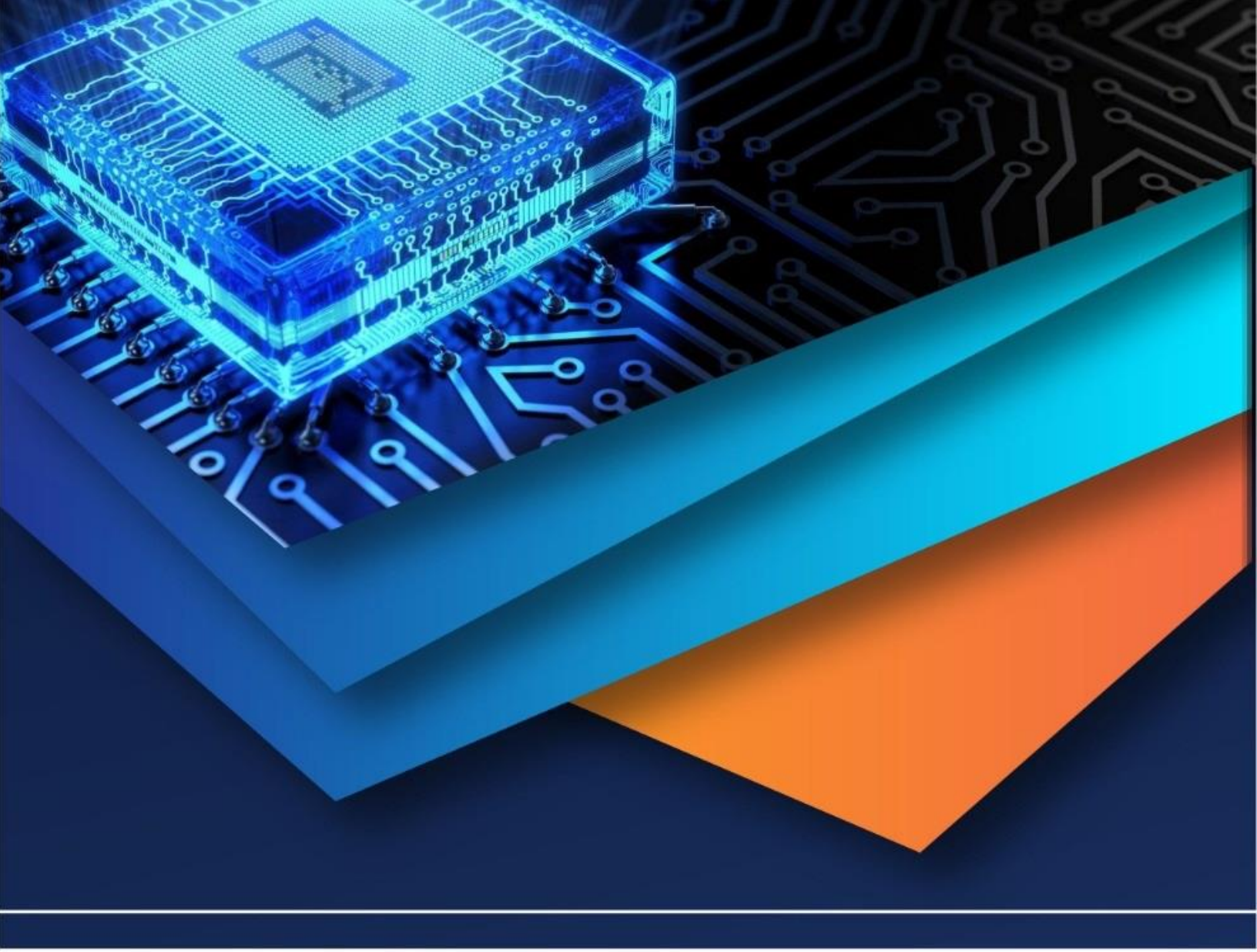
VIII. CONCLUSION

In conclusion, the SmartFarm project represents a significant advancement in precision agriculture, leveraging machine learning to address key challenges in soil classification and crop recommendation. The systematic methodology, encompassing data collection, preprocessing, and the integration of machine learning models, has demonstrated the potential to revolutionize decision-making in agriculture. By providing accurate soil classifications and personalized crop recommendations, SmartFarm empowers farmers with actionable insights, promoting more sustainable and efficient farming practices. The success of SmartFarm is underscored by its user-friendly interface, ensuring accessibility for farmers, and its scalability, allowing adaptation to diverse agricultural landscapes. The project's commitment to rigorous evaluation metrics and a feedback loop for iterative improvement enhances its reliability and effectiveness in real-world farming scenarios. As agriculture stands at the intersection of tradition and technological innovation, SmartFarm serves as a beacon for the future of precision agriculture. The transformative potential of machine learning, as demonstrated in this project, holds promise for enhancing global food production, mitigating environmental impact, and fostering a more sustainable and resilient agricultural sector. While SmartFarm has made significant strides, the journey towards precision agriculture is ongoing. Future research and developments could explore the integration of additional data sources, advanced machine learning techniques, and collaborative efforts to address regional nuances and evolving agricultural practices. As we move forward, SmartFarm stands as a testament to the synergies between technology and agriculture, paving the way for a future where data-driven decisions underpin a more resilient and sustainable global food system.

REFERENCES

[1] Athani, S., Tejeshwar, C. H., Patil, M. M., Patil, P., & Kulkarni, R. (2017, February 1). Soil moisture monitoring using IoT enabled arduino sensors with neural networks for improving soil management for farmers and predict seasonal rainfall for planning future harvest in North Karnataka — India
[2] Biradar, H. B., & Shabadi, L. (2017). Review on IOT based multidisciplinary models for smart farming.
[3] Mehta, P., Shah, H., Kori, V., Vikani, V., Shukla, S., & Shenoy, M. (2015). Survey of unsupervised machine learning algorithms on precision agricultural data. IEEE Xplore.
[4] Sandhiya, D., & Thiyaneswaran, B. (2017). Extraction of dorsal palm basilic and cephalic hand vein features for human authentication system. 2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET). IEEE.
[5] Radhika, Y., & Shashi, M. (2009). Atmospheric Temperature Prediction using Support Vector Machines. International Journal of Computer Theory and Engineering, 55–58.
[6] Ahmad, S., Kalra, A., & Stephen, H. (2010). Estimating soil moisture using remote sensing data: A machine learning approach. Advances in Water Resources, 33(1), 69–80.
[7] Kumar, R., Singh, M. P., Kumar, P., & Singh, J. P. (2015). Crop Selection Method to maximize crop yield rate using machine learning technique. IEEE Xplore.
[8] Thiyaneswaran, B., Anguraj, K., Kumarganesh, S., & Thangaraj, K. (2020). Early detection of melanoma images using gray level co-occurrence matrix features and machine learning techniques for effective clinical diagnosis. International Journal of Imaging Systems and Technology, (ima.22514). doi:10.1002/ima.22514
[9] Shah, N. P., & Bhatt, P. (2017). Greenhouse Automation And Monitoring System Design And Implementation. International Journal of Advanced Research in Computer Science, 8(9), 468–471.
[10] Jayanthi, J., and J. Selvakumar. "A novel framework to facilitate personalized web search in a dual mode." Cluster Computing 20.4 (2017): 3527-3535.

[11] Parker, C. (2012). Unexpected challenges in large scale machine learning. Proceedings of the 1st International Workshop on Big Data, Streams and Heterogeneous Source Mining Algorithms, Systems, Programming Models and Applications - BigMine ’
[12] B Thiyaneswaran, K Anguraj, M Sindhu, N. S Yoganathan, J Jayanthi. (2020). Development of Iris Biological Features Extraction for Biometric Based Authentication to Prevent Covid Spread. International Journal of Advanced Science and Technology, 29(3), 8266–8275.
[13] Pudumalar, S., Ramanujam, E., Rajashree, R. H., Kavya, C., Kiruthika, T., & Nisha, J. (2017). Crop recommendation system for precision agriculture. IEEE Xplore.
[14] Anitha, P., & Chakravarthy, T. (2018). Agricultural Crop Yield Prediction using Artificial Neural Network with Feed Forward Algorithm. International Journal of Computer Sciences and Engineering, 6(11), 178–181. https://doi.org/10.26438/ijcse/v6i11.178181
[15] B Thiyaneswaran, V Bhuvaneshwaran, M.S Dharun, K Gopu, T Gowsikan (Ed.). (2020). Breathing level monitoring and alerting by using embedded IOT (Vol. 10). Journal of Green Engineering.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)