A Survey on Soil Fertility and Crop Recommendation Using Machine Learning Technique

Manish Kumar[1], Anurag Shrivastava[2]

[1]MTech Scholar, CSE, Department, NIIST, Bhopal

[2]Associate Professor, Department of CSE NIIST, Bhopal

Abstract—Soil fertility plays a critical role in agricultural productivity and food security. With increasing global population and diminishing arable land, optimizing crop yield through intelligent soil management has become imperative. This survey explores the integration of machine learning (ML) techniques in assessing soil fertility and recommending suitable crops, aiming to enhance sustainable farming practices. The study reviews a broad spectrum of recent works employing supervised, unsupervised, and ensemble learning methods, including decision trees, random forests, support vector machines, and deep learning models. These models utilize soil parameters such as pH, nitrogen, phosphorus, potassium levels, and moisture content to classify soil types and suggest optimal crops for cultivation. The survey highlights the strengths and limitations of various ML algorithms in terms of accuracy, interpretability, and scalability. Additionally, it discusses the challenges of data availability, feature selection, and real-time implementation in field conditions. By analyzing existing approaches, this survey offers insights into future directions and the potential of ML in revolutionizing precision agriculture through automated crop recommendation systems.

Index Terms—Deep Learning, Fertilizer Recommendation, Crop Recommendation, soil data; soil analysis

I. INTRODUCTION

Agriculture forms the backbone of many economies, particularly in developing countries, where it sustains livelihoods and ensures food security. A fundamental component of agricultural productivity is soil fertility, which refers to the soil's ability to supply essential nutrients to crops in adequate amounts. Traditionally, soil fertility has been assessed through manual testing and expert consultation. While effective, these conventional methods are often labor-intensive, time-consuming, and regionally inconsistent. In the context of increasing population pressure, climate change, and dwindling arable land, it has become imperative to adopt data-driven approaches for sustainable farming practices. One promising direction lies in the application of machine learning (ML) techniques for soil fertility analysis and crop recommendation.

Machine learning offers intelligent algorithms capable of identifying patterns in large, complex datasets, making it well-suited for analyzing multidimensional agricultural data. By leveraging historical soil test results, weather conditions, crop yield records, and sensor data, ML models can predict soil nutrient levels and recommend optimal crops with a high degree of accuracy. Algorithms such as Decision Trees, Support Vector Machines (SVM), Random Forests, and Artificial Neural Networks (ANNs) have been widely explored for their potential in predicting soil fertility indices and classifying suitable crops based on varying agro-ecological zones. The integration of machine learning in agriculture goes beyond mere prediction. It allows real-time monitoring, site-specific fertilizer application, and precision farming that minimizes input waste while maximizing yield. Additionally, ML can help detect imbalances in soil nutrients (like nitrogen, phosphorus, potassium—NPK), analyze soil pH, and assess organic content, all of which influence crop performance. This is particularly valuable for smallholder farmers who often lack access to tailored agricultural advice.

Despite its potential, the practical implementation of machine learning in soil and crop management faces several challenges. These include the availability of quality datasets, computational constraints in rural settings, and the need for domain-specific model training. Moreover, interpretability and scalability of

models remain critical concerns in ensuring widespread adoption among farmers and agronomists. This paper presents a comprehensive survey of machine learning applications in soil fertility assessment and crop recommendation. It reviews recent literature, compares the effectiveness of different ML algorithms, and evaluates real-world case studies where such systems have been deployed. The goal is to understand current trends, identify research gaps, and explore how ML can be leveraged for data-driven, sustainable, and smart agriculture. As the agricultural sector continues to evolve, embracing intelligent technologies like machine learning may be key to enhancing productivity and resilience in farming systems worldwide.

II. LITRETURE REVIEW

Recent studies have demonstrated the efficacy of machine learning in soil fertility prediction and crop recommendation. Algorithms such as Random Forest, Decision Trees, and Support Vector Machines have been widely applied to analyze soil parameters like pH, NPK levels, and moisture content. Researchers highlight improved accuracy and efficiency compared to traditional methods. Several works emphasize real-time data integration from sensors, enabling site-specific recommendations. However, challenges like data quality, regional generalization, and model interpretability remain areas of ongoing research and development.

Authors [1] developed a model to recommend crop and fertilizer using two machine learning algorithms. The RF algorithm, which has shown high level of accuracy in many different agricultural applications, is used for recommending crops, while the hierarchical Clustering algorithm is used for fertilizer recommendation. The models used Crop nutrient requirement and soil sample data for training and testing. The RF and hierarchical algorithm were trained to recommend crop and fertilizer on the basis of multiple biophysical variables and soil nutrients. The system was found effective in recommending crop and fertilizer with an accuracy of 99.70%. The results showed that the model performed effectively and it is versatile machine-learning model for recommending crop and fertilizer due to the high accuracy and precision values. This research pointed out various steps in which a crop and fertilizer

recommendation system was achieved using a random forest and hierarchical Clustering algorithms. Authors [2] applied an ensemble feature selection approach to identify critical predictors. To address the class imbalance, Generative Adversarial Networks (GANs) were used to generate synthetic data, ensuring the model's robustness in identifying underrepresented cases. Additionally, a hybrid loss function combining cross-entropy and focal loss was implemented to improve classification, especially for hard-to-detect instances. Our results show that the attention-based DBN model, augmented with synthetic data from GANs and optimized with a hybrid loss function, achieves an AUC of 1.00, F1-score of 0.97, precision of 0.98, and recall of 0.95, outperforming several baseline models. This research offers a novel and effective approach for early diabetes detection, demonstrating potential for use as a clinical tool in preventive healthcare settings.

Author's [2] main aim of this research is to determine whether the soil is fertile based on soil properties like N, P, K, Ph, nutrient level, moisture levels, temp rainfall, and topography. Material/Method: We used the dataset from Kaggle, where N, P, K, and pH values are input into the model, and the ML determines whether it is fertile or not. In this paper, four machine learning classifiers are trained, and determine the best classifier based on the performance metrics. Result: The results demonstrated that the machine learning classifier significantly improves prediction accuracy. Authors used LR, KNN, NB, and DT classifiers to increase the accuracy, as well as to increase the efficiency of the soil fertility assessment. The DT classifier exhibited well in comparison to other classifiers. The DT classifier's accuracy was 89%, but the performance metrics precision, LR, and KNN, was 90%.

Author's [3] reviews the current state of digital technologies in agriculture and discusses future research directions to advance data-driven decision making on farms. Digital technologies are revolutionizing agriculture by enabling data-driven decision making. A combination of sensors, satellite imagery, and AI analytics is providing farmers with unprecedented insights to optimize crop management. Sensors monitor soil moisture, temperature, and nutrient levels in real-time. High-

resolution satellite images track crop health, growth stages, and yield potential. Machine learning algorithms process this data to generate actionable recommendations on irrigation, fertilization, pest control, and harvest timing. Case studies demonstrate how these technologies have increased yields, reduced inputs, and improved sustainability on farms worldwide. However, challenges remain in technology adoption due to high costs, lack of digital literacy, and data privacy concerns. Overcoming these barriers will be crucial to harnessing the full potential of digital farming.

Authors [4] apply the science of machine learning in the field of agriculture, by carrying soil fertility analysis using most accurate algorithm. The fertility of soil plays a principal role in determining the suitability of cultivating a particular crop on a given soil type. Analysis is carried out by the examination of various properties of the soil like the pH value, Electrical Conductivity, Moisture content, Temperature and (N)Nitrogen (P)Phosphorous (K) Potassium levels, followed up by soil type classification. Finally, a recommendation for the most suitable crop is provided in real time.

III. FINDINGS OF THE SURVEY

The survey conducted on Soil Fertility and Crop Recommendation Using Machine Learning Techniques reveals a dynamic and evolving field, where digital technologies and data-driven models are significantly transforming traditional agricultural practices. The findings of the survey can be categorized into six key areas:

1. Dominance of Supervised Learning Models: The most widely used machine learning techniques for soil fertility and crop recommendation are supervised algorithms, particularly Decision Trees, Random Forests, Support Vector Machines (SVM), Naïve Bayes, and Artificial Neural Networks (ANNs). These models have been successful in classifying soil fertility levels and predicting suitable crops based on multiple soil parameters such as pH, nitrogen (N), phosphorus (P), potassium (K), organic matter, and moisture content. Among these, Random Forests frequently outperformed others in terms of accuracy and robustness, especially when applied to noisy or incomplete datasets.

2. Real-Time Monitoring through IoT and Sensors: Many systems are now integrating IoT-based soil sensors to provide real-time data on soil health indicators. These sensors feed continuous streams of data into ML models, allowing dynamic recommendations for crop rotation, irrigation scheduling, and fertilizer dosing. The synergy of sensors with ML models has proven effective in enabling precision agriculture and reducing human error in soil assessment.

3. Satellite Imagery and Remote Sensing Applications:
High-resolution satellite data is increasingly being used in conjunction with ML algorithms to assess crop growth stages, vegetation indices, and yield forecasts. Remote sensing technologies enhance the spatial scalability of ML systems, allowing crop recommendation not just at field-level, but across regions and agro-climatic zones.

4. Increased Accuracy and Efficiency: Machine learning models, particularly ensemble and deep learning methods, have consistently demonstrated higher accuracy in predicting soil fertility status and crop suitability compared to traditional statistical or rule-based systems. Moreover, these models reduce the time and cost associated with manual soil testing and expert consultations.

5. Challenges in Implementation: Despite technological advancements, significant barriers to adoption persist. These include the lack of standardized datasets, limited digital infrastructure in rural areas, high deployment costs, and the digital literacy gap among farmers. Moreover, model interpretability remains a concern, as many ML techniques are viewed as "black-box" systems by stakeholders.

6. Sustainability and Input Optimization: Several case studies confirm that ML-driven recommendations lead to lower input costs, optimized fertilizer usage, and enhanced crop yields. This contributes not only to economic gains for farmers but also to environmental sustainability by minimizing nutrient leaching and soil degradation.

Finally the survey highlights that machine learning is revolutionizing soil fertility assessment and crop recommendation by offering scalable, accurate, and real-time solutions. However, to unlock its full potential, challenges related to infrastructure,

usability, and trust must be addressed through interdisciplinary research and policy support.

IV. CONCLUSION

The integration of machine learning techniques into soil fertility assessment and crop recommendation systems marks a significant advancement in precision agriculture. This survey has demonstrated that machine learning models, especially supervised learning algorithms such as Decision Trees, Random Forests, and Neural Networks, offer considerable improvements in accuracy, efficiency, and scalability over traditional methods. By leveraging real-time data from sensors and satellite imagery, these models enable dynamic, location-specific recommendations that enhance crop productivity and optimize resource usage.

Machine learning applications in agriculture not only improve yield but also contribute to sustainable farming by minimizing excess fertilizer application and reducing environmental degradation. Moreover, the combination of soil nutrient analysis, climatic data, and remote sensing has enabled holistic crop planning that aligns with ecological and economic goals. However, the widespread adoption of these technologies is still constrained by challenges such as high implementation costs, lack of digital literacy, data privacy concerns, and limited access to reliable datasets. Addressing these issues through infrastructure development, education, and policy reforms is essential for fully realizing the benefits of digital agriculture. In conclusion, machine learning holds transformative potential for modern farming. With continued research and practical implementation, it can play a pivotal role in securing global food security and promoting sustainable land use.

REFERENCES

[1] S. S. Nandom1, *, G. T. Abe2, I. P. Gambo3 "Application of random forest and hierarchical clustering models for crop and fertilizer recommendation to farmers" Nigerian Journal of Technology (NIJOTECH) Vol. 44, No. 1, March, 2025

[2] Ashutosh Sarangi et al "Enhancing Process Control in Agriculture: Leveraging Machine Learning for Soil Fertility Assessment" https://doi.org/10.3390/engproc2024067031, MDPI, 2025

[3] Arijit Ghosh et al "Data Driven Decision Making in Agriculture with Sensors, Satellite Imagery and AI Analytics by Digital Farming" https://doi.org/10.9734/acri/2025/v25i51186 , 2025

[4] Muneshwara M .S, "Soil Fertility Analysis and Crop Prediction using Machine Learning" IJITEE) ISSN: 2278-3075 (Online), Volume-9 Issue-6, April 2020

[5] Jagruti Raut et al "Soil Fertility and Crop Recommendation using Machine Learning and Deep Learning Techniques: A Review" Turkish Journal of Computer and Mathematics Education 2020

[6] Balakrishnan, N., & Muthukumarasamy, G. (2016). Crop Production - Ensemble Machine Learning Model for Prediction. International Journal of Computer Science and Software Engineering, 5(7), 148–153.

[7] Benedet, L., "Rapid soil fertility prediction using X-ray fluorescence data and machine learning algorithms. Catena, 197(April 2020). https://doi.org/10.1016/j.catena.2020.105003

[8] Benke, K. K., Norng, S., Robinson, N. J., Chia, K., Rees, D. B., & Hopley, J. (2020). Development of pedotransfer functions by machine learning for prediction of soil electrical conductivity and organic carbon content. Geoderma, 366(January). https://doi.org/10.1016/j.geoderma.2020.114210

[9] Ghosh, S., & Koley, S. (2014). Machine Learning for Soil Fertility and Plant Nutrient Management using Back Propagation Neural Networks. International Journal on Recent and Innovation Trends in Computing and Communication, 2(2), 292.

[10] Hassan Hayatu, I., Mohammed, A., Ahmad Isma'eel, B., & Yusuf Ali, S. (2020). K-Means Clustering Algorithm Based Classification of Soil Fertility in North West Nigeria. Fudma Journal of Sciences, 4(2), 780–787. https://doi.org/10.33003/fjs-2020-0402-363

[11] Rachna kumara et al "a new hybrid deep learning model for diabetic retinopathy detection" Journal of Theoretical andApplied Information Technology 30th September 2024. Vol.102. No. 1

[12] od_Security_in_Africa. [2] Nair, P. K. R., Kumar, B. M., Nair, V. D. "Soil Organic Matter (SOM) and Nutrient Cycling. In: An Introduction to Agroforestry", Springer, Cham. 2021.

[13] Yemefack, M. "Preserving Soil Fertility for Ensuring Food Security in Africa", in

proceedings of the American Association for the Advancement of Science, Chicago, 2014.

[14] Yange, T. S., Egbunu, C. O., Rufai, M. A., Onyekwere, O., Abdulrahman, A. A., Abdulkadri, I. "Using Prescriptive Analytics for the Determination of Optimal Crop Yield", In: International Journal of Data Science and Analysis, 6(3), 72; 2020.

[15] Gamage, A., Et. Al "Role of organic farming for achieving sustainability in agriculture", Farming System, vol. 1, no. 1, p. 100005, 2023.

[16] Khatri, P., Et al "Understanding the intertwined nature of rising multiple risks in modern agriculture and food system," Environment, Development and Sustainability, vol. 26, pp. 24107–24150, 2024. doi: 10.1007/s10668-023-03638-7.