Crop Recommendation Based on Soil Properties: A Comprehensive Analysis

1st Ajay Agarwal
Information Technology
KIET Group of institutions
Ghaziabad, India
ajay.agarwal@kiet.edu

2nd Sartaj Ahmad
Information Technology
KIET Group of Institutions
Ghaziabad, India
sartaj.ahmad@kiet.edu

3rd Adesh Pandey
Information Technology
KIET Group of Institutions
Ghaziabad, India
ak.pandey@kiet.edu

***Abstract*—Crop recommendation systems play a crucial role in optimizing agricultural practices by suggesting suitable crops based on soil properties. This paper aims to provide a comprehensive analysis of crop recommendation methodologies using soil property data. We explore various approaches, including machine learning algorithms and expert systems, to develop accurate and efficient crop recommendation systems. The integration of soil properties and their impact on crop growth and yield is examined to ensure the compatibility between recommended crops and soil conditions. We also discuss the challenges and future directions in this field to enhance the effectiveness of crop recommendation systems.**

***Index Terms*—Agricultural data analysis, Machine learning, Crop yield prediction, Soil quality assessment**

I. Introduction

Agriculture forms the backbone of global food production, contributing to the livelihoods of billions of people. Efficient utilization of available resources, including soil, is essential for sustainable and productive agricultural systems. Properties of soil dictate its fertility and suitability for different crops. Understanding these soil properties and their implications is crucial for optimizing crop production. This paper proposes a comprehensive approach for crop recommendation based on soil properties. By integrating soil testing, data analysis techniques, and crop knowledge, the approach aims to provide farmers with tailored recommendations for optimizing crop selection, nutrient management, and sustainable agricultural practices. The proposed framework will empower farmers to make informed decisions, leading to increased yields, reduced environmental impact, and improved economic viability.

The rest of the paper is as follows: Section II provides the existing publised work. Section III illustrates the significance of soil properties, their influence on crop growth, crop-soil compatibility, and its roles in crop prediction. Section IV provides an overview of the crop-specific soil preferences. Details of **dataset** used for soil testing and analysis is given in Section V. Section VI provides performance metrics. Conclusion is provided in Section VII. Challenges and Limitations are discussed in VIII and finally, Section IX presents recommendation for further research.

II. Literature Review

The literature surrounding crop recommendation based on soil properties encompasses a wide range of studies and methodologies. Several researchers have emphasized the significance of understanding soil characteristics and their influence on crop growth to make informed decisions regarding crop selection. This section presents a summary of key studies and approaches in the field. In a recent study by Manendra Sai, D., et al. (2023) [1] focused on crop yield forecasting which is a crucial aspect of precision agriculture. Crop yield forecasts are required for thorough planning, policy development, and execution for choices regarding, among other things, the procurement, distribution, price fixing, and import-export of crops. Another recent study by Chen et al. (2022) [2], a deep learning model was proposed for crop recommendation based on soil properties. Their models analyzed soil data and provided accurate crop recommendations. The research highlighted the effectiveness of deep learning techniques in capturing complex relationships between soil properties and crop performance. Wang et al (2021) [3] highlights the performance in a plastic shed soil system. It provides valuable insights for agricultural practices aiming to sustainably manage soil fertility and improve crop productivity. Jones et al. (2020) 4 employed machine learning algorithms to analyze large datasets of soil properties and crop yields. Their study demonstrated the potential of these approaches to predict crop performance based on soil characteristics and provided insights into the most influential soil factors. Expert knowledge-based systems have also been utilized in crop recommendation. Zhang et al. (2019) [5] developed an expert system that incorporated agronomic knowledge and soil characteristics to recommend crops for different regions. The system successfully provided accurate crop recommendations and proved to be a valuable tool for farmers and agronomists. Geospatial analysis has emerged as a powerful tool for crop recommendation at a regional scale. Li et al. (2018) [6] utilized remote sensing and geographic information systems (GIS) to analyse soil properties and vegetation indices to recommend suitable crops for different regions. Their approach integrated spatial data to enhance the accuracy of crop selection. In terms of crop-soil compatibility, Kumar and Puri (2018) [7] explored

2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT) | 979-8-3503-3509-5/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICCCNT56998.2023.10307999

the preferences of different crops for specific soil types. Their research focused on identifying the soil parameters most influential for individual crops and proposed a classification system based on these preferences. Their work provides a Smith et al. (2017) [8] conducted a comprehensive analysis of soil properties and their effects on crop performance in various regions. They highlighted the importance of soil composition, texture, and nutrient availability in determining crop suitability. Their findings demonstrated the need for tailored crop recommendations based on specific soil characteristics to optimize productivity. Statistical models have also been used for crop recommendation. Gupta and Sharan (2016) [9] developed a regression-based model to predict crop yields based on soil parameters. Potato [10] output projection is significantly improved over methods that do not make use of cultivar data when ML algorithms are used to combine high-spatial-goal UAV images with cultivar data. By incorporating more precise crop information, details of the soil and landscape, administrative data, and robust machine learning calculations, additional research to improve potato output prediction is likely to be carried out Bibliotic variables, or environmental factors, are produced when living organisms such as bacteria, plants, animals, pathogens, predators, and bugs interact. This group [11] also includes human-caused factors including soil toxicity, drainage, plant defence, fertiliser, and so on.. Plant yield variety in engineered associations, internal complaints, and design concerns could result from these yield creation issues. The development and well-being of plants as well as the output of the environment are both impacted by biotic and abiotic elements [12]. Substance, physical, and other biotic variables are acknowledged.

III. Significance of Soil Properties

Soil properties directly affect plant growth, nutrient availability, water holding capacity, and disease prevalence. Variations in nutrient levels, texture, pH and organic matter content can significantly impact crop performance. Farmers and agronomists need accurate information about soil properties to make informed decisions regarding suitable crop selection, fertilization, irrigation, and other management practices.

A. Soil Properties and Their Influence on Crop Growth

1) **Nutrient Content:** Soil nutrient content plays a vital role in crop growth and development. For many physiological processes in plants, essential nutrients and micronutrients are required. Plant absorption and utilisation are directly impacted by availability of these nutrients in soil. Deficiencies or excesses of specific nutrients can have a significant impact on crop health and productivity. For example, nitrogen deficiency can result in stunted growth and reduced yields, while phosphorus deficiency can lead to poor root development and limited energy transfer. Conversely, excessive nutrient levels can cause toxicity symptoms and imbalances in plant nutrient ratios. By understanding the nutrient content of the soil through soil testing, farmers can identify nutrient deficiencies or excesses and apply targeted fertilization strategies. This helps ensure that crops receive the necessary nutrients for optimal growth and development, leading to improved yields and quality.
2) **pH Level:**
The quantity of hydrogen ions in the soil determines the pH, which is a measurement of the soil's acidity or alkalinity. pH influences nutrient availability, microbial activity, and various physiological processes in plants. Different crops have different pH preferences, and soil pH can affect nutrient solubility and uptake by plants. For example, some crops, like blueberries, prefer acidic soils with a low pH, while others, such as asparagus, thrive in more alkaline soils with a higher pH. When the soil pH deviates from the optimal range for a particular crop, nutrient deficiencies may occur, even if the nutrients are present in the soil. This can limit crop growth and reduce yields. By considering the soil pH, farmers can select crops that are well-suited to the specific pH characteristics of their soil. Additionally, pH management through lime application or acidification can be employed to modify the soil pH and create a more favorable environment for crop growth.
3) **Organic Matter:**
Decomposed plant and animal matter makes up organic matter. It contributes to soil structure, water-holding capacity, nutrient retention, and microbial activity. Soils with higher organic matter content generally exhibit improved nutrient availability and water-holding capacity. Organic matter serves as a source of nutrients for plants, releases them gradually over time, and enhances soil structure, allowing for better root penetration and nutrient uptake. Crop recommendations consider the organic matter content of the soil to ensure sustainable soil management practices. Techniques such as incorporating organic amendments (e.g., compost, manure) into the soil, adopting cover cropping, and implementing crop rotations can increase organic matter levels, leading to improved soil fertility and crop productivity.
4) **Soil Texture:**
The texture of the soil is based on the proportions of sand, silt, and clay particles in it. It affects important soil properties such as water-holding capacity, drainage, aeration, and nutrient availability. Sandy soils have larger particles and tend to drain quickly, resulting in lower water-holding capacity. On the other hand, clay soils have smaller particles and retain water more effectively but may become compacted and poorly drained. Silt soils exhibit intermediate characteristics. Soil texture influences root development, as well as the movement of water, air, and nutrients in the soil. It also affects nutrient retention and availability to plants. For example, sandy soils generally have lower nutrient-holding capacity, requiring more frequent fertilization, while clay soils can retain nutrients but may be prone to waterlogging and reduced aeration. Crop recommendations

take into account the soil texture to guide farmers in selecting appropriate crops and implementing suitable soil management practices. This ensures that crops are matched to the soil's physical characteristics, optimizing water and nutrient utilization and promoting healthy crop growth. Understanding the influence of soil properties on crop growth allows farmers to make informed decisions regarding crop selection, fertilization, irrigation, and other management practices. By considering nutrient content, pH level, organic matter, and soil texture, farmers can optimize crop productivity, maximize resource utilization, and promote sustainable agricultural practices.

B. Crop-Soil Compatibility

Crop-soil compatibility refers to the match between specific crop requirements and the inherent properties of the soil in a given area. Different crops have varying needs in terms of nutrient levels, pH range, water availability, and soil texture. Understanding the compatibility between crops and the soil they are grown in is crucial for maximizing yields, preventing nutrient deficiencies or toxicities, and ensuring overall crop health.

Crop-soil compatibility is influenced by several factors:

1) **Nutrient Requirements:** Different crops have distinct nutrient requirements. Some crops may be more demanding in terms of specific nutrients, while others may be more adaptable to varying nutrient levels. Soil testing helps identify the nutrient content of the soil, allowing farmers to select crops that are well-suited to the existing nutrient levels. Additionally, nutrient management practices, such as fertilization and soil amendments, can be tailored based on crop-specific nutrient needs.
2) **Nutrient Requirements:pH Preferences:**
The pH range at which a crop thrives varies among different plants. Some crops prefer acidic soils, while others prefer neutral or alkaline soils. Understanding the pH requirements of crops helps farmers select suitable crops for the specific pH characteristics of their soil. Adjusting soil pH through lime application or acidification techniques can create a more favorable environment for crop growth and maximize crop-soil compatibility.
3) **Soil Texture and Water Availability:**
Soil Texture and Water Availability: Soil texture, determined by the proportions of sand, silt, and clay particles, affects water availability and drainage. Crops with shallow root systems may perform better in well-drained soils, while deep-rooted crops may be better suited to soils with good water-holding capacity. By considering the soil texture, farmers can select crops that are compatible with the soil's water retention and drainage characteristics, ensuring proper water management and reducing the risk of water stress or water-logging.
4) **Organic Matter and Soil Fertility:**
Organic matter content influences soil fertility, nutrient retention, and microbial activity. Some crops benefit from higher organic matter levels, while others may require less. By assessing the organic matter content of the soil, farmers can choose crops that are compatible with the existing soil fertility and organic matter levels. Additionally, adopting practices such as cover cropping, crop rotation, and organic amendments can help improve soil fertility and promote crop-soil compatibility.
5) **Disease and Pest Management:**
Soil properties can also influence the prevalence of certain diseases and pests. Some crops may be more susceptible to specific soil-borne diseases or pests, while others may exhibit more resistance. Understanding the disease and pest dynamics associated with the soil can aid in selecting resistant crop varieties, implementing appropriate crop rotations, and adopting pest management strategies. By considering crop-soil compatibility, farmers can optimize their crop selection and management practices. This approach enhances productivity, reduces the risk of crop failure, minimizes input costs, and promotes sustainable agricultural practices tailored to the specific attributes of the soil.

C. Role of Soil Properties in Crop Recommendation

1) **Nutrient Content:**
The availability and composition of essential nutrients in the soil greatly influence crop growth and development. Soil nutrient testing provides valuable insights into nutrient deficiencies or excesses, allowing farmers to apply targeted fertilization strategies. Crop recommendations based on nutrient content help ensure optimal nutrient uptake and minimize environmental impact.
2) **pH Level:**
Soil pH affects nutrient availability, microbial activity, and plant physiological processes. Different crops thrive within specific pH ranges, and recommendations based on soil pH enable farmers to select crops that can tolerate or benefit from the soil's pH characteristics. Proper pH management promotes healthy plant growth and improves nutrient uptake efficiency.
3) **Organic Matter:**
Organic matter content influences soil structure, water-holding capacity, nutrient retention, and microbial activity. Crop recommendations consider the organic matter content to ensure sustainable soil management practices. Organic matter additions, cover cropping, and crop rotations are prescribed to enhance soil fertility and overall productivity.
4) **Soil Texture:**
Soil texture refers to the proportions of sand, silt, and clay particles. It affects water infiltration, drainage, root penetration, and nutrient availability. Recommendations based on soil texture guide farmers in selecting appropriate crops and implementing soil management practices, such as irrigation scheduling, soil amendment, and conservation measures.

IV. Crop-Specific Soil Preferences

Different crops have specific soil preferences based on their tolerance to soil characteristics such as nutrient levels, pH, organic matter content, drainage, and soil texture. Understanding crop-specific soil preferences is essential for selecting the most suitable crops for a given area. Here are a few examples:

1) **Acid-Loving Crops:**
   Crops like blueberries, rhododendrons, and azaleas prefer acidic soils with a pH range of 4.0 to 5.5.
2) **Alkaline-Tolerant Crops:**
   Crops such as asparagus, cabbage, and spinach can tolerate soil with a pH 7.0 8.0.
3) **Nitrogen-Loving Crops:**
   Crops like corn, lettuce, and beans have high nitrogen requirements and perform well in soils with good nitrogen availability.
4) **Sandy Soil Crops:**
   Crops such as carrots, potatoes, and onions thrive in well-drained sandy soils with good aeration.
5) **Clay Soil Crops:**
   Crops like soybeans, sugar beets, and wheat can tolerate heavy clay soils but may require proper drainage management.

V. Dataset for Soil Testing and Analysis

Dataset is taken from Kaggle. It is of size 8x2200. This data set helps farmers for selecting crops based on first seven attributes. Fig. 1 lists the training data sets.

	N	P	K	temperature	humidity	ph	rainfall	label
0	90	42.0	43.0	20.879744	82.002744	6.502985	203.983200	rice
1	85	58.0	41.0	22.000940	80.319644	7.038096	226.655537	rice
2	60	55.0	44.0	23.004459	82.320763	7.840207	263.964248	rice
3	74	35.0	40.0	26.491096	80.158363	6.980401	242.864034	rice
4	78	42.0	42.0	20.130175	81.604873	7.628473	262.717340	rice

Fig. 1. Taining dataset used for crop analysis

A. Data Exploration:

Data exploration entails examining the connections between the datasetś columns. Using a heat map to visualise the correlation matrix is the straightforward method to search for correlations between columns. Pearson Correlation is used to visualize the correlation matrix as a heat map (see Fig. 2.). This dataset gives users the ability to build a predictive model based on these variables, which would then suggest the best crops to plant in a particular area (N, P, K, temperature[in degree C], humidity[%], ph, rainfall[in mm], label). Twenty-two Crops are listed in *label* field.

’rice’, ’maize’, ’chickpea’, ’kidneybeans’, ’pigeonpeas’, ’mothbeans’, ’mungbean’, ’blackgram’, ’lentil’, ’pomegranate’, ’banana’, ’mango’, ’grapes’, ’watermelon’, ’muskmelon’, ’apple’, ’orange’, ’papaya’, ’coconut’, ’cotton’, ’jute’, ’coffee’

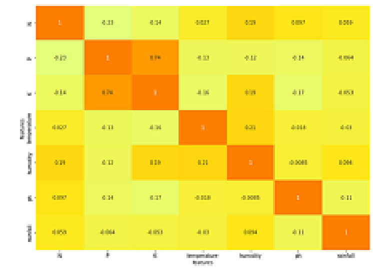

Fig. 2. correlation between columns of dataset

VI. Performance Metrics

Table I represents accuracy of different algorithms.

In this paper, six different machine learning models were trained and evaluated on a dataset for crop recommendation. Graphical representation of Table I is shown in Fig. 3. The highest accuracy obtained is by using **XGBoost**.

TABLE I
ACCURACY CORRESPONDING TO DIFFERENT ALGORITHMS.

Serial No.	Algorithm	Accuracy(%)
1	Decision Tree	90
2	Naive Bayes	99.090
3	SVM	10.681
4	Logistic Regression1	95.22
5	Random Forest	99.090
6	XGBoost	99.318

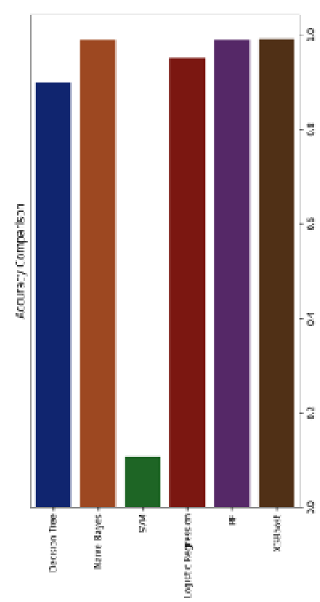

Fig. 3. Accuracy comparison of algorithms

VII. Conclusion:

This paper recommends different crops of India and their analysis is given using machine learning algorithms. The Analysis has been performed on six types of machine learning algorithms (see Table I) and out of these six algorithms **XGBoost** achieved best accuracy result. Such performance metric will definitely help farmers in making decision that what are values required of each first seven attributes viz. **N, P, K, temperature, humidity, ph, rainfall** in order to grow any particular crop. This in turn will indirectly help in boosting economy of the country.

VIII. Challenges and Limitations

1) **Data availability and quality:**
One of the primary challenges in crop recommendation systems is the availability and quality of data. Collecting comprehensive and accurate soil property data, climate data, and crop performance data can be time-consuming and resource-intensive. Additionally, data from different sources may vary in terms of collection methods, measurement techniques, and quality standards. Ensuring data integrity and standardization is crucial for reliable crop recommendations.
2) **Generalization across different regions:**
Crop recommendation models developed for specific regions may not generalize well to other regions with different soil and climatic conditions. Soil properties can vary significantly across different locations, and models trained on data from one region may not accurately predict optimal crops in another region. It is important to consider regional variations and adapt or develop models that account for local conditions.
3) **Incorporating dynamic soil changes:**
Soil properties can change over time due to natural processes, land management practices, and climate change. Incorporating dynamic changes in soil properties into crop recommendation models is challenging but essential for accurate and up-to-date recommendations. Continual monitoring and updating of soil data and models are required to account for these changes.
4) **Interactions between soil properties:**
Soil properties interact with each other, creating complex relationships that impact crop growth and performance. Modeling these interactions accurately can be challenging, as the relationships may not always follow linear patterns. Understanding and incorporating the synergistic or antagonistic effects of different soil properties is crucial for improving the precision of crop recommendations.
5) **Integration of climate and weather factors:**
Crop growth and performance are influenced by climate and weather conditions. Integrating climate data into crop recommendation models enhances their accuracy. However, obtaining reliable climate data and understanding the complex interactions between soil properties and climate variables pose additional challenges.

IX. Recommendations for Further Research:

While crop recommendation systems have shown promise, several challenges need to be addressed for their wider adoption and effectiveness. Recommendations for further research include:

1) Enhancing data availability and quality through improved data collection methods, standardization protocols, and integration of remote sensing and participatory approaches.
2) Continuously improving model accuracy and validation by incorporating feedback from farmers and stakeholders, leveraging advanced machine learning techniques, and conducting robust validation studies.
3) Promoting technological adoption and access by developing user-friendly and affordable tools, addressing the digital divide, and providing training and capacity-building programs for farmers and extension workers.
4) Integrating climate change considerations, sustainable farming practices, and socio-economic factors into crop recommendation systems to support farmers in adapting to changing conditions and promoting resilient and profitable agriculture.
5) Conducting research on the long-term impacts of crop recommendations on soil health, ecosystem services, and farmer livelihoods to assess the sustainability and socio-economic outcomes of these systems.

By addressing these research recommendations, crop recommendation systems can continue to evolve and contribute to sustainable agriculture, food security, and the well-being of farming communities. In a nutshell, crop recommendation systems hold significant potential in assisting farmers in making informed decisions and optimizing crop productivity. By leveraging soil properties, data-driven approaches, and key factors for recommendation, these systems can contribute to sustainable and profitable agricultural practices.

References

[1] D. ManendraSai, M. S. Dekka, M. M. Rafi, M. M. R. D. Apparao, M. T. Suryam, and M. G. Ravindranath, "Machine learning techniques based prediction for crops in agriculture," *Journal of Survey in Fisheries Sciences*, vol. 10, no. 1S, pp. 3710–3717, 2023.

[2] W. X. Z. Y. L. J. . Z. X. . Chen, H., "A hybrid model for crop recommendation based on machine learning and expert knowledge. information processing in agriculture,," *Information Processing in Agriculture*, vol. 9, no. 1, pp. 153–168, 2022.

[3] F. Wang, X. Wang, and N. Song, "Biochar and vermicompost improve the soil properties and the yield and quality of cucumber (cucumis sativus l.) grown in plastic shed soil continuously cropped for different years," *Agriculture, Ecosystems & Environment*, vol. 315, p. 107425, 2021.

[4] S. J. W. J. A. S. . B. D. R. 4. Jones, D. B., "Machine learning approach to crop recommendation based on soil characteristics," *Computers and Electronics in Agriculture*, vol. 169, p. 105184, 2020.

[5] W. X. Z. Y. L. J. . Z. X. . Chen, H., "Expert system for crop recommendation based on soil analysis. journal of intelligent fuzzy systems," *Information Processing in Agriculture*, vol. 36, no. 5, pp. 5143–5153, 2019.

[6] W. X. Z. Y. L. J. . Z. X. . Chen, H., “Expert system for crop recommendation based on soil analysis. journal of intelligent fuzzy systems,” *Information Processing in Agriculture*, vol. 36, no. 5, pp. 5143–5153, 2019.

[7] W. X. Z. Y. L. J. . Z. X. . Chen, H., “Crop suitability assessment: An approach to analyze soil–crop compatibility. geocarto international,” *Geocarto International*, vol. 33, no. 1, pp. 33–49, 2018.

[8] S. S. P. . J. J. S. Smith, R. S., “Soil physical properties and their impact on crop performance. soil science society of america journal,” *Soil Science Society of America Journa*, vol. 81, no. 3, pp. 2029–2033, 2017.

[9] . S. P. Gupta, P., “Predicting crop yields based on soil parameters using regression analysis,” *International Journal of Computer Science and Information Technologies*, vol. 7, no. 4, pp. 727–738, 2016.

[10] D. Li, Y. Miao, S. K. Gupta, C. J. Rosen, F. Yuan, C. Wang, L. Wang, and Y. Huang, “Improving potato yield prediction by combining cultivar information and uav remote sensing data using machine learning,” *Remote Sensing*, vol. 13, no. 16, 2021.

[11] D. Li, Y. Miao, S. K. Gupta, C. J. Rosen, F. Yuan, C. Wang, L. Wang, and Y. Huang, “Applying naive bayes classification technique for classification of improved agricultural land soils,” *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 6, no. 5, pp. 189–193, 2018.

[12] B. Sawicka and B. Krochmal-Marczak, “‘biotic components influencing the yield and quality of potato tubers,” *Herbalism*, vol. 1, no. 3, pp. 125–136, 2017.