e-ISSN:2582-7219

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 7, Issue 7, July 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.521

6381 907 438 6381 907 438 ijmrset@gmail.com www.ijmrset.com

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707062 |

Soil Analysis and Crop Recommendation using Machine Learning

Aishwarya, Dr. Madhwaraj K G

PG Student, Department of MCA, Mangalore Institute of Technology & Engineering Moodabidri, Karnataka, India

Professor & Head, Department of MCA, Mangalore Institute of Technology & Engineering Moodabidri, Karnataka, India

ABSTRACT: In today's agricultural context, crop yield optimization and sustainable farming methods are critical. Soil quality is crucial to crop health. Conventional techniques for suggestions for crops derived from evaluation of soil are sometimes time-consuming, labour-intensive and inaccurate. To deal with these issues, the research we conducted suggests a machine learning (ML) based soil examination and crop suggestion system to address these problems. In this study, a comparative study was taken up with different ML algorithms like Random Forest, SVM, KNN and MobileNetV2. A dataset was taken and used on all these algorithms. It is found that MobileNetV2 gives higher correctness in comparison with the other algorithms. Further, a user-friendly web application that is easy to use was developed for farmers to help them know the type of soil and the type of crops that can be cultivated in that soil.

KEYWORDS: Convolutional Neural Network (CNN), Random Forest, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), MobileNetV2.

I.INTRODUCTION

Many economies are built on agriculture, which is also essential to ensure the world's food security. Almost half of India's population depends on agriculture for their livelihoods, making it the foundation of the nation's economy. India grows a vast range of crops due to its diverse range of soil types and temperatures, including crops like cotton, tea, and sugarcane as well as staples like rice, wheat, and maize. The use of high-yielding variety seeds, chemical fertilizers, and contemporary irrigation techniques during the Green Revolution of the 1960s and 1970s significantly changed Indian agriculture and improved food production. The need for food is rising due to population expansion, so it is critical to boost agricultural productivity while upholding environmental practices. The country's economy relies on agriculture, which employs millions of farmers and contributes significantly to the Gross National Product. Indian agriculture yields a wide variety of crops because of varied climatic conditions and different kinds of soil. Traditional farming practices have been maintained alongside modern techniques. Conventional techniques for soil analysis and crop recommendation mostly rely on expert knowledge, historical data, and manual procedures. Although these techniques can be somewhat successful, they are frequently costly, time-consuming, and have limited scalability. The emergence of ML offers a chance to improve existing procedures by offering more precise, effective, and expandable solutions which can lead to ineffective agricultural practices and low yields. The use of precision farming and digital platforms for market access are two examples of how modern technology and creativity are being integrated to ensure food security and boost economic development in Indian agriculture. The use of soil analysis and crop recommendation system can raise agricultural output, reduce resource waste, and improve sustainability, resource consumption and environmental damage. Precision agriculture, backed by modern technologies, provides a possible answer to these difficulties by allowing for data-driven decision-making based on soil type. Soil health is a key factor in crop productivity, however traditional soil analysis procedures are labour-intensive, time-consuming, and frequently lack the precision required for decision-making. Incorporating ML into soil analysis can transform the process by giving accurate, real-time insights and personalized crop suggestions. These innovations have great potential for the future of Indian agriculture. This study is to show how the soil analysis and crop recommendation can be used in practice to improve agricultural sustainability and production. The paper is organized as follows. Section II provides an overview of the related work on soil analysis and crop recommendation. Section III shows research methodology. Section IV consists of results and observations. Section V includes conclusion and reference.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707062 |

II. BACKGROUND AND RELATED WORK

Pudumalar S et al. [1] used the ML technology and data analysis to monitor and manage field variability in crops. The algorithms used are Random Forest, KNN and Naïve Bayes This study proposed a recommendation system where the data was collected from a soil testing lab with the soil attributes namely, its depth, soil color. The best outcome comes with KNN model. Neelam Singh et al. [2] conducted a comparative study on algorithms namely KNN, Random Forest, K-Star, and Naive Bayes. The most effective algorithm for crop prediction technique for the Uttarakhand region, in particular the Almora district outcome comes with KNN. Vishal Kumar et al. [3] explores a system designed to assist farmers in predicting suitable crops based on soil type. This study utilizes datasets with pictures of several soil types and details on different crops, the system applies a CNN to identify the type of soil and suggest suitable crops. By ensuring that farmers apply the optimum amount of nutrients and fertilizers according on the characteristics of the soil, the aim is to increase crop output. This strategy will help farmers to make decisions that will increase their revenue and lower their risk of financial distress while also attempting to mitigate the negative consequences of problems like drought and poor soil on agricultural productivity. Devdatta A. Bondre and Mr. Santosh Mahagaonkar [4] addresses the issues of modern agriculture by utilizing ML techniques to improve crop yield prediction and fertilizer recommendation. The system incorporates a variety of variables, including soil nutrients, crop yield statistics, and location data, to help farmers choose the best crops and fertilizers for optimal yield results. Shivani Girme et al. [5] described a web-based application that helps farmers by automating soil classification and crop suggestion using advanced ML techniques. The system uses CNN to classify soil images based on their features, and SVM algorithms to recommend crops for the different soil types. The authors stress the necessity of specific soil classifying for efficient crop production and recommend a system that comprises taking soil images, analyzing them with CNN, and proposing crops using SVM. Ayesha Barvin and T. Sampradeepraj [6] conducted the comparative study of two graph-based crop recommendation algorithms, namely Graph Convolutional Network (GCN) and Graph Neural Network (GNN), aimed at optimizing crop selection based on various environmental factors such as soil nutrients, temperature, humidity, soil PH, and rainfall.

III. RESEARCH METHODOLOGY

The research methodology adopted in this study consists of steps like data gathering, data pre-processing, model selection and implementation. Figure 1 shows the workflow of the performed study. It starts with a soil image, which is then cleaned and processed before being used to train the model. The pre-processed data is used to train the model, allowing it to learn and detect essential features from soil photos. After training, the model is tested with new photos to determine its performance and ability to generalize to previously unseen data. The model then generates expected results that classify the soil types. These forecasts are combined into the final results, which provide useful insights for applications such as soil type and crop recommendations. This procedure guarantees an organized and successful approach to soil categorization via ML

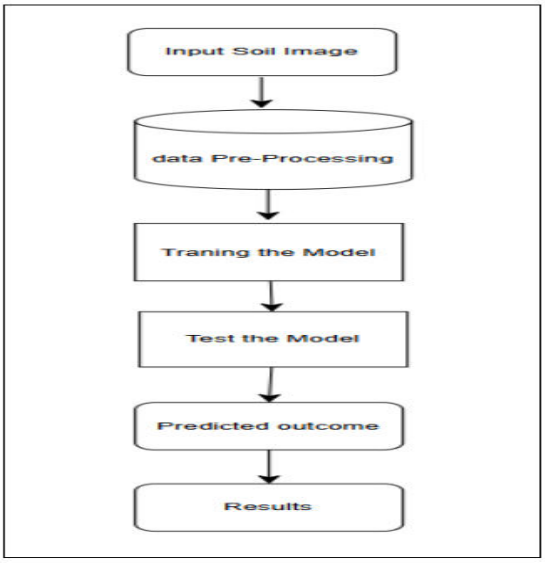

Figure 1: Workflow of the study

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707062 |

A. Data Collection

The first step is data collection where, here the dataset is collected from online available source like Kaggle https://www.kaggle.com/datasets/gosegomenwe/soil-data-v3. In this data set there are eight types of different soil images as shown in the below figure 2 namely alluvial soil, black soil, clay soil, red soil, sand soil, marshy soil and silt soil.

Figure 2: Different types of Soils

B. Data Preprocessing and Model selection

After collecting the required data set the next step is data pre-processing. The data that has been pre-processed properly is guaranteed to be reliable, consistent, and clean enough to train the strong model. Before model building model, selection is to be done by doing a comparative study between the multiple models, choose the model which is giving higher accuracy. The models Includes Random Forest, SVM, KNN and MobileNetV2.

1. Random Forest

The Random Forest model is one of the flexible and effective ML approach for classification and regression applications is the Random Forest model. During training, it creates a large number of decision trees, from which it outputs the mean prediction for each tree. A random subset of the training data is used to build each tree in the forest, and splits are created based on random subsets of characteristics within each tree. This unpredictability aids in lowering overfitting and enhancing the model's ability to generalize to new data. The Random Forest model achieves good prediction performance and robustness to noise and outliers in the dataset through the aggregation of numerous decision trees. The result of this model is shown in below figure 3 as a confusion matrix.

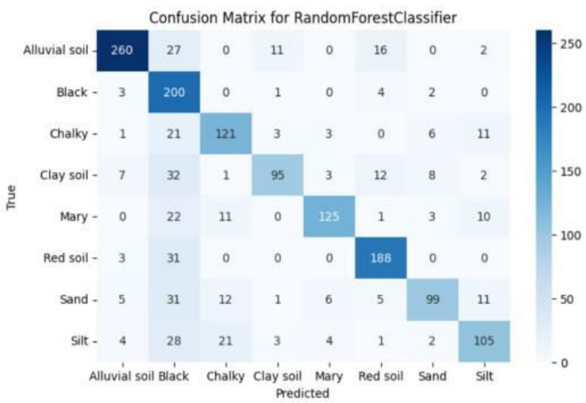

Figure 3: Confusion Matrix for Random Forest

2. K-Nearest Neighbors

The KNN model is a straightforward non-parametric instance-based learning strategy that may be used to both classification and regression applications. KNN classifies a data point according to the class majority of its k closest neighbors in the feature space, where k is a user-defined number. The procedure computes the difference between the data point and its neighbors, usually using Euclidean distance, and assigns the label that is most common among the k closest neighbors. Despite its simplicity and minimalism of implementation, KNN is computationally demanding on

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707062 |

big datasets and may be sensitive to irrelevant or redundant features. It works well in scenarios with complex and non-linear decision boundaries. This makes it an effective tool for various ML techniques. This model yields the following outcome appears in below figure 4.

Figure 4: Confusion Matrix for KNN

3. Support Vector Machine

The SVM model is a highly effective supervised learning methodology for regression classification and tasks. SVM works by determining the most effective hyperplane to divide data into various classes in a high-dimensional feature space. The approach seeks to increase the margin between the nearest information points of distinct classes, known as support vectors, hence improving the model's generalizability. SVM is capable of manage both non-linear and linear classification problems by transforming input information into an area with more dimensions with a linear separator using kernel functions. It is especially successful in high-dimensional environments and when sample sizes are out of proportion to dimensions in the sample Despite being computationally intensive and sensitive. The result of this model is demonstrated in below figure 5.

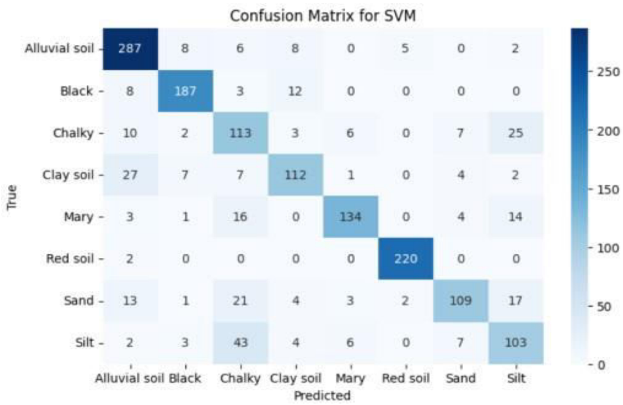

Figure 5: Confusion Matrix for SVM

4. MobileNetV2

MobileNetV2 is a cutting-edge architecture of neural network built for high-performance mobile and embedded vision applications. It extends the original MobileNets by incorporating inverted residual structures and linear bottlenecks as illustrated in below figure 5. The main concept is to use thin layers to save computing expenses without sacrificing model accuracy. Depth-wise separable convolutions are used by MobileNetV2 in which break down traditional convolutions into two simpler operations: depth wise and pointwise convolutions. This division greatly decreases the size of parameters and computations. The inverted residuals employ shortcut connections, enabling for more effective training and inference. Furthermore, linear bottlenecks help to conserve information by avoiding nonlinearities from deleting too much data. This architecture strikes a compromise between performance and efficiency, making it works

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707062 |

well for real-time applications on resource constrained hardware. The resultant of this algorithm is presented in below Figure 6 as a confusion matrix.

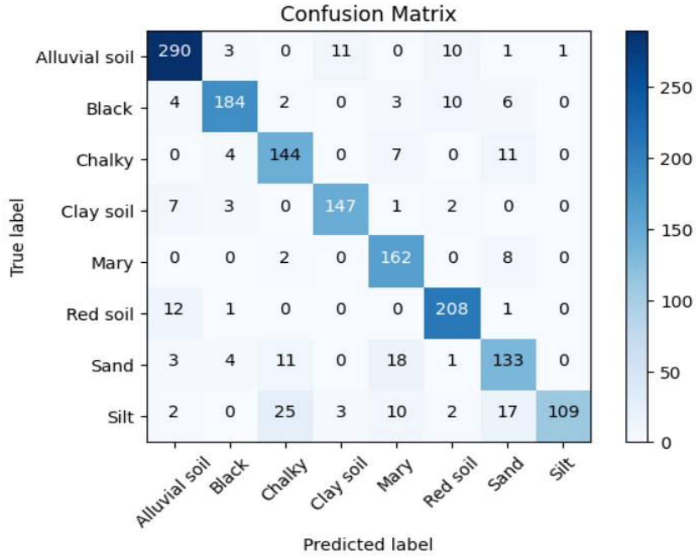

Figure 6: Confusion Matrix for MobileNetV2

IV. RESULTS AND OBSERVATION

In this study the information is gathered from the Kaggle and other internet sources provided by a dataset of soil image data set. The data generally consists of images of eight different types of soils. The collection of data consists of total 5258 of soil images from the data set where Various models, including Random Forest, KNN, and SVM, are applied to the same dataset, and the models' outputs are displayed as confusion matrices, as seen in figures 3, 4, 5, and 6.

Table 1. Performance and Analysis of ML models

Model	Accuracy%	F-1 Score%	Recall%
Random Forest	75.44	75.92	79.69
K-Nearest Neibhours	54.8	56.84	55
SVM	79.36	79.37	79.61
MobileNetV2	87	85.8	87

As mentioned in the above table 1 the mobileNetV2 is giving the higher accuracy so we can say this model is most accurate for the chosen dataset.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707062 |

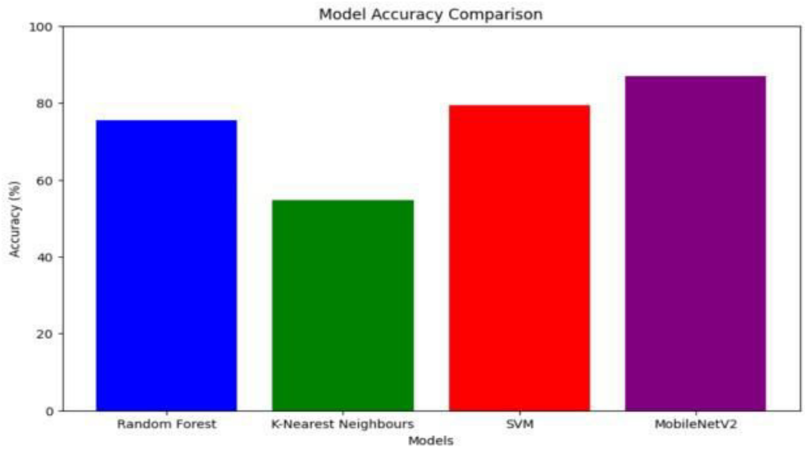

Figure 7: Bar graph for accuracy Comparison

The above figure 8 represents the precision of four different ML models. Each bar shows one model, with the height of the bar denoting the accuracy percentage. The graph reveals that MobileNetV2 is the most precise with 87%, followed by SVM at 79.36%, Random Forest at 75.44%, and KNN at 54.8%. This visual comparison emphasizes the performance differences across the models, with MobileNetV2 surpassing the others, demonstrating that it is the most reliable model on the market for this specific need.

V. CONCLUSION

This study conducted a comparative comparison of various models and compared their results and also it is cantered on building an easily navigable web application using flask frame work where the user can easily upload the image of the soil and when they click on the submit button then they can get the results as what type of soil it is and according to soil type it recommends suitable crop. The study examines strong performance of MobileNetV2 is greater in real-time application than other models, as indicated in Table 2. This study emphasizes on how cutting-edge technology have transformed India's agricultural environment. This model considering a several different kinds of variables including soil types by seeing its texture and color, which are necessary to enhance crop choices. This research illustrates how farmers may easily access cutting-edge technology via user friendly web application. The benefits of utilizing ML in crop recommendation, including precision, customization, scalability and flexibility emphasize the potential of these technologies to completely change conventional agricultural methods, with its innovative technology and farmer focused approach, the suggested crop selection system is a big step toward reviving the agriculture industry and providing hope for a better future

REFERENCES

1. Pudumalar, S., Ramanujam, E., Rajashree, R. H., Kavya, C., Kiruthika, T., & Nisha, J. (2017, January). Crop recommendation system for precision agriculture. In 2016 eighth international conference on advanced computing (ICoAC) (pp. 32-36). IEEE.
2. singh, Neelam, et al. "Crop prediction method to maximize crop yield rate using machine learning technique: A case study for Uttrakhand Region." Journal of Critical Reviews 7.12 (2020): 4603-08.
3. Kumar, Vishal, et al. "Agriculture soil analysis for suitable crop prediction." PP, Agriculture Soil Analysis for Suitable Crop Prediction (May 24, 2021) (2021).
4. Bondre, Devdatta A., and Santosh Mahagaonkar. "Prediction of crop yield and fertilizer recommendation using machine learning algorithms." International Journal of Engineering Applied Sciences and Technology 4.5 (2019): 371-376.
5. Shivani Grime, Shivani Mutke, Srushti Kokare, Akanksha Khobragade, Prof. Kailash Tambe Soil Classification Using Convolutional Neural Network and Crop Recommendation based on Support Vector Machine International Journal of Advanced Research in Science, Communication and Technology (IJARSCT) Volume 4, Issue 2, April 2021
6. Ayesha Barvin, P., & Sampradeepraj, T. (2023). Crop Recommendation Systems Based on Soil and Environmental Factors Using Graph Convolution Neural

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707062 |

7. Network: A Systematic Literature Review. Engineering Proceedings, 58(1), 97. Motwani, A., Patil, P., Nagaria, V., Verma, S., & Ghane, S. (2022, January). Soil analysis and crop recommendation using machine learning. In 2022 International Conference for Advancement in Technology (ICONAT) (pp. 1-7). IEEE.
8. Kumar, V., Kumar, R., Kumar, S., & Jorvekar, P. P. (2021). Agriculture soil analysis for suitable crop prediction. PP, Agriculture Soil Analysis for Suitable Crop Prediction (May 24, 2021).
9. Singh, N., Pant, D., Singh, D. P., & Pant, B. (2020). Crop prediction method to maximize crop yield rate using machine learning technique: A case study for Uttrakhand Region. Journal of Critical Reviews, 7(12), 4603-08.
10. Babu, S. (2013, August). A software model for precision agriculture for small and marginal farmers. In 2013 IEEE Global Humanitarian Technology Conference: South Asia Satellite (GHTC-SAS) (pp. 352-355). IEEE.
11. Rajak, R. K., Pawar, A., Pendke, M., Shinde, P., Rathod, S., & Devare, A. (2017). Crop recommendation system to maximize crop yield using machine learning technique. International Research Journal of Engineering and Technology, 4(12), 950-953.
12. Reddy, A. K. M., Chithra, S., Hemashree, H. M., & Thanu, K. (2020). Soil classification and crop suggestion using machine learning. Int. J. for Res. in Appl. Sci. and Eng. Technol., 8(7), 1625-1628.
13. Manjula, A., & Narsimha, G. (2015, January). XCYPF: A flexible and extensible framework for agricultural Crop Yield Prediction. In 2015 IEEE 9th international conference on intelligent systems and control (ISCO) (pp. 1-5). IEEE.
14. Savla, A., Israni, N., Dhawan, P., Mandholia, A., Bhadada, H., & Bhardwaj, S. (2015, March). Survey of classification algorithms for formulating yield prediction accuracy in precision agriculture. In 2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS) (pp. 1-7). IEEE.
15. Anguraj, K., Thiyaneswaran, B., Megashree, G., Shri, J. P., Navya, S., & Jayanthi, J. (2021). Crop recommendation on analyzing soil using machine learning. Turkish Journal of Computer and Mathematics Education, 12(6), 1784-1791.

INNO SPACE

SJIF Scientific Journal Impact Factor

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com