

Soil Analysis and Crop Recommendation Using Machine Learning

Dr. Sandhya G¹, Himanshu Bhushan², Charan N P³, Dayakar Reddy⁴

¹Prof & HOD, Dept of AI&ML, M S Engineering College

^{2,3,4} B. Tech, (AI&ML) M S Engineering College

Abstract—Agriculture is the backbone of Indian economy and livelihood to many people. Farmers often choose crops for their field based on their own experience and instinct. This sometimes leads to loss and less yield. If the selection of crop is done with the productivity data of the entire region, it may lead to better results. However, all the crops cannot be cultivated in a particular soil. So, the soil must be analysed crops must be suggested based on the type of soil. This work suggests an idea that is useful and easily accessible to all farmers of India without any need of hardware. A list of crops with their success rate will be suggested to the farmer when the region of agriculture and soil image are given as inputs. This list of crops is both profitable and produce more yield in that region.

Index Terms—CNN (Convolution Neural Network), soil pH, sustainable agriculture, crop prediction.

I. INTRODUCTION

Agriculture is the primary source of livelihood for about 58% of the population of India. Continuous efforts have been taken to develop this sector as the whole nation depends on it for food. For thousands of years, we have been practicing agriculture but still, it remained underdeveloped for a long time. After the green revolution, we became self-sufficient and started exporting our surplus to other countries. Earlier we used to depend completely on monsoon for the cultivation of food grains but now we have constructed dams, canals, tube-wells, and pump-sets. Also, we now have a better variety of fertilizers, pesticides, and seeds, which help us to grow more food in comparison to what we produce during old times. With the advancement of technology, advanced equipment, better irrigation facilities agriculture started improving. Furthermore, our agriculture sector has grown stronger than many countries and we are the largest exporter of many food grains. In recent years,

farmers are suffering financially and are facing many hardships. This is due to various reasons such as urbanisation, globalisation, pollution, water scarcity, less rainfall, low fertility of soil, drastic climatic changes, political and economic reasons, poverty, lack of technological assistance etc. Though we have very less to contribute to improvise the natural factors to help agriculture, we have a lot to contribute to this sector through computer science and technology. Internet of Things (IoT), Artificial Intelligence, smart agriculture, Agricultural Engineering, Irrigation Engineering are some of the fields that contributed to the development of agriculture in recent years.

There are many ways to suggest crops suitable for a farm land. It can be based on the climate or soil or the crop that produces high profit in that region. We want to suggest crops considering all these factors.

II. PROPOSED SYSTEM METHODOLOGY

The project involves collecting soil-related image data from reliable sources like Kaggle, focusing on four soil types: Alluvial Soil, Red Soil, Clay Soil, and Black Soil. Data preprocessing is carried out using Python, with libraries such as TensorFlow, Pandas, NumPy, Matplotlib, Keras (for model building), scikit-learn (for data splitting), and PIL (for image processing). The images are retrieved along with their labels, resized to a uniform size of 128x128 pixels, and converted into NumPy arrays to ensure consistency. Finally, the dataset is split into training (80%) and testing (20%) subsets for model development and evaluation.

III. SYSTEM ARCHITECTURE

The overall Architecture diagram for the proposed system is shown in Fig 1. The proposed work is split into different processing phases namely Soil Classification, Suitable Crop Suggestion and Best Crop Prediction. These working phases execute in the depicted flow to produce the list of crops with success rate as output from the input soil image and region. The Soil Classification module is designed to classify the different types of soil using a deep learning model. This model inputs soil images from the user and states the type of the soil as output. The output is one of the following: Alluvial soil, red soil, Black soil, Clay soil.

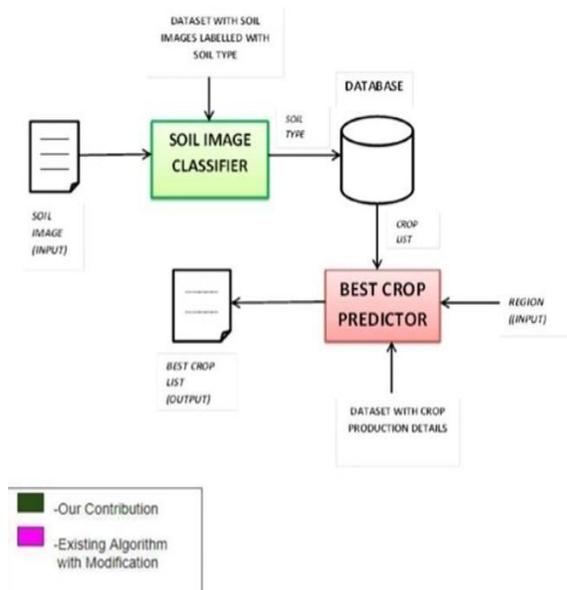


Fig 1 Overall Architecture Diagram

III. SYSTEM IMPLEMENTATION

SOIL CLASSIFICATION - In this model, the aim is to classify the different types of soil using a deep learning model. This model inputs soil images from the user and states the type of the soil as output. We used SVM and CNN architectures like LeNet, AlexNet, VGG 16, ResNet for soil image classification and evaluated the accuracy of each of the classifiers. The CNN model that produced the highest accuracy was chosen for the soil classification. The models are trained with four soil types, namely red, alluvial, black and clay.

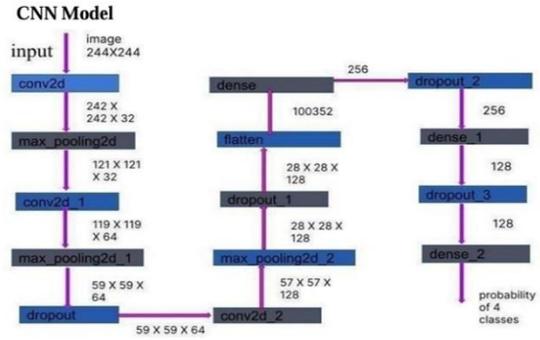


Fig 2 Architecture of Custom CNN model

The CNN architecture depicted in Fig 2 is built with conventional layer by layer feature extraction techniques. There are three convolutional layers with ReLU activation function followed by max pooling. Then the feature map is flattened. Finally, there are three fully connected layers with ReLU activation function. Dropouts are added to avoid overfitting. The final dense layer has softmax activation function.

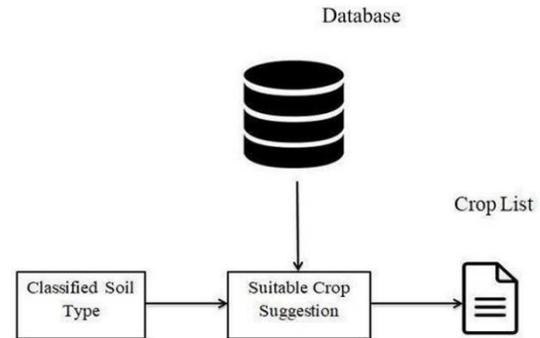


Fig 3 Suitable Crop suggestion model

As shown in Fig 3, the type of soil from the previous model is used to decide the suitable crops cultivable in that area. This module provides a list of suitable crops for the soil type fetched from local storage. This list was collected from authorised sources. Table 5.1 shows the list of crops for the four types of soil.

BEST CROP PREDICTION The aim of this model is to find the crops that are best for their region, so that the farmers can get a maximum profit by cultivating these crops. This model is fed with the list of crops from the previous model, the region as input and it will output a list of best crops and success rate of those crops. The model is trained using data for the past 10 years collected from various trusted sources.

Algorithms used will be customized using multiple linear regression and customized K fold method.

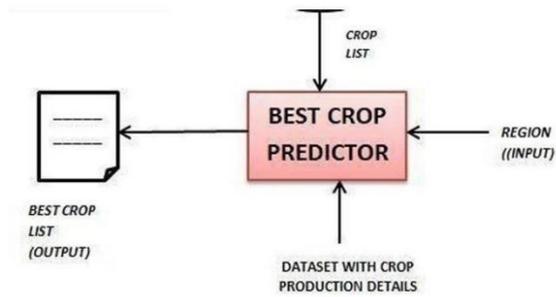


Fig 4 Best crop prediction Architecture

Ten different multiple linear regressions are done to predict various parameters like imports, exports, gross production, production per unit area and production. These regressions are done beforehand and the predicted values are stored in separate csv files. These values are used in future calculations of success rate.

IV SYSTEM TESTING AND DASHBOARD

The detection of Testing is done to find flaws. The objective behind testing is to locate every potential flaw or weakness in a work element. It offers a plan for testing the convenience of individual segments, subassemblies, gatherings, or possibly final products. It was the most considered normal approach to testing programming to satisfies client suppositions and needs and doesn't breakdown in a way that is unsuitable. Different test sorts exist. Each test type answers a specific examination need.

Test cases are the absolute experiments that were engaged for testing the application to build it virus free however much as could be expected. The application is tested by the engineers using these test cases. Those test cases are preferred from a range of testing methods.

Every field entry needs to function correctly. You have to click the designated link to activate the pages. There shouldn't be any delays in the entry screen, messages, or answers.

Test Case ID	Title of test Cases	Explanation	Actual Outcome	Result
1	Load Dataset	Loading the dataset	Shows in output screen	Pass
2	Preprocessing Data	Preprocessing of data to detect missing values	Display the missing values	Pass
3	Splitting of Data	Encoding before and after index of data	Encoding data is going to display of the dataset	Pass
4	Features Extraction	It extracts the features	Shows the features of the datasets	Pass
5	CNN Architecture	Soil analysis and crop Recommendation	Soil analysis and crop Recommendation	Pass
6	User Login	Valid Login Credentials	Soil analysis and crop recommendation Home page	Pass
7	Login	Invalid Details	Login Page	Fail

Table I: Test Case Report

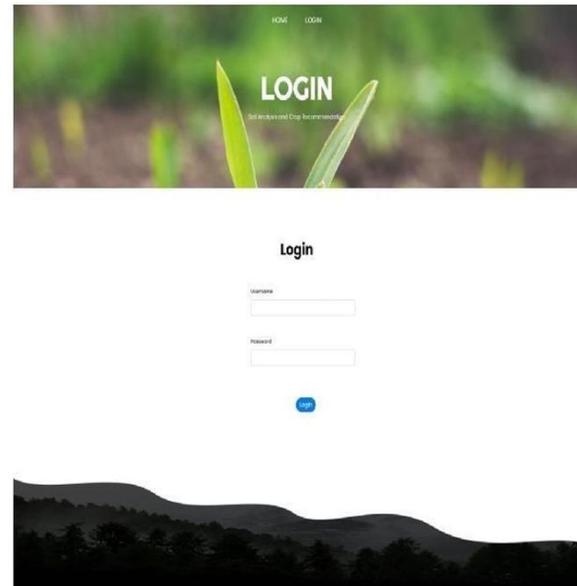


FIG 5 LOGIN PAGE

The login page in above contains a username and password that the administrator can access with valid login credentials.

The image provided show a page of the system Soil Analysis and Crop Recommendation which allows an administrator to securely login to the system. It's interface design is appealing with the background

color in green which confirms that the application is nature-themed. Also, at the top of the page, the HOME button and the LOGIN button have been added for ease of access and usability. Centrally, the section that needs to be filled in with the credentials is located also, the username and password input fields are labeled appropriately. There is a blue “Login” button below the input fields which allows an unregistered user to submit their credentials. This page makes sure that only users with appropriate login credentials can access the backend features of the system which guarantees data security as well as systems integrity. The system's design and arrangement improve the overall experience of the user significantly and focuses on the aim of the application, which is smart agriculture.

V RESULTS



FIG 5 PREVIEW

Figure 5 above shows, uploaded image of another type of soil in preview page.

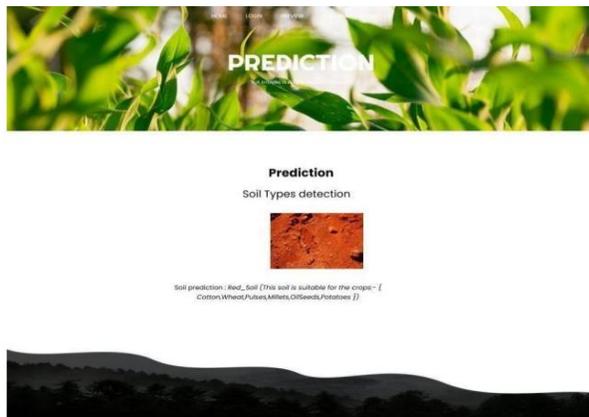


FIG 6 PREDICTION

Figure 6 above shows predict the type of soil and recommend crops.

VI CONCLUSION

Soil images are classified accurately. Soil image classification works well for real time images. Crops with success rate are calculated taking all the mentioned parameters like export, import, production per unit area etc into account. The developed website is extremely user-friendly with simple and clear migrations. Most of the calculations are done beforehand to reduce the latency to the users. We strongly believe that the developed system solves the problem of choosing suitable crops for their fields by farmers.

A flourishing agricultural sector is key for India’s sustained economic growth. The goal was to empower farmers with lesser land holdings by increasing profitability and maximizing crop yield. In this experiment CNN Architecture gave significantly better results to the image classification for selected soil classification dataset. The final CNN Architecture had an amazing Training accuracy of 97.34% and Validation accuracy of 99.21%.

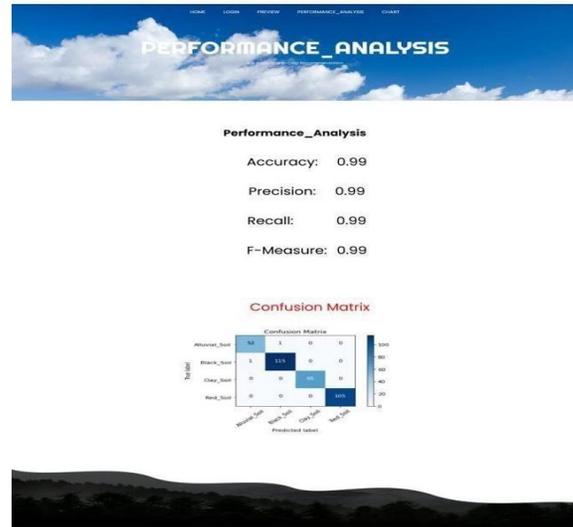


FIG 7 PERFORMANCE ANALYSIS

Figure 7 above shows the performance analysis of soil with the help of confusion matrix.

The image that has been provided corresponds with the Chart Page of the Soil Analysis and Crop Recommendation system, which is highlighted in Figure 8. This page is pivotal in portraying and graphically summarizing the various types of soils contained in the dataset integrated into the system.

There is also a menu at the top of the page with a clear and organized header including Home, Login, Preview, Performance Analysis, Chart, and Logout, which are all easily accessible and clickable links to other modules of the application. A prominent pie chart divides the dataset into segments of different Soil Types, Alluvial, Black, Red, and Clay Soils, where each type is represented by a color for distinction. So, for example, the chart as it stands summarizes the percentage composition of each soil type in the dataset and gives an overview to users and administrators about the dominant or rare soil types available in the dataset. The chart enables users to quickly identify the predominance or scarcity of certain soil types, and this primary step of the analysis helps to illustrate insights into the soil distribution patterns, which greatly influences the crop recommendation logic contained in the application. Furthermore, the layout and background image illustrate the application's agricultural theme, thus making the interface of the application informative and visually attractive.

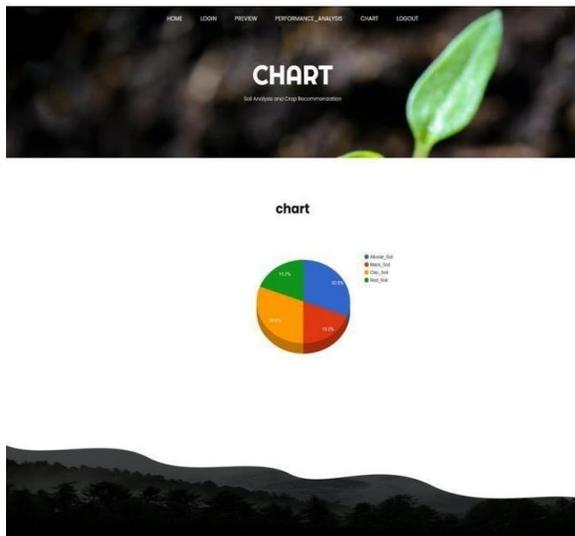


FIG 8 CHART

Figure 8 above shows the structure with percentage of types of soil in given dataset.

VII FUTURE SCOPE/ENHANCEMENT

While the developed system takes only soil type to determine the crops suitable, it might be more realistic if the weather and climatic conditions are also considered to make the decision. Instead of manual entry of a region, GPS technology can be used to

determine the location. With the availability of the type of soil in a particular region, the usage of images to find the type of soil can be eliminated. The website can be extended as a complete guide to farmers including the fertilizers, pesticides to be used etc. Stand on the Weather Factors, this revise explains how data mining can help predict crop yield. The webpage we created is super easy to use! Plus, it has an accuracy of over 75% for all the crops & districts included. Just by entering climate info for their location, anybody can find out about the crop yields they're interested in.

REFERENCES

- [1] S. " A s/w replica for precision agriculture for marginal and small farmers " is what Babu presents in Proc. Doi: 10.1109/GHTCSAS.2013.6629944 South Asia Satellite, IEEE Global Humanitarian Technology Conference pages 352-355
- [2] S. Pudumalar et al.'s "Crop suggestion system for exactitude agriculture" in Proc. Eighth World Conference Adv., Comput., pages 32-36, doi:10.1109/ICoAC.2017.7951740
- [3] R. K. "Crop Proposal Framework to Augment Harvest Yield Using AI Method," International, Rajak et al. Res. J. of Eng. And Technol., December 2017, volume 4, issue 12, pages 950-953
- [4] S. Reddy. H.M. Chithra, "Soil Classification and Crop Suggestion Using Device wisdom," worldwide quarterly of functional trade and Science Technology, vol. 8, no. 7, july 2020, pages 1625-1628
- [5] A. In Proc., Savla et al.'s "Survey of cataloguing algorithm for prepare acquiesce calculation precision in precision agriculture" 2015 Int. Conf. Communication innovation and embedded systems doi:10.1109/ICIECS.2015.7193120, Syst., pp. 1-7
- [6] A. Manjula and G. Narsimha's article, "XCYPF: bendable and Extensible scaffolds for farming collect Yield Prediction," appeared in Proc. 2015 Conf. Intell. doi:10.1109/ISCO.2015.7282311, Syst. what's more, Control, pp. 1-5
- [7] R. Kumar, M.P. P. Singh. "Harvest Determination Strategy to Augment Harvest Yield Rate Using AI Method," by J.P. and Kumar Singh, in Proc. 2015 Int. Conf. Splendid Technol. additionally, Make

- due. For Comput., Commun., Controls, Energy and Mater., pp. 138-145, doi: 10.1109/ICSTM.2015.7225403
- [8] Abimala, T., Sashya, S.F. and Sripriya, K., (2020), 'Soil Classification and Crop Suggestion using Image Processing', International Journal on Computing, pp. 3544.
- [9] Barman, U. and Choudhury, R.D., (2020), 'Soil texture classification using multi class support vector machines. Information Processing in Agriculture, 7(2), pp.318-332.
- [10] Bhattacharya, Biswanath, and Dimitri P. Solomatine., (2006), 'Machine learning in soil classification', IEEE International Joint Conference on Neural Networks, pp.186-195.
- [11] Devi, M. P. K., Anthiyur, U., and Shenbagavadivu, M. S., (2016). 'Enhanced Crop Yield Prediction and Soil Data Analysis Using Data Mining'. International Journal of Modern Computer Science, 4(6).
- [12] Rahman, S.A.Z., Mitra, K.C. and Islam, S.M., (2018), 'Soil classification using machine learning methods and crop suggestions based on soil series. 21st IEEE International Conference of Computer and Information Technology, pp. 1-4.
- [13] Ramesh, D., Vardhan, B. V., (2013), 'Data mining techniques and applications to agricultural yield data'. International journal of advanced research in computer and communication engineering, 2(9), pp.3477-3480.
- [14] Saranya, N., Mythili, A., (2020), 'Classification of Soil and Crop Suggestion Using Machine Learning Techniques'. International Journal of Engineering Research & Technology, pp.671-673.
- [15] Sruthika, Dr.S.Padmavathi, (2016), 'Performance of SVM Classifier for Image Based Soil Classification'. International conference on Signal Processing, Communication, Power and Embedded System, pp.411-414.
- [16] Suresh, G., (2021), 'Efficient Crop Yield Recommendation System Using Machine Learning for Digital Farming'. International Journal of Modern Agriculture, 10(1), pp. 906-914.
- [17] Zhang, X., Younan, N.H., and O'Hara, C.G., (2005), 'Wavelet domain statistical hyperspectral soil texture classification'. IEEE Transactions on geoscience and remote sensing, 43(3), pp.615-618.