METHODOLOGY

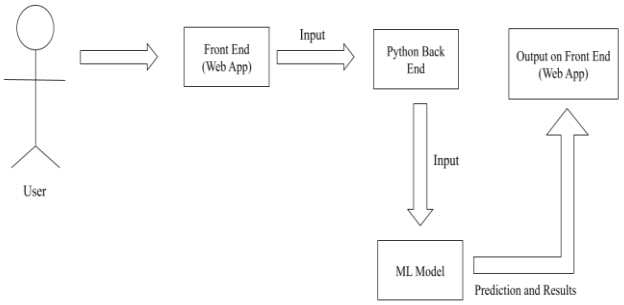

Figure 1: Architecture Diagram

We propose a system with a simple, cost-effective, user-friendly User Interface that is also time efficient. Our proposed approach assists farmers and users in achieving their objectives. This method recommends crops and fertilizers while also forecasting plant problems. In this proposed system, we will collect factors such as nitrogen, phosphorus, potassium, and others and recommend crops or fertilizers using methods such as Decision Tree, Random Forest, Naive Bayes, Support Vector Machine, and Logistic Regression, which will aid in accurate prediction. In addition, in this proposed system, we will take a picture of the plant then the algorithm will predict the disease using the ResNet algorithm. As a result, this approach will make farming easier while also increasing customer satisfaction.

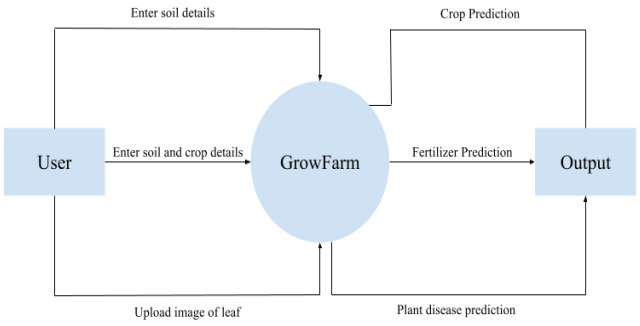

Figure 2: Data Flow Diagram

Step 1:

Loading the dataset, the data which was collected from the study Crop will be used in the system to optimize Crop Production Using Machine Learning Algorithms, which included a crop recommendation dataset. The accuracy of a machine learning model is determined by the quality of the data.

Step 2:

Pre-processing of the input dataset, the most essential or time-consuming task in any machine learning project is data pre-processing. During the pre-processing step, missing values are filled using techniques such as mean, mode, and median, scaling or transforming values in a certain range, cleaning the data, encoding categorical data, and checking for variable correlation so that the accuracy can be increased.

Step 3:

Analyzing exploratory data, before getting hands dirty with model construction, univariate, bivariate, and multivariate analysis are carried out to uncover hidden patterns in the data and to try to interpret the data. A few examples of the univariate analytic plot are PDF and CDF while multivariate plots are pair plot, box plot, and Heatmaps.

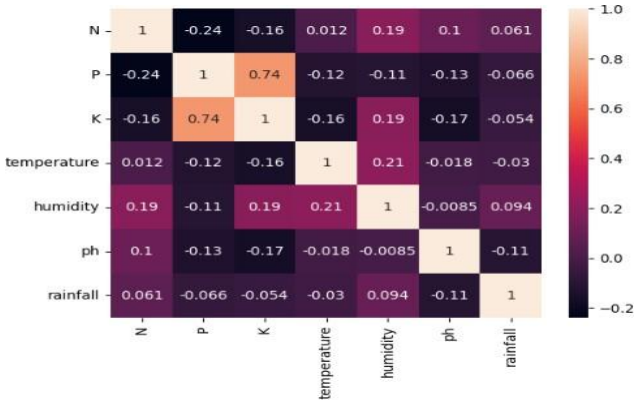

Figure 3: Correlation Metrics on the croprecommendation dataset

Step 4:

Splitting data into training and testing in this step, the pre-processed dataset is divided into training and testing groups based on 80:20 ratios, which indicates that 80% of the data is used for training and 20% for testing on the unseen dataset and cross-validation to discover the best hyper parameter.

Step 5:

Creating a classification model based on the training data, the training dataset is delivered to the individual classifier in this stage, and the model is trained on top of it.

Decision Tree:

A decision tree is a supervised machine-learning technique that can be applied to both classification and regression. A decision tree has a structure similar to a flowchart, with attributes and class labelsdisplayed by a tree.

Logistic Regression:

Logistic regression is any other efficient supervised ML set of guidelines used for binary categorization problems (while the goal is categorical). Logistic regression uses a

logistic characteristic mentioned below to model a binary output variable. The primary distinction between linear regression and logistic regression is that logistic regression has a range of zero to one. Furthermore, logistic regression, as opposed to linear regression, no longer requires a linear connection between input and output variables. This is due to the employment of a non-linear log transformation on the chance ratio.

Random Forest:

Random Forest is a machine-learning system that is built on ensembles. Ensemble approaches are a type of method that allows us to mix independent or similar algorithms to construct a powerful model. A random forest is a collection of multiple decision trees that have the highest depth until the nodes can separate with the least variability and bias.

Gaussian Naive Bayes:

Gaussian Naive Bayes is a simple and straightforward machine learning technique.According to the Naive Bayes hypothesis,qualities must be independent of one another. Internally, the Bayes theorem is applied in Binary Classification, which is a statistical based technique. If the characteristics of the dataset have a Gaussian distribution, the model is known as Gaussian Naive Bayes.

Support Vector Machine:

SVM stands for Support Vector Machine algorithm which is a machine learning technique. First, it plots each data element inN Dimension space and then selects the hyper-plane that best segregates the two classes with the greatest margin in the Linear Kernel. It is tough to find the optimal hyper-plane in a Support vector machine.

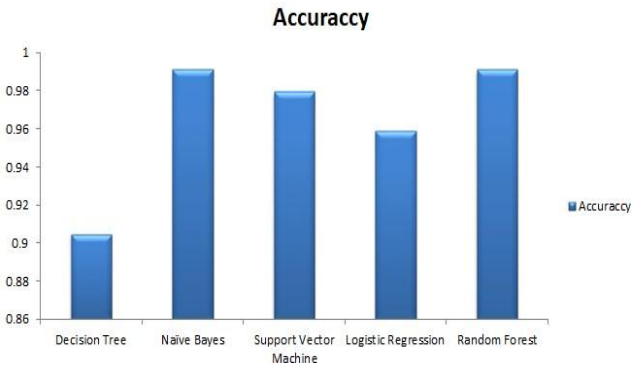

Figure 4: Accuracy comparison chart of differentalgorithms used in the classification model

ResNet :

This network's topology is designed to allow vast amounts of fully connected layers to feature efficiently. Yet, adding a number of deep layers to a network frequently causes output degradation. This is known as the vanishing gradient problem, in which neural networks, while learning via reduced back propagation, rely on gradient descent, descending the loss feature to identify the minimizing weights. Because of the presence of several layers, the repeated multiplication effects inside the gradient get less and smaller, "vanishing," leading to saturation within the community's overall performance or maybe worsening the overall performance. Theprimary principle of ResNet is the use of leaping connections, also known as shortcut connections or identification connections. These connections are typically formed by jumping over one or two layers, forming shortcuts between those layers. The goal ofestablishing those shortcut connections was totackle the major problem of vanishing gradient encountered by deep networks. These identification mappings, for starters, no longer perform anything more than bypass the connections, resulting in the employment of previous layer activation.

Future Work

To improve results and support, the systemcan be expanded further by adding the following functionality:

The focus of future work will be on upgrading datasets on a regular basis to create reliable forecasts, and the process can be automated without modifying the dataset manually. Linking the system to physical devices to transform it into an IOT device that can check soil components without human intervention and select crops to grow depending on the results. Providing users with real-time crop market rates. Multilingual communication is possible, allowing people from all over the world to use this system.

CONCLUSION

This system developed with machine learning known as Crop Recommendation or Prediction system will assist in recommending the best crop to grow inland, as well as which fertilizer to use, and provide plant disease detection based on images, which will be easily available and used by users in order to make a decision on which crop to grow based on the soil nutritional values and climate in that region. The model proposed in the research can be expanded in the future to include crop and fertilizer recommendations as well as plant disease detection in a mobile app. Consequently, our website will assist farmers in sowing the appropriate seed based on soil requirements and increasing their cropyields in order to boost production and profit in their operations (if any) from suchtechniques.

REFERENCES

1. S. Pudumalar, E. Ramanujam, R. H. Rajashree, C. Kavya, T. Kiruthika and J. Nisha, "Crop recommendation system for precisionagriculture," 2016 Eighth International

Conference on Advanced Computing (ICoAC), 2017, pp. 32-36, doi:10.1109/ICoAC.2017.7951740. IEEE

2. S. M. PANDE, P. K. RAMESH, A. ANMOL, B. R. AISHWARYA, K. ROHILLA and K. SHAURYA, "Crop Recommender System Using Machine Learning Approach," 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), 2021, pp.1066-1071, doi: 10.1109/ICCMC51019.2021.9418351. IEEE

3. S. Bangaru Kamatchi, R. Parvathi, Improvement of Crop Production Using Recommender System by Weather Forecasts, Procedia Computer Science, Volume 165, 2019, Pages 724- 732, ISSN 1877-0509. ScienceDirect

4. N. H. Kulkarni, G. N. Srinivasan, B. M. Sagar and N. K. Cauvery, "Improving Crop Productivity Through A Crop Recommendation System Using Ensembling Technique," 2018 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions (CSITSS), 2018, pp. 114-119, doi: 10.1109/CSITSS.2018.8768790. IEEE

5. Akulwar, Pooja. (2020). A RecommendedSystem for Crop Disease Detection and Yield Prediction Using Machine Learning Approach. 10.1002/9781119711582.ch8. ResearchGate

6. S. Vaishnavi., M. Shobana., R. Sabitha. and S. Karthik., "Agricultural Crop Recommendations based on Productivity and Season," 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), 2021, pp. 883-886, doi: 10.1109/ICACCS51430.2021.9441736. IEEE

7. Dr.. G.Suresh, Dr. A.Senthil Kumar, Dr.S.Lekashri, Dr. R.Manikandan, “Efficient Crop Yield Recommendation System Using Machine Learning For Digital Farming”, IJMA, vol. 10, no. 1, pp. 906 - 914, Mar. 2021.

8. Sameer Panat , Acropolis Institute Of Technology And Research , Indore; Siddhant Tripathi, Acropolis Institute Of Technology And Research , Indore; Sarthak Bhaiji, Acropolis Institute Of Technology And Research, Indore; Vandana Kate, Acropolis Institute Of Technology And Research , Indore; Kavita Namdev, Acropolis Institute Of Technology And Research , Indore. IJSRD.

9. S. S. Kale and P. S. Patil, "A Machine LearningApproach to Predict Crop Yield and Success Rate," 2019 IEEE Pune Section International Conference (PuneCon), Pune, India, 2019, pp. 1-5, doi: 10.1109/PuneCon46936.2019.9105741.

10. R. Kumar, M. P. Singh, P. Kumar and J. P. Singh, "Crop Selection Method to maximize crop yield rate using machine learning technique," 2015 International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM), Avadi, India, 2015, pp. 138-145, doi: 10.1109/ICSTM.2015.7225403.

11. V. Kumar, H. Arora, Harsh and J. Sisodia, "ResNet-based approach for Detection and Classification of Plant Leaf Diseases," 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2020, pp. 495-502, doi: 10.1109/ICESC48915.2020.9155585.

12. Ramkumar, M., Catharin, S. S.,Ramachandran, V., & Sakthikumar, A. (2021). Cercospora Identification in Spinach Leaves Through Resnet-50 Based Image Processing. Journal of Physics, 1717(1), 012046, doi: 10.1088/1742-6596/1717/1/012046.