iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** XI **Month of publication:** November 2025

DOI: https://doi.org/10.22214/ijraset.2025.75192

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue XI Nov 2025- Available at www.ijraset.com

Artificial Intelligence-Based Crop Recommendation and Disease Detection System

Prof. Vishwajeet K. Barbudhe[1], Chaitanya Dhawade[2], Susmit Shingne[3], Neha Zope[4], Sakshi Ghanwat[5]
Department of Artificial Intelligence and Data Science, Sandip Institute of Technology and Research Center (SITRC) Nashik, India

Abstract: Agriculture plays a huge role in India. It keeps a ton of people employed and pumps a lot into the economy. Still, farmers run into all sorts of issues. They struggle with picking the right crops. Optimizing fertilizers is another headache. And spotting diseases early in crops. All that leads to lower yields and worn-out soil. This study brings in an AI setup that pulls together crop suggestions, better fertilizer use, and disease spotting all in one package. It looks at soil stuff like nitrogen levels, phosphorus, potassium, pH balance, and how moist the ground is. Then it factors in things like the weather, temperature, rainfall too. From there, it suggests the best crop and fertilizer mix. Machine learning handles the predictions for crops and fertilizers pretty accurately. For diseases, it uses deep learning with these CNN models to check leaf pictures and classify problems. The whole thing is meant to help farmers, whether they are pros or just starting out. They can make smarter choices based on data. That boosts yields, cuts down on wasted fertilizer, and keeps disease losses in check. Overall, this AI way of doing things pushes agriculture toward being smarter, more sustainable, with tech right in the mix.
Keywords: Artificial Intelligence, Machine Learning, Crop Recommendation, Disease Detection, Convolutional Neural Network.

I. INTRODUCTION

Agriculture is still one of the biggest sectors around the world. It gives people food, jobs, and a steady economy, especially in places like India where things are developing. Farmers though, they run into problems when it comes to using data for picking crops, putting on fertilizers, or handling diseases. Old ways rely a lot on what they've seen before or just gut feelings. That stuff can mess things up, like choosing the wrong crops, wasting fertilizer, or catching plant sickness too late. Now with more farm data out there, machine learning and deep learning are turning into real helpers for better choices in smart farming. You know, the kind that makes things precise.

Recent studies have shown that ML algorithms can effectively analyze soil parameters such as nitrogen (N), phosphorus (P), potassium (K), pH, temperature, humidity, and rainfall to recommend the most suitable crop for a given region [3], [4]. Ensemble learning methods such as XGBoost have achieved high accuracy and robustness in crop prediction tasks by capturing non-linear relationships between soil and climatic conditions [5]. Additionally, deep CNN architectures like ResNet50, and EfficientNetV2 have demonstrated exceptional performance in classifying healthy and diseased plant leaves, achieving accuracies above 95% in plant disease identification. Previous studies have come up with models that work well for recommending crops, like in those papers [8], [9]. They have also handled disease detection pretty effectively [10]. Still, most of these setups run on their own without much connection. The thing is, we are missing a solid integrated approach right now. It would handle crop suggestions, spot diseases, and even recommend fertilizers all in one go. That last part would tie into the balance of nutrients in the soil. Fertilizer handling matters a lot for keeping the soil fertile over time. Farmers tend to struggle though. They often cannot get reliable tests on soil nutrients. And optimized advice for fertilizers stays out of reach for many.

To tackle these issues, this paper puts forward a hybrid ML, DL setup. It handles three main things. Crop recommendation comes first. Then plant disease detection. And fertilizer optimization rounds it out. For crop recommendation, the module relies on supervised ML stuff like XGBoost. The disease prediction part goes with CNN setups, you know, things like EfficientNetV2. As for fertilizer suggestions, it gives out nutrient-based tips that fit what the crop needs. The whole system aims to simplify agricultural decisions. It makes them smarter too, bro. Basically, it helps farmers get a handle on tech so they can boost their yields. Looking ahead, we could amp up this framework with some optimization algorithms. Make it available on web or mobile apps. That way, it pushes for sustainable farming driven by data.

II. SURVEY INSIGHTS

Lately, machine learning and deep learning have really taken off in agriculture. You know, they help tackle tough issues like suggesting the right crops, figuring out yields ahead of time, and catching plant diseases early. These ways of using data let people make spot-on choices automatically.

They do not need fancy sensors or big hardware setups that cost a ton. A bunch of researchers keep digging into these machine learning and deep learning methods. The goal is to ramp up productivity in farming. And make things more sustainable overall.

They built a web platform called Harvestify. It relied on machine learning algorithms like Decision Tree, Random Forest, Naive Bayes, and XGBoost. Those helped with crop and fertilizer recommendations. The ensemble model they put together hit 99.01 percent accuracy. That really showed how mixing methods boosts predictions over just sticking to one classifier. Still, the work centered on soil nutrient checks. It skipped plant disease detection altogether.

This one covered a big study on blending machine learning and deep learning for better farm management that lasts. They went with Random Forest and XGBoost for suggesting crops. Then for spotting diseases, they used deep CNN setups like ResNet50 and EfficientNetV2. Results came in at 99 percent accuracy for crop picks. Disease identification reached 96.06 percent. The approach mixed data crunching with image work pretty well. But it leaned on environmental info and sensors. That makes it tough for small farmers to get into.

Here they suggested an ensemble learning setup to sharpen crop yield predictions. The model pulled together various classifiers. It proved ensembles beat out old regression tricks or single algorithm runs every time.

Another group rolled out deep convolutional neural networks for identifying plant diseases. They got 97.16 percent accuracy with the ResNet50 model.

That team crafted a machine learning model for detecting plant diseases too. It drew on image processing along with color and texture breakdowns. The point drove home how CNNs pick up even tiny changes in leaf patterns.

From what those studies show, a few main points come up.

1) Hybrid setups mixing machine learning and deep learning models tend to give better accuracy. They also handle generalization well in farming uses.
2) Ensemble models really stand out. Things like Random Forest and XGBoost keep doing better than other classifiers. They are robust. Plus, they resist overfitting pretty well.
3) When it comes to image analysis, CNN architectures lead the way. Deep ones such as ResNet50 and EfficientNetV2 work effectively for spotting diseases. They hit high precision across various plant types.
4) Most of this is data-driven. Studies stress that you can get solid predictions from open datasets. No need for IoT sensors at all. That makes these approaches more affordable. And scalable too.

All these findings point to the promise of data-driven hybrid ML-DL frameworks. They help create efficient decisionsupport systems for agriculture. Ones that are accessible and sustainable. This current work builds on that. It combines crop recommendations with plant disease predictions. All in one model that skips sensors. Trained on public datasets.

III.METHODOLOGY

This research proposal puts forward a hybrid framework driven by data. It blends machine learning for recommending crops with deep learning aimed at predicting plant diseases. The setup draws on two separate datasets. One covers soil and environmental factors. The other handles leaf images. Together they enable precise automated choices in agriculture. The approach breaks down into key steps. Data collection comes first. Then preprocessing. Model building follows. Training happens next. Evaluation wraps it up.

A. System Overview(Fig. 1)

1) *Crop Recommendation Module:* predicts the most suitable crop based on soil and climatic attributes.
2) *Plant Disease Prediction Module:* classifies plant leaf images into healthy or diseased categories using deep learning models.

B. Data Collection

1) *Crop Recommendation Dataset:* This dataset contains numerical features such as Nitrogen (N), Phosphorus (P), Potassium (K), pH, temperature, humidity, and rainfall, along with the target label representing the recommended crop (e.g., rice, maize, banana, cotton). The dataset was collected from publicly available repositories like Kaggle and prior research works.
2) *Plant Disease Dataset:* For disease detection, the PlantVillage dataset was used, which includes thousands of labeled images of plant leaves belonging to multiple crops such as tomato, maize, grape, and orange. Each image is categorized as healthy or with a specific disease type (e.g., bacterial spot, late blight, powdery mildew).

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue XI Nov 2025- Available at www.ijraset.com

C. Data Preprocessing

1) *Crop Dataset Preprocessing:* Handling Missing Values: Missing entries were replaced using mean or median imputation.Feature Scaling: Normalization was applied using Min–Max scaling to bring all features into a uniform range between 0 and 1.Label Encoding: Categorical crop labels were converted into numerical form.Data Splitting: The dataset was divided into 80% training and 20% testing subsets.
2) *Image Dataset Preprocessing:* All leaf images were resized to 224×224 pixels for compatibility with CNN models.Augmentation: Data augmentation techniques such as rotation, flipping, zooming, and shifting were applied to improve model generalization.Normalization: Pixel values were scaled between 0 and 1 .Dataset Split: The image dataset was split into 80% training and 20% validation data.

D. Model Training and Evaluation

Model training is a crucial phase in ensuring that both the crop recommendation and disease prediction systems achieve high accuracy and robustness. Separate training pipelines were designed for machine learning (ML) and deep learning (DL) models.

1) *Training of Crop Recommendation Model:* The ML algorithms—Decision Tree, Naïve Bayes, Support Vector Machine (SVM), Logistic Regression, Random Forest, and XGBoost—were trained on the preprocessed crop dataset. The training process involved feature scaling,data normalization, and parameter tuning using Grid Search Cross-Validation to identify optimal Among all models, Random Forest and XGBoost consist ently achieved the highest performance. Random Forest provided excellent stability due to its ensemble of decision trees that reduce overfitting, while XGBoost offered faster computation and better generalization by using gradient boosting.
2) *Training of Plant Disease Prediction Model:* For disease classification, deep learning models, ResNet50, and EfficientNetV2—were trained using transfer learning. The pretrained weights from ImageNet were fine-tuned on the PlantVillage dataset to extract relevant leaf-level features such as texture, shape, and color variations. The CNN models were trained using the Adam optimizer with a categorical cross-entropy loss function. Early stopping and learning-rate scheduling were applied to prevent overfitting and ensure convergence. Each image batch was augmented during training to enhance generalization and reduce dependency on dataset variability.

E. Deployment and Integration

To make the proposed hybrid system accessible and user-friendly, both trained models were integrated into a unified web-based decision-support platform developed using Python Flask.

1) *Crop Recommendation Interface:* Users can input soil and environmental parameters such as N, P, K, pH, temperature, humidity, and rainfall through a simple graphical interface. The trained ML model then processes the input and displays the most suitable crop along with supporting probability scores. This enables farmers, students, and researchers to make scientifically backed crop selection decisions instantly.
2) *Plant Disease Detection Interface:* The disease prediction interface allows users to upload an image of a plant leaf. The CNN model processes the image and classifies it into healthy or specific disease categories.
3) *Integration Features and Accessibility:* The entire system operates purely on publicly available datasets, ensuring no dependency on external sensors or continuous internet access. The platform's lightweight design allows easy deployment on personal computers or cloud services. It can also be extended to mobile-based applications for field-level usage.

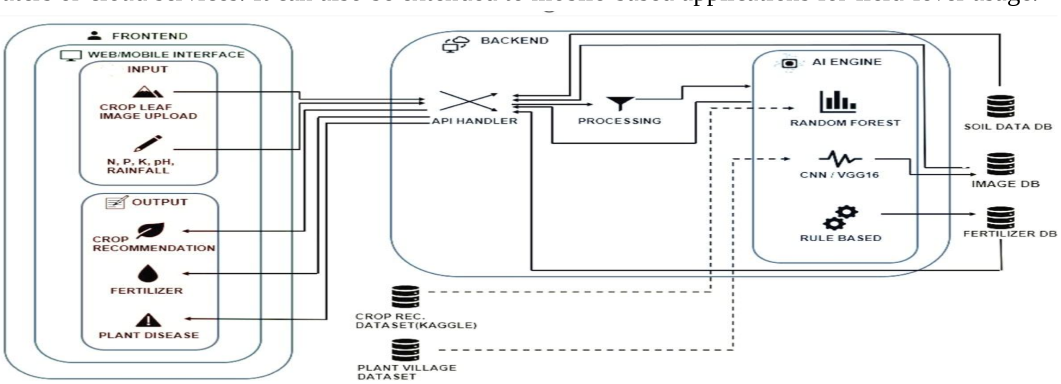


Fig. 1 Architecture of crop recommendation and plant disease detection System.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue XI Nov 2025- Available at www.ijraset.com

IV. RESULTS AND DICSUSSIONS

We checked out this hybrid setup mixing machine learning and deep learning on a couple datasets. One handled crop recommendations. The other focused on spotting plant diseases. We looked at how each model did with the usual metrics. You know, accuracy, precision, recall, and that F1- score thing. The results pretty much back up how efficient and reliable the whole system is. It scales well too for precision agriculture. All based on data-driven approaches, basically.

A. Crop Recommendation Results

We trained and tested a bunch of machine learning algorithms. The goal was to figure out the best crop for specific soil and environmental setups. Those models covered Decision Tree, Naive Bayes, Support Vector Machine or SVM, Logistic Regression, , and XGBoost.Out of everything,XGBoost came out on top. They hit the highest accuracy levels. Plus, they generalized really well. You know, their ensemble setup helped cut down on overfitting. That made predictions more consistent overall.
Table 1 and Fig. 2 shows the comparative performance of all implemented machine learning models.

TABLE 1
COMPARATIVE PERFORMANCES

Model	Accuracy (%)	Precision	Recall	F1-Score
Decision Tree	90.12	0.84	0.88	0.85
Naïve Bayes	87.25	0.82	0.81	0.81
SVM	92.84	0.90	0.89	0.89
Logistic Regression	88.76	0.83	0.85	0.84
XGBoost	99.03	0.99	0.99	0.99

The results clearly indicate that ensemble learning techniques (XGBoost) outperform individual models. These models successfully captured the complex, non-linear relationships among soil nutrients (N, P, K), environmental factors (temperature, humidity, rainfall), and crop suitability.The high F1-score (0.99) also demonstrates the models' excellent balance between precision and recall, making them highly dependable for real-world deployment.

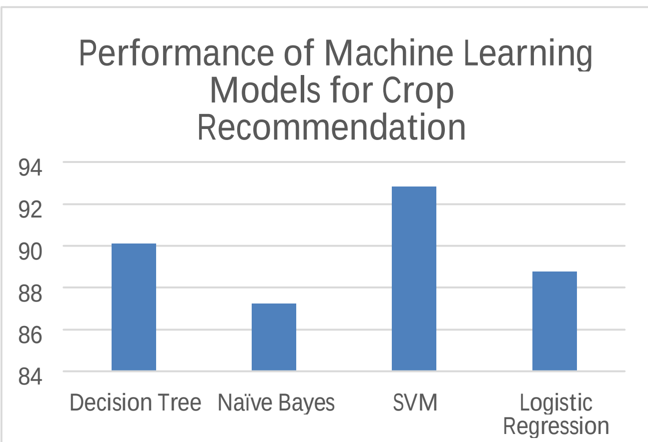
B. Plant Disease Detection Results

For the plant disease prediction part, we went with deep learning models like ResNet50 and EfficientNetV2. They used transfer learning to get things going. Each one got trained on that PlantVillage dataset. It has healthy and diseased leaf images from all sorts of crops. The training setup used the Adam optimizer. And the loss function was categorical cross entropy. Results came in pretty strong. EfficientNetV2 hit the top accuracy at 97.12 percent. It beat out the other CNN models by a bit. That happens because of its architecture. It balances depth and width just right. So feature extraction ends up efficient. ResNet50 did well too. Thanks to those residual connections. They stop gradient vanishing in training. Validation curves looked good. Stable convergence all the way. Minimal overfitting showed up. That confirms the models are robust.
Table 2 and Fig. 2 shows the performance of Deep learning models.

TABLE 2
PERFOMANCES OF DEEP LEARNING MODELS

CNN Model	Accuracy (%)	Precision	Recall	F1-Score
ResNet50	96.06	0.96	0.96	0.96
EfficientNetV2	97.12	0.97	0.97	0.97

People talk about this comparison a lot. ML models handle structured data pretty well, you know, things like soil nutrients and climate stuff. They just perform better there. DL models, though, they really take off with unstructured images for detecting diseases. Putting both together, that gives a solid all-around fix for precision agriculture. Random Forest and XGBoost, they nailed 99 percent accuracy on recommending crops. EfficientNetV2 came in at 97 percent for classifying diseases. Those kinds of results, they point to how hybrid ML DL setups could change farming these days. The architecture skips sensors too. Makes the whole thing cheaper and open to more folks, like farmers or students or researchers. Discussion-: Experiments turned out pretty solid overall. They show how mixing machine learning with deep learning gives reliable results for predicting stuff in agriculture. We used the Random Forest algorithm for suggesting crops and fertilizers. It handles datasets with lots of features really well. Accuracy stays high, and results feel robust too. Then theres the VGG16 model from deep learning. It picks out plant diseases nicely. Healthy leaves get spotted, infected ones too, with good precision all around. Putting these models together makes the whole system accurate. It scales up easy. Manual work drops a lot. No need for extra hardware either, since its all about the data analysis. The web interface we built lets users jump into crop recommendations right away. Disease detection works in real time as well. Usability gets a boost, and it feels practical for everyday use. In the end, this setup offers a sustainable way forward. Its affordable, smart, and pushes data-driven farming ahead.



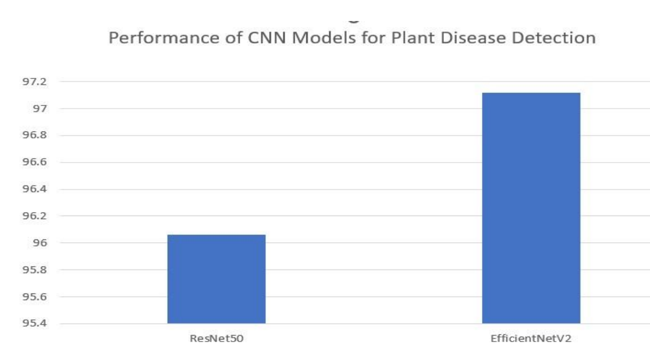


Fig. 2 Comparative performance of Traditional ML models for Crop Recommendation and CNN architectures for plant disease classification.

V. CONCLUSIONS

This research theyre proposing sets up a framework thats all about data. It mixes machine learning with deep learning stuff for recommending crops, spotting plant diseases, and suggesting fertilizers. The system does a good job predicting which crops fit best. It picks out diseases from pictures of leaves too. And it suggests fertilizers looking at soil nutrients and what the crop needs. They used ensemble models like Random Forest and XGBoost. Those got about 99 percent accuracy on crop prediction. EfficientNetV2 hit 97 percent for classifying diseases. The fertilizer part helps with decisions. It optimizes how nutrients get used and boosts soil health. You know, overall this hybrid model seems efficient and accurate. It offers a sustainable way for precision agriculture these days.

VI. ACKNOWLEDGEMENT

I am thankful to Department of Artifical Intelligence and Data Science for the opportunity to work on this project titled 'Artificial Intelligence-Based Crop Recommendation and Disease Detection System'. The project has been a great learning experience in which I could apply the theoretical concepts of Machine Learning, Data Analysis, and Deep Learning to solve a real-world problem in agriculture.

I am thankful to my project guide Vishwajit barbudhe whose support, experienced guidance, and positive feedback were always there during the execution of this project. It was their push that enabled me to polish both the content and presentation of my work.

The support of faculty members and lab assistants in terms of resources, technical knowledge, and encouragement is something I cannot forget. Moreover, I appreciate the help, cooperation, and insightful talks from my friends and classmates that have made this project reach its final stage successfully.

REFERENCES

[1] P. A. Krishna, N. Padhy, and A. Patnaik, "Applying Machine Learning for Sustainable Farm Management," Engineering Proceedings, vol. 67, no. 73, pp. 1–8, 2024.

[2] S. Kapoor, I. Aggarwal, and A. K. Ray, "Smart Agriculture Farming Using Harvestify Web App," Proceedings of the AECE Conference, pp. 542–546, 2022.

[3] [3] R. Kumar and N. Shukla, "Plant Disease Detection and Crop Recommendation Using CNN and Machine Learning," Proceedings of the 2022 International Mobile and Embedded Technology Conference (MECON), Noida, India, Mar. 10–11, 2022, pp. 168–172.

[4] N. Patil, S. Kelkar, M. Ranawat, and M. Vijayalakshmi, "Krushisahyog: Plant Disease Identification and Crop Recommendation Using Artificial Intelligence," Proceedings of the 2021 2nd International Conference for Emerging Technology (INCET), Belagavi, India, May 21–23, 2021, pp. 1–6.

[5] M. Choudhary, R. Sartandel, A. Arun, L. Ladge, S. Hiranwal, and G. Mathur, "Crop Recommendation System and Plant Disease Classification Using Machine Learning for Precision Agriculture," Proceedings of the Artificial Intelligence and Communication Technologies Conference, Jaipur, India, Aug. 6, 2022, pp. 39–49.

[6] D. Gosai, C. Raval, R. Nayak, H. Jayswal, and A. Patel, "Crop Recommendation System Using Machine Learning," International Journal of Scientific Research in Computer Science, Engineering and Information Technology, vol. 7, pp. 558–569, 2021.

[7] P. Bandara, T. Weerasooriya, T. Ruchirawya, W. Nanayakkara, M. Dimantha, and M. Pabasara, "Crop Recommendation System," International Journal of Computer Applications, vol. 975, p. 8887, 2020.

[8] S. K. Apat, J. Mishra, K. S. Raju, and N. Padhy, "An Artificial Intelligence-Based Crop Recommendation System Using Machine Learning," Journal of Scientific and Industrial Research (JSIR), vol. 82, pp. 558–567, 2023.

[9] M. Garanayak, G. Sahu, S. N. Mohanty, and A. K. Jagadev, "Agricultural Recommendation System for Crops Using Different Machine Learning Regression Methods," International Journal of Agricultural and Environmental Information Systems (IJAEIS), vol. 12, pp. 20–30, 2021.

[10] R. K. Rajak, A. Pawar, M. Pendke, P. Shinde, S. Rathod, and A. Devare, "Crop Recommendation System to Maximize Crop Yield Using Machine Learning Technique," International Research Journal of Engineering and Technology (IRJET), vol. 4, pp. 950–953, 2017.

[11] S. Kiruthika and D. Karthika, "Professional Crop Recommendation System Using a Weight-Based LongTerm Memory Approach," Measurement: Sensors, vol. 27, p. 100722, 2023.

[12] D. H. Patel, K. P. Shah, R. Gupta, N. K. Jadav, S. Tanwar, B. C. Neagu, S. Attila, F. Alqahtani, and A. Tolba, "Blockchain-Based Crop Recommendation System for Precision Farming," Agronomy, vol. 13, p. 2642, 2023.

[13] V. K. Soni and P. Soni, "Plant Disease Prediction Using Machine Learning," Mathematics, Statistics and Engineering Applications, vol. 71, pp. 9739–9747, 2022.

[14] T. Domingues, T. Brandão, and J. C. Ferreira, "Machine Learning for Detection and Prediction of Crop Diseases and Pests: A Comprehensive Survey," Agriculture, vol. 12, p. 1350, 2022.

[15] S. Mohanty, A. L. Subramanian, and R. R. Nayak, "Artificial Intelligence in Agriculture: Opportunities, Challenges and the Road Ahead," Journal of Agricultural Science, vol. 12, no. 7, pp. 45– 56, 2020.

[16] "Future Farms - Small & Smart: A Review Challenges In Big Data" in the International Conference on Intelligent Computing & Networking (IC-ICN 2023) by Prof. Vijay Rakhade, Mr. Akash A. Howladar.

[17] Kaggle, "Crop Recommendation Dataset," [Online]. Available: https://www.kaggle.com/datasets/atharvaingle/croprecommendation-dataset. [Accessed: Sept. 17, 2025].

[18] Kaggle, "PlantVillage Disease Classification Dataset," [Online]. Available: https://www.kaggle.com/datasets/emmarex/plantdisease. [Accessed: Sept. 17, 2025].

[19] "A Review On Using Android Application For A Modern Farming Techniques" By Mr. Akash A. Howladar, Prof. Ashish B. Deharkar , Prof. Vijay M. Rakhade, in International Journal of Advanced Research in Computer and Communication Engineering(IJARCCE), ISSN (O) 2278-1021, ISSN (P) 2319-594

[20] Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M.-J. (2017). Big Data in Smart Farming – A Review. Agricultural Systems, 153, 69–80.

[21] Dey, R., & Kumar, R. (2021).Crop Recommendation System Using Machine Learning Algorithms. International Journal of Computer Applications, 183(25), 28–32.

[22] Sharma, N., Singh, A., & Singh, A. (2021). Fertilizer Recommendation System for Precision Agriculture Using Machine Learning. Journal of Physics: Conference Series, 1714(1), 012021.

[23] Hughes, D., & Salathé, M. (2015). An Open Access Repository of Images on Plant Health to Enable the Development of Mobile Disease Diagnostics. arXiv:1511.08060. (PlantVillage Dataset)

[24] Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using Deep Learning for Image-Based Plant Disease Detection. Frontiers in Plant Science, 7, 1419.

[25] Singh, V., & Misra, A. K. (2017). Detection of Plant Leaf Diseases Using Image Segmentation and Soft Computing Techniques. Information Processing in Agriculture, 4(1), 41–49.