AI-Based Agriculture application for crop recommendation and guidance for farmers

Harsh Verma
Department of CSE
ABES Engineering College
Ghaziabad, India
Harsh.20B0101139@abes.ac.in

Amrit Singh
Department of CSE
ABES Engineering College
Ghaziabad, India
amrit.20B0101106@abes.ac.in

Sandhya Avasthi
Computer Science and Engineering
ABES Engineering College
Ghaziabad, India
sandhya_avasthi@yahoo.com

Tanushree Sanwal
KIET School of Management
KIET Group of Institutions Delhi-NCR,
Ghaziabad, India
tanushree.sanwal@kiet.edu

***Abstract*— Agriculture plays a crucial role in safeguarding food security and economic stability, especially in countries where a significant population relies on farming. However, traditional farming practices are often marred by challenges such as suboptimal crop selection, inefficient fertilizer management, disease outbreaks, and market price uncertainties. This paper compares different algorithms like Gaussian Naive Bayes, SVM, Decision tree regression, XGBoost, random forest, logistic regression, etc., and based on the result generated the best algorithm will be used. The proposed smart agriculture system includes applications for crop recommendation, fertilizer guidance, plant disease prediction, and crop price forecasting and will incorporate these applications into a single website which later can also be developed and connected through the servers. This research provides the farmers with new and advanced information regarding their soil and field data. Through binding the power of data-driven decision-making, this project aims to enhance agricultural productivity, encourage sustainability, and empower farming communities. Continuous improvement, user feedback, and interdisciplinary collaboration are integral to the project's success.**

Keywords- Smart agriculture, machine learning, deep learning, recommending crops based on different soil conditions, fertilizer recommendation, predicting plant diseases, predicting the price of crop based on market price.

I. INTRODUCTION

Agriculture plays an important role in the economic development of a country. Agriculture is the only source of living for various families living in villages. Still, agriculture faces various problems like changing weather and soil conditions and it also includes various crop diseases. So, to tackle these problems, we have innovative solutions like we have started using Machine Learning and Deep Learning. This project involves the usage of cutting-edge technologies to overcome the obstacles that come in the way of farming.

Earlier, when there was a practice of traditional agriculture, farmers faced various challenges in the selection of crops, management of fertilizer, detection of various crop diseases, and predicting the price of the market to sell the crops appropriately. So to tackle these problems, this project aims to provide a smart agriculture system to farmers. The main features that this project will include are the detection of crop diseases, offering fertilizer recommendations related to specific crops and soil conditions, and predicting crop prices.

So, this research paper focuses on the research of different modern technologies to allow farmers to make decisions accordingly.

II. RELATED WORKS

In this literature review, we delve into a collection of research papers that explore the dynamic intersection of artificial intelligence (AI) and agriculture, particularly in the context of smart farming. These papers investigate the transformative potential of machine learning, deep learning, and AI-driven applications in addressing agricultural challenges, enhancing crop yield predictions, optimizing farming practices, and providing practical solutions to farmers. Through a concise examination of each paper's key findings and contributions, this review sheds light on the promising advancements and implications of AI technology in modern agriculture.

A research paper from 2022 explores the transformative potential of the machine and deep learning methods in smart farming. It delves into the application of CNNs for image-based crop disease detection, time-series study with RNNs and LSTM networks, Support Vector Machines (SVMs) for classification, then ensemble approaches like Random Forest and Gradient Boosting [1]. Results related to disease detection accuracy and crop yield prediction are presented, highlighting the significant role of AI in optimizing farming practices.

Another paper from 2022 introduces a machine learning-based application, emphasizing practical solutions to agricultural challenges [2]. While not deeply technical, it underscores the importance of data preprocessing techniques and the application of supervised learning algorithms like decision trees, k-nearest neighbors, and neural networks. It likely discusses user feedback and real-world applications of the AI-driven solution, showcasing the practical influence of AI in agriculture.

A paper from 2020 offers a comprehensive exploration of AI adoption in Indian agriculture, emphasizing its transformative power [3]. Although it may not delve into specific algorithms, it discusses AI and machine learning applications such as computer vision, predictive analytics, and recommendation systems. The paper highlights the importance of data collection methods like IoT sensors and remote sensing, and it addresses challenges and opportunities specific to Indian agriculture.

Research from 2019 focuses on the estimation of crop yield and fertilizer recommendations using ML algorithms [4]. It

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

employs regression models, clustering techniques, and decision support systems to optimize agricultural practices. This paper presents results related to yield prediction accuracy and the effectiveness of fertilizer recommendations, showcasing the practicality of machine learning in precision agriculture.

Work from 2021 explores the potential of smart farming techniques integrating machine learning to enhance productivity [5]. While not providing specific metrics, it emphasizes sensor data integration, time-series analysis, and geospatial analysis using GIS for precision agriculture, highlighting the potential for increased efficiency in farming.

A paper from 2022 introduces the Harvestify web app for smart agriculture, discussing its functionalities, user interface, and integration with machine learning models [6]. While specific results may depend on user engagement, it showcases the practical application of AI in providing agricultural insights through a user-friendly interface.

Another paper from 2022 presents a specific implementation of CNNs for crop disease detection and fertilizer recommendations within the Project system [7]. It also discusses the accuracy of disease detection and the effectiveness of fertilizer suggestions, demonstrating the practicality of the proposed system in addressing key agricultural challenges.

A systematic review provides a comprehensive overview of AI technologies in agriculture [8]. It summarizes findings from various research papers, offering insights into the state of the art, trends, and challenges of AI in agriculture, making it valuable for understanding the broader landscape of AI applications in farming.

A paper discusses the notion of climate-smart agriculture and the subject of scientific basis [9]. Although not directly related to AI, it emphasizes the importance of sustainable agricultural practices and offers insights into the role of technological advancements like AI in addressing climate-related challenges in farming, providing context for the broader agricultural landscape and the importance of adopting innovative approaches to ensure agricultural sustainability.

These papers collectively contribute to the increasing knowledge in the field of AI and machine learning in agriculture, emphasizing the potential for increased efficiency, sustainability, and productivity in farming practices. Researchers and practitioners in the field can draw valuable insights and practical applications from these works to inform their efforts in the domain of smart agriculture.

TABLE I. SUMMARY OF THE ABOVE PAPERS

Year	Research Focus	Key Findings
2022 [1]	Application of machine and deep learning methods in smart farming.	- CNNs for image-based crop disease detection. - RNNs or LSTMs for time-series data analysis. - SVMs for classification tasks. - Ensemble methods for predictive modeling. - Enhanced disease detection accuracy. - Improved crop yield prediction. - Optimization of farming practices.
2022 [2]	Development of a machine learning-based application for agricultural problem-solving.	- Data preprocessing techniques. - Supervised learning algorithms for problem-solving. - Possible NLP for text analysis (if applicable). - Practical solutions to agricultural challenges. - User feedback and real-world applications of the application.
2020 [3]	Examination of AI's role in transforming Indian agriculture, with a focus on applications and challenges	- AI and ML characteristics in crop monitoring, yield estimation, and decision support. - Data collection methods (e.g., IoT sensors, remote sensing). - Discussion of challenges and opportunities in Indian agriculture. - Potential for the transformative impact of AI.
2019 [4]	Prediction of crop yield and fertilizer recommendations using machine learning algorithms.	- Regression models for yield prediction. - Clustering techniques for soil data analysis. - Decision support systems for fertilizer recommendations. - Enhanced accuracy in yield predictions. - Effective fertilizer recommendations for optimizing farming practices.
2021 [5]	Exploration of smart farming techniques integrating machine learning for enhanced productivity.	- Integration of sensor data. - Time-series analysis for precise planning. - Geospatial analysis using GIS for spatial optimization. - Potential advantages of data-driven decision-making in agriculture.
2022 [6]	Development and application of the Harvestify web app for smart agriculture.	User-friendly web app interface. - Integration with machine learning models for data-driven decision support. - Real-world application of AI for providing agricultural insights.
2022 [7]	CNN implementation for crop disease detection and fertilizer recommendation.	Presented accuracy of disease detection and the effectiveness of fertilizer suggestions
2022 [8]	Systematic literature review of AI methods in agriculture.	Offers a comprehensive study of AI techniques in agriculture, summarizing findings and trends from existing research papers.
2014 [9]	Discussion of climate-smart agriculture and its scientific basis.	Provides insights into sustainable agricultural practices and the role of technological advancements like AI in addressing climate-related challenges in farming.

III. PROPOSED FRAMEWORK AND METHODOLOGY

The field of smart farming is vast and based on which different data and information are also available from various sources. To find suitable and perfect techniques and algorithms, we decided to implement and compare various algorithms to the more accurate and stable ones finalized. The information for such studied algorithms has also been mentioned.

A. Gaussian Naïve Bayes

This algorithm is based on probability and it classifies things based on probability. It is based on Bayes' theorem, which measures the probability of a data point of a particular class. We will be using this algorithm for crop disease prediction and crop recommendation according to different soils. When a new data point is provided, the algorithm will calculate its probability with different classes. The section with maximum probability is predicted as the likely outcome.

B. Decision Tree Regression

We are using this algorithm for crop recommendation and disease prediction. It works by dividing data into different subsets and data is divided based on useful features. After that, a tree-like structure contains decision rules. Let's take the case of crop recommendation in which trees can make decisions based on soil characteristics and climate factors and can suggest suitable crops. If we take the case of disease prediction, plant leaf attributes play an important role in the identification of potential diseases

C. Support Vector Machine

This algorithm is a supervised learning algorithm that is proficient in classification tasks. In this algorithm, different classes are classified into a single dataset by a hyperplane. In this project, we will be using this algorithm for crop disease prediction and crop classification. The model will be trained on labeled data, because of which it can differentiate between healthy and diseased crops, and also based on input features, it can classify crops.

D. Logistic Regression

This algorithm is used for crop recommendation based on different soil characteristics and is also used for different classification tasks in this project. It is a supervised learning algorithm that is proficient in classification tasks. Based on historical data and soil conditions, this algorithm will calculate the probability of whether a crop is suitable for given soil conditions or not.

Logistic function:

$$P\left(y = \frac{1}{x}\right) = \frac{1}{1 + e^{-z}} \tag{1}$$

Logit function:

$$\log(\text{odds}) = \ln\left(\frac{P\left(y = \frac{1}{x}\right)}{1 - P\left(y = \frac{1}{x}\right)}\right) = w * x + b \tag{2}$$

E. Random Forest

In this project, we will be using this algorithm for different tasks, such as crop recommendation and disease prediction. It depends on numerous decision trees to make estimations and predictions. Each decision tree is being trained on a dissimilar subset of the dataset. Then, this algorithm aggregates the decisions made by individual decision trees, and then the final decision is taken. Its capabilities of managing high-dimensional data and reducing overfitting make it a valuable tool for refining agricultural outcomes and improving resource consumption.

F. XG Boost

This algorithm is used in this project to improve model accuracy and predictive performance. This method joins the results of different decision trees and also refines estimations iteratively with each new tree. It is proficient in different applications, including crop yield estimations, and disease detection, because of its capabilities to manage complex datasets, minimize overfitting, and optimize model training. By using gradient boosting and regularization techniques, it efficiently learns patterns.

IV. IMPLEMENTATION

In the process of comparing different algorithms within a framework, it is imperative to establish a systematic methodology. This involves delineating clear objectives, selecting representative datasets, and executing experiments with well-defined parameters. Rigorous statistical analyses are then employed to ascertain the validity of observed distinctions between algorithms. This structured approach ensures the reliability and validity of comparisons made regarding algorithm performance.

A. Problem Statement

In agriculture, farmer faces numerous challenges related to crop selection and fertilizer management and to address these issues and empower farmers with data-driven solutions, this paper aims to develop a comprehensive smart agriculture system. The system will incorporate machine learning and deep learning techniques to provide crop recommendations based on soil data and offer fertilizer recommendations tailored to specific crops and soil conditions.

B. Data

The basis for the paper's functionalities mainly relies on the gathered data. For the success of the research, custom-built data has been put together with the efforts of different peers, which ultimately helps in providing a powerful understanding of the crops and the used fertilizer in the fields. The collection of the data also involved the cleaning process to ensure its reliability and accuracy. Additionally, other data sources, like weather data, contribute to the system's effectiveness by providing essential information for decision-making. The data contain some specific tuples that would help determine the crop and fertilizers and these tuples are rainfall, pH, N, P, K, temperature, humidity, etc. It is custom-built data and has information from various reliable and accurate sources.

C. Application Interface

The application interface of the concerned paper provides the farmers and data scientists with a user-friendly web page that helps them to communicate with the updated and labeled data from the core functionalities of the system. The required framework has been developed in such a way that it asks the user about every detail related to their field, soil, types of crops they want to plant, and information related to fertilizer. All the needed information has been carefully divided into different web pages which are later linked together to give a better approach to understanding the system. Based on the input provided by the user the website tries to provide the best possible result according to the gathered data.

D. *Project Architecture*

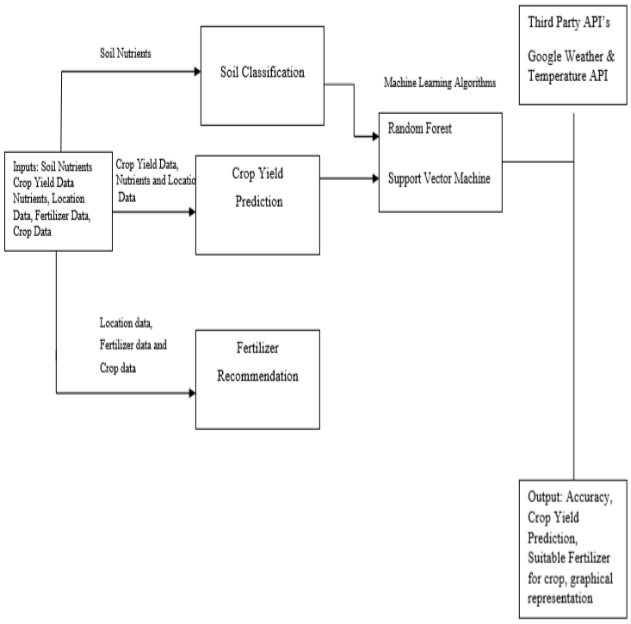


Fig. 1. Architecture of Application

E. *Modelling*

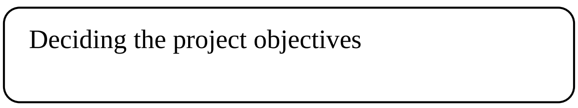
The backbone of the research lies in the modeling and structuring; of the core components of the researched subjects. It provides its insights based on users' mentioned data as the input and based on the data calculated it gives the result. Machine learning and Deep learning are the key components of the modeling process as they help in providing insightful recommendations regarding crops and fertilizer to the user. These models are continuously improved and refined to enhance their accuracy and reliability. These technologies and methods help drive the system by analyzing, processing, and interpreting the complex nature of agricultural data and provide a better understanding to the user.

By incorporating these techniques in the research the paper aims to provide farmers with new technologies for improving their agricultural methods through its data-driven methods. It is also important to note that the acquired data is not perfect and is only a papered speculation. Nevertheless, it demonstrates the immense potential of leveraging Machine Learning and Deep Learning in agriculture, provided that large-scale implementation is supported by authentic and verified data sources.

V. EXPERIMENTAL SETUP

In this study, we have provided the step-by-step process and data flow of every upcoming event. The flow of the model starts with the problem statement and objective while moving forward to acquiring the data, building, and training with a user application.

A. *Step to Create Model*



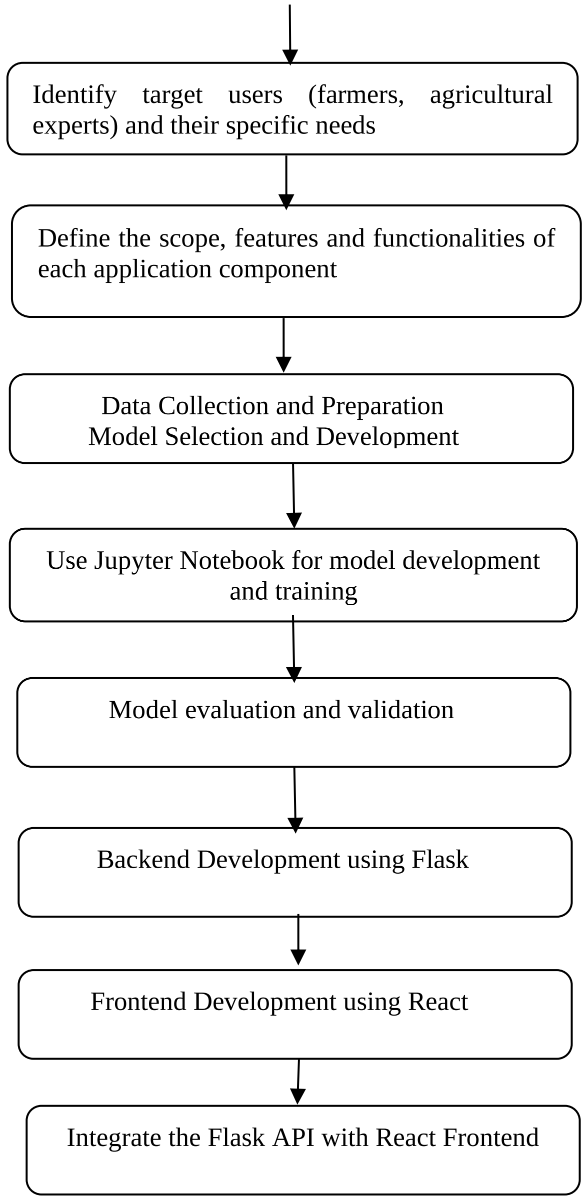


Fig. 2. Flow diagram for creating the model

B. *Tools and Technologies Used*

Flask, a web framework in Python, serves as the backbone for creating web applications. It offers a minimalistic and flexible structure that helps developers to define and express their routes and logic. When a Flask application is run, it launches a server that listens for incoming HTTP requests, handling them with the specified Python code. In contrast, Python is a versatile and high-level programming language known for its clean syntax and dynamic semantics. Its widespread use extends beyond web development to encompass fields such as data and business analysis, artificial intelligence, machine and deep learning, and scientific computing.

HTML is an important part of any development process as it uses and incorporates its various tools and tags to give structure to the content which helps us to build proper schema for any project. While HTML helps in providing the structure to the content CSS helps in beautifying it, It helps by providing various new options for applying color and textual formats so that the information and content look more appealing and easy to understand. CSS also helps in providing various types of responsive webpages through which they can concert themselves according to any screen size.

Anaconda, a renowned open-source distribution, brings together Python and R, two powerful languages for scientific computing, data science, and machine learning. It simplifies the setup by offering a comprehensive suite of pre-installed libraries and tools, making it a go-to choice for professionals in these fields. One of the standout tools within Anaconda is Jupyter Notebook, a versatile platform for data virtualization, management, analysis, and training. It provides various programming languages (like Python and R), making it an ideal environment for building machine learning models, exploring datasets, and evaluating model performance. It has a user-friendly interface which also helps in data preprocessing and cleaning which is a vital step in any data analysis or ML projects.

Collectively, these technologies provide a robust ecosystem for developers and data scientists to create web applications, conduct data-driven research, and develop machine learning models efficiently. They empower individuals and teams to utilize the full potential of the web and data for various applications, from building websites and web applications to solving complex data-related challenges.

VI. RESULT

For the crop recommendation model, a bar graph is plotted between various given algorithms, and the results are noted down. By comparing and calculating various factors among the algorithms their accuracies come as follows:

TABLE II. ACCURACIES OF DIFFERENT ALGORITHMS

S. No	Algorithms	Precision/Accuracies
1.	Decision Tree	0.9113
2.	Naïve Bayes	0.9909
3.	SVM	0.9818
4.	Logistic Regression(LR)	0.9568
5.	Random Forest	0.9909
6.	XGBoost	0.9886

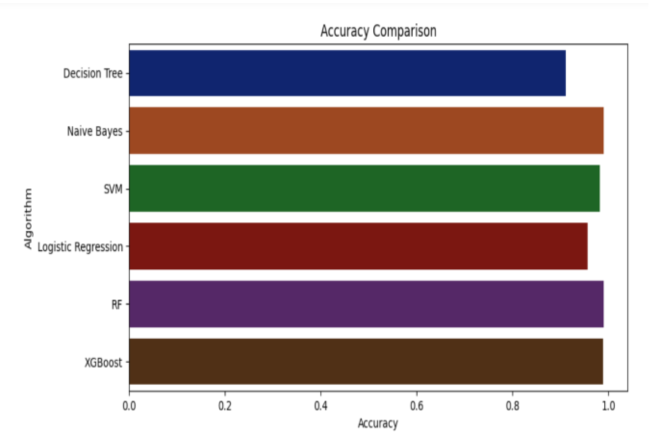


Fig. 3. Accuracy Comparison

It is visible that Random Forest and Naïve Bayes have the greatest accuracy with 0.9909 followed by XGBoost with 0.9986 and then the rest of the algorithms. The cross-validation score is also calculated and based on the results Random Forest is selected as the best algorithm for crop recommendation and as per it the prediction is made. The flow diagram of the crop recommendation model is shown in Fig. 4.

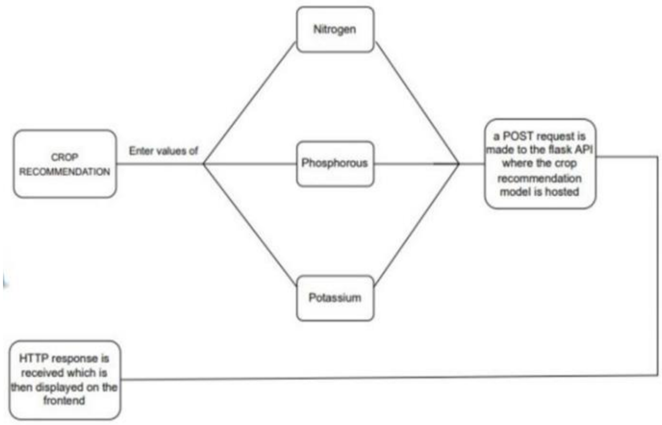


Fig. 4. Flow Diagram of Crop Recommendation Model

For the Fertilizer recommendation model, we refined the raw and abundant data into more manageable and processed data. With the help of mergeable data which is made from crop and fertilizer csv files some of the new features are introduced like NPK (Nitrogen, Phosphorus, Potassium), temperature, and pH values. Based on the values entered in the system the required fertilizer and its quantity will appear. The flow diagram of fertilizer recommendation is shown in Fig. 5.

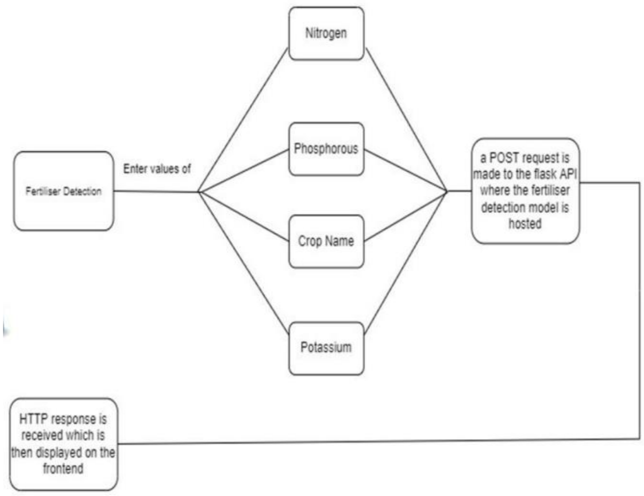


Fig. 5. Flow Diagram of Fertilizer Recommendation Model

VII. CONCLUSION

This agriculture-focused machine learning web application represents a pivotal advancement in the farming industry, providing an intelligent and intuitive platform for farmers to optimize their agricultural practices. By harnessing the capabilities of various ML techniques such as Decision Trees, XGBoost, SVM, Naive Bayes, and Random Forest, this tool revolutionizes farming methods by offering tailored recommendations across crucial aspects.

Firstly, it assists in crop selection, enabling farmers and users to make appropriate decisions about which crops to cultivate based on factors like soil type, weather patterns, historical data, and market demand. These methods help in increasing the yield, and profit margin while simultaneously decreasing the risk of any crop failure. Secondly, the research provides important and valuable insight into the use of fertilizers and improving nutrient usage based on the soil's requirements and its composition. This exactitude decreases the wastage of resources as well as environmental impact yet ensuring the ideal and best crop growth and health. Overall these wide-ranging tools incorporate most of the cutting-edge technologies and functionalities with agricultural expertise which inspires farmers and people with its data-driven insights. Its versatile approach not only helps in increasing productivity and profitability but also encourages the use of sustainable farming techniques while contributing to be more resilient and efficient agricultural ecosystem approach.

VIII. REFERENCES

[1] Durai, S. K. S., & Shamili, M. D. (2022). Smart farming using machine learning and deep learning techniques. Decision Analytics Journal, 3, 100041.

[2] Gupta, S., Chopade, A., Jain, N., & Bhonde, A. (2022). Farmer's Assistant: A Machine Learning Based Application for Agricultural Solutions. arXiv preprint arXiv:2204.11340.

[3] Kumar, T., & Prakash, N. (2020). Adoption of AI in agriculture: the game-changer for Indian farmers. In Proceedings of the 13th IADIS International Conference ICT, Society and Human Beings 2020, ICT 2020 and Proceedings of the 6th IADIS International Conference Connected Smart Cities 2020, CSC 2020 and Proceedings of the 17th IADIS International Conference Web Based Communities and Social Media 2020, WBC 2020-Part of the 14th Multi Conference on Computer Science and Information Systems, MCCSIS 2020 (pp. 204-208).

[4] Bondre, D. A., & Mahagaonkar, S. (2019). Prediction of crop yield and fertilizer recommendation using machine learning algorithms. International Journal of Engineering Applied Sciences and Technology, 4(5), 371-376.

[5] Pawar, S., Dere, S., Akangire, A., Kamble, H., & Shrawne, S. (2021). Smart farming using machine learning. Smart Comput.

[6] Kapoor, S., Aggarwal, I., & Kumar Ray, A. (2022). Smart Agriculture Farming Using Harvestify Web App. Available at SSRN 4157630.

[7] Sachin Adulkar, Vivek Pawar, Aniket Choudhari, Shubham Kothekar, Shruti Agrawal."Harvestify - Crop Disease Detection and Fertilizer Suggestion using CNN", Volume 11, Issue IV, International Journal for Research in Applied Science and Engineering Technology (IJRASET) Page No: 3596-3604, ISSN: 2321-9653, www.ijraset.com

[8] Elbasi, E., Mostafa, N., AlArnaout, Z., Zreikat, A. I., Cina, E., Varghese, G., ... & Zaki, C. (2022). Artificial intelligence technology in the agricultural sector: a systematic literature review. IEEE Access.

[9] Steenwerth, K. L., Hodson, A. K., Bloom, A. J., Carter, M. R., Cattaneo, A., Chartres, C. J., ... & Jackson, L. E. (2014). Climate-smart agriculture global research agenda: scientific basis for action. Agriculture & Food Security, 3(1), 1-39.