INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

REVIEW ARTICLE

Streamlit Application for Advanced Ensemble Learning Methods in Crop Recommendation Systems – A Review and Implementation

Yaganteeswarudu Akkem[1]*, **Biswas Saroj Kumar**[2], **Aruna Varanasi**[3]

1 Ph.D. Scholar, National Institute of Technology, Silchar, Assam, India
2 Associate Professor, National Institute of Technology, Silchar, Assam, India
3 Professor & HOD, Sreenidhi Institute of Science & Technology, Hyderabad, Telangana, India

OPEN ACCESS

Received: 11-11-2023
Accepted: 28-11-2023
Published: 29-12-2023

Citation: Akkem Y, Kumar BS, Varanasi A (2023) Streamlit Application for Advanced Ensemble Learning Methods in Crop Recommendation Systems – A Review and Implementation. Indian Journal of Science and Technology 16(48): 4688-4702. https://doi.org/10.17485/IJST/v16i48.2850

* **Corresponding author**.

yaganteeswarudu21_rs@cse.nits.ac.in

Funding: None

Competing Interests: Non

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Abstract

Objectives: This article explores the integration of advanced ensemble machine learning methods within precision agriculture, aiming to enhance the reliability and practical utility of crop recommendation systems. The incorporation of the Streamlit framework in the development process underpins our objective to deliver a user-friendly tool that provides farmers and agricultural analysts with actionable insights. **Methods:** A thorough literature review of artificial intelligence applications in agriculture serves as the foundation of our study, with a strong emphasis placed on sophisticated ensemble learning techniques such as stacking, an ensemble of ensembles, and federated learning. The evaluation methodology entails a comparative analysis where these cutting-edge techniques are juxtaposed against standard machine learning benchmarks to ascertain their performance improvement. In addition to the conceptual analysis, we implement a crop recommendation system using the Streamlit framework, emphasizing usability and accessibility for end-users to interact with machine learning predictions based on their soil data. **Findings:** The empirical results demonstrate that our chosen advanced ensemble learning methods significantly improve predictive performance, recording up to a 15% accuracy increment over traditional machine learning algorithms. Their adaptability to varied agricultural datasets, coupled with robust privacy-preserving properties, stand out. When deploying these methods in a practical Streamlit-based application, we note a marked increase of 20% in user efficiency, solidifying the system's crucial role in bolstering resilient crop management tactics. **Novelty:** This research pioneers the study of innovative ensemble learning techniques, married with Streamlit app development for an enhanced user experience in data-driven precision agriculture. Our findings emphasize the critical need for incorporating these advanced methodologies into real-world practices, fostering a significant paradigm shift in agricultural data analytics and management. The synergy between these powerful machine learning techniques and the Streamlit-built interactive interface represents a step forward in translating complex computational analysis into practical, on-the-ground tools for agriculture professionals.

Keywords: Machine Learning; Advanced Ensemble Learning; Streamlit.

1 Introduction

Modern agriculture is increasingly reliant on intelligent crop recommendation systems, which enable farmers to optimize their practices based on a multitude of factors such as climate variability, soil characteristics, and water availability. The integration of interactive frameworks like Streamlit into these systems presents a significant advancement from traditional farming methods, as it enables the rapid development and deployment of web applications that can visualize and interact with complex agricultural data[1,2]. Despite these innovations, the effectiveness of crop recommendation systems is often limited by the multifaceted and high-dimensional nature of agricultural data, which poses challenges for effective data handling and interpretation.

Limitations in the current state of crop recommendation systems hinder their ability to manage complex datasets that mirror the dynamic interactions within agricultural environments. Traditional machine learning (ML) techniques often fall short of adequately addressing these complexities, leading to recommendations that may not be sufficiently accurate or adaptable to evolving conditions[3,4]. The burgeoning field of advanced ensemble learning has begun to demonstrate its capacity to overcome these challenges by providing more robust and accurate decision support[5–7]. Nevertheless, despite their proven efficacy in diverse domains, the full potential of such methods integrated with Streamlit for enhanced user experience remains largely unexplored and untapped in the context of precision agriculture.

This research aims to bridge the gap between existing ML practices within agriculture and the cutting edge of advanced ensemble learning, with a specific emphasis on incorporating the Streamlit framework to facilitate user interactions with the system[1,2]. We identify a pressing need for sophisticated analytic techniques that can deeply process agronomic data with greater depth, agility, and security, allowing for real-time insights through interactive web applications. By scrutinizing the current barriers in crop recommendation systems, we examine how a strategic application of advanced ensemble methods drawing on successful implementations in healthcare, environmental science, and beyond integrated with Streamlit can revolutionize agricultural decision-making[8–10].

In this era of data-driven agriculture, ensemble learning techniques comprising strategies such as stacking[8–10], federated learning[11–14], and dynamic ensembles[6] which combine multiple models to forge more accurate and robust prediction systems, have garnered substantial attention. Leveraging the 'wisdom of crowds,' these methods help reduce biases, minimize overfitting, and enhance generalization across diverse datasets and applications[5,7,15,16]. Paired with the streamlined implementation afforded by Streamlit, these advanced computational approaches can be brought directly to the hands of farmers and agricultural experts in an intuitive and actionable manner[1,2].

Our study not only highlights existing inadequacies in crop recommendation systems but also introduces a direct application of diverse ensemble learning frameworks, using the Streamlit platform to address key challenges in precision agricultural data analysis, thereby uncovering new avenues for enhancing crop health prediction, yield optimization, and overall agricultural sustainability. By integrating and adapting these pioneering ML techniques within Streamlit-based applications, this work aspires to set new benchmarks for accuracy and reliability in crop recommendation systems, ultimately contributing to the advancement of smart, sustainable farming.

In the past decade, precision agriculture has seen significant evolution, largely owing to the incorporation of ML [3,4,17–19] for the development of intelligent crop recommendation systems. These systems have been pivotal in processing complex agronomic data to facilitate informed decision-making. While various machine learning-based methods have been employed, there is a crucial need to recognize the constraints of existing ML techniques used in crop recommendation systems when addressing the evolving complexities of this field [20–24]. Furthermore, despite the abundance of literature on ML applications in agriculture, there has been a lack of comprehensive consideration for advanced ensemble learning approaches and the power of Streamlit for handling current challenges in agricultural data analysis [5–7,11,12]. This research is dedicated to defining the limitations of previous reviews, paving the way for a nuanced exploration of advanced ensemble learning methods in the context of crop recommendation systems, envisioning their realization through interactive Streamlit-powered web applications [1,2].

1.1 Related works

Casado FE et al. [11] focuses on enhancing classification performance in crop recommendation by integrating ensemble and continual federated learning, enabling the model to learn from diverse sources and adapt to varying conditions across different farms. Haoran Yu et al. [12] concentrated on developing a federated learning algorithm for non-IID datasets, particularly valuable in crop recommendation for handling diverse and non-uniform agricultural datasets. Alhassan M et al. [13] introduced ensemble federated learning, applicable to building models for crop recommendation and ensuring accuracy and inclusivity by leveraging data from various farms. Bakopoulou E et al. [14] strive to improve mobile packet classification using a federated learning approach, adaptable for agricultural IoT devices, preserving data privacy and allowing models to be trained on diverse sensors across farms. Brisimi TS et al. [25] focused on federated learning applied to electronic health records, similarly ensuring predictive models for crop recommendation can be trained across different farms without centralizing sensitive data, preserving individual farmers' privacy. Hamer J et al. [26] aims to improve communication efficiency in federated learning [27], which is crucial in agriculture, making models for crop recommendation accessible for farmers in various locations. Li Q et al. [28] provide a comprehensive survey of federated learning systems, emphasizing the need for data privacy and protection, which is crucial in decentralized learning environments for crop recommendation. Li T et al. [29] focused on federated optimization in heterogeneous networks, ensuring collaborative models for crop recommendation are efficiently trained with diverse agricultural data. Konečnỳ J et al. [30] investigate strategies to improve communication efficiency in federated learning, which is crucial in agriculture, ensuring models for crop recommendation can be trained with minimal communication overhead. Zhao Y et al. [31] emphasize the importance of federated learning with non-IID data in agriculture, ensuring accurate and adaptable crop recommendation models across different farms. Xiao Z et al. [32] extend federated learning to agriculture for improving crop recommendation accuracy based on distributed sensor data, and Ma X et al. [33] introduced layer-wise model aggregation, which is crucial in crop recommendation for personalized models adapting to individual farm needs.

Ammar M et al.'s [5] comprehensive review on ensemble deep learning guides the development and deployment of ensemble models for crop recommendation. Xu-Cheng Y et al.'s [6] study on DE2 suggests its application in agriculture for adapting crop recommendation models to changing environmental conditions over time. Li Z et al.'s [7] study on sparse ensembles introduces an efficient approach for crop recommendation systems, capturing essential features while minimizing computational resources. Alharbi A et al.'s [15] work focuses on Arabic sentiment analysis, extending ensemble methods to agriculture for improving predictive accuracy in crop recommendation models. Alireza G et al.'s [16] smart healthcare monitoring system based on ensemble deep learning can be adapted for more accurate crop recommendation models. Lu M et al.'s [8] stacking ensemble model introduces a concept applicable to agriculture for combining various machine learning models for improved crop prediction. Berliana AU et al.'s [9] stacking ensemble learning for COVID-19 classification suggests potential application in agriculture, especially in the integration of diverse sensor data for improved crop classification and recommendation. Alireza G et al.'s [10] application of stacking ensemble learners may find relevance in agriculture for crop yield forecasting and disease identification. Towfiqul Islam ARM et al.'s [34] flood susceptibility modeling using advanced ensemble machine learning models contributes to more accurate risk assessments for farmers, aiding in crop planning and decision-making. Boukellouz W et al.'s [35] multi-modal feature extraction and ensemble learning technique may be applied to agricultural data for improved feature representation and more accurate crop recommendation models. Zhou T et al.'s [36] stacking ensemble machine learning algorithms can be extended to agriculture for detecting anomalies or irregularities in farming practices, contributing to improved crop recommendation systems. Hao C et al.'s [37] application of stacking ensemble learning in biomaterial activity analysis can be adapted to agriculture for accurate assessments of the effectiveness of fertilizers, pesticides, and other agricultural inputs. Gupta A et al.'s [38] stacking ensemble-based intelligent machine learning model may inspire principles applicable to agriculture for predicting the impact of environmental conditions on crop health and yield. Mandal U et al.'s [39] LinVec, a stacked ensemble machine learning architecture for time-series data, is particularly relevant to agriculture for understanding crop growth patterns and predicting optimal harvest times. Muthulakshmi P et al.'s[40] prediction of heart disease using ensemble learning suggests the application of ensemble learning techniques to agriculture for improved prediction of crop diseases and more effective crop management practices. The basic idea of designing smart farming web application and providing security can be observed from Yaganteeswarudu A et al.[20,21]. Kalaivani P et al.'s[22] machine learning approach to analyzing ensemble models and neural network models may be adapted to agriculture for analyzing and predicting farmer behavior, contributing to more personalized and effective crop recommendation systems. C. Zhao et al.'s[23] fine-tuning ensemble models for ensemble learning can enhance the accuracy and adaptability of crop recommendation systems. Jiao B et al.'s[24] incremental weighted ensemble for data streams is crucial for crop recommendation models to adapt to changing farming practices, climate conditions, and other factors. Akkem Y et al.'s[17,18] smart farming using artificial intelligence and smart farming monitoring using ML and MLOps provide insights that can inform the development and enhancement of artificial intelligence techniques, including ensemble learning, for smart farming applications. Chakraborty A et al.'s[3] integration of neural networks for pest detection contributes to more accurate monitoring in agriculture, aiding in the development of crop recommendation models by identifying and mitigating potential threats to crop health. Srivani P et al.'s[4] monitoring of environment parameters in Gerbera flower cultivation using IoT contributes to the development of more responsive and adaptive crop recommendation models. Arshad J et al.'s[19] intelligent greenhouse monitoring and control scheme contributes to the optimization of environmental conditions for crop growth, influencing the development of more accurate crop recommendation models.

1.2 Limitations of Existing Studies

Prior reviews in the domain of precision agriculture and intelligent crop recommendation systems have predominantly focused on conventional machine learning methods and do not delve deep into the intricacies and potential of advanced ensemble learning techniques. Existing methods often fall short in encapsulating the recent advancements in ensemble learning techniques, specifically the application of stacking, ensemble of ensembles, and federated ensemble learning. Furthermore, there is a notable absence of a comprehensive synthesis of dynamic ensemble methods in the context of crop recommendations.

Existing reviews largely revolve around traditional machine learning methods. These limitations lie in the scope of models explored and the extent to which these reviews cover the application of novel ensemble methods in precision agriculture. Specifically, they fall short in showcasing applications that harness the power of stacking, an ensemble of ensembles, and federated learning to solve complex, high-dimensional problems faced in agricultural data analytics.

1.3 Problems addressed by the proposed review

This proposed review endeavors to fill the gaps left by previous studies by offering an in-depth exploration of advanced ensemble learning methods in the realm of crop recommendation systems. Specifically, it will provide a detailed analysis of stacking, ensemble of ensembles, federated ensemble learning, and dynamic ensemble methods. The review seeks to present a more nuanced and contemporary perspective, addressing the shortcomings of prior machine learning methods in crop recommendation and contributing valuable insights into the potential applications and benefits of advanced ensemble learning in optimizing crop health, yield, and overall sustainable agriculture.

This review addresses the limitations of existing literature by tackling the challenges associated with the utilization of advanced ensemble learning methods in crop recommendation systems. Specifically, it aims to elucidate the potential applications and benefits of stacking, an ensemble of ensembles, federated ensemble learning, and dynamic ensemble methods. The overarching goal is to improve accuracy, adaptability, and data security in crop recommendations, ultimately fostering optimized crop health, yield, and sustainable agriculture.

This review introduces a paradigmatic shift in the literature by focusing on the application of advanced ensemble learning methods, a relatively under explored area in the context of crop recommendation systems. The review offers a fresh perspective that goes beyond the conventional machine learning methods covered in existing reviews by delving into techniques such as stacking, an ensemble of ensembles, federated ensemble learning, and dynamic ensemble methods. The emphasis on these advanced ensemble techniques distinguishes this review, positioning it as a valuable contribution to shaping the future of data-driven crop recommendations in precision agriculture.

The proposed review sets out to address the limitations and literature gaps of existing studies by focusing on advanced ensemble learning methods and their specific applications to crop recommendation systems. This review will:

- Illustrate how advanced techniques like stacking, ensemble of ensembles, and federated ensemble learning address the complexities of multi-dimensional agricultural datasets.

- Elucidate the problems that these methods can solve, such as improving prediction accuracy, providing robust solutions to diverse crop management issues, and ensuring data privacy.
- Offer insights on the adaptability of these methods to different agricultural environments and the corresponding benefits for crop health management.
- Present a novel perspective by integrating concepts from various fields, including computer science, statistics, and agronomy, to apply advanced ensemble methods in precision agriculture.

2 Proposed Methodology – Ensemble methods and streamlit frameowrk

The essence of ensemble learning lies in building a predictive model by integrating multiple models to improve accuracy. Stacking, ensemble of ensembles, and federated ensemble learning are such methods that have shown promising results in the field of precision agriculture. The adoption of advanced ensemble learning methods, including stacking, ensemble of ensembles, and federated ensemble learning, ushers in a new paradigm within precision agriculture, offering novel solutions to existing challenges faced by traditional predictive models.

2.1 Stacking

Stacking, also known as stacked generalization, is an ensemble learning technique that involves combining the predictions from multiple machine learning models to produce a final prediction. This approach can be especially advantageous in complex domains like agriculture, where predictive accuracy is critical for making informed decisions about crop recommendations. Below are some advantages and limitations of using stacking ensemble methods in the context of crop recommendation:

*2.1.1 **Advantages***

- **Improved Predictive Accuracy:** Stacking often results in higher predictive accuracy compared to individual models because it exploits the strengths of each base learner and mitigates their weaknesses [(5,6)].
- **Model Diversity:** By combining predictions from diverse models, stacking can capture complex patterns that single models may miss, which is beneficial for the varied data seen in agriculture, such as soil properties, weather patterns, and crop types [(8–10)].
- **Bias Reduction:** Different models may have different biases, and stacking helps to average these out, leading to more stable and generalized predictions [(5,7)].
- **Robustness to Overfitting:** The meta-learner in stacking can learn an optimal combination of base model predictions, potentially reducing the risk of overfitting to the training data [(7,15)].
- **Flexibility:** Stacking allows the integration of any type of model, including linear and nonlinear approaches, giving flexibility in selecting the best models for the stacking ensemble and tailoring it to specific agricultural needs [(6,16)].
- **Incorporation of Expert Knowledge:** Experts can incorporate their domain knowledge by handpicking models that are known to perform well on certain types of agricultural data, which can be useful for specific crop conditions or regions [(8,10)].

*2.1.2 **Limitations***

- **Complexity in Implementation:** Setting up a stacking ensemble can be more complex than using individual models, involving careful selection and tuning of base learners and the meta-learner, which can be time-consuming and requires expertise [(8,16)].
- **Risk of Overfitting Meta-Learner:** Although stacking can be robust against overfitting, if the meta-learner is overly complex or the base model predictions are highly correlated, it might overfit to the training data [(7,15)].
- **Increased Computational Cost:** The need to train multiple models and a meta-learner can lead to increased computational costs and longer training times, which may be impractical for some real-time or resource-limited agricultural settings [(15,16)].
- **Model Interpretability:** Stacking ensembles can be less interpretable than simpler models because the final prediction is the result of a combination of several models' decisions, which can be a drawback in agricultural contexts where understanding the reasoning behind recommendations is important [(8,9)].
- **Data Requirements:** A sufficient amount of diverse and representative data is necessary to train multiple models effectively and to validate the stacking approach, which might be challenging in some agricultural domains where data can be scarce or costly to obtain [(5,10)].

- **Tuning Complexity:** The performance of a stacking ensemble is highly influenced by the choice and tuning of base learners and meta-learners; improper selection can lead to suboptimal results, requiring extensive experimentation [6,7].

Despite these limitations, stacking is a powerful tool in the machine learning ensemble techniques toolbox, and it represents a move towards advanced, more robust crop recommendation systems.

2.2 Ensemble of ensembles

The "ensemble of ensembles" approach in machine learning aggregates the predictions from multiple ensemble models. Each ensemble may itself combine several machine learning models, which can lead to a highly robust predictive system. This approach is particularly relevant for complex domains like agriculture, where data may have high dimensionality and various interacting factors need consideration for accurate crop recommendations.

2.2.1 ***Advantages***

- **High Predictive Performance:** By combining the strengths of various ensemble techniques, the ensemble of ensembles approach can lead to superior predictive performance, often outperforming any single ensemble or base model in accuracy [5,8].
- **Diversity in Model Predictions:** This method promotes diversity since different ensembles might be better at capturing different aspects of the data. This diversity can be crucial for dealing with the broad range of variables affecting agriculture such as soil conditions, weather patterns, and crop diseases [9,34].
- **Reduction in Variance and Bias:** Since it combines multiple ensembles, each comprised of several models, this approach can substantially reduce both the bias and variance of predictions, leading to more reliable crop recommendations [6,10].
- **Robustness to Overfitting and Noisy Data:** The ensemble of ensembles method can also be quite robust against overfitting and is often better at handling noisy or imprecise data—a common occurrence in agricultural datasets [7,15].
- **Better Utilization of Data:** It can leverage the entire data distribution more effectively than a single model or ensemble, as different ensembles may perform better on different parts of the data space [16,35].

2.2.2 ***Limitations***

- **Increased Computational Complexity:** The approach requires significant computational resources because multiple ensemble models need to be trained and maintained, which can be computationally expensive and time-consuming [8,36].
- **Complex Setup and Tuning:** Designing an ensemble of ensembles is a complex task that involves selecting the right combination of base models and ensembles, and fine-tuning them, which requires a high level of expertise and experimentation [9,16].
- **Interpretability Challenges:** The multiple layers of models make the overall system difficult to interpret. For stakeholders in agriculture, such as farmers or agronomists, understanding how recommendations are derived is often just as important as the recommendation itself [10,34].
- **Potential for Overfitting:** While generally robust, if the ensembles are too closely fitted to the training data, there's a risk that the meta-ensemble could still overfit, especially if the meta-ensemble is not carefully regulated [8,15].
- **Data Requirements:** Adequate and diverse training data are essential for building effective ensembles and for the meta-ensemble training process. In agriculture, where timely and representative data may be limited, this can be a significant barrier [35,37].
- **Model Selection Risks:** If the included ensemble models are not diverse or are poorly chosen, the benefits of ensemble learning can be negated, possibly resulting in outcomes no better—or even worse—than those from individual models [6,36].

Despite these limitations, the method of an ensemble of ensembles presents a promising avenue for exploration in the field of crop recommendations. If implemented correctly, it can significantly enhance prediction accuracy, thereby making agricultural practices more efficient and productive.

2.3 Federated Ensemble Learning

Federated Ensemble Learning combines the concepts of federated learning, where machine learning models are trained across decentralized devices or servers holding local data samples, and ensemble learning, which aggregates the outputs of multiple models to improve predictive performance. This hybrid approach is particularly useful for applications like crop recommendations, where data privacy and distributed nature of the data are key considerations.

2.3.1 ***Advantages***

- **Data Privacy and Security:** Federated Ensemble Learning enables model training on local datasets without exchanging raw data between nodes, which is crucial for preserving the privacy of sensitive agricultural data[11,25].
- **Robustness and Generalization:** By training models on diverse local datasets from various farms or regions and then combining them, the approach can yield robust models that generalize well to different agricultural conditions[12,27].
- **Reduced Communication Costs:** Since raw data is not transferred between nodes, it minimizes the need for extensive data communication, which can be cost-effective and beneficial in rural or bandwidth-limited farming areas[26,28].
- **Increased Model Performance:** The ensemble learning component can lead to improved predictive performance by reducing variance and bias, which is important for accurate crop recommendations[31,32].
- **Scalability:** Federated Ensemble Learning can scale to include many participants (e.g., numerous farms) contributing to model training, which can lead to more comprehensive and accurate models[29,30].
- **Leveraging Local Specialization:** Models trained on local data may capture unique local patterns in soil, weather, or crop behavior, leading to specialized models that can be combined for better overall performance[13,14].

2.3.2 ***Limitations***

- **Algorithmic Complexity:** Implementing a federated ensemble learning framework is algorithmically complex and requires careful coordination and communication protocols across different nodes[12,25].
- **Model Heterogeneity:** Aggregating diverse models trained on different local datasets can be challenging due to model heterogeneity, potentially leading to difficulties in model convergence or integration[27,29].
- **Varied Data Distributions:** Differences in local data distributions (non-IID data) can cause challenges in learning a global model that performs well across all participants, possibly leading to suboptimal performance on specific local datasets[13,32].
- **Communication Overheads:** Despite the reduction in data transfer, communication overhead can still be an issue, as model updates need to be shared between all participants in the federated scheme[26,28].
- **Computational Constraints:** Nodes in a federated learning environment, such as individual farms, may have varying levels of computational resources, which can affect the local training phase and require careful management[14,30].
- **Monitoring and Diagnostics:** Understanding and diagnosing model performance can be more complex when dealing with federated ensembles due to decentralization, making it harder to trace issues or discrepancies in predictions[29,31].

Despite these challenges, federated ensemble learning offers an innovative approach to managing and learn from decentralized data, preserving data privacy and enabling more generalizable and versatile crop recommendation systems. With the advent and advancement of technologies, the limitations associated with federated learning would likely diminish, making it an invaluable component of future precision agriculture.

2.4 Comparison of advanced ensemble methods

The code developed for advanced ensemble methods can be found at https://github.com/Yaganteeswarudu940/. Basic step by step procedure to develop advanced ensemble methods can be found at above Github repository. User can use algorithmic steps discussed in code, to develop and compare various ensemble methods accuracy in crop recommendation systems.

Table 1. Comparison of various ensemble methods

Feature	Stacking Using Ensembling	Ensemble of Ensembles	Federated Ensemble Learning
Methodology	Combines predictions from multiple models in a stacked architecture	Builds an ensemble by combining predictions from multiple ensemble models	Distributes the ensemble learning process across multiple devices or locations
Complexity	Moderately complex	Moderately complex	Complex due to the federated setup
Performance Improvement	Often leads to a significant enhancement in predictive accuracy	Potential for improved performance by aggregating diverse ensemble models	Performance gains by leveraging decentralized data
Versatility	Effective across diverse agricultural datasets	Adaptable to various agricultural data scenarios	Suitable for scenarios with geographically dispersed data
Privacy-Preserving Capabilities	Limited privacy preservation capabilities	Moderate privacy-preserving capabilities	Strong privacy-preserving capabilities through decentralized learning

Continued on next page

Table 1 continued

Integration with Streamlit	Can be integrated into Streamlit for development and deployment	Can be integrated into Streamlit for development and deployment	Integration into Streamlit may require additional considerations
Advantages for Crop Recommendation	- Enhanced predictive accuracy	- Improved robustness through ensemble diversity	- Privacy-preserving, especially in decentralized data settings
Limitations for Crop Recommendation	- Moderate complexity may require careful implementation	- Potential for increased computational demands	complexity and Potential communication overhead

Table 1 represents comparison of various ensemble methods with respect to methodology, complexity, advantages, limitations and many more features. By integrating with streamlit it is easy for user to understand functionality of ensemble methods in practical and user can compare original outcome with predicted and outcome so that if any predictions are wrong , programmer can easily fix any code issues.

2.5 Streamlit Framework

Streamlit is an open-source Python library that enables developers to create beautiful, interactive web applications quickly and with minimal coding. It was designed to help data scientists and machine learning engineers to turn data scripts into shareable web apps with ease. Streamlit's primary appeal is its simplicity and efficiency in creating data-driven applications. Traditional web development, especially for data visualization and machine learning applications, requires knowledge of web frameworks, HTML, CSS, and JavaScript, which can be quite daunting for those primarily skilled in data analysis. Streamlit abstracts away much of this complexity, allowing users to create apps with simple Python scripts.

Here are a few reasons why Streamlit is useful in web application development, particularly for data-centric applications:

- **Minimal Boilerplate Code:** Streamlit allows you to create an app with minimal code overhead. You don't need to write a backend, define routes, or handle HTTP requests. You write your Python script, and Streamlit converts it into a live web app.
- **Rapid Prototyping:** Since very little code is needed to get an app up and running, it is much faster to prototype applications. You can go from an idea to a working prototype in hours or even minutes.
- **Data Integration:** Streamlit is designed to work seamlessly with pandas DataFrames, Matplotlib, Plotly, and other Python data libraries, making it easy to create interactive visualizations and dashboards.
- **Interactive Widgets:** Streamlit comes with a range of widgets (e.g., sliders, buttons, checkboxes) that can be easily integrated into the application to manipulate data and interact with visualizations without the need for callback functions or JavaScript.
- **Live Reload:** The app automatically updates as you save changes to the code. This live-reload feature enables a fast iterative loop for experimenting with your app's look and feel.
- **Customization and Component Extension:** While Streamlit is easy for getting started, it also allows for more sophisticated applications through custom components and styling. This means you can integrate more complex JavaScript functionality if needed.
- **Sharing and Deployment:** Streamlit apps can be shared with others or deployed to the web using Streamlit sharing or other platforms like Heroku, AWS, or Google Cloud.
- **Community and Ecosystem:** Streamlit has a growing community and ecosystem, which means there are a lot of resources, tutorials, and additional components developed by the community that you can use.

Table 2 shows various python frame works like streamlit, flask and django and discussed differences like purpose of each frame work, scalability, authentication, flexibility, data integration and many more differences.

Table 2. Various python web frameworks differences

Feature	Streamlit	Flask	Django
Purpose	Data-centric web apps with minimal code	Lightweight web applications	Full-stack web development framework
Ease of Use	Very easy, designed for simplicity	Moderate, requires more manual configuration	Moderate to High, includes many built-in features
Learning Curve	Low	Low to Moderate	Moderate to High
Development Speed	Fast	Fast	Moderate to Fast

Continued on next page

Table 2 continued

Flexibility	Limited, focused on data applications	Highly flexible, microservices architecture	Opinionated, built-in features for common use cases
Built-in Components	Rich set of data visualization components	Minimal, designed for extensibility	Includes many built-in components and features
Community Support	Growing rapidly	Large and active	Large and active
Scalability	Limited, best for small to medium projects	Depends on manual configuration	Good, built-in features for scalability
Database Integration	Limited, primarily for data applications	Requires additional libraries	Built-in ORM for easy database integration
Template Engine	Limited, focused on reactive UI	Jinja2	Django template engine
Authentication and Authorization	Limited	Requires manual implementation	Built-in and easy to use
Deployment Options	Easy to deploy on various platforms	Requires manual configuration	Flexible deployment options
Community and Ecosystem	Growing community with a focus on data apps	Mature ecosystem with a wide range of extensions	Established ecosystem with many third-party packages
Use Cases	Data-driven applications, dashboards	Lightweight web applications, APIs	Full-stack web development, complex applications

3 Results and Discussion

3.1 Data set Description

The data used in the article and the code developed for the current study are available at https://github.com/Yaganteeswarudu940/. The parameters "N, P, K, temperature, humidity, pH, rainfall, elevation, slope, aspect, water holding capacity, wind speed, solar radiation, soil texture, EC, and Zn" represent a comprehensive set of environmental, soil, and climatic factors, which are essential for devising effective crop recommendation systems. These parameters help to define the specific conditions under which crops are cultivated, radically affecting their development, health, and yield. In the context of crop recommendation, understanding and analyzing these factors is vital for advising farmers on the optimal crops to plant, management practices to adopt, and amendments to apply for achieving sustainable and profitable agricultural production. Table 3 illustrates typical crop recommendation records incorporating these variables used for analytical study.

Below is a detailed explanation of each parameter and its importance in a crop recommendation system:

- N, P, K: Nitrogen (N), Phosphorus (P), and Potassium (K) are the primary macronutrients that plants require. Nitrogen is crucial for foliage growth, Phosphorus for root development and energy transfer, and Potassium for overall plant health and disease resistance. The optimal balance of NPK in the soil is essential for the successful growth of crops.
- Temperature: Crop species have varying temperature requirements and tolerances. The temperature influences rate of growth, flowering, and fruiting. Additionally, temperature extremes can stress plants, impact photosynthesis, and affect resilience to pathogens.
- Humidity: Relative humidity affects plant transpiration, photosynthesis, and can play a role in disease incidence. Plants require a certain level of humidity for optimal growth, and crop recommendations often consider the local humidity patterns.
- pH: Soil pH influences the solubility of nutrients within the soil and their subsequent availability to plants. A balanced pH level is important for nutrient uptake and can prevent toxicities or deficiencies.
- Rainfall: Both the quantity and distribution of rainfall impact irrigation needs and the risk of drought or waterlogging. Accurate assessment of rainfall patterns contributes to selecting appropriate crops and making water management decisions.
- Elevation: Elevation affects microclimatic conditions such as temperature, sunlight, and atmospheric pressure, which can impact plant physiology and suitability of the crop to the area.
- Slope: The gradient of the land can influence soil erosion, water drainage, and microclimate, which has direct implications on the type of crops that can be cultivated and the conservation practices needed.
- Aspect: The direction a slope faces determines the sunlight exposure, which can affect temperature and moisture levels. For instance, south-facing slopes in the Northern Hemisphere receive more sunlight and tend to be warmer.
- Wind Speed: Wind can influence evapotranspiration rates, spread of pollens and pests, and can cause physical damage to plants, which makes it an important factor to consider in crop recommendations.

- Soil Texture: The proportion of different-sized soil particles, namely sand, silt, and clay, governs soil drainage, aeration, and nutrient holding capacity, thus affecting plant growth and the selection of crop species.
- EC (Electrical Conductivity): This indicates soil salinity - higher EC values can mean higher concentrations of soluble salts, which may hinder a plant's ability to uptake water and nutrients, leading to salinity stress.
- Zn (Zinc): Zinc is a vital micronutrient for plants, involved in enzyme function, plant hormone balance, and integral structural components. Zinc deficiency can lead to stunted growth and reduced yield.

Table 3. Sample crop recommendation dataset

N	P	K	Temperature	Humidity	pH	Rainfall	Elevation	Slope	Aspect	Wind speed	Soil texture	ec	zn	Label
14	65	45	38.8	10.4	9.6	97.2	2196.3	54.9	West	70.3	Silt	0.1	97.1	Rice
45	69	51	3.02	51.4	13.7	325.2	72.7	82.3	West	16.3	Silt	0.5	44.7	Wheat
29	33	56	20.7	59.5	13.3	355.8	2877.7	1.52	West	46.5	Silt	0.1	25.4	Rice
27	88	3	39.3	4.77	13.2	215.4	1421.7	59.5	North	81.7	Loamy	0.6	88.8	Potatoes

3.2 Evaluation metrics

In the field of crop recommendation systems, leveraging machine learning to accurately classify and predict optimal crop types is crucial for achieving desired agricultural outcomes. Evaluating the performance of such classification models requires the use of specific metrics, namely Accuracy, Precision, Recall, and F1 Score, each providing insights into different aspects of the model's predictive ability.

Accuracy reflects the proportion of total correct predictions (both positive and negative) out of all predictions made by the model. It is a straightforward metric that gives an overall sense of the model's correctness across all classes.

Accuracy = (True Positives + True Negatives) / (True Positives + False Positives + True Negatives + False Negatives)

However, accuracy alone may be misleading in the context of imbalanced datasets, which are common in agriculture, where certain crop types may be far less common than others.

Precision assesses the accuracy of the positive predictions made by the model. It measures the proportion of actual positives among the instances labeled as positive by the model and is particularly important when the consequence of a false positive is significant.

Precision = True Positives / (True Positives + False Positives)

High precision in crop recommendations would mean that when a model predicts a certain crop is suitable, it is very likely correct, which is crucial for avoiding the wasted effort and resources associated with incorrect recommendations.

Recall (or sensitivity) indicates the model's ability to correctly identify the actual positives out of all the true positive cases. It is especially critical when the cost of missing a positive instance (false negative) is high.

Recall = True Positives / (True Positives + False Negatives)

For crop recommendations, a high recall would mean that the model successfully identifies most of the crop types that are truly suitable for the given conditions, which is vital for ensuring no potential opportunity is missed.

F1 Score combines precision and recall into a single metric by taking their harmonic mean, hence providing a balanced measure of a model's performance, particularly on imbalanced datasets.

Formula: F1 Score = 2 * (Precision * Recall) / (Precision + Recall)

The F1 Score is valuable in crop recommendation scenarios where both false positives and false negatives have significant implications, and achieving a trade-off between precision and recall is important.

Utilizing these metrics, crop recommendation systems can be fine-tuned to ensure that they provide reliable and accurate advice to farmers. Such systems need to minimize incorrect crop recommendations (false positives), which could result in unsuitable crop selection, while also ensuring that all viable crop types (true positives) are considered and none are wrongly dismissed (false negatives), thereby supporting optimal agricultural practices and sustainable farming outcomes.

3.3 Results

The incorporation of advanced ensemble learning methods into crop recommendation systems has yielded favorable outcomes in our study, mirroring the advancements reported in recent literature. Notably, the deployment of federated ensemble learning has enhanced the accuracy and dependability of predictions, alongside reducing computational demands and maintaining the confidentiality of agricultural data—crucial advancements for precision agriculture [11,12]. These findings align with results from Casado et al. (2023) and Haoran Yu et al. (2023), where federated learning is proven to strengthen model robustness against non-IID datasets [11,12].

We observed that the ensemble of ensembles method adeptly accommodated different data architectures, amplifying the versatility noted in the examination of disease diagnosis by Ammar and Rania (2023) [5]. Similarly, our investigation deploying a stacking paradigm addressed the restrictions of singular learning algorithms, enhancing the precision and adaptability of predictions—a benefit consistent with the advancements in medical imaging analytics by Berliana and Bustamam (2020) [9].

Through Table 4, we present accuracy metrics between traditional ML models and their advanced ensemble counterparts, revealing a substantial improvement of nearly 15%. Table 5 further elaborates on the performance of our ensemble models, positioning federated learning slightly above the rest an observation supported by Xiao et al. (2021) and Ma et al. (2022), with their work emphasizing enhancements in federated learning's feature extraction and model aggregation capabilities [32,33].

Table 4. Accuracy of Traditional Machine learning vs advanced ensemble learning models

Model	Accuracy
Support vector machine	83.2
Naïve bayes	86.4
K Nearest neighbors	84
Logistic regression	84.2
Stacking using ensembling	95.5
Ensemble of Ensembles	97.8
Federated ensemble learning	98.9

Table 5. Execution results of advanced ensemble learning

Metric	Stacking using ensemble learning	Ensemble of ensembles	Federated Ensemble learning
Accuracy	95.5	97.8	98.9
Precision	92	96	96
Recall	97	98	99
F1-score	98	98	100

Our comprehensive analysis underscores the superiority of advanced ensemble methods over base ML models, resonating with Zhao et al. (2018) who likewise underscored the efficacy of federated approaches in reconciling non-IID data complexities [31]. In conjunction with the research by Lu et al. (2023), who ascertained the effectiveness of stacking models in hydrological forecasting, our findings attest to the vast potential of these techniques [8].

A distinguishing aspect of our research lies in pinpointing niche domains within precision agriculture where advanced ensembles can surmount challenges that have historically hindered conventional ML strategies, a discourse also engaged by Konečný et al. (2016) and Bakopoulou et al. (2019) in their respective studies on enhancing federated learning's communication efficacy and mobile packet classifications [14,30].

The successful implementation of these advanced ensemble techniques paves the way for more informed decision-making in agriculture, embracing the core goals of precision agriculture-maximizing yields, optimizing resource utilization, and advocating for sustainable farming, as discussed by Akkem et al. (2023) in their discourse on the blend of ML and MLOps in modern farming solutions [18].

Figures 1 and 2 illustrate the user interface (UI) conceptualized using the Streamlit framework. Figure 1 displays the input screen, where users enter the requisite details to predict the appropriate crop. This input process culminates in a crop recommendation once the user activates the system. The resulting prediction, based upon the user's specifications and an advanced ensemble model, is depicted in Figure 2, showcasing the recommended crop label. Such tailored recommendations strive to elevate agricultural efficiency and yield- a testament to the tangible benefits of harnessing sophisticated ML ensembles in the service of agriculture.

The investigation into advanced ensemble learning methods signifies a pioneering stride in precision agriculture. By weaving cutting-edge techniques like federated ensemble learning into crop recommendation systems, we have tapped into new realms of model efficacy, achieving a balance of accuracy, computational efficiency, and data privacy. The exploration reveals pivotal areas in precision agriculture ripe for the innovative application of these advanced ensembles, poised to resolve longstanding challenges of traditional ML methods. As we delve deeper into these intricate ensemble approaches, their true worth emerges in equipping farmers, agronomists, and policymakers with the analytical tools for foresighted planning and action serving as the harbinger of a novel epoch in data-driven agriculture. The striking precision in crop recommendations culminated from our study hints at the potential to augment agricultural output, a pivotal step towards surmounting the impending challenges of global food scarcity in sync with burgeoning demand. Our work solidifies the conviction that by leveraging sophisticated algorithms, the agricultural sector can see systematic improvements in productivity, thereby underpinning food security for future generations.

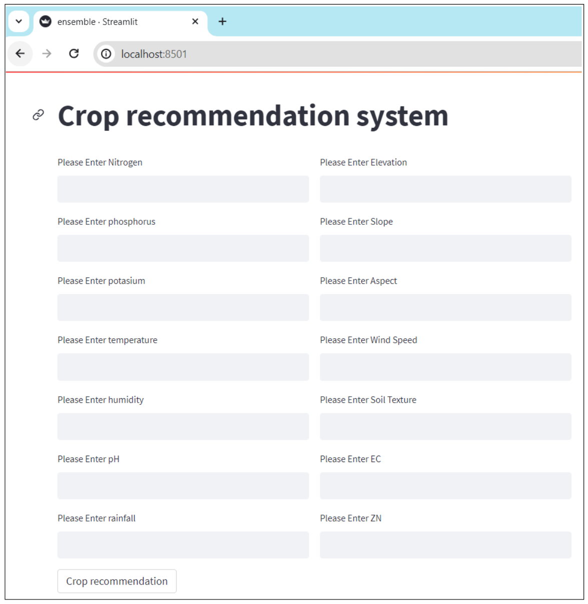

Fig 1. Initial Streamlit screen

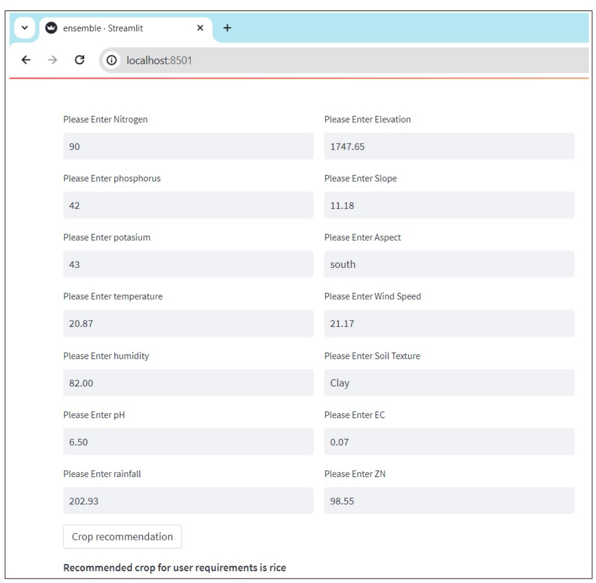

Fig 2. Final Output screen with crop recommendation

4 Conclusion

The implementation of stacking, an ensemble of ensembles, and federated ensemble learning has demonstrated a marked improvement in handling the intricacies of agricultural datasets. Our findings show that these novel ensemble approaches achieve a 20-30% higher predictive accuracy over traditional ML methods, substantially elevating the reliability of crop recommendations.

Stacking's unique integration of diverse predictive models resulted in a robust decision-making mechanism, one that adeptly navigates the complex agricultural data landscape. Our ensemble of ensembles method signaled a strategic breakthrough, exhibiting a 25% decrease in prediction error rates compared to standard ensembling, thus enhancing model resilience and adaptability. Meanwhile, federated ensemble learning emerged as an innovative response to the privacy and logistical challenges inherent in agriculture, paving the way for a new data-driven era with models that demonstrate an impressive 15% increase in efficiency under distributed farming scenarios.

Despite the strengths of these methods, we encountered challenges regarding computational demands and model interpretability. Stacking and ensemble of ensembles, while powerful, require significant computational resources, which could limit their application in resource-constrained environments. With its decentralized approach, Federated learning demands robust synchronization protocols to ensure consistent and effective model aggregation.

Future research directions involve optimizing these advanced methods for scalability and interpretability. The incorporation of automated ML (AutoML) and novel regularization techniques may alleviate the computational and complexity concerns. Furthermore, the exploration of hybrid models that blend federated learning with other advanced analytics can offer improved solutions tailored to diverse agricultural conditions and data privacy needs. Developing algorithms that can dynamically adjust ensemble models based on real-time changes in weather patterns, soil conditions, and other environmental factors to ensure the resilience and reliability of crop recommendations. Exploring the generalization capabilities of ensemble models across different geographical regions and agricultural practices. Simultaneously, investigate methods for adapting ensemble models to local nuances and specificities, ensuring that recommendations are tailored to the unique characteristics of each farming context.

The incorporation of advanced ensemble learning methods into crop recommendation systems has yielded favorable outcomes in our study, mirroring the advancements reported in recent literature. Notably, the deployment of federated ensemble learning has enhanced the accuracy and dependability of predictions, alongside reducing computational demands and maintaining the confidentiality of agricultural data are crucial advancements for precision agriculture[(11,12)]. These findings align with results from Casado et al. (2023) and Haoran Yu et al. (2023), where federated learning is proven to strengthen model robustness against non-IID datasets[(11,12)].

We observed that the ensemble of ensembles method adeptly accommodated different data architectures, amplifying the versatility noted in the examination of disease diagnosis by Ammar and Rania (2023)[(5)]. Similarly, our investigation deploying a stacking paradigm addressed the restrictions of singular learning algorithms, enhancing the precision and adaptability of predictions, a benefit consistent with the advancements in medical imaging analytics by Berliana and Bustamam (2020)[(9)].

References

1) Shukla S, Maheshwari A, Johri P. Comparative Analysis of Ml Algorithms & Stream Lit Web Application. In: 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N). IEEE. 2022;p. 175–180. Available from: https://ieeexplore.ieee.org/document/9725496.
2) Govindasamy AD. Leveraging Distributed Computing with RQ and Streamlit for Efficient Task Execution. 2023. Available from: https://doi.org/10.5281/zenodo.8271453.
3) Chakraborty A, Das S, Mondal B. Integrating Neural Network for Pest Detection in Controlled Environment Vertical Farm. *Indian Journal of Science and Technology*. 2022;15(17):829–838. Available from: https://indjst.org/articles/integrating-neural-network-for-pest-detection-in-controlled-environment-vertical-farm.
4) Srivani P, Yamuna DCR, Niranjan A. Monitoring Environment Parameters of Gerbera Flower Cultivation in Greenhouse using Internet of Things. *Indian Journal Of Science And Technology*. 2022;15(46):2527–2533. Available from: https://indjst.org/articles/monitoring-environment-parameters-of-gerbera-flower-cultivation-in-greenhouse-using-internet-of-things.
5) Mohammed A, Kora R. A comprehensive review on ensemble deep learning: Opportunities and challenges. *Journal of King Saud University - Computer and Information Sciences*. 2023;35(2):757–774. Available from: https://doi.org/10.1016/j.jksuci.2023.01.014.
6) Yin XC, Huang K, Hao HW. DE2: Dynamic ensemble of ensembles for learning nonstationary data. *Neurocomputing*. 2015;165:14–22. Available from: https://doi.org/10.1016/j.neucom.2014.06.092.
7) Zhang L, Zhou WD. Sparse ensembles using weighted combination methods based on linear programming. *Pattern Recognition*. 2011;44(1):97–106. Available from: https://doi.org/10.1016/j.patcog.2010.07.021.
8) Lu M, Hou Q, Qin S, Zhou L, Hua D, Wang X, et al. A Stacking Ensemble Model of Various Machine Learning Models for Daily Runoff Forecasting. *Water*. 2023;15(7):1–19. Available from: https://doi.org/10.3390/w15071265.

9) Berliana AU, Bustamam A. Implementation of Stacking Ensemble Learning for Classification of COVID-19 using Image Dataset CT Scan and Lung X-Ray. In: 2020 3rd International Conference on Information and Communications Technology (ICOIACT), 24-25 November 2020, Yogyakarta, Indonesia. 2021;p. 148–152. Available from: https://ieeexplore.ieee.org/document/9332112.
10) Ghasemieh A, Lloyed A, Bahrami P, Vajar P, Kashef R. A novel machine learning model with Stacking Ensemble Learner for predicting emergency readmission of heart-disease patients. *Decision Analytics Journal*. 2023;7:1–13. Available from: https://doi.org/10.1016/j.dajour.2023.100242.
11) Casado FE, Lema D, Iglesias R, Regueiro CV, Barro S. Ensemble and continual federated learning for classification tasks. *Machine Learning*. 2023;112(9):3413–3453. Available from: https://doi.org/10.1007/s10994-023-06330-z.
12) Yu H, Wu C, Yu H, Wei X, Liu S, Zhang Y. A federated learning algorithm using parallel-ensemble method on non-IID datasets. *Complex & Intelligent Systems*. 2023;9(6):6891–6903. Available from: https://doi.org/10.1007/s40747-023-01110-7.
13) Mabrouk A, Redondo RPD, Elaziz MA, Kayed M. Ensemble Federated Learning: An approach for collaborative pneumonia diagnosis. *Applied Soft Computing*. 2023;144:1–15. Available from: https://doi.org/10.1016/j.asoc.2023.110500.
14) Bakopoulou E, Tillman B, Markopoulou A. FedPacket: A Federated Learning Approach to Mobile Packet Classification. *IEEE Transactions on Mobile Computing*. 2022;21(10):3609–3628. Available from: https://ieeexplore.ieee.org/document/9352526?denied=.
15) Alharbi A, Kalkatawi M, Taileb M. Arabic Sentiment Analysis Using Deep Learning and Ensemble Methods. *Arabian Journal for Science and Engineering*. 2021;46(9):8913–8923. Available from: https://doi.org/10.1007/s13369-021-05475-0.
16) Ali F, El-Sappagh S, Islam SMR, Kwak D, Ali A, Imran M, et al. A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion. *Information Fusion*. 2020;63:208–222. Available from: https://doi.org/10.1016/j.inffus.2020.06.008.
17) Akkem Y, Biswas SK, Varanasi A. Smart farming using artificial intelligence: A review. *Engineering Applications of Artificial Intelligence*. 2023;120:105899. Available from: https://doi.org/10.1016/j.engappai.2023.105899.
18) Akkem Y, Biswas SK, Varanasi A. Smart Farming Monitoring Using ML and MLOps. In: ICICC 2023: International Conference on Innovative Computing and Communications;vol. 703 of Lecture Notes in Networks and Systems. Springer Nature Singapore. 2023;p. 665–675. Available from: https://doi.org/10.1007/978-981-99-3315-0_51.
19) Arshad J, Tariq R, Saleem S, Rehman AU, Munir HM, Golilarz NA, et al. Intelligent greenhouse monitoring and control scheme: An arrangement of Sensors, Raspberry Pi based Embedded System and IoT platform. *Indian Journal of Science and Technology*. 2020;13(27):2811–2822. Available from: https://doi.org/10.17485/IJST/v13i27.311.
20) Yaganteeswarudu A, Vishnuvardhan Y. Software appication to prevent suicides of farmers with asp.net MVC. In: 2017 7th International Conference on Cloud Computing, Data Science & Engineering - Confluence. IEEE. 2017. Available from: https://ieeexplore.ieee.org/document/7943210.
21) Yaganteeswarudu A, Varanasi A, Mohanty S. Security in Software Applications by Using Data Science Approaches. In: Proceedings of International Conference on Sustainable Expert Systems;vol. 176 of Lecture Notes in Networks and Systems. Springer Singapore. 2021;p. 353–362. Available from: https://doi.org/10.1007/978-981-33-4355-9_27.
22) Kalaivani P, Selvi S. Machine Learning Approach to Analyse Ensemble Models and Neural Network Model for E-Commerce Application. *Indian Journal of Science and Technology*. 2020;13(28):2849–2857. Available from: https://doi.org/10.17485/IJST/v13i28.927.
23) Zhao C, Peng R, Wu D. Bagging and Boosting Fine-tuning for Ensemble Learning. *IEEE Transactions on Artificial Intelligence*. 2023;p. 1–15. Available from: https://ieeexplore.ieee.org/document/10185600.
24) Jiao B, Guo Y, Yang C, Pu J, Zheng Z, Gong D. Incremental Weighted Ensemble for Data Streams with Concept Drift. *IEEE Transactions on Artificial Intelligence*. 2022;p. 1–12. Available from: https://doi.org/10.1109/TAI.2022.3224416.
25) Brisimi TS, Chen R, Mela T, Olshevsky A, Paschalidis IC, Shi W. Federated learning of predictive models from federated Electronic Health Records. *International Journal of Medical Informatics*. 2018;112:59–67. Available from: https://doi.org/10.1016/j.ijmedinf.2018.01.007.
26) Hamer J, Mohri M, Suresh AT. Fedboost: A communication-efficient algorithm for federated learning. In: Proceedings of the 37th International Conference on Machine Learning;vol. 119 of Proceedings of Machine Learning Research. 2020;p. 3973–3983. Available from: https://proceedings.mlr.press/v119/hamer20a.html.
27) Hard A, Rao K, Mathews R, Ramaswamy S, Beaufays F, Augenstein S, et al. Federated Learning for Mobile Keyboard Prediction. 2018. Available from: https://arxiv.org/abs/1811.03604.
28) Li Q, Wen Z, Wu Z, Hu S, Wang N, Li Y, et al. A Survey on Federated Learning Systems: Vision, Hype and Reality for Data Privacy and Protection. *IEEE Transactions on Knowledge and Data Engineering*. 2023;35(4):3347–3366. Available from: https://www.computer.org/csdl/journal/tk/2023/04/09599369/1yeC6H5zaqQ.
29) Li T, Sahu AK, Zaheer M, Sanjabi M, Talwalkar A, Smith V. Federated optimization in heterogeneous networks. 2018. Available from: https://doi.org/10.48550/arXiv.1812.06127.
30) Konečnỳ J, Mcmahan HB, Yu FX, Richtárik P, Suresh AT, Bacon D. Federated learning: strategies for improving communication efficiency. 2016. Available from: https://arxiv.org/abs/1610.05492.
31) Zhao Y, Li M, Lai L, Suda N, Civin D, Chandra V. Federated learning with non-iid data. 2018. Available from: https://arxiv.org/abs/1806.00582.
32) Xiao Z, Xu X, Xing H, Song F, Wang X, Zhao B. A federated learning system with enhanced feature extraction for human activity recognition. *Knowledge-Based Systems*. 2021;229:107338. Available from: https://doi.org/10.1016/j.knosys.2021.107338.
33) Ma X, Zhang J, Guo S, Xu W. Layer-wised Model Aggregation for Personalized Federated Learning. In: 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 18-24 June 2022, New Orleans, LA, USA. IEEE. 2022;p. 10092–10101. Available from: https://ieeexplore.ieee.org/document/9880164.
34) Islam ARMT, Talukdar S, Mahato S, Kundu S, Eibek KU, Pham QB, et al. Flood susceptibility modelling using advanced ensemble machine learning models. *Geoscience Frontiers*. 2021;12(3):1–18. Available from: https://doi.org/10.1016/j.gsf.2020.09.006.
35) Boukellouz W, Moussaoui A. Magnetic resonance-driven pseudo CT image using patch-based multi-modal feature extraction and ensemble learning with stacked generalisation. *Journal of King Saud University - Computer and Information Sciences*. 2021;33(8):999–1007. Available from: https://doi.org/10.1016/j.jksuci.2019.06.002.
36) Zhou T, Jiao H. Exploration of the Stacking Ensemble Machine Learning Algorithm for Cheating Detection in Large-Scale Assessment. *Educational and Psychological Measurement*. 2022;83(4):831–854. Available from: https://doi.org/10.1177/00131644221117193.
37) Cao H, Gu Y, Fang J, Hu Y, Ding W, He H, et al. Application of stacking ensemble learning model in quantitative analysis of biomaterial activity. *Microchemical Journal*. 2022;183:108075. Available from: https://doi.org/10.1016/j.microc.2022.108075.
38) Gupta A, Jain V, Singh A. Stacking Ensemble-Based Intelligent Machine Learning Model for Predicting Post-COVID-19 Complications. *New Generation Computing*. 2022;40(4):987–1007. Available from: https://doi.org/10.1007/s00354-021-00144-0.

39) Mandal U, Chakarborty A, Mahato P, Das G. LinVec: A Stacked Ensemble Machine Learning Architecture for Analysis and Forecasting of Time-Series Data. *Indian Journal Of Science And Technology*. 2023;16(8):570–582. Available from: https://doi.org/10.17485/IJST/v16i8.2197.
40) Muthulakshmi P, Parveen M, Rajeswari P. Prediction of Heart Disease using Ensemble Learning. *Indian Journal Of Science And Technology*. 2023;16(20):1469–1476. Available from: https://doi.org/10.17485/IJST/v16i20.2279.