Machine Learning Based Precision Agriculture using Ensemble Classification with TPE Model

[1]Latha M, [2]Mandadi Vasavi, [3]Chunduri Kiran Kumar, [4]Balamanigandan R, [5]John Babu Guttikonda, and [6]Rajesh Kumar T
[1]Department of Information Technology, SRM Institute of Science and Technology, Ramapuram, Chennai, India.
[2]Department of Computer Science and Engineering, R V R and J C College of Engineering, Andhra Pradesh, India.
[3]Department of Computer Science and Application, Koneru Lakshmaiah Education Foundation Deemed to be University Vaddeswaram, Andhra Pradesh, India.
[4,6]Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (SIMATS), Chennai, India.
[5]Department of Computer Science and Engineering, Vijaya Engineering College, Tanikella, Telangana, India.
[1]latham@srmist.edu.in, [2]vasavilahari@gmail.com, [3]kirancvl17@gmail.com, [4]balamanigandanr.sse@saveetha.com, [5]johnbabug@gmail.com, [6]t.rajesh61074@gmail.com

Correspondence should be addressed to Balamanigandan R : balamanigandanr.sse@saveetha.com

Article Info
Journal of Machine and Computing (http://anapub.co.ke/journals/jmc/jmc.html)
Doi: https://doi.org/10.53759/7669/jmc202404025
Received 10 March 2023; Revised from 25 August 2023; Accepted 16 December 2023.
Available online 05 January 2024.

Abstract – Many tasks are part of smart farming, including predicting crop yields, analysing soil fertility, making crop recommendations, managing water, and many more. In order to execute smart agricultural tasks, researchers are constantly creating several Machine Learning (ML) models. In this work, we integrate ML with the Internet of Things. Either the UCI dataset or the Kaggle dataset was used to gather the data. Effective data pretreatment approaches, such as the Imputation and Outlier (IO) methods, are necessary to manage the intricacies and guarantee proper analysis when dealing with data that exhibits irregular patterns or contains little changes that can have a substantial influence on analysis and decision making. The goal of this research is to provide a more meaningful dataset by investigating data preparation approaches that are particular to processing data. Following the completion of preprocessing, the data is classified using an average approach based on the Ensemble of Adaptive Neuro-Fuzzy Inference System (ANFIS), Random Neural Network (PNN), and Clustering-Based Decision Tree (CBDT) techniques. The next step in optimising the hyperparameter tuning of the proposed ensemble classifier is to employ a new Tree-Structured Parzen Estimator (TPE). Applying the suggested TPE based Ensemble classification method resulted in a 99.4 percent boost in accuracy.

Keywords – Smart Farming, Machine Learning, Data Preprocessing, Ensemble Classification, Tree-Structure Parzen Estimator.

I. INTRODUCTION

IoT (Internet of Things) integration in agriculture is a revolutionary strategy that combines cutting-edge technology with time-tested farming techniques to completely transform farming practices [1]. Using smart sensors, connected devices, and data analytics, IoT in agriculture monitors, controls, and optimizes a range of farming operations [2]. Farmers can now make informed decisions by using real-time data from these networked devices to gather vital information about soil moisture, temperature, crop health, and environmental conditions [3]. Through the utilization of IoT, farmers can optimize resource utilization, monitor livestock health, automate irrigation and fertilization procedures, increase crop yields, and obtain valuable insights for accurate agricultural planning—all of which are critical components of sustainable and effective farming practices [4,5].

Precision agriculture, which is frequently heralded as the farming of the future, transforms conventional agricultural methods by utilizing state-of-the-art technologies to maximize crop production's sustainability, productivity, and efficiency [6]. This creative method combines cutting-edge instruments like GPS, sensors, drones, and data analytics to carefully examine and control variability in fields [7]. Precision agriculture maximizes resource utilization, reduces environmental impact, and increases crop yields by precisely adjusting irrigation, fertilization, and pesticide application to specific areas [8]. Precision agriculture enables informed decision-making by giving farmers comprehensive insights into crop health, soil conditions, and growth patterns [9]. This enables the adoption of tailored strategies to meet the particular requirements of each field segment and allows for targeted interventions [10]. The importance of soil is found in its many functions.

Soil provides vital nutrients, water, and support for crop growth, thereby fostering and maintaining agriculture [11]. In order to maximize crop yield and soil health, machine learning (ML) in soil agriculture uses sensor technology, data-driven insights, and predictive analytics to transform farming practices [12]. ML algorithms enable accurate decision-making for crop management, irrigation, and fertilization by analyzing soil composition, moisture levels, nutrient content, and environmental data gathered by IoT sensors. These algorithms, which interpret intricate data patterns and aid in the early detection of disease, nutrient deficiencies, or soil degradation, include neural networks, decision trees, and regression models [13].

Main Contributions

- Effective Data Preprocessing Techniques
    - A focus on handling anomalous trends and minute modifications in the gathered data to guarantee precise analysis and judgment, investigation and application of particular IO data preprocessing methods designed to increase the dataset's informativeness.
- Classification Using Multiple ML Models
    - Application of Ensemble of PNN, CBDT, and ANFIS techniques for data classification of soil to detect fertile or not.
- Hyperparameter Tuning with Novel TPE
    - The proposed ensemble classifier's hyperparameters can be tuned by TPE in order to get accurate result.
- Performance Evaluation
    - Evidence of notable improvements in accuracy when using the suggested TPE-based Ensemble classification.

Organization of the Work

Residual sections of the paper are organised as: In Section 2, the essential literatures are reviewed, and in Section 3, the proposed model is briefly discussed. A summary and conclusion are provided in Section 5 to finish. The results and an overview of the validation process are included in Section 4.

II. RELATED WORKS

The goal of [14] study was to create new machine learning techniques for better agricultural prediction. Nonetheless, there appeared to be a mismatch between agricultural machine learning and the field's core research. The predicament was made worse by the current problems with agricultural data. The impact of these data issues on different machine learning techniques used in agriculture was examined in this study. The paper investigated the use of KNN and naive bayes classification algorithms for precision agriculture. After the data was analyzed, several factors were taken into account to determine the best classification method for precision agriculture.

Data were gathered by a sensor network dispersed throughout the soil of a commercial strawberry farm in the framework by [15] in order to infer the final physicochemical properties of the fruit at harvest points close to the sensor locations. Neural networks and Gaussian process regression models are two examples of empirical and statistical models that were jointly studied to predict important physicochemical characteristics of strawberries. For example, color could be predicted to be within 9% and 14% of their expected value ranges, respectively, when combined with soluble solids content (sweetness). With this degree of precision, the next stage of managing soil conditions for long-term, premium strawberry production was made possible.

Three stages were involved in the work by [16] pre-processing, feature selection (FS), and classification. After pre-processing the dataset, FS was carried out using the Variance Inflation Factor algorithm (VIF) and Correlation-based FS (CBFS). A proposed IoT-based smart agriculture system calls for a two-tier machine learning model. Based on input soil properties, the Adaptive k-Nearest Centroid Neighbour Classifier (aKNCN) divided soil samples into classes and estimated soil quality in the first tier. With the help of the Extreme Learning Machine algorithm (ELM), the second tier forecast crop yield. To improve ELM's performance accuracy, weights were updated in an optimized strategy using a modified Butterfly Optimization Algorithm (mBOA). The suggested system was evaluated using PYTHON as an implementation tool, which assessed performance using a variety of metrics using soil datasets.

A Chaotic Jaya farming was introduced in the study by [17]. The described method produced feature vectors automatically by combining CV with metaheuristic algorithms and the SqueezeNet model for soil classification. Through hyperparameter tuning, the CJO algorithm enhanced the performance of the SqueezeNet model. Furthermore, soil types were classified using the Elman technique, with parameters modified by the chicken swarm algorithm (CSA). The

Kaggle dataset was used to evaluate the soil classification performance of the CJOCV-STC method, and the results showed that it performed better than previous methods, with an accuracy increase of 98.47%.

For crop prediction, ML models like DT, SVM, RF were used in the study by [18]. With an accuracy of 99.24%, RF performed better than the other models among them. Thus, by including parameters such as temperature, humidity, pH, rainfall, phosphorus, nitrogen, and potassium, the suggested system was able to suggest the ideal crop for these types of land. Thus, the goal of this system was to support farmers, the government, and other agriculture industry stakeholders in making important decisions.

The stage of an end-to-end system integrating contemporary tools for accurate soil condition monitoring and control was described in the paper by [19]. A sensor network in the soil of a commercial strawberry farm collected data for the proposed framework, which inferred the physicochemical properties of the fruit at harvest points. Neural regression models are two examples of empirical and statistical models that were jointly studied to predict important physicochemical characteristics of strawberries. With this precision, the next stage of managing soil conditions for long-term, premium strawberry production was made possible.

Research Gap

The extant literature has presented diverse methodologies for the assimilation of ML into agricultural practices. These methodologies have focused on domains such as precision farming, crop prediction, and soil condition monitoring. To ensure accurate analysis and decision-making, a research gap is evident in the area of IoT and ML integration, particularly with regard to addressing irregularities in agricultural data. While some studies investigated sensor-based predictive models for monitoring soil conditions and predicting crop quality, none specifically addressed the difficulties resulting from irregularities in data gathered from sources such as UCI or Kaggle datasets, a gap that the proposed paper fills. By examining data preprocessing strategies designed to address irregular patterns in agricultural data gathered from various sources, this research seeks to close this gap. The paper also introduces a new method that combines IoT and ML. It utilizes a combination of PNN, CBDT and ANFIS techniques, which are then optimized through hyperparameter tuning using a TPE. Improved accuracy is the goal of the suggested methodology, a direction not specifically addressed by the previous research.

III. PROPOSED METHODOLOGY

This research suggests TPE based Ensemble Classification for soil data classification. **Fig 1** depicts the flow of the suggested approach's steps. The preliminary step is data preprocessing with IO method, which is followed by Ensemble classification and hyperparameter tuning using TPE.

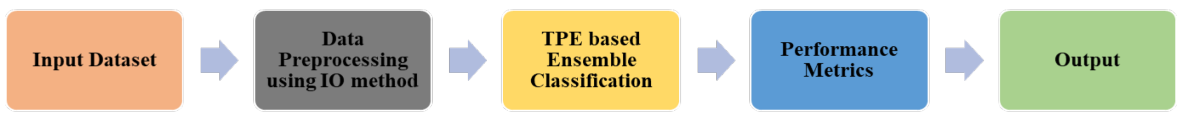

Fig 1. Workflow of the Proposed Model.

Dataset Description

Data preparation is the initial stage in smart farming [20]. Databases such as Oracle, relational databases, or even simple spreadsheets like Excel could include data pertaining to agriculture. The dataset used in this work is sourced from open-source sources such as UCI Machine Learning Repository or Kaggle. To determine if soil is fertile, **Table 1** displays its fertility relative to a number of soil factors, such as potassium, nitrogen, sand, and others.

Table 1. Soil Fertility Dataset

pH	EC	OC	OM	K	Mn	Sand	Silt	CEC	Output
7.75	0.4	0.02	0.01	275	4.6	84.3	6.8	7.81	Fertile
8.38	1.09	0.03	0.06	96	4.2	91.6	4.2	7.21	Non-Fertile

Data Preprocessing

In this case, we used a systematic approach [21] to preprocess the data that was left over after eliminating correlations. We looked at each data set separately and adopted the optimal methodology. Because many of the data points had unique characteristics, we were able to evaluate and analyse them separately. Not every data followed the same procedure since every variable has its own distinct qualities that must be carefully considered.

Imputations
In MATLAB, the ismissing(A) function produces a logical array that indexes the points in the input data A that contain missing values. This function may be used to programmatically find missing values on multiple platforms. The dataset's administrators use their best judgement to decide what counts as missing values, which can take the form of not-a-number (NaN) values, not-a-datetime (NaT) values, undefined values, or blank spaces. Mean imputation, regression imputation, , DBScan, and Isolation Forest are some of the approaches that can be used to deal with missing values and manage outliers before imputation. These measures might be taken by the proposed study to guarantee that data imputations are correct and dependable..

Outlier Filtering
Outliers in data may be shown in two ways: first, by displaying the data points in contradiction of time, which shows values that differ from the average; and second, by using a histogram, which shows values that are farther away from the mean. A programmed method is used to eliminate outliers by eliminating numbers that do not fall inside a given range..

Data Classification Using Ensemble Classification
The classifier receives the preprocessing output when data preprocessing has finished. In order to classify data, this study used a mixture of probabilistic neural networks, clustering-based decision trees, and adaptive neuro-fuzzy inference schemes. Averaging the model is how ensemble learning is carried out..

Adaptive Neuro-Fuzzy Inference System (ANFIS)
Input to the ANFIS is provided by the data. The ANFIS network is one kind of neural network that uses the network. Adaptable nodes make up every single node in the top layer. The outputs of layer 1 represent the inputs' fuzzy membership grade [22].

In the second layer, there are nodes that are fixed. The fact that they bear the letter M implies that their primary function is that of a simple layer might be represented by these. On the third layer, there are fixed nodes as well. They normalise the firing intensity from the previous layer; the symbol N indicates this. In the fourth layer, the nodes are able to adjust. The output of each node in this layer is just the normalised firing intensity multiplied by a first-order polynomial. There is a single permanent node in the fifth layer that contains the letter S. Every time a signal comes in, this node adds them up.

Probabilistic Neural Network
Multiple smaller networks, each representing an approximation of the pdf for a different class, comprise a Probabilistic Neural Network. The input nodes contain the input data. The second functions, with the points in the train set acting as centres. The third layer takes an regular of the second layer's results for each class. The greatest value is picked in the fourth stage of voting [23,24].

Clustering-Based Decision Tree
Training data set $(X_i, Y_i), i = 1,2,3, \dots \dots, N$ where Xi refers to a continuous-valued vector in n dimensions, and $Y_i = \{0,1\}$ The corresponding class identification is denoted by "0" for normal and "1" for abnormalities. The proposed method consists of two parts: training and testing [25]. Following method, the training space is partitioned into k separate clusters C_1, C_2, C_3,..., C_K. For the C4.5 decision tree, the cases in each k-Means cluster are utilised for training purposes. According to the k-Means method, each training instance is only linked to one cluster. The C4.5 decision tree, however, fine-tunes its decision bounds when subgroups or overlaps are present inside a cluster by partitioning the instances across the feature space using an if-then rule set. This training is done on that cluster fine.

Ensembling
The idea behind ensemble-based systems is that various classifiers or features could lead to different types of errors, and that combining models can help reduce these errors by averaging them. Ensemble learning is commonly used to improve classification or forecast performance when a single model is inadequate, especially in cases involving several classes..

Hyperparameter tuning using TPE
Every machine learning model relies on its hyperparameters to determine how effective it is. In this study, TFE is employed to adjust the hyperparameters of the ensemble classifier that is suggested. They have command over the structure of the model or the learning process. Choosing hyperparameters is not a standardised process, though, in reality. Alternatively, optimisation search algorithms are used to adjust hyperparameters via trial and error, or they are left at their default settings on occasion. By recasting the problem as an optimisation one, hyperparameter optimisation offers a methodical way to tackle this issue: ideally, a set of hyperparameters would minimise the discrepancy between the expected and actual values.

The hyperparameter tuning in the proposed study was done using the TPE technique. When it comes to finding the hyperparameters of a machine learning model, the TPE method is one of the most successful sequential model-based global optimisation algorithms. The TPE technique generates functions in a space by using the Parzen-window density estimation. There are three possible distributions that may be used to construct the search space: uniform, discrete, and logarithmic.

The distribution is initialised during the startup iterative procedure through a random search that selects the response surface {θ^((i))),y^((i))~,i=1,2,…,N_"int " }, where θ is the set of hyperparameters, y is the surface, and N_"int " is the total number of iterations. Parzen window estimators [26,27] are the foundational building blocks of the TPE method. Commonly known as the kernel density estimator, the Parzen window estimator is using the Parzen window estimators. Using a quantile threshold value y*, which is completely subjective, the calculated hyperparameters are divided into two groups. The estimator p(θ■y) is distinct in Equation (1) as the product of the algorithm's configuration space and the samples of excellent and bad hyperparameters.

$$p(\theta \mid y) = \begin{cases} pr_{\text{good}}(\theta) & \text{if } y < y* \\ pr_{\text{bad}}(\theta) & \text{if } y \geq y* \end{cases} \tag{1}$$

in which y<y^* denotes a function value below the threshold. Equation (2) explains that two hyperparameter distributions can be obtained: one with a function value less than the threshold value (pr_good (θ)) and another with a value greater than the threshold value (pr_bad (θ)). Equation (2) shows how to find the best hyperparameter setup:

$$\theta *= \operatorname{argmin} \frac{pr_{\text{bad}}(\theta)}{pr_{\text{good}}(\theta)} \tag{2}$$

The TPE algorithm chooses the optimum hyperparameters according to a set of the best observations and respective distributions, in addition to choosing the best observations.

IV. RESULTS AND DISCUSSION

Experimental Setup

The experimental setup for the proposed model is clearly laid out in **Table 2** which includes details like the software utilised and the amount of data used for training and testing.

Table 2. Experimental Setup

Component Description	**Software Component**
MATLAB 2013a	Coding Language
Component Description	Hardware Component
512 Mb	RAM
Intel Core Processor	Classification
40GB	Hard Disk
44 Mb	Floppy drive
15 VGA Colour	Monitor

Performance Metrices

This paper calculated four main analytical metrics: true positive (TP), true negative (TN), $false\ positive\ (FP), and\ false\ negative\ (FN)$ to assess the efficacy of the classification system created using the datasets.

When evaluating the efficacy of a classification model, ACC is described as the ratio of correct assumptions to total assumptions made in equation (3):

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \tag{3}$$

PR, also known as positive predictive value in equation (4), measures the proportion of correctly detected positive examples to altogether positive examples:

$$Precision = \frac{TP}{TP+FP} \tag{4}$$

RC, also known as rate or sensitivity, is the percentage of appropriately identified positive cases among all positive instances, as shown in equation (5).

$$Recall = \frac{TP}{TP+FP} \tag{5}$$

The F1 in equation (6) is an integrated metric that incorporates PR and RC into a single numerical value:

$$F1 = \frac{Precision*Recall}{Precision+Recall} \tag{6}$$

Classification Analysis

Table 3. Classification Analysis with Existing Models and Proposed Model

Models	ACC (%)	PR (%)	RC (%)	F1 (%)
AdaBoost	91.1	90.2	91.0	90.9
XGBoost	92.3	91.4	91.8	91.6
GBM	93.4	92.6	92.6	92.4
Random Forest	94.2	93.5	93.7	94.6
Proposed TPE based Ensemble Classification model	95.8	94.6	95.2	95.3

From **Table 3** and **Fig 2,** performance metrics for different machine learning models are shown. AdaBoost (ACC: 91.1, PR: 90.2, RC: 91.0, F1: 90.9), XGBoost (ACC: 92.3, PR: 91.4, RC: 91.8, F1: 91.6), Gradient Boosting Machines (GBM) (ACC: 93.4, PR: 92.6, RC: 92.6, F1: 92.4), Random Forest (ACC: 94.2, PR: 93.5, RC: 93.7, F1: 94.6), and the Proposed Tree-Structured Parzen Estimator (TPE) based Ensemble Classification model (ACC: 95.8, PR: 94.6, RC: 95.2, F1: 95.3). These models were assessed according to how well they performed on each of the designated metrics. The consequences showed that the projected TPE based Ensemble Classification model was the best; it had the highest values on every metric and was more predictive than the other models that were looked at.

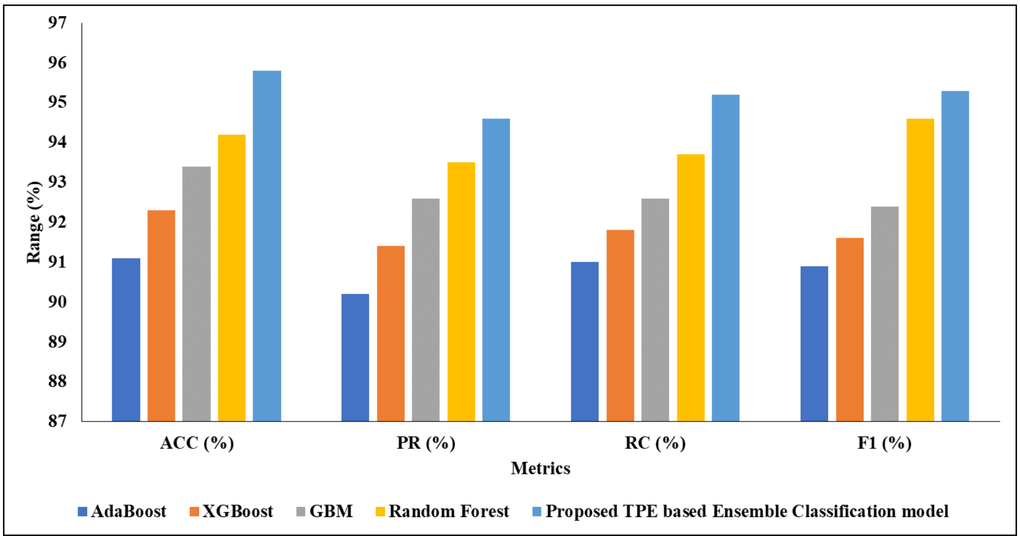

Fig 2. Classification Analysis

Table 4. Classification Analysis with existing models and proposed model

Models	ACC (%)	PR (%)	RC (%)	F1 (%)
AdaBoost	97.2	96.1	95.6	95.4
XGBoost	97.8	96.2	95.9	96.3
GBM	98.3	97.3	96.4	97.1
Random Forest	98.7	97.4	97.1	97.9
Proposed TPE based Ensemble Classification model	**99.4**	**98.5**	**98.4**	**98.6**

Table 4 and **Fig 3** shows the performance metrics for different machine learning models. AdaBoost's results included a 97.2% accuracy rate, 96.1% precision rate, 95.6% recall rate, and 95.4% F1-score. With an accuracy of 97.8%,

precision of 96.2%, recall of 95.9%, and an F1-score of 96.3%, XGBoost demonstrated its performance. GBM showed a 98.3% accuracy, a 97.3% precision, a 96.4% recall, and a 97.1% F1-score. With a 98.7% accuracy rate, 97.4% precision rate, 97.1% recall rate, and 97.9% F1-score, Random Forest performed better. With an accuracy of 99.4%, precision of 98.5%, and an F1-score of 98.6%, the suggested TPE-based Ensemble Classification model showed exceptional performance. Together, these metrics show how effective and competitively successful these models are in the given soil classification, which supports smart farming practices.

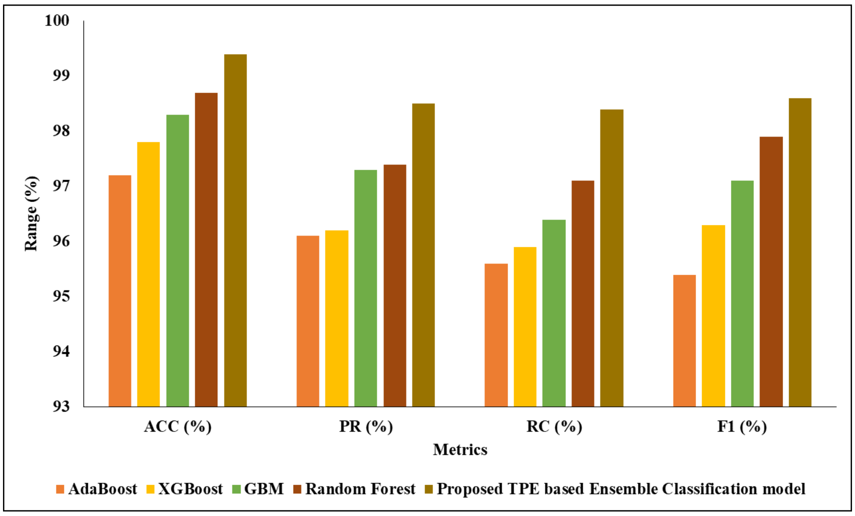

Fig 3. Classification Analysis with TPE

V. CONCLUSION

The convergence of IoT and ML in the context of smart farming offers a promising path toward improving diverse agricultural operations. The integration of these technologies, which is the main focus of this paper, represents a significant advancement in smart farming practices. This study highlights the importance of managing irregularities and minute data changes, which are essential for accurate analysis and decision-making in agricultural contexts. It does this by utilizing data from well-known repositories like UCI or Kaggle. Through the implementation of an IO technique for efficient data preprocessing, this study aims to reduce complexity and guarantee the production of a clean dataset that can be used to derive insightful agricultural insights. A new approach in this field is illustrated by the subsequent classification utilizing an Ensemble approach that combines the PNN, CBDT and ANFIS techniques, all combined via an averaging mechanism. In addition, the implementation of the TPE for hyperparameter tuning in this ensemble classifier represents a novel attempt at maximum efficiency. The final results indicate significant improvements in accuracy highlighting the effectiveness and promise of the suggested TPE-based Ensemble classification for intelligent farming towards more accurate and dependable agricultural decision-making procedures. Accuracy, precision, recall, and f-measure are all attained by the suggested model, which stands at 99.4 percent. Findings indicate that the suggested model outperforms the state-of-the-art models. Nevertheless, computing the suggested model takes more time, hence dimensionality reduction is necessary. Deep learning may be used to further enhance this model.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interests

The author(s) declare(s) that they have no conflicts of interest.

Funding

No funding agency is associated with this research.

Competing Interests

There are no competing interests.

References

[1]. R. Alfred, J. H. Obit, C. P.-Y. Chin, H. Haviluddin, and Y. Lim, "Towards Paddy Rice Smart Farming: A Review on Big Data, Machine Learning, and Rice Production Tasks," IEEE Access, vol. 9, pp. 50358–50380, 2021, doi: 10.1109/access.2021.3069449.

[2]. E. M. B. M. Karunathilake, A. T. Le, S. Heo, Y. S. Chung, and S. Mansoor, "The Path to Smart Farming: Innovations and Opportunities in Precision Agriculture," Agriculture, vol. 13, no. 8, p. 1593, Aug. 2023, doi: 10.3390/agriculture13081593.

[3]. S. I. Saleem, S. R. M. Zeebaree, D. Q. Zeebaree and A. M. Abdulazeez, "Building Smart Cities Applications based on IoT Technologies: A Review," Technology Reports of Kansai University, vol. 62, no. 3, pp. 1083-1092, 2020.

[4]. A. Zervopoulos et al., "Wireless Sensor Network Synchronization for Precision Agriculture Applications," Agriculture, vol. 10, no. 3, p. 89, Mar. 2020, doi: 10.3390/agriculture10030089.

[5]. D. R. Vincent, N. Deepa, D. Elavarasan, K. Srinivasan, S. H. Chauhdary, and C. Iwendi, "Sensors Driven AI-Based Agriculture Recommendation Model for Assessing Land Suitability," Sensors, vol. 19, no. 17, p. 3667, Aug. 2019, doi: 10.3390/s19173667.

[6]. W.-S. Kim, W.-S. Lee, and Y.-J. Kim, "A Review of the Applications of the Internet of Things (IoT) for Agricultural Automation," Journal of Biosystems Engineering, vol. 45, no. 4, pp. 385–400, Nov. 2020, doi: 10.1007/s42853-020-00078-3.

[7]. O. Friha, M. A. Ferrag, L. Shu, L. Maglaras, and X. Wang, "Internet of Things for the Future of Smart Agriculture: A Comprehensive Survey of Emerging Technologies," IEEE/CAA Journal of Automatica Sinica, vol. 8, no. 4, pp. 718–752, Apr. 2021, doi: 10.1109/jas.2021.1003925.

[8]. H. K. Adli et al., "Recent Advancements and Challenges of AIoT Application in Smart Agriculture: A Review," Sensors, vol. 23, no. 7, p. 3752, Apr. 2023, doi: 10.3390/s23073752.

[9]. K. Paul et al., "Viable smart sensors and their application in data driven agriculture," Computers and Electronics in Agriculture, vol. 198, p. 107096, Jul. 2022, doi: 10.1016/j.compag.2022.107096.

[10]. M. Brambilla et al., "From Conventional to Precision Fertilization: A Case Study on the Transition for a Small-Medium Farm," AgriEngineering, vol. 3, no. 2, pp. 438–446, Jun. 2021, doi: 10.3390/agriengineering3020029.

[11]. D. Radočaj, M. Jurišić, M. Gašparović, I. Plaščak, and O. Antonić, "Cropland Suitability Assessment Using Satellite-Based Biophysical Vegetation Properties and Machine Learning," Agronomy, vol. 11, no. 8, p. 1620, Aug. 2021, doi: 10.3390/agronomy11081620.

[12]. S. Gokool et al., "Crop Monitoring in Smallholder Farms Using Unmanned Aerial Vehicles to Facilitate Precision Agriculture Practices: A Scoping Review and Bibliometric Analysis," Sustainability, vol. 15, no. 4, p. 3557, Feb. 2023, doi: 10.3390/su15043557.

[13]. S. Parez, N. Dilshad, N. S. Alghamdi, T. M. Alanazi, and J. W. Lee, "Visual Intelligence in Precision Agriculture: Exploring Plant Disease Detection via Efficient Vision Transformers," Sensors, vol. 23, no. 15, p. 6949, Aug. 2023, doi: 10.3390/s23156949.

[14]. S. B. Kasturi, CH. Ellaji, D. Ganesh, K. Somasundaram and B. Sreedhar, "IoT and Machine Learning Approaches for Classification in Smart Farming," Journal of Survey in Fisheries Sciences, vol. 10, no. 4S, pp. 3373-3385, 2023.

[15]. A. Haldorai, B. Lincy R, S. M, and M. Balakrishnan, "An improved single short detection method for smart vision-based water garbage cleaning robot," Cognitive Robotics, vol. 4, pp. 19–29, 2024, doi: 10.1016/j.cogr.2023.11.002.

[16]. A. Gupta and P. Nahar, "Classification and yield prediction in smart agriculture system using IoT," Journal of Ambient Intelligence and Humanized Computing, vol. 14, no. 8, pp. 10235–10244, Jan. 2022, doi: 10.1007/s12652-021-03685-w.

[17]. H. Alshahrani et al., "Chaotic Jaya Optimization Algorithm With Computer Vision-Based Soil Type Classification for Smart Farming," IEEE Access, vol. 11, pp. 65849–65857, 2023, doi: 10.1109/access.2023.3288814.

[18]. P. Kathiria, U. Patel, S. Madhwani, and C. S. Mansuri, "Smart Crop Recommendation System: A Machine Learning Approach for Precision Agriculture," Machine Intelligence Techniques for Data Analysis and Signal Processing, pp. 841–850, 2023, doi: 10.1007/978-981-99-0085-5_68.

[19]. Y. Akkem, S. K. Biswas, and A. Varanasi, "Smart Farming Monitoring Using ML and MLOps," Lecture Notes in Networks and Systems, pp. 665–675, 2023, doi: 10.1007/978-981-99-3315-0_51.

[20]. I. V. Mboweni, D. T. Ramotsoela, and A. M. Abu-Mahfouz, "Hydraulic Data Preprocessing for Machine Learning-Based Intrusion Detection in Cyber-Physical Systems," Mathematics, vol. 11, no. 8, p. 1846, Apr. 2023, doi: 10.3390/math11081846.

[21]. S. R. B. S, M. G, and E. Sherly, "Kidney Stone Detection from CT images using Probabilistic Neural Network(PNN) and Watershed Algorithm," 2023 International Conference on Advances in Intelligent Computing and Applications (AICAPS), Feb. 2023, doi: 10.1109/aicaps57044.2023.10074562.

[22]. R. Nagi and S. S. Tripathy, "Plant disease identification using fuzzy feature extraction and PNN," Signal, Image and Video Processing, vol. 17, no. 6, pp. 2809–2815, Jan. 2023, doi: 10.1007/s11760-023-02499-x.

[23]. A. K. Thakur, A. Mukherjee, P. K. Kundu, and A. Das, "Classification and Authentication of Induction Motor Faults using Time and Frequency Feature Dependent Probabilistic Neural Network Model," Journal of The Institution of Engineers (India): Series B, vol. 104, no. 3, pp. 623–640, Mar. 2023, doi: 10.1007/s40031-023-00872-5.

[24]. A. B. Tufail et al., "3D convolutional neural networks-based multiclass classification of Alzheimer's and Parkinson's diseases using PET and SPECT neuroimaging modalities," Brain Informatics, vol. 8, no. 1, Nov. 2021, doi: 10.1186/s40708-021-00144-2.

[25]. J. Liang et al., "Intelligent fault diagnosis of rotating machinery using lightweight network with modified tree-structured parzen estimators," IET Collaborative Intelligent Manufacturing, vol. 4, no. 3, pp. 194–207, Sep. 2022, doi: 10.1049/cim2.12055.

[26]. S. Watanabe, N. Awad, M. Onishi and F. Hutter, "Speeding Up Multi-Objective Hyperparameter Optimization by Task Similarity-Based Meta-Learning for the Tree-Structured Parzen Estimator," arXiv 2022, arXiv:2212.06751.