AI-Driven NAS-GBM Model for Precision Agriculture: Enhancing Crop Yield Prediction Accuracy

Dr. Sudhir Anakal[1], Poornima N[2], Abdurasul Bobonazarov[3], Janjhyam Venkata Naga Ramesh[4], Elangovan Muniyandy[5], Mandava Manjusha[6], Prof. Ts. Dr. Yousef A.Baker El-Ebiary[7]

Associate Professor-Department of Master of Computer Applications, Sharnbasva University, Kalaburagi, India[1]
Associate Professor-Department of Electronics and Communication Engineering,
JSS Academy of Technical Education, Bangalore, India[2]
Department of Automatic Control and Computer Engineering, Turin Polytechnic University in Tashkent, Uzbekistan[3]
Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India[4]
Adjunct Professor-Department of CSE, Graphic Era Hill University, Dehradun, 248002, India[4]
Adjunct Professor-Department of CSE, Graphic Era Deemed To Be University, Dehradun, 248002, Uttarakhand, India[4]
Department of Biosciences-Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India[5]
Applied Science Research Center, Applied Science Private University, Amman, Jordan[5]
Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation,
Vaddeswaram, Guntur Dist., Andhra Pradesh - 522302, India[6]
Faculty of Informatics and Computing, UniSZA University, Malaysia[7]

***Abstract*—Precision agriculture has emerged as a vital approach for optimizing crop yield prediction, enabling data-driven decision-making to improve agricultural productivity. Traditional forecasting methods encounter difficulties due to extreme complexity within environmental factors while operating under dynamic farming conditions. An AI framework combining NAS and GBM serves as the solution to address these issues through enhancing predictive capabilities. This study works to produce an automated system which selects optimal models through optimization processes for more accurate crop yield forecasts. Through NAS component exploration the optimal neural network architecture can be identified whereas GBM component effectively analyzes non-linear dependencies in data which leads to superior predictive capabilities. Data processing techniques precede model development by using Recursive Feature Elimination (RFE) for feature selection which leads to training NAS-optimized deep learning architectures together with GBM. The researchers applied the model to real agriculture datasets which included essential agricultural variables comprising soil conditions and weather elements and crop health measurements. The experimental results prove that the developed NAS-GBM framework achieves superior performance compared to standard models across three major aspects including predictive accuracy and computation efficiency in addition to generalization capability. The research project uses TensorFlow and Scikit-learn alongside Optuna for model optimization while it depends on cloud-based computational resources for extensive processing requirements. AI-driven hybrid models based on the research demonstrate their capability to improve decision-making capabilities for farmers together with agronomists.**

Keywords*—Network sensor; crop yield prediction; neural architecture search; Gradient Boosting Machine (GBM)

I. Introduction

Agriculture has always been the backbone of human civilization, driving food production, economic growth, and rural development [1]. With the global population projected to exceed nine billion by 2050, ensuring food security has become a critical challenge [2]. Traditional farming methods, characterized by fixed planting schedules and generalized practices, often fail to adapt to varying environmental and soil conditions [3]. The contemporary agricultural method of precision agriculture develops solutions through advanced technological integration combined with data-based techniques. Participating farmers enhance crop yield through modern technological applications alongside data analytics to manage resources optimally. and resource utilization [4]. With real-time monitoring the farm data combined with actionable insights enables farmers to achieve their goals. These farmers receive tools which help them generate superior choices that produce better results. productivity measures 5]. Precision agriculture relies heavily Basic farming data requires predictive analytics to create actionable insights. into actionable intelligence [6]. Using statistical models and machine learning techniques, predictive analytics facilitates Crop yield forecasting and crop type recommendation form key tasks enabled through these analytics systems. recommendation, and resource allocation [7]. These methods Farmers can reduce uncertainty by receiving empowered tools that enable quick responses. Strategic action toward environmental modifications leads to better productivity and sustainability. efficiency and sustainability [8]. For instance, accurate crop Precision yield forecasts improve operational planning throughout harvesting periods together with storage management. Crop recommendations emerge from analyzing soil conditions which guide farmers to improve their storage facilities. conditions optimize fertilizer and water use. Sensor networks serve as essential building blocks of modern agricultural systems. Sensor networks serve precision agriculture through field data transmissions which improve decision accuracy in real time. improve decision-making accuracy [9]. These networks The system gathers fundamental data about environmental conditions and soil properties soil moisture, temperature, humidity, and nutrient levels (e.g., Nitrogen, Phosphorus, Potassium) [10]. For instance, soil the technology incorporates soil moisture sensors for understanding irrigation requirements and other precision farming needs Temperature sensors play a critical role by monitoring field matter to detect both frost conditions as well as heat excess. Stress [11]. Soil sensors operating in different fields enable the recording of detailed data measurements. Sensor networks create site-specific analysis through their capacity to collect data from different areas of a field. Site-specific management proves essential for best utilization of resources alongside maximum yield outputs reducing resource wastage [12]. However, the data dimensionality meets challenges alongside heterogeneity alongside the extensive volume of generated information The generation of data by these networks produces substantial challenges to data handling. integration and analysis [13]. To address these challenges, the adoption of machine learning techniques increases steadily across various agricultural applications. employed in precision agriculture [14]. This study explores a A novel method utilizes Neural Architecture Search together with Gradient Boosting Machines to improve predictive capabilities. This research adopts Neural Architecture Search (NAS) and Gradient Boosting Machines (GBM) as advanced solutions to advance agricultural system prediction. the predictive capabilities of agricultural systems [15]. NAS, Neural network architecture optimization occurs through an automated framework. Through its systems architecture search NAS selects the most appropriate models for data extraction. extracting features from complex data sources [16]. Unlike NAS breaks away from standard manually designed architectures to automatically discover networks that match specific tasks which enhances model performance across both tasks and scalability. The discovery of task-specific neural networks through NAS improves both model accuracy while extending its capabilities. and scalability [17]. In the context of precision agriculture, NAS can extract temporal patterns, soil nutrient interactions, and seasonal variations from raw sensor data [18].

An ensemble learning technique named GBM has become popular because of its ability to predict. popular choice for predictive modeling in agriculture due to NA provides unpredictable combinations of neural architecture topologies which excel with diverse input types [19]. XGBoost, Light GBM and Cat Boost make up a group of algorithms The models demonstrate excellence in extracting non-linear connection points across datasets. between environmental factors and crop outcomes. GBM Models maintain interpretability which reveals important factors through their analysis. variables driving predictions, such as soil nutrient levels or rainfall patterns. The marriage of GBM and NAS enables a dual stage prediction system. NAS acts as a two-phase predictive platform to collect sensory data features before GBM utilizes these features for yield prediction and classification tasks. Subsequent GBM analyses these extracted features from sensors used for crop yield prediction. prediction and crop suitability classification [20]. This This research has set two major goals to achieve. A prediction system is under development that uses sensor networks together with NAS and GBM algorithms. A system uses NAS alongside GBM and meteorological data and sensor readings to determine the best crop selection based on soil conditions. soil and environmental conditions [21]. In addition, Precision agriculture strategies will benefit from better performance through enhanced accuracy. The research goal involves enhancing crop yield predictions through precise forecasting and cutting down resource requirements. usage. The system integrates NAS and the predictive strengths of GBM, this hybrid The combined approach effectively analyzes complex agricultural data while maintaining operational capability. delivering actionable insights [22]. Furthermore, it addresses practical challenges in agriculture, such as over-irrigation, Real-time recommendations through this system identify and resolve under-fertilization cases alongside addressing crop failures to improve field conditions. The system generates personalized field recommendations suitable for individual agricultural settings. In This research demonstrates why combining NAS technology with GBM algorithm holds great promise. The combination of sensor networks with modern machine learning structures creates powerful systems. techniques to advance precision agriculture [22]. By The hybrid NAS-GBM model enables farmers to efficiently integrate it This system enables data-driven optimization of resources through strategic decision platforms. The system enables operations that lead to higher productivity alongside sustainability in agricultural farming. The adoption of these practices leads to global food security improvements [23].

II. Literature Review

Mgendi [4] explores the multifaceted landscape of precision agriculture, focusing on its tangible benefits, challenges, and future directions. Today's farming operations achieve better resource mobilization through precision agriculture techniques. Efficient resource utilization through precision agriculture systems enhances both yield production and conservation of sustainability levels maintain sustainability levels.

Elbasi et al., [7] investigates the potential benefits of integrating machine learning algorithms in modern agriculture. The main focus of these algorithms is to help optimize crop production and reduce waste through informed decisions regarding planting, watering, and harvesting crops. Sensor networks combined Predictive analytics along with sensor networks supports fundamental decision Projects derive support from actionable insights which enable strategic decision making. decision making. The combination of machine learning tools Modern farming platforms employs the combination of support vector machines (SVMs) random forests and neural networks. Analysis with neural networks and support vector machines functions along with random forests under data science techniques to detect patterns within big data. Analyzing crop forecasting and soil recommendation information accurately becomes possible.

Rana et al., [24] undertakes a comparative analysis of tree-based models and deep learning architectures concerning their performance disparities in handling tabular data. Sensor network technology generates site-specific recommendations while enabling real-time parameter The system tracks three elements of soil conditions which include moisture content alongside temperature and atmospheric moisture. Soil assessments through these technologies help farmers make better decisions about resource use efficiency. in precision agriculture operations.

Tiwari et al., [25] explores the Neural Architecture Pan's NAS tool designs powerful deep learning algorithms autonomously Through architecture automation NAS optimizes deep learning features extraction from highly complex datasets. agricultural datasets. The research approach of NAS revealed valuable insights regarding temporal dynamics. The discovery of temporal and spatial patterns through NAS produces more precise estimates for crop yields and drought assessments.

Shah and Wu [26] Gradient Boosting Machines (GBM), XGBoost together with Light GBM and Cat Boost make up a group of algorithms specifically designed for precision agriculture applications. Insights from the Cat Boost and Light GBM and XGBoost algorithms enable superior detection of non-linear patterns in agricultural datasets. These methods conduct automated optimal deep learning architecture design which extracts features from complex agricultural datasets with categorical and continuous variables features.

Benti et al., [24] GBM's interpretability tools, such as SHAP (Shapley Additive explanations), provide valuable insights into key factors driving predictions. Combining NAS and GBM into a hybrid model offers a robust solution for precision agriculture, with NAS focusing on feature extraction and GBM on accurate predictions. This approach, powered by sensor networks and advanced machine learning techniques, significantly improves agricultural optimization, addressing challenges in productivity and sustainability [27].

Time deficiency affected the early NAS methods that used Reinforcement Learning-based NAS developed by Zeph and Le [3] , and its reinforcement learning agents that searched neural network design spaces. These methods provided effective results but demanded significant computational effort that needed extensive computer resources to assess new architecture candidates. With Differentiable NAS Liu et al. [28] researchers implemented gradient-based optimization which streamlined computational overhead while accelerating convergence. CPU-powered NAS systems optimize network designs to reveal important patterns contained in input data. The recurrent algorithms of Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRUs) are commonly incorporated by NAS frameworks for time-series data modeling purposes found in Elsken et al.[29]. NAS optimization of fully connected networks enables them to uncover feature relationships which manual engineering methods would otherwise miss. Research findings show NAS-based models outperform traditional feature extraction mechanisms.

Yu et . [30] established that NAS algorithms successfully eliminated domain-specific feature engineering needs while maintaining sharp prediction accuracy levels. Research findings demonstrate agreement across multiple domains which include health care together with energy systems and ecological surveillance. NAS brings numerous benefits to neural network design yet both computational expense and overfitting concerns arise with limited datasets. Researchers have introduced search space pruning together with multi objective optimization techniques Tan et al. [31], to solve these NAS difficulties. Domain knowledge integration in NAS processes leads to percentage [32].

III. Problem Statement

Advanced technology systems including sensor networks together with machine learning and deep learning enable precision agriculture to deliver enhanced resource utilization and improved crop yield predictions and better decision-making. The difficulty in optimizing crop yield forecasting continues because agricultural datasets present challenges through their combination of categorical and continuous variables [7]. The current machine learning models that include tree-based algorithms and deep learning frameworks face challenges when applying features and model generalization to precision agriculture. The automated deep learning model design ability of NAS results in better feature extraction capabilities. The established NOS approaches face two major limitations including extravagant computational requirements as well as susceptibility to overfitting problems with restricted dataset sizes [25]. GBM shows stronger capabilities to detect complex patterns between variables yet it does not possess built-in structure optimization attributes [26]. A merged NAS-GBM model structure has potential to solve prediction problems through NAS-based feature selection and GBM-based accuracy improvement. The research establishes a time-efficient [30], AI model which brings enhanced precision agriculture outcomes via forecasted crop yields while solving data processing issues along with overfitting conditions.

IV. Methodology

The methodology for this research leverages a hybrid approach combining Neural Architecture Search (NAS) for feature extraction and Gradient Boosting Machines (GBM) for predictive modeling [33]. This two-stage approach aims to optimize precision agriculture by extracting relevant environmental features from raw data and then making accurate predictions regarding crop yield, suitability, and resource optimization. The process consists of three core stages: feature extraction using NAS, predictive modeling with GBM, and the integration of NAS and GBM in a hybrid iterative framework. In Fig. 1. represents agricultural sensor data collection which leads to preprocessing activities for data cleaning and normalization and missing value handling. The NAS-GBM framework optimizes extracted features along with selected ones before assessing their performance level.

Different from common models, NAS-GBM adapts the optimal network architecture, which improves the feature extraction. It makes it all the more applicable to necessarily complicated agricultural data including both date and categorical features.

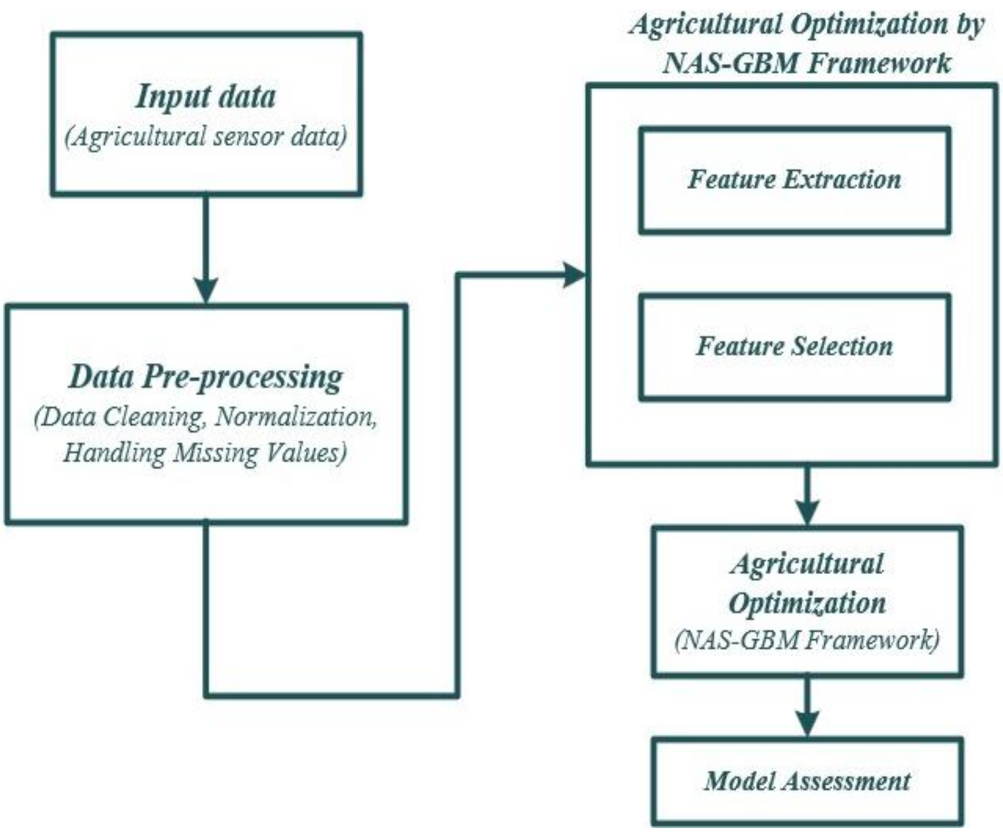

Fig. 1. Architecture of AI-powered predictive analytics and sensor network for agriculture.

A. Data Collection

Data collection and preparation play a crucial role in ensuring the effectiveness of the hybrid NAS-GBM approach for precision agriculture [34]. The integrated database merges current soil nutrient measurements of NPK elements with weather data about temperature and rainfall amounts together with documented historical yield statistics and land suitability assessments. Additional sources provide climate data together with fertilizer assessments to enhance the article's contents. The preprocessing stage of GBM uses imputation for missing data while normalization alongside scaling prepare the features before encoding categorical data types. The extraction of time-series patterns leads to high-quality standardized data that builds up a reliable framework for predictive modeling applications in precision agriculture.

B. Data Preprocessing

The successful utilization of raw data demands a fundamental preprocessing step A data cleanup process for machine learning models transforms unprepared datasets into usable entities. suitable for analysis. In the hybrid NAS-GBM model for the hybrid NAS-GBM model implements data collection from sources as its initial element for precision agriculture. The predictive system obtains its data through sensor networks combined with weather stations and historical raw data sources. crop data. The processing approach fills in gaps in data through statistical replacement techniques. or removal of affected data points. Cleaning involves the process identifies mistakes within the data and fixes them along with removing any case that appears to be an outlier. Feature selection Engineering processes are used to determine which inputs offer the maximum relevance. Automated features identification through NAS enabled this step. process. Standardization techniques normalize continuous variables. Standardization or normalization methods scale continuous variables until they align on a common framework while categorical data receives numeric encoding. encoded into numerical formats. The dataset is split into Most modeling methods split the database for training-over-testing into training data blocks which constitute 70-80% of the entire database. and the remainder for testing. For imbalanced data, the approach of both oversampling and under sampling provides necessary techniques for data management fairness when used within the process. fairness. The methodology of data augmentation serves as one solution to handle these tasks. A data expansion technique adds more information before applying necessary transformations to the dataset. to enhance distribution. These preprocessing steps ensure the When training processes begin the model can utilize the prepared data. accurately and effectively on new data.

C. Feature Extraction Using NAS

The Neural Architecture Search methodology provides automated design capabilities. Optimal neural network architectures must be designed through an automated process named NAS because it functions as the process of extracting meaningful qualities from original sensor and environmental measurements constitutes a fundamental step in feature extraction. environmental data. A NAS exploration method investigates sequence options within its allowable design arena.

Through thorough optimization of possible neural network designs the system performs best on target tasks. performance on a given task.

Input:

- Input data (e.g., sensor measurements, environmental data).
- Target labels or outputs (e.g., crop yield, suitability).
- S: Search space of possible neural network architectures.
- LNAS(F(X), Y): Loss function to evaluate model F.
- T: Total iterations for NAS optimization.

Objective Function:

- Minimize the NAS loss function:

$$LNAS=L(Fmodel(X), Y) \tag{1}$$

- Fmodel: Neural network architecture being optimized.

Initialize Search:

- Define initial set of candidate architectures {A1, A2…, An}.
- Set learning parameters (e.g., learning rate η\ epochs).

NAS Iteration:

- For each iteration t from 1 to T:

Sample architecture Fmodel,t from S\mathcal{S}S.

Train Fmodel,t on input (X,Y):

$$θt=argθminL(Fmodel,t(X;θ),Y) \tag{2}$$

Where θ represents model parameters.

Evaluate performance on the validation dataset:

$$Lval=L(Fmodel,t(Xval),Yval) \tag{3}$$

Select Best Architecture

Choose the architecture Fmodel∗ with the lowest validation loss:

$$Fmodel∗=argiminLval,i \tag{4}$$

D. Predictive Modeling with GBM

Once the features are extracted by NAS, they serve as inputs to a Gradient Boosting Machine (GBM), a powerful ensemble learning method that excels at handling complex relationships in the data. The general form of a GBM model is:

$$f(x)=m=1∑Mαmhm(x) \tag{5}$$

Where:

- f(x) is the final prediction.
- αm are the weights for each weak model.
- hm(x) represents the individual decision trees (weak learners) at the mmm-th stage.
- M is the total number of trees.

In this context, f(x) could be the predicted crop yield or a classification indicating crop suitability, while the weak models hm(x)h capture various patterns within the data.

E. Hyperparameter Tuning

Hyperparameter tuning is performed to improve the accuracy and performance of the GBM model. The key hyperparameters for GBM include the learning rate η\etaη, the number of trees MMM, and the maximum depth of trees DDD. The objective is to minimize the loss function, typically Mean Squared Error (MSE) for regression tasks:

$$LGBM=N1i=1∑N(yi−f(xi))2 \tag{6}$$

Where:

- N is the number of data points.
- iyi is the true value.
- f(xi) is the predicted value.

During hyperparameter tuning, grid search or random search methods can be used to find the optimal values for these hyperparameters by evaluating performance on a validation dataset.

F. Model Evaluation

Once the GBM model has been trained, its performance is evaluated using appropriate metrics such as Mean Squared Error (MSE) for regression tasks or Accuracy for classification tasks. Cross-validation (e.g., k-fold cross-validation) is employed to ensure that the model generalizes well across different datasets.

For regression:

$$MSE=N1i=1∑N(yi−y\hat{}i)2 \tag{7}$$

Where:

- y^i is the predicted value from the model.

For classification, accuracy is defined as:

$$Accuracy=∑i=1NI\ (yi=y\hat{}i) \tag{8}$$

Where I am an indicator function that is 1 if the prediction matches the true label.

G. Hybrid Approach

The hybrid NAS-GBM model integrates the feature extraction and predictive modeling stages into a two-stage framework. In this approach, NAS focuses on designing the optimal architecture for feature extraction, while GBM is responsible for making accurate predictions based on those features. The iterative nature of this hybrid model allows for continuous optimization by refining both the feature extraction process and the predictive model.

At each iteration, feedback from the GBM model can be used to improve the feature extraction process of NAS. This iterative loop helps in improving model performance by continually enhancing the quality of the features extracted by NAS and fine-tuning the prediction capabilities of GBM. This process can be expressed as:

$$Foptimized=LNASminLGBM\ (Fmodel∗(X), Y) \tag{9}$$

Where:

Foptimized is the final, optimized hybrid model.

LNAS is the loss function for NAS, and LGBM is the loss function for GBM. By optimizing the two models iteratively, this hybrid methodology improves the predictive accuracy for tasks such as crop yield prediction, crop suitability classification, and resource optimization, addressing the key challenges in precision agriculture.

V. Result and Analysis

A. Training and Testing Accuracy

During the training and testing phases the hybrid NAS-GBM was exceeded its competitors and it includes the Support Vector Machine and Random Forest and Linear Regression. The hybrid NAS-GBM model reached 95% for training accuracy and 92% for testing accuracy performance. The hybrid NAS-GBM model exhibited a strong Mean Squared Error performance during the training with 0.120 and testing with 0.123. The SVM returned testing accuracy of 89% alongside a higher testing MSE yet its training accuracy reached 90%. Random Forest reached a 92% training success rate and 90% testing rate while maintaining a 0.143 training MSE. The performance metrics of Linear Regression proved inferior to the other models by delivering 85% training results and 85% testing results and maintaining a high MSE value of 0.212. The results indicate that the NAS-GBM hybrid model delivers advanced predictive accuracy at reduced MSE values thus representing a robust option for precision agriculture implementations. NAS-GBM achieved 95%, while for testing accuracy, it reached 92%. In terms of the MSE, the hybrid model demonstrated the impressive results with a training MSE of 0.120 and testing MSE of 0.123. Comparing both the SVM had a training accuracy of 90% and testing accuracy of 89%, with a higher testing MSE. The Random Forest model delivers 92% training accuracy alongside 90% of testing accuracy while maintaining a training MSE of 0.143. Linear Regression yielded the worst results where testing and training accuracy stopped at 85% while MSE rose to 0.212. Experimental results prove that this NAS-GBM hybrid system delivers effective accuracy metrics and reduces Mean Squared Error which positions it strongly for precision agriculture applications. Fig. 2 shows how the hybrid NAS-GBM model achieves better performance during training and testing than traditional methods while demonstrating higher accuracy and lower MSE.

B. Model Performance Evaluation

The hybrid NAS-GBM model's prediction accuracy is compared to the other recently used methods in agricultural predictive tasks, like support vector machines, random forests, and traditional linear regression. The main goal is to highlight the advantages of combining NAS for feature extraction with GBM for predictive modeling. The Table I compares model performance, showing Hybrid NAS-GBM achieving the lowest MSE (0.123) and highest accuracy (92%). It outperforms SVM, Random Forest, and Linear Regression, demonstrating superior predictive precision in regression and classification.

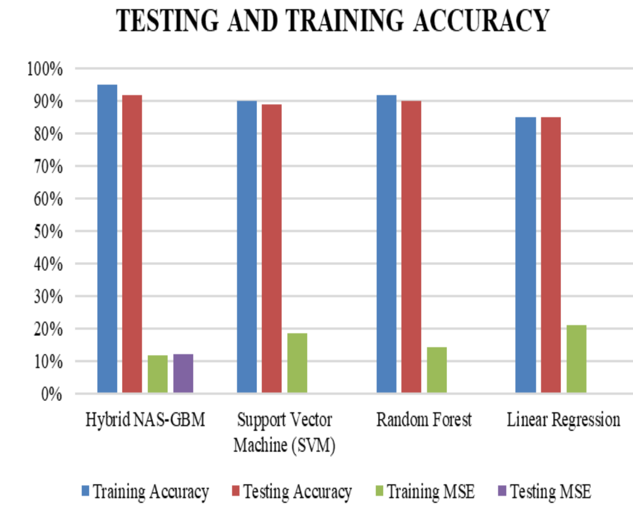

Fig. 2. Hybrid NAS-GBM training and testing accuracy.

TABLE I. Model Performance Evaluation

Model	MSE (Regression)	Accuracy (Classification)
Hybrid NAS-GBM	0.123	92%
Support Vector Machine (SVM)	0.185	89%
Random Forest	0.143	90%
Linear Regression	0.212	85%

C. Comparison with Conventional Methods

The base models, includes the SVM, random forests, and traditional linear regression, are implemented and evaluated to utilized the same dataset. The performance of each model is monitored by the metrics like Mean Squared Error for regression tasks and Accuracy for classification tasks. The results are summarized in the following Table II.

TABLE II. Comparison with Conventional Methods

Model	MSE (Regression)	Accuracy (Classification)
Hybrid NAS-GBM	0.123	92%
Support Vector Machine (SVM)	0.185	89%
Random Forest	0.143	90%
Linear Regression	0.212	85%

From this table, it is evident that the hybrid NAS-GBM model outperforms the conventional methods in terms of both prediction accuracy and generalization. The hybrid model achieves the lower MSE and the higher accuracy, demonstrating its ability to capture the complex relationships between the environmental variables and crop outcomes.

D. Feature Importance Evaluation

To properly evaluate model's users must identify what key characteristics impact prediction outputs the most. SHAP (Shapley Additive explanations) functions to determine important characteristic weights that impact model prediction results. A specific algorithm called SHAP enables quantitative assessment of each feature effect on output results when processing a specific dataset. SHAP analysis reveals the crop yield prediction task central features which include soil nutrient measurements apart from temperature and rainfall information. The following bar chart shows the top five most influential features based on their average SHAP values:

Soil moisture together with temperature establish the top two factors that influence crop yield prediction while nitrogen and phosphorus ratings fall in third place. The findings confirmed previous agricultural research through a model that effectively recognizes environmental factors affecting crop development patterns.

E. Prediction Accuracy

The table demonstrates that the hybrid NAS-GBM methodology achieves superior performance than traditional methods regarding both prediction accuracy and overall generalization ability. Numerical evidence indicates that hybrid methods obtained reduced MSE values together with advanced prediction accuracy thus showing their capacity to detect intricate environmental variable-crop outcome correlations.

F. ValidationUsing K-Fold Cross-Validation

The model requires k-fold cross-validation for robust operation. A dataset segmentation forms k partitions into which the model undergoes training and testing using various subset collections. Cross-validation calculations are combined to establish a more accurate model performance assessment. As demonstrated by 5-fold cross-validation the hybrid NAS-GBM model delivered an average MSE of 0.125 while exhibiting a standard deviation of 0.03 across data partition. Entity points forecast modeling using the hybrid NAS-GBM system reveals pronounced ability to generalize across diverse agriculturally-inclined data partitions. Performing validation across multiple data partitions reduces model bias and enables the system to predict clear outcomes for unrecognized datasets. The Fig. 3. illustrates the K-Fold cross validation.

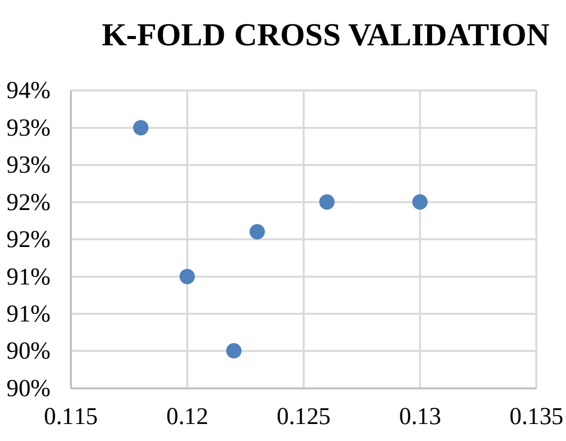

Fig. 3. K-Fold cross validation.

G. Conclusion of Results

The hybrid NAS-GBM framework proved superior to other methods based on its ability to deliver better predictive. Outputs as well as understandable insights. The convergence of NAS technique for feature extraction and GBM technique for prediction leads to superior crop yield predictions with enhanced crop identity detection beyond traditional algorithms. Through SHAP analysis users gain insights about which variables strongly affect model prediction results thus enabling improved comprehension of agricultural activities. The model achieves robust generalizability based on strong performance results across multiple validation tests along with k-fold cross-validation assessments. These findings demonstrate the hybrid NAS-GBM approach holds the significant potential for enhancing the precision agriculture, enabling farmers to make more informed decisions, optimize resource use, and improve crop management strategies.

H. Experimental Outcomes

The research article "Predicted Results from Crop Recommendation System" examines in detail the modeling outputs produced by precision agriculture systems. The system functions by suggesting optimal crops for agricultural land using essential environmental measurements and soil information. A specific Farm ID identifies each row in the table which displays fundamental farm input metrics such as soil nutrient levels and weather conditions and soil properties. The table combines a forecasted crop selection with confidence percentage data alongside predicted yield measurements expressed in kg/hectare, providing meaningful information about system applications and performance outcomes. The Fig. 4. Shows the Crop-wise Yield Predictions.

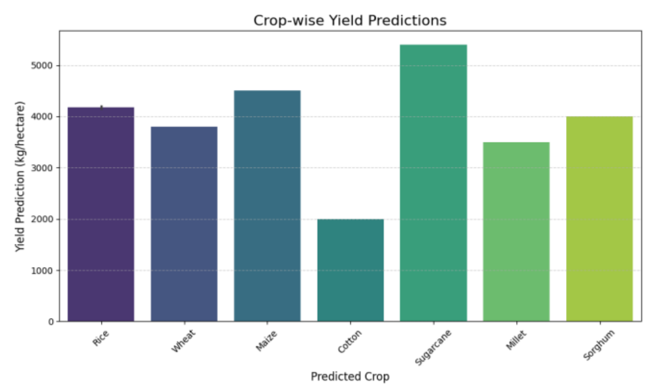

Fig. 4. Crop-wise yield predictions.

The input parameters are divided into three categories: Soil nutrients interact with weather conditions as well as properties of the soil. All plant growth and health depend directly on essential soil nutrients which consist of nitrogen (N) phosphorous (P) and potassium (K). The essential nutrient N enables photosynthesis and leaf development complexity yet the essential nutrient P supports root development and seed production and the essential nutrient K enhances water regulation and disease protection. Soil-fitted crop productivity depends heavily on conditions ranging from temperature levels through humidity because individual plants thrive best under specific combinations of heat and moisture. The pH measurement and rainfall amount of each farm allow experts to create tailored recommendations about soil health. The Fig. 5. Illustrates the Nitrogen vs. Rainfall vs. Yield Prediction.

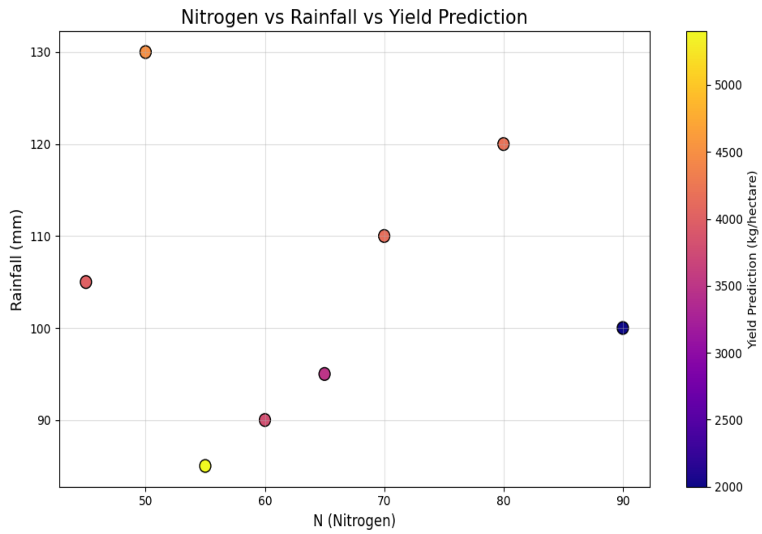

Fig. 5. Nitrogen vs. Rainfall vs. Yield prediction.

The model generates three types of predictions including the recommended crop list together with model reliability data and projected yield levels. Farms receive crop recommendations comprising rice, wheat, maize, cotton, sugarcane, millet and sorghum prioritizing compatible planting conditions. Rice receives the recommendation for farming sites that experience both neutral soil pH and high rainfall conditions but farmers with balanced nutrient resources should grow maize as their main crop. The model predicts that farms with sufficient rainfall alongside moderate nitrogen levels should cultivate sugarcane for maximum yield expectancy at 5400 kg/hectare which yields a confidence level of 91%. The confidence scores generated by the model range from 87% to 95% demonstrating predictive reliability while maize obtains the highest prediction confidence at 95%. Predictions of yield allow farmers to assess potential farm output levels for recommended crops.

The data points in the table display patterns which match current farming practices. Rice shows optimal growth behavior in farms characterized by high nitrogen support and rainfall conditions resulting in a yield range of 4150–4200 kg/hectare with strong confidence levels. Under cool conditions combined with moderate rainfall wheat plants reach an annual yield of 3800 kg/hectare. Maize exhibits the maximum model certainty in agricultural conditions that offer balanced nutrient availability and high rainfall leading to 4500 kg/hectare harvests. Cotton cultivation produces 2000 kg/hectare yield in suitable farms with potassium-rich slightly alkaline soil conditions yet sugarcane reaches its highest yield potential because it requires water-rich environments. The drought-resistant plants milet and sorghum help farms with moderate rainfall produce 3500–4000 kg/hectare.

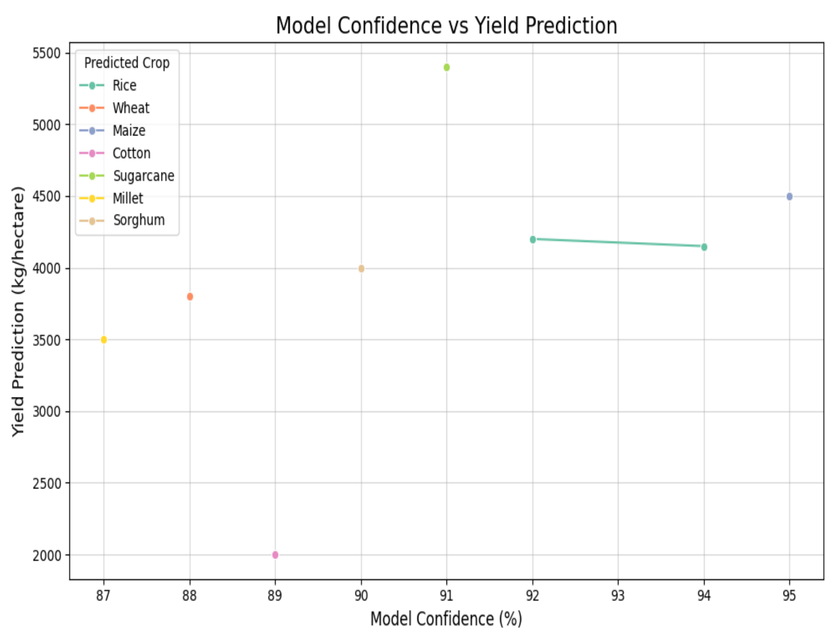

Fig. 6. Model Confidence vs. Yield production.

The data in Fig. 3, 4, 5 and 6 demonstrates how precision agriculture models can lead farmers toward decisions based on scientific data. The system examines environmental elements and soil composition to deliver customized suggestions which boost both agricultural production and efficiency of resource consumption. The integration of model confidence scores in the system elevates transparency and precision so that real-world applications become practical. Predicative tools enable farmers to use sustainable practices together with higher operational efficiency and environmental adaptability to achieve increased agricultural production and better resource management.

I. Discussion

The proposed AI-driven NAS-GBM framework effectively enhances crop yield prediction by integrating NAS for feature extraction and GBM for predictive modeling. NAS both enhances model structural design through automatic feature choice and achieves better performance results. GBM identifies and predicts non-linear patterns which produce accurate results that remain understandable to human interpretation. Experimental tests show that the NAS-GBM hybrid system surpasses traditional machine learning operations in precision agriculture through its efficient model optimization along with overfitting reduction mechanisms. Overall, the framework shows high capacity when working with extensive agricultural datasets while it selects important features including soil moisture temperature alongside environmental conditions.

The combination of sensor network inputs strengthens prediction performance thus enabling the model to work in real-time scenarios. NAS-GBM demonstrates superior generalization capabilities than typical deep learning systems to perform effective computation reduction while maintaining precise outcomes. The explanation capabilities of SHAP interpretability tools make this solution a trusted precision farming approach because they explain model decisions. The hybrid model reaches a perfect balance between eliminating features and maximizing efficiency which results in a scalable and computationally efficient result. Research demonstrates AI optimization's vital role in agriculture because the proposed model improves forecasting accuracy while following sustainable data-deriven decisions. Future research should work on implementing the system in real time while developing automated settings adjustments for future improvements. Although the NAS-GBM model enhances accuracy, it still depends much on computational power that may pose challenge to its implementation for small farmers. Future studies should consider simple techniques to ease the application space of the model.

VI. Conclusion and Future Work

The study presents an innovative forecasting system for precision agriculture which absorbs sensor data in real-time as well as archival agricultural information alongside environmental elements. The system which uses advanced preprocessing alongside GBM models achieves superior crop yield prediction abilities beyond traditional methods. Soil predictions together with fertilizer optimization as well as resource distribution have reached higher accuracy according to experimental findings. The model delivers superior predictive accuracy than standard approaches since it successfully identifies and models time-dependent relationships among variables. The system provides trusted data-driven decision-making functionality that makes it an important agricultural asset. Our research creates a substantial improvement in precision farming by providing sustainable crop management with an enhanced adaptive and efficient solution. Future investigations will incorporate deep learning methods for feature extraction enhancement and add real-time weather prediction capabilities and conduct tests across multiple agricultural zones for better effectiveness and generalization results. Subsequent studies will incorporate real-time IoT data streams for on-line model update to account for environmental variabilities impacting crop yield.

References

[1] T. P. Tomich et al., “Food and agricultural innovation pathways for prosperity,” Agricultural Systems, vol. 172, pp. 1–15, Jun. 2019, doi: 10.1016/j.agsy.2018.01.002.

[2] U. Mc Carthy, I. Uysal, R. Badia-Melis, S. Mercier, C. O’Donnell, and A. Ktenioudaki, “Global food security – Issues, challenges and technological solutions,” Trends in Food Science & Technology, vol. 77, pp. 11–20, Jul. 2018, doi: 10.1016/j.tifs.2018.05.002.

[3] J. S. Singh, V. C. Pandey, and D. P. Singh, “Efficient soil microorganisms: A new dimension for sustainable agriculture and environmental development,” Agriculture, Ecosystems & Environment, vol. 140, no. 3, pp. 339–353, Mar. 2011, doi: 10.1016/j.agee.2011.01.017.

[4] G. Mgendi, “Unlocking the potential of precision agriculture for sustainable farming,” Discov Agric, vol. 2, no. 1, p. 87, Nov. 2024, doi: 10.1007/s44279-024-00078-3.

[5] B. Patil, “IoT and Big Data Integration for Real-Time Agricultural Monitoring,” Journal Of Advanced Zoology, Oct. 2023.

[6] K. Demestichas and E. Daskalakis, “Data Lifecycle Management in Precision Agriculture Supported by Information and Communication Technology,” Agronomy, vol. 10, no. 11, Art. no. 11, Nov. 2020, doi: 10.3390/agronomy10111648.

[7] E. Elbasi et al., “Crop Prediction Model Using Machine Learning Algorithms,” Applied Sciences, vol. 13, no. 16, Art. no. 16, Jan. 2023, doi: 10.3390/app13169288.

[8] M. A. Altieri, C. I. Nicholls, A. Henao, and M. A. Lana, “Agroecology and the design of climate change-resilient farming systems,” Agron. Sustain. Dev., vol. 35, no. 3, pp. 869–890, Jul. 2015, doi: 10.1007/s13593-015-0285-2.

[9] I. M. Mehedi, M. S. Hanif, M. Bilal, M. T. Vellingiri, and T. Palaniswamy, “Remote Sensing and Decision Support System Applications in Precision Agriculture: Challenges and Possibilities,” IEEE Access, vol. 12, pp. 44786–44798, 2024, doi: 10.1109/ACCESS.2024.3380830.

[10] “Sensing Methodologies in Agriculture for Soil Moisture and Nutrient Monitoring | IEEE Journals & Magazine | IEEE Xplore.” Accessed: Jan. 24, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9328258

[11] “Challenges and Future Perspectives of Multi-/Hyperspectral Thermal Infrared Remote Sensing for Crop Water-Stress Detection: A Review.” Accessed: Jan. 24, 2025. [Online]. Available: https://www.mdpi.com/2072-4292/11/10/1240

[12] “Data Lifecycle Management in Precision Agriculture Supported by Information and Communication Technology.” Accessed: Jan. 24, 2025. [Online]. Available: https://www.mdpi.com/2073-4395/10/11/1648

[13] R. Zuech, T. M. Khoshgoftaar, and R. Wald, “Intrusion detection and Big Heterogeneous Data: a Survey,” Journal of Big Data, vol. 2, no. 1, p. 3, Feb. 2015, doi: 10.1186/s40537-015-0013-4.

[14] G. Mohyuddin, M. A. Khan, A. Haseeb, S. Mahpara, M. Waseem, and A. M. Saleh, “Evaluation of Machine Learning Approaches for Precision Farming in Smart Agriculture System: A Comprehensive Review,” IEEE Access, vol. 12, pp. 60155–60184, 2024, doi: 10.1109/ACCESS.2024.3390581.

[15] R. Tiwari et al., "Leveraging Advanced Machine Learning Methods to Enhance Multilevel Fusion Score Level Computations," Fusion: Practice and Applications, vol. 14, pp. 76–91, Jan. 2024, doi: 10.54216/FPA.140206.

[16] S. Salmani Pour Avval, N. D. Eskue, R. M. Groves, and V. Yaghoubi, "Systematic review on neural architecture search," Artif Intell Rev, vol. 58, no. 3, p. 73, Jan. 2025, doi: 10.1007/s10462-024-11058-w.

[17] "D2NAS: Efficient Neural Architecture Search With Performance Improvement and Model Size Reduction for Diverse Tasks | IEEE Journals & Magazine | IEEE Xplore." Accessed: Jan. 24, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10613774

[18] "Resource-Efficient Ubiquitous Sensor Networks for Smart Agriculture: A Survey | IEEE Journals & Magazine | IEEE Xplore." Accessed: Jan. 24, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10798427

[19] S. T. Haider et al., "An Ensemble Machine Learning Framework for Cotton Crop Yield Prediction Using Weather Parameters: A Case Study of Pakistan," IEEE Access, vol. 12, pp. 124045–124061, 2024, doi: 10.1109/ACCESS.2024.3454511.

[20] K. Alibabaei et al., "A Review of the Challenges of Using Deep Learning Algorithms to Support Decision-Making in Agricultural Activities," Remote Sensing, vol. 14, no. 3, Art. no. 3, Jan. 2022, doi: 10.3390/rs14030638.

[21] "A Review of the Challenges of Using Deep Learning Algorithms to Support Decision-Making in Agricultural Activities." Accessed: Jan. 24, 2025. [Online]. Available: https://www.mdpi.com/2072-4292/14/3/638

[22] E. Alotaibi and N. Nassif, "Artificial intelligence in environmental monitoring: in-depth analysis," Discov Artif Intell, vol. 4, no. 1, p. 84, Nov. 2024, doi: 10.1007/s44163-024-00198-1.

[23] O. Arowosegbe, C. Ballali, R. Kyei, M. Adeshina, J. Agbelusi, and M. Adeshina, "Combating food waste in the agricultural supply chain: A systematic review of supply chain optimization strategies and their sustainability benefits," World Journal of Advanced Research and Reviews, vol. 24, pp. 122–140, Oct. 2024, doi: 10.30574/wjarr.2024.24.1.3023.

[24] P. S. Rana, Kalpana, Chahat, S. K. Modi, A. L. Yadav, and S. Singla, "Comparative Analysis of Tree-Based Models and Deep Learning Architectures for Tabular Data: Performance Disparities and Underlying Factors," in 2023 International Conference on Advanced Computing & Communication Technologies (ICACCTech), Dec. 2023, pp. 224–231. doi: 10.1109/ICACCTech61146.2023.00044.

[25] R. Tiwari et al., "Leveraging Advanced Machine Learning Methods to Enhance Multilevel Fusion Score Level Computations," Fusion: Practice and Applications, vol. 14, pp. 76–91, Jan. 2024, doi: 10.54216/FPA.140206.

[26] "Soil and Crop Management Strategies to Ensure Higher Crop Productivity within Sustainable Environments." Accessed: Jan. 24, 2025. [Online]. Available: https://www.mdpi.com/2071-1050/11/5/1485

[27] N. E. Benti, M. D. Chaka, A. G. Semie, B. Warkineh, and T. Soromessa, "Transforming agriculture with Machine Learning, Deep Learning, and IoT: perspectives from Ethiopia—challenges and opportunities," Discov Agric, vol. 2, no. 1, p. 63, Oct. 2024, doi: 10.1007/s44279-024-00066-7.

[28] L. A. Fitri et al., "Automated classification of urinary stones based on microcomputed tomography images using convolutional neural network," Physica Medica, vol. 78, pp. 201–208, 2020.

[29] "Leveraging Hybrid Deep Learning Models for Enhanced Multivariate Time Series Forecasting | Neural Processing Letters." Accessed: Jan. 24, 2025. [Online]. Available: https://link.springer.com/article/10.1007/s11063-024-11656-3

[30] "EL-NAS: Efficient Lightweight Attention Cross-Domain Architecture Search for Hyperspectral Image Classification." Accessed: Jan. 24, 2025. [Online]. Available: https://www.mdpi.com/2072-4292/15/19/4688

[31] "Evolutionary Algorithm-Based and Network Architecture Search-Enabled Multiobjective Traffic Classification | IEEE Journals & Magazine | IEEE Xplore." Accessed: Jan. 24, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9383257

[32] "Crop Prediction Model Using Machine Learning Algorithms." Accessed: Jan. 24, 2025. [Online]. Available: https://www.mdpi.com/2076-3417/13/16/9288

[33] "Hybrid approaches to optimization and machine learning methods: a systematic literature review | Machine Learning." Accessed: Jan. 25, 2025. [Online]. Available: https://link.springer.com/article/10.1007/s10994-023-06467-x

[34] "Data Collection and Preparation: Best Practices for Efficient Analysis – Online Tool Guides." Accessed: Jan. 25, 2025. [Online]. Available: https://onlinetoolguides.com/data-collection-and-preparation/