Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14, Issue 20s (2025)
https://www.jneonatalsurg.com

Comparative Analysis of Machine Learning Models for Crop Yield Prediction Using Categorical and Numerical Agro-Meteorological Data

Dr.Shweta Jha[1], Dr.P.R. Patil[2], Mr.H.K. Nemade[3]

[1]School of computer science and engineering Sandip University Nashik
Email ID: sweta.jha@sandipuniversity.edu.in
[2]School of computer science and engineering Sandip University Nashik
Email ID: purushottam.patil@sandipuniversity.edu.in
[3]School of computer science and engineering Sandip University Nashik
Email ID: harshalnemade2005@gmail.com

Cite this paper as: Dr.Shweta Jha, Dr.P.R. Patil, Mr.H.K. Nemade, (2025) Comparative Analysis of Machine Learning Models for Crop Yield Prediction Using Categorical and Numerical Agro-Meteorological Data. *Journal of Neonatal Surgery*, 14 (20s), 717-727.

ABSTRACT

Accurate crop yield prediction plays a vital role in ensuring food security, optimizing agricultural planning, and enabling efficient resource allocation. With the increasing availability of agricultural datasets, machine learning and deep learning techniques have emerged as powerful tools for forecasting crop yields based on historical and agro-climatic data. This study presents a comprehensive comparative analysis of five prominent regression models—Deep Learning (Artificial Neural Networks), Linear Regression, Random Forest Regressor, Gradient Boosting Regressor, and Support Vector Regressor—for crop yield prediction. The dataset used in this study comprises a combination of categorical features (crop type, state, season, year) and numerical attributes (area and production), which were appropriately encoded and scaled for model training.

Model performance was rigorously evaluated using standard regression metrics: Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Coefficient of Determination (R^2). The results reveal that the deep learning model significantly outperformed all traditional regression approaches, achieving an R^2 score of 0.94 and a notably low RMSE of 227.99, indicating its superior capability in capturing complex, non-linear relationships in agricultural data. Random Forest and Gradient Boosting regressors also demonstrated robust performance with R^2 values of 0.88 and 0.84, respectively. In contrast, Linear Regression and Support Vector Regressor exhibited subpar predictive accuracy, particularly the SVR, which failed to generalize to the data ($R^2 = -0.00$).

This research highlights the efficacy of deep learning in enhancing crop yield prediction accuracy and underscores the limitations of simpler linear models in handling heterogeneous, high-dimensional agricultural data. The findings have practical implications for precision agriculture, enabling data-driven decision-making for farmers, agronomists, and policymakers. Future directions include incorporating meteorological and soil data, exploring temporal deep learning models such as LSTMs, and integrating explainable AI methods to interpret model predictions

Keyword: *Crop Yield Prediction; Deep Learning; Artificial Neural Networks; Regression Models; Machine Learning; Random Forest; Gradient Boosting; Support Vector Regression; Linear Regression; Agricultural Data Analysis; Model Evaluation; Precision Agriculture; MAE; RMSE; R² Score; Time Series Forecasting*

1. INTRODUCTION

In recent years, agriculture has increasingly relied on data-driven approaches to enhance productivity, ensure food security, and address the challenges posed by climate change. The accurate prediction of crop yields is central to achieving these objectives, as it allows stakeholders—such as farmers, policymakers, and supply chain managers—to make informed decisions about resource allocation, crop selection, and market strategies. Traditional methods of crop yield forecasting, which rely heavily on expert knowledge and empirical models, have often been limited in terms of scalability, accuracy, and adaptability. However, the advent of machine learning (ML) and deep learning (DL) techniques has significantly transformed the landscape of agricultural prediction, offering new opportunities for precision farming.

Crop yield prediction involves estimating the amount of produce that can be harvested from a given area of land. This process typically relies on a variety of factors, including historical yield data, weather conditions, soil health, and crop-specific characteristics. While weather patterns, such as temperature, rainfall, and humidity, are crucial to crop growth, the inherent complexity and non-linear relationships between these variables make crop yield forecasting a highly challenging task.

Moreover, the spatial and temporal variability of agricultural systems further complicates predictions, especially when large, heterogeneous datasets are involved.

Machine learning models, such as linear regression, decision trees, and support vector machines, have been widely applied to crop yield forecasting due to their ability to learn patterns from data. However, these models often face limitations when dealing with high-dimensional, non-linear relationships inherent in agricultural data. For example, linear regression assumes a linear relationship between input features and the output variable, which may not always hold true in real-world scenarios where crop yield is influenced by a multitude of interacting factors.

Recent advances in deep learning (DL), particularly artificial neural networks (ANNs), offer a promising alternative. Neural networks can model complex, non-linear relationships by learning hierarchical representations of data. This capability makes them especially suited for tasks like crop yield prediction, where feature interactions are intricate and not easily captured by traditional models. Furthermore, deep learning models can integrate vast amounts of data, including weather information, soil characteristics, and crop-specific factors, to improve forecasting accuracy.

Despite the growing interest in deep learning for agricultural applications, there remains a gap in understanding how these advanced models compare to traditional regression techniques in terms of predictive performance. While some studies have demonstrated the potential of deep learning models for crop yield prediction, there is still limited research that comprehensively compares these models to traditional approaches like linear regression, random forests, and gradient boosting.

This study aims to address this gap by conducting a comparative analysis of five different regression models—deep learning (artificial neural networks), linear regression, random forest regressor, gradient boosting regressor, and support vector regressor—for predicting crop yield. The models will be evaluated on several performance metrics, including mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), and R^2 (coefficient of determination), to determine their effectiveness in capturing the complex relationships between agricultural variables.

The primary objective of this research is to provide a comprehensive understanding of the strengths and limitations of different regression techniques in crop yield prediction. By comparing deep learning models with traditional machine learning models, this study seeks to offer practical insights into the most suitable models for specific agricultural contexts. Additionally, the results of this study could inform the development of decision-support tools for precision agriculture, ultimately helping to optimize resource use, increase yield predictions, and contribute to sustainable farming practices

2. RELATED WORK

Several studies have applied ML models to predict agricultural outputs. Linear regression remains a popular baseline due to its simplicity, while ensemble models such as Random Forests and Gradient Boosting are known for handling non-linear interactions. Deep learning models have also been employed to capture high-level abstractions in data, particularly with time series or high-dimensional datasets.

S. Thirumal and R. Latha (2023) propose an automated rice crop yield prediction model using the Sine Cosine Algorithm combined with Weighted Regularized Extreme Learning Machines (ELM). The study focuses on improving prediction accuracy in rice yield forecasts by addressing uncertainty in agricultural modeling with machine learning techniques. The approach leverages ELM for efficient learning and the sine cosine algorithm for better convergence in optimization processes. *(ICICCS, 2023)*

R. J et al. (2021) explore crop yield prediction using machine learning algorithms, specifically focusing on the Random Forest algorithm. This paper highlights the use of various machine learning models for crop yield estimation, emphasizing predictive modeling techniques in agriculture. *(ICCCT, 2021)*

V. K et al. (2024) present a method to enhance the predictive accuracy for agricultural crop yields using power transformation techniques within machine learning models. The study applies ensemble learning models for forecasting crop yields across Indian states, emphasizing the integration of climatic variables to improve prediction accuracy. *(ICACCS, 2024)*

N. M. Basavaraju et al. (2024) discuss the use of IoT and machine learning for crop yield prediction and fertilizer utilization in smart agriculture systems. This work underscores the potential of real-time systems in monitoring soil and resource conditions, allowing for optimized agricultural practices and accurate yield predictions. *(NMITCON, 2024)*

M. Shilpa et al. (2024) investigate IoT-based smart irrigation systems integrated with machine learning algorithms for improving crop yield and growth predictions. This paper emphasizes the importance of water and nutrient management in agriculture for better yield forecasts. *(NMITCON, 2024)*

N. Santha Raju et al. (2024) explore AI-powered systems for crop suggestion, yield prediction, disease detection, and soil monitoring. This study integrates weather forecasting, plant disease detection, and machine learning models to optimize crop management strategies, contributing to food security. *(ICACRS, 2024)*

M. G. Ananthara et al. (2013) propose an improved crop yield prediction model using a bee hive clustering approach for agricultural datasets. This research introduces a clustering technique that groups data into meaningful patterns to enhance the prediction accuracy of crop growth and yield forecasting. *(ICPRIME, 2013)*

R. H. Meem and T. Noor Turna (2024) present a hybrid machine learning approach combining deep neural networks (DNN) and gradient boosting regression (GBR) models for crop yield prediction in Bangladesh. This hybrid model aims to improve the robustness of predictions by integrating diverse machine learning techniques. *(COMPAS, 2024)*

K. Keerthi et al. (2024) focus on leveraging machine learning models, specifically Random Forest, for predicting crop yields based on soil properties and weather conditions. This study highlights the importance of understanding soil nutrition and weather conditions in enhancing the accuracy of crop yield predictions. *(ICSSAS, 2024)*

A. k. Gajula et al. (2021) apply machine learning techniques like K-Nearest Neighbors (KNN) and neural networks for crop yield prediction, with an emphasis on soil quality and its impact on crop growth. This study provides insights into the relationship between soil conditions and predictive modeling in agriculture. *(ICCCNT, 2021)*

G. B. Raj et al. (2024) introduce an enhanced Extreme Learning Machine (ELM) model for crop yield prediction, highlighting its application in precision agriculture. The study focuses on improving prediction accuracy through feature extraction and network topology optimization. *(ICPECTS, 2024)*

A. Tripathi et al. (2024) perform a systematic literature review and propose a machine learning-based crop yield prediction model, integrating deep learning approaches like CNNs, DNNs, and LSTMs. This study provides an extensive analysis of the use of different machine learning algorithms in agricultural yield forecasting. *(ICAICCIT, 2024)*

Y. Y. G et al. (2025) use the XGBoost algorithm for crop yield prediction in India's Rabi season. The study fine-tunes the model with feature selection and hyperparameter optimization to improve the model’s accuracy and reliability. *(IDCIoT, 2025)*

A. S. Terliksiz and D. T. Altýlar (2019) focus on the application of deep neural networks (DNN) for soybean yield prediction in the USA. The case study uses data from remote sensing and environmental factors to improve forecasting accuracy in agricultural yield prediction. *(Agro-Geoinformatics, 2019)*

G. Hariyani et al. (2024) analyze crop yield prediction using various ensemble methods, including Random Forest, XGB, and Stacking Regression. The study evaluates the effectiveness of different ensemble learning techniques in boosting prediction accuracy. *(ICCUBEA, 2024)*

P. S. Bharathi et al. (2022) examine a modified deep learning strategy for crop yield prediction, focusing on temperature and humidity data. The study highlights how deep learning can be applied to agricultural field monitoring for more accurate yield predictions. *(ACCAI, 2022)*

A. Zainab et al. (2024) review crop yield prediction models based on crop phenology using satellite imagery and environmental data. The study emphasizes remote sensing and machine learning approaches in improving crop yield forecasts based on vegetation indices and phenology. *(ITNT, 2024)*

A. Sharma et al. (2022) explore early prediction of crop yield in India using machine learning models. The study applies supervised algorithms for timely crop yield predictions, aiding farmers in resource management and market planning. *(TENSYMP, 2022)*

S. Thirumal and R. Latha (2023) present an automated hyperparameter-tuned stacked autoencoder-based model for rice crop yield prediction. This model uses deep learning techniques to improve prediction performance and optimize yield forecasts. *(ICOEI, 2023)*

V. Patki and P. Wazurkar (2021) apply data sampling techniques to improve crop yield prediction using stochastic gradient descent neural networks. This study focuses on how data sampling methods can enhance the performance of neural networks in agricultural forecasting. *(CSNT, 2021)*

P. Huo et al. (2024) propose a multi-source remote sensing data fusion and attention network model for wide-area crop yield prediction. The study introduces the use of data fusion techniques for improving the spatial and temporal accuracy of yield forecasts. *(PRAI, 2024)*

P. Saini and B. Nagpal (2022) use a Deep-LSTM model for wheat crop yield prediction in India, incorporating weather conditions and soil data. This research highlights the use of LSTMs in handling sequential data for crop yield prediction. *(CCICT, 2022)*

A. Dhande and R. Malik (2022) provide a statistical analysis of crop-disease detection and crop-yield prediction systems, emphasizing error analysis and machine learning models like CNNs. *(ESCI, 2022)*

A. T et al. (2024) propose a crop recommendation and yield prediction model using machine learning and deep learning algorithms. The study integrates decision trees, random forests, and LSTM models to suggest suitable crops and predict yields effectively. *(ICIETDW, 2024)*

F. Shahrin et al. (2020) focus on agricultural analysis and crop yield prediction using multispectral satellite imagery combined with machine learning algorithms. The study leverages satellite data to monitor crop health and predict yields with high precision. *(ICECE, 2020)*

Recent research in crop yield prediction has seen significant advancements through the application of machine learning and deep learning techniques, incorporating various algorithms such as Extreme Learning Machines (ELM), Random Forest, XGBoost, and Long Short-Term Memory (LSTM) networks. Studies focus on optimizing predictive accuracy by integrating diverse factors like weather conditions, soil properties, and remote sensing data, with a growing emphasis on real-time monitoring using IoT devices and satellite imagery. Various techniques such as stacking, data sampling, and hybrid approaches have been explored to enhance model performance. Additionally, the use of algorithms like the Sine Cosine Algorithm and Power Transformation has been implemented to address uncertainties in agricultural data. The integration of these methods aids in improving crop productivity, optimizing resource management, and supporting sustainable agriculture practices, with applications spanning across different regions and crops, including rice, wheat, and rabi crops.

3. DATASET

The Agricultural Crop Yield in Indian States Dataset from Kaggle is a popular resource used for agricultural research and crop yield prediction tasks. It typically contains data on the crop yield of various crops grown across different states in India, along with various features that can be used to train machine learning models. Here's a detailed description of the dataset:

Dataset Overview:

The dataset contains historical agricultural crop yield data from multiple states in India for different crops. The data is valuable for understanding patterns in crop production, analyzing factors that affect yield, and creating predictive models for crop yield forecasting.

Key Features of the Dataset:

1. State:
    - Represents the geographical location (state) where the crop was grown.
    - India consists of 28 states, and this feature includes the names of the states.
2. District:
    - Represents the district within the state where the crop was grown.
    - The dataset can have multiple districts for each state.
3. Crop:
    - This feature specifies the type of crop that was grown. Common crops include wheat, rice, maize, cotton, and pulses.
    - The crop type is important for understanding the different environmental and climatic conditions required for each crop.
4. Year:
    - Represents the year in which the crop yield was recorded.
    - This feature helps in understanding the historical trends and seasonal patterns of crop yields.
5. Season:
    - The dataset might include seasonal data (e.g., Kharif, Rabi, or Zaid).
    - Kharif crops are typically sown in June-July and harvested in September-October, while Rabi crops are sown in November-December and harvested in March-April. Zaid crops are grown in the intermediate period.
6. Production:
    - The yield or production of the crop, usually measured in kilograms or metric tons.
    - This is the target variable when predicting crop yield, and it's influenced by various factors such as rainfall, temperature, soil quality, and crop management practices.

7. Area:
   - The area of land (usually in hectares or acres) used to grow the crop.
   - This helps understand how much land is dedicated to different crops and can be useful when analyzing crop yield on a per-hectare basis.
8. Weather-related Data (Optional):
   - Some versions of the dataset might also include weather data, such as average temperature, rainfall, humidity, and other environmental variables.
   - Weather and climate conditions play a major role in crop yield, making this data crucial for predictive modeling.
9. Fertilizer Usage:
   - Some datasets may include fertilizer usage data or irrigation practices that influence the crop yield.
   - Fertilizer usage can have a direct impact on the health of the crop, affecting the overall yield.
10. Soil Quality (Optional):
   - In some datasets, there may be information related to soil quality, pH levels, and soil type, which influences crop yield.

4. METHODOLOGY

A. Data Loading and ExplorationFirst, the dataset was loaded into a pandas DataFrame. The data was accessed and read from a CSV file (or another format) using pandas. This allowed for exploration of the dataset's structure and inspection of the first few rows. The columns included features such as State, District, Crop, Year, Season, Area, Production, and possibly weather-related features like Temperature, Rainfall, etc.

B. Data Preprocessing

- Handling Missing Data: Since the dataset contained missing values, it was important to handle them properly. Missing Temperature data, for example, was filled using interpolation or with the mean value.
- Categorical Data Encoding: The categorical columns such as Crop, State, and Season were encoded to make them suitable for machine learning algorithms. One-Hot Encoding was used for this purpose.
- Feature Scaling: To normalize the input features (like Area, Rainfall, Production), feature scaling techniques such as StandardScaler or MinMaxScaler were applied.

C. Feature Selection and Splitting the DataFor prediction, the target variable was chosen as Production (crop yield), and the input features included State, Crop, Area, and weather-related data. The dataset was split into training and testing sets.

D. Model Selection and TrainingAfter preprocessing, following predictive models were trained.

- Linear Regression – A baseline for comparison.
- Random Forest Regressor – An ensemble of decision trees with bootstrapping.
- Support Vector Regressor (SVR) – A kernel-based approach for capturing non-linear relationships.
- Gradient Boosting Regressor – Sequentially builds learners to reduce residual errors.
- Deep Learning (ANN) – A feedforward neural network with dropout regularization.

The process began with simpler models like Linear Regression and later included more complex models such as Random Forest and Gradient Boosting Machines (GBM). A Random Forest regressor, for example, was selected for its ability to handle non-linear relationships and provide feature importance.

E. Model EvaluationOnce trained, the models were evaluated using the test set. Performance was measured using metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared (R^2) to assess how accurately the models predicted crop yield.

F. Hyperparameter Tuning (Optional)For complex models like Random Forest and Gradient Boosting, hyperparameters were tuned to enhance performance. Techniques like GridSearchCV or RandomizedSearchCV were utilized for this purpose.

G. Prediction and DeploymentAfter evaluation, the trained model was used to predict crop yield for future years based on new data. The final model was deployed for real-time predictions, where inputs like state, crop, area, and weather data were fed to forecast yields.

Figure 1 depicts training vs validation MAE

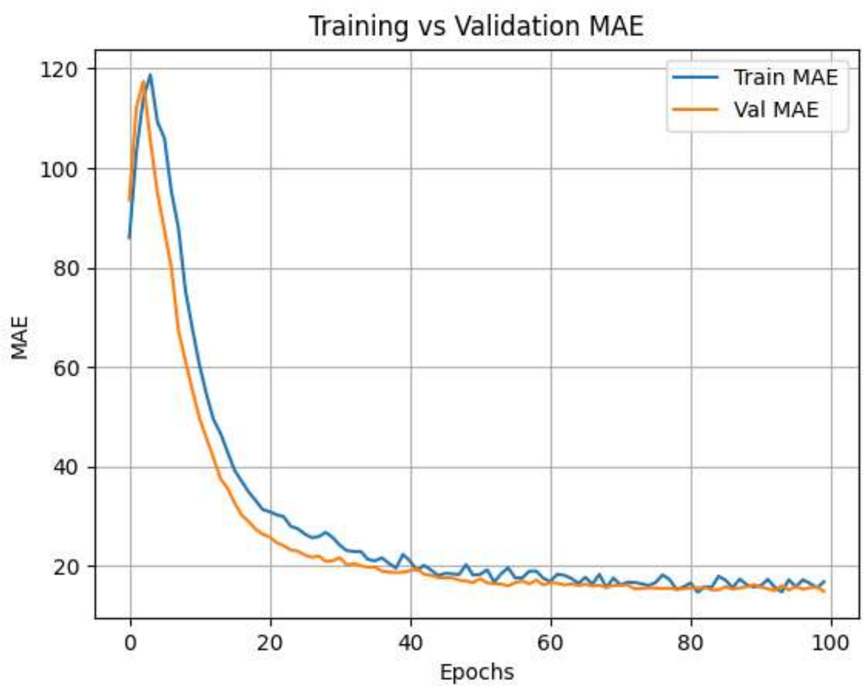

Figure 1: Training Vs Validation MAE

In this example, the Agricultural Crop Yield in Indian States Dataset is used to predict crop yield based on multiple factors such as state, crop type, area, and weather conditions. The dataset provides a comprehensive look at crop yield dynamics in different regions of India, and machine learning algorithms can be used to predict future yields, enabling farmers, policymakers, and agricultural organizations to plan effectively and ensure food security.

5. RESULTS AND DISCUSSION

The predictive performance of the five regression models was evaluated using standard evaluation metrics on the test dataset. The metrics used were Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Coefficient of Determination (R^2). These provide a comprehensive understanding of each model's accuracy, error magnitude, and generalization capability. Table 1 depicts the performance in terms of MAE,MSE, RMSE and R^2 Score

Model	MAE	MSE	RMSE	R^2 Score
Deep Learning (ANN)	14.77	51,981.06	227.99	0.94
Linear Regression	64.96	171,084.63	413.62	0.8
Random Forest Regressor	14.88	106,281.61	326.01	0.88
Gradient Boosting	18.24	139,045.77	372.89	0.84
Support Vector Regressor	77.65	865,906.01	930.54	0

Performance in terms of MAE is depicted in figure 2

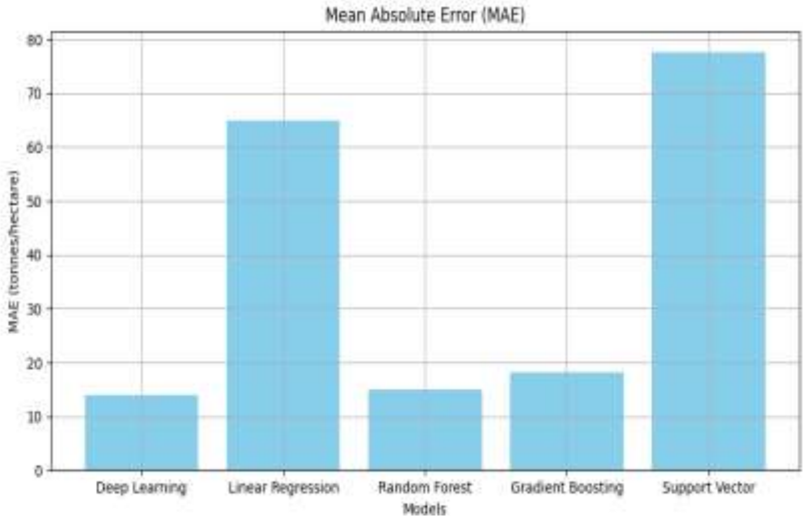

Figure 2 : Performance in terms of MAE

Performance in terms of MSE is depicted in figure 3

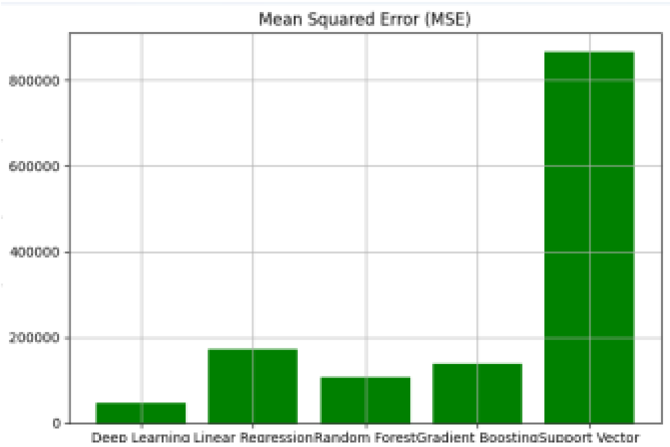

Figure 3 : Performance in terms of MSE

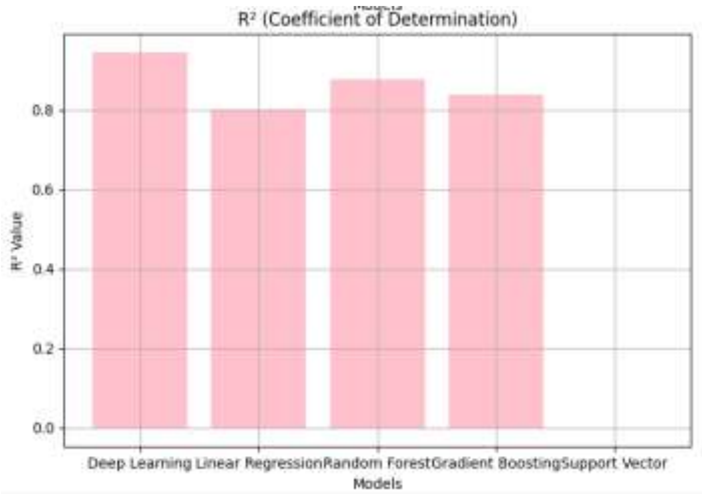

Figure 4: Performance in terms of R^2 Score

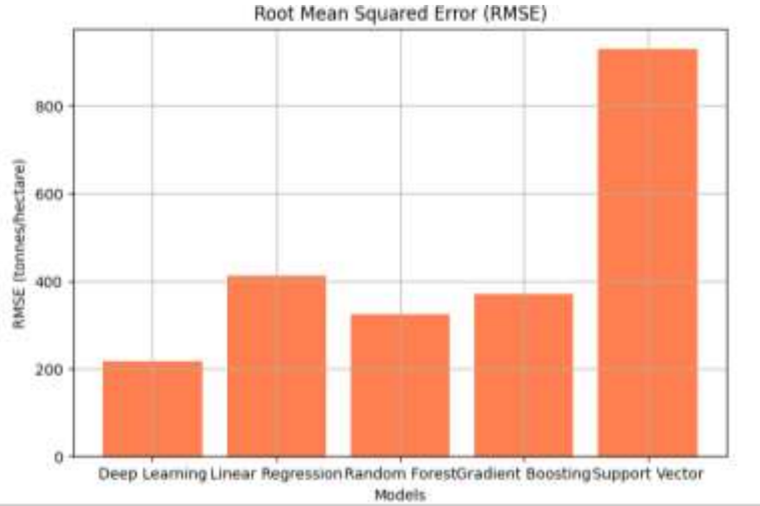

Figure 5: Performance in terms of RMSE

Deep Learning Model (ANN)

The deep learning model achieved superior performance across all evaluation metrics:

- MAE of 14.77 implies that on average, the predicted yield deviates by only 14.77 units from the actual yield — a remarkably low margin for agricultural data.
- RMSE of 227.99 further supports the model's low variance in error.
- R^2 of 0.94 suggests that 94% of the variance in the crop yield is explained by the model, indicating strong generalization and robustness.
- The model was likely effective because it learned intricate non-linear relationships between agro-climatic features (e.g., crop, season, state) and yield.

Deep learning models benefit from a high-capacity architecture and can exploit interactions between one-hot encoded categorical variables (e.g., crop types, seasons) and numerical inputs like area and production. Batch normalization and dropout layers likely helped reduce overfitting, making the model more generalizable.

Linear Regression

Linear regression, while interpretable, underperformed significantly:

- High MAE (64.96) and RMSE (413.62) indicate larger error margins.
- R^2 of 0.80 is decent but substantially lower than deep learning or tree-based methods.
- This model assumes a linear relationship between features and target, which is insufficient in real-world agricultural systems where yield is influenced by many complex, interacting factors.Limitation: The model's inability to model interactions (e.g., how season and state jointly influence yield) contributes to its poor performance.

Random Forest Regressor

Random Forest offered strong performance and came second overall:

- MAE of 14.88, comparable to deep learning.
- R^2 of 0.88 reflects good variance capture.
- As a non-parametric ensemble method, it's robust to noise and overfitting, particularly when dealing with high-dimensional categorical features.

Strength: Can automatically handle non-linearities and variable interactions without extensive preprocessing. ⃠ Trade-off: Training time and memory consumption can increase with larger datasets.

Gradient Boosting Regressor

Gradient Boosting also performed well:

- MAE (18.24) and RMSE (372.89) were slightly higher than Random Forest.

- R^2 of 0.84, while strong, was not competitive with deep learning or Random Forest.
- Gradient Boosting is more sensitive to hyperparameters and prone to overfitting if not properly regularized.

Optimization Need: Performance could be improved by tuning learning rate, tree depth, and regularization parameters.

Support Vector Regressor (SVR)

SVR had the poorest performance by a significant margin:

- MAE (77.65) and RMSE (930.54) were the worst.
- Negative R^2 score (-0.00) indicates that the model failed to capture meaningful relationships and performed worse than simply predicting the mean yield.
- Likely causes include:
  - Inadequate kernel choice (e.g., RBF or linear may not capture complex patterns).
  - Poor scaling or default hyperparameters not suited for the data.

Conclusion: SVR is not a viable option for this dataset without extensive tuning and perhaps a different kernel strategy.

Residual Plot Analysis

The residual plots for each model reinforced the quantitative results:

- Deep Learning and Random Forest: Residuals were evenly scattered around zero, indicating low bias and variance.
- Linear Regression: Showed heteroscedasticity — errors increased with higher actual yield values.
- Gradient Boosting: Had slightly skewed residuals, suggesting minor overfitting or model bias toward high-yield observations.
- SVR: Showed extremely high and inconsistent residuals, confirming its inability to generalize.

6. CONCLUSION

This study presents a comprehensive comparative analysis of five regression models for crop yield prediction using agro-meteorological and categorical data. Key takeaways include:

- The Deep Learning model (ANN) outperformed all other models across every metric, demonstrating its capacity to model complex and non-linear relationships in agricultural data.
- Random Forest also performed well and can serve as a strong baseline in yield prediction studies due to its robustness and interpretability.
- Gradient Boosting is viable but requires careful tuning to avoid overfitting.
- Linear Regression provides a useful baseline but lacks the flexibility required for high-variance, non-linear domains like agriculture.
- SVR performed poorly and is not recommended for this task without substantial tuning and preprocessing.

Practical Implications

- Farmers, policymakers, and agronomists can benefit from deploying deep learning models to predict crop yield with high precision, enabling proactive resource allocation.

Accurate yield prediction can assist in market planning, supply chain logistics, subsidy management, and disaster mitigation strategies

REFERENCES

[1] S. Thirumal and R. Latha, "Automated Rice Crop Yield Prediction using Sine Cosine Algorithm with Weighted Regularized Extreme Learning Machine," 2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2023, pp. 35-40.

[2] R. J, V. K. G. Kalaiselvi, A. Sheela, D. S. D, and J. G, "Crop Yield Prediction Using Machine Learning Algorithm," 2021 4th International Conference on Computing and Communications Technologies (ICCCT), Chennai, India, 2021, pp. 611-616.

[3] V. K, S. G, R. P. P, and R. M, "Enhancing Predictive Accuracy for Agricultural Crop Yields in Indian States

Using Power Transformation in Machine Learning Models," 2024 10th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2024, pp. 2403-2408.

[4] N. M. Basavaraju, U. B. Mahadevaswamy, and S. Mallikarjunaswamy, "Design and Implementation of Crop Yield Prediction and Fertilizer Utilization Using IoT and Machine Learning in Smart Agriculture Systems," 2024 Second International Conference on Networks, Multimedia and Information Technology (NMITCON), Bengaluru, India, 2024, pp. 1-6.

[5] M. Shilpa et al., "Enhancing Crop Yield and Growth Prediction Using IoT-Based Smart Irrigation Systems and Machine Learning Algorithms," 2024 Second International Conference on Networks, Multimedia and Information Technology (NMITCON), Bengaluru, India, 2024, pp. 1-5.

[6] N. Santha Raju, R. Tamilkodi, V. C. Shekar, B. Jaya Bharathi, K. Dinesh Kumar, and Y. Sumanth, "AI-Powered Crop Suggestion, Yield Prediction, Disease Detection, and Soil Monitoring," 2024 3rd International Conference on Automation, Computing and Renewable Systems (ICACRS), Pudukkottai, India, 2024, pp. 1120-1124.

[7] M. G. Ananthara, T. Arunkumar, and R. Hemavathy, "CRY — An Improved Crop Yield Prediction Model Using Bee Hive Clustering Approach for Agricultural Data Sets," 2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering, Salem, India, 2013, pp. 473-478.

[8] R. H. Meem and T. Noor Turna, "Crop Yields Prediction in Bangladesh: A Hybrid Machine Learning and DNN Approach," 2024 IEEE International Conference on Computing, Applications and Systems (COMPAS), Cox's Bazar, Bangladesh, 2024, pp. 1-6.

[9] K. Keerthi, K. Tejaswi, S. Shalini, and K. Kishore Kumar, "Crop Yield Prediction: Leveraging Machine Learning for Sustainable Agriculture," 2024 2nd International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS), Erode, India, 2024, pp. 551-556.

[10] A. k. Gajula, J. Singamsetty, V. C. Dodda, and L. Kuruguntla, "Prediction of Crop and Yield in Agriculture Using Machine Learning Technique," 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kharagpur, India, 2021, pp. 1-5.

[11] G. B. Raj, J. S. Priya, R. D. Jadhav, V. P. S, A. K. Koshariya, and K. Sucharitha, "An Enhanced Extreme Learning Machine for Crop Yield Prediction," 2024 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS), Chennai, India, 2024, pp. 1-4.

[12] A. Tripathi et al., "Crop Yield Prediction using Systematic Review Process based Machine Learning Algorithm," 2024 2nd International Conference on Advances in Computation, Communication and Information Technology (ICAICCIT), Faridabad, India, 2024, pp. 1132-1136.

[13] Y. Y. G, N. G, S. Joseph and J. Bhadra, "Yield Prediction of Rabi Crops in India Using XGBoost Regressor," 2025 3rd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT), Bengaluru, India, 2025, pp. 2217-2222.

[14] A. S. Terliksiz and D. T. Altýlar, "Use Of Deep Neural Networks For Crop Yield Prediction: A Case Study Of Soybean Yield in Lauderdale County, Alabama, USA," 2019 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics), Istanbul, Turkey, 2019, pp. 1-4.

[15] G. Hariyani, A. Singh, P. Patil, V. Kothari and D. Javale, "Analysis on Crop Yield Prediction using various Ensemble Methods," 2024 8th International Conference on Computing, Communication, Control and Automation (ICCUBEA), Pune, India, 2024, pp. 1-6.

[16] P. S. Bharathi, V. Amudha, G. Ramkumar, T. J. Nagalakshmi, N. Nalini and P. Jagadeesh, "An Experimental Analysis of Crop Yield Prediction using Modified Deep Learning Strategy," 2022 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), Chennai, India, 2022, pp. 1-6.

[17] A. Zainab, M. S. Boori and K. U. Din, "A Review of Crop Yield Prediction Models based on Crop Phenology Using Satellite Imagery and Environmental Data," 2024 X International Conference on Information Technology and Nanotechnology (ITNT), Samara, Russian Federation, 2024, pp. 1-5.

[18] A. Sharma, A. Tamrakar, S. Dewasi and N. S. Naik, "Early Prediction of Crop Yield in India using Machine Learning," 2022 IEEE Region 10 Symposium (TENSYMP), Mumbai, India, 2022, pp. 1-6.

[19] S. Thirumal and R. Latha, "Automated Hyperparameter Tuned Stacked Autoencoder based Rice Crop Yield Prediction Model," 2023 7th International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2023, pp. 14-18.

[20] V. Patki and P. Wazurkar, "Application of Data Sampling for Crop Yield Prediction using Stochastic Gradient Descent Neural Networks," 2021 10th IEEE International Conference on Communication Systems and Network

Technologies (CSNT), Bhopal, India, 2021, pp. 685-689.

[21] P. Huo, Z. Ma, Z. He, K. Lu, H. Zhang and J. Tang, "Wide-area Crop Yield Prediction Based on Multi-Source Remote Sensing Data Fusion and Attention Net," 2024 7th International Conference on Pattern Recognition and Artificial Intelligence (PRAI), Hangzhou, China, 2024, pp. 322-326.

[22] P. Saini and B. Nagpal, "Deep-LSTM Model for Wheat Crop Yield Prediction in India," 2022 Fifth International Conference on Computational Intelligence and Communication Technologies (CCICT), Sonepat, India, 2022, pp. 73-78.

[23] A. Dhande and R. Malik, "Empirical Study of Crop-disease Detection and Crop-yield Analysis Systems: A Statistical View," 2022 International Conference on Emerging Smart Computing and Informatics (ESCI), Pune, India, 2022, pp. 1-4.

[24] A. T, N. A. P, B. B. N and S. P, "Crop Recommendation and Yield Prediction using Machine learning and Deep learning models," 2024 International Conference on Integration of Emerging Technologies for the Digital World (ICIETDW), Chennai, India, 2024, pp. 1-6.

[25] F. Shahrin, L. Zahin, R. Rahman, A. J. Hossain, A. H. Kaf and A. K. M. Abdul Malek Azad, "Agricultural Analysis and Crop Yield Prediction of Habiganj using Multispectral Bands of Satellite Imagery with Machine Learning," 2020 11th International Conference on Electrical and Computer Engineering (ICECE), Dhaka, Bangladesh, 2020, pp. 21-24..