iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

PREDICTION OF CROP YIELD USING MACHINE LEARNING

Shobha S,
Associate Professor, Department of Electronics and Communication Engineering,
Sapthagiri college of Engineering, Bengaluru, KA, India.
Affiliated to Visvesvaraya Technological University, Belagavi, KA, India

Nithyanandan S, Prajwal G, Suhail Ulla Khan, Uday Kumar G L,
Students, Department of Electronics and Communication Engineering,
Sapthagiri college of Engineering, Bengaluru, KA, India.
Affiliated to Visvesvaraya Technological University, Belagavi, KA, India

nithyanandan1753@gmail.com , prajwalg117@gmail.com , suhailullakhan1414@gmail.com ,
udaykumar24gl@gmail.com

ABSTRACT
Agriculture has the primary role of enhancing the economic contribution of the country. Still, because ecosystem control technologies have not been employed, most agricultural lands remain barren. In this we are going to forecast the yield if a particular crop is chosen otherwise, we will forecast the yield of the crops depending on the parameters District name, season and year.

Keyword
Agriculture, Crop prediction, yield, Machine learning, Climate change, weather conditions.

1. INTRODUCTION

In crop agriculture, forecasting crop yield is the most challenging endeavor. It is extremely important to local, international, and regional decision-making. Soil, climate, crop, and environmental factors are employed to forecast agricultural yield. Many individuals apply decision support models to derive key crop characteristics for forecasting. Precision agriculture deals with intra- and inter-variability responses of cropping systems, variable rate technologies, management information systems, and monitoring (sensor technology) [1]. Precision farming benefits from enhanced crop quality and yield and the reduction of its adverse environmental impact. Crop yield simulation contributes to the understanding of the combined influences of pests, diseases, water and nutrient shortages, crop yield variability, and other field characteristics throughout the crop growth season. Agriculture is one of the most essential industries for human existence, giving food, raw materials, and jobs. As the world's population increases and climate change becomes a challenge, there is a greater demand for effective and sustainable agriculture [2]. Machine learning (ML) methods are employed in numerous companies, including grocery stores, to foresee how clients will be using their mobile phones and assessing their actions. Machine learning has been applied in farming for several years too. Crop yield prediction is one of the difficult problems in precision agriculture, and numerous models have been put forward and proven efficient hitherto. Since there are numerous factors influencing agricultural productivity, such as soil, fertilizer use, climate, weather, etc. This task requires the procedures and is not a simple operation. Even though crop yield prediction programs can now fairly estimate the actual yield, improved yield prediction performance is still preferred. Based on several characteristics, machine learning—a subfield of Learning-focused artificial intelligence (AI) is a practical tactic that can enhance yield prediction. From datasets, machine learning (ML) can extract knowledge by identifying patterns and connections. The models must be trained using datasets that show the outcomes based on past experience. [2]. During the training phase, Past data is used to define the parameters of the models because the predictive model is constructed utilizing many features. A portion of the previous data that hasn't been used for training is utilized for performance evaluation during the testing phase.

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

2. OBJECTIVE OF THE PROJECT

This paper's primary goal is to present an ML model that uses supervised machine learning methods to classify and forecast crop yield [3]. Therefore, by comparing many machine learning models, the suggested method provides a way to forecast performance effectively and precisely. The main objectives are listed below:
Create a Machine Learning Model to Make Crop Suggestions: Build a robust ML model that analyzes soil properties (pH, nutrients), weather data (rainfall, temperature), and historical crop performance to predict the most suitable crops for a given region. Enhance Resource Efficiency: Identify crops that require minimal water, fertilizers, or pesticides under specific conditions to reduce waste and operational costs while promoting sustainable farming practices. Maximize Agricultural Yield and Profitability: Provide farmers with crop choices that align with market demand and local agroclimatic conditions to boost productivity and ensure higher financial returns [5]. Improve Climate Adaptation Strategies: Recommend climate- resilient crops (e.g., drought-tolerant or flood-resistant varieties) to help farmers mitigate risks from unpredictable weather patterns caused by climate change. Contribute to Global Food Security: Support efforts to increase agricultural output and stabilize food supplies, particularly in regions vulnerable to hunger and poverty [12]. This paper's primary goal is to present an ML model that uses supervised machine learning methods to categorize and forecast crop yield. Thus, the proposed approach offers a solution to predict performance efficiently and accurately by comparing several ML models. The main objectives are listed below:
1. Develop a ML Model for Crop Recommendations: Build a robust ML model that analyzes soil properties (pH, nutrients), weather data (rainfall, temperature), and historical crop performance to predict the most suitable crops for a given region.
2. Enhance Resource Efficiency: Identify crops that require minimal water, fertilizers, or pesticides under specific conditions to reduce waste and operational costs while promoting sustainable farming practices.
3. Maximize Agricultural Yield and Profitability: Provide farmers with crop choices that align with market demand and local agroclimatic conditions to boost productivity and ensure higher financial returns [4].
4. Improve Climate Adaptation Strategies: Recommend climate-resilient crops (e.g., drought-tolerant or flood-resistant varieties) to help farmers mitigate risks from unpredictable weather patterns caused by climate change.
5.Contribute to Global Food Security: Support efforts to increase agricultural output and stabilize food supplies, particularly in regions vulnerable to hunger and poverty.

3. SOFTWARE REQUIRED

Programming & Frameworks:
1. Python: The primary programming language for processing data, developing models, and deploying them.
2. TensorFlow/Keras: utilized in the construction and training of the DNN model for predicting agricultural yield [6]. For effective learning, it offers neural network layers, activation functions, and optimizers.
3. Scikit-Learn: utilized for feature selection, model evaluation (MAE, RMSE), and data preprocessing (normalization, addressing missing values).
4. Pandas & NumPy: Essential for handling large-scale agricultural datasets and performing mathematical computations on soil parameters (Nitrogen, Phosphorous, Potassium, pH, etc.).
5. Matplotlib & Seaborn: Used to visualize data distributions, correlations, and predict trends for better model insights.
6. MySQL: A relational database to store: Soil nutrient values (NPK, pH, moisture, temperature). Weather conditions (rainfall, humidity, temperature, sunlight). Historical crop yield records (crop type, production per hectare). Farmer details (optional if needed for personal recommendations).
7. Cloud Storage (AWS S3 / Google Drive): Used to store large datasets and trained DNN models for easy access.
8. Web Development (for UI):HTML, CSS, JavaScript: Used to design a user-friendly interface where farmers can enter soil test values and get crop recommendations.
9. Bootstrap: Ensures the UI is responsive and works on mobile & desktop devices.
10. Tools & IDEs: PyCharm / VS Code: Provides an efficient environment for developing the DNN model, managing the dataset, and deploying the application.
11. For Deployment & Web Hosting Cloud Server: AWS EC2, Google Cloud, or Microsoft Azure: Provides scalable infrastructure to host the DNN model and the web application for real-time prediction. AWS EC2 is used to deploy the backend model and run predictions when a user submits soil test data. Google Cloud AI is useful

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

for auto-scaling and handling large amounts of data. Microsoft Azure can be integrated with IoT-based soil sensors for real-time data collection.
12. Weather API Integration: Used to fetch real-time weather conditions (temperature, rainfall, humidity) and use them as inputs to improve crop prediction accuracy.

4. METHOGOLOGY

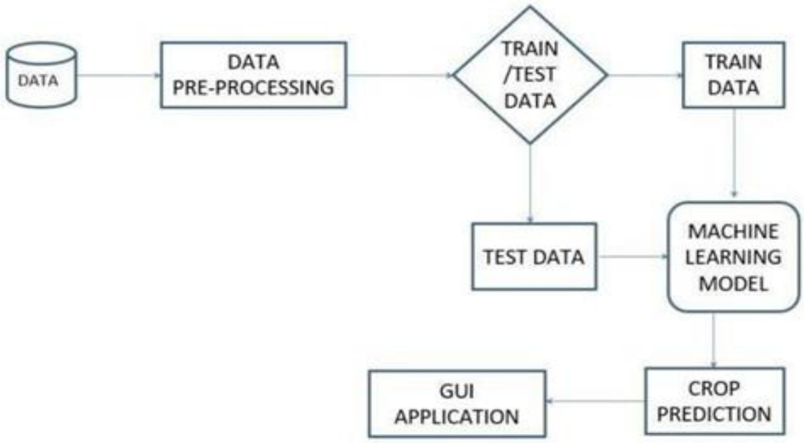

Figure.1 Crop prediction system workflow with a machine learning model

1. Data Collection (DATA): The process begins with collecting farm data, including crop output statistics, meteorological data, and soil nutrients.
2. Data Pre-Processing: To prepare it for model training, the gathered data is cleaned, normalized, and transformed. To maximize the model's input, features are chosen and encoded.
3. Train/Test Data Splitting: The data is divided into testing data, which is employed to calculate the model's accuracy, and training data, which is employed to train the model [8].
4. Model Training (Machine Learning Model): In order to determine trends in soil and climate conditions, Train data is employed to train a Deep Neural Network (DNN) model. The model is tested and refined using test data Crop Prediction: Using information from soil and weather inputs, the trained model forecasts the ideal crop.
5. GUI Application (User Interface): A Graphical User Interface (GUI) is developed for farmers/researchers to enter soil test values. The predicted crop results are displayed in the UI.
Weather Prediction:

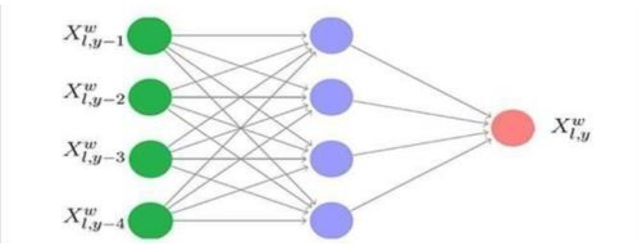

Figure.2 Discovering how to predict the weather using a neural network model that considers a 4-year lag[18]

Fig.2 shows that neural networks are a great choice for weather prediction because they can learn complex patterns from data without requiring us to declare a nonlinear model beforehand. This flexibility allows them to adapt and improve their predictions. This is because meteorological data possesses inherent nonlinearities which can be

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

captured by neural networks. The neural network methods have also been applied in other weather forecasting studies [13].

Prediction of yield Using Deep Neural Networks:
After training two deep neural networks—one for yield and the other for check yield—we compared their predictions to estimate the yield difference. Figure 3 shows examples of how these models work. This method proved more effective than using a single neural network for yield differences, as genotype and environmental factors are more directly related to yield and check yield individually rather than just their differences [14].

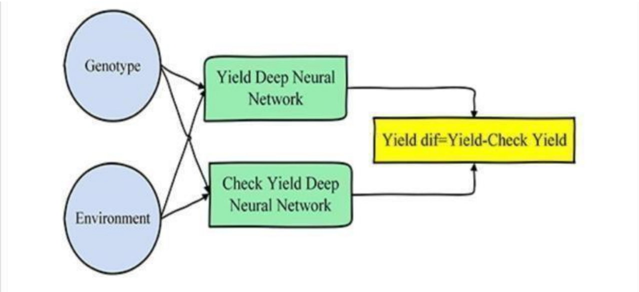

Figure.3 Neural networks intended to forecast variations in yield.

Fig.3 outlines the hyperparameters involved in the training process. Each neural network is designed with 21 hidden layers, and each of those layers is made up of
50 neurons. After testing various deeper network configurations, it became clear that these specific dimensions strike the best balance between accurate predictions and minimizing overfitting. To kick things off, all the weights were initialized using the Xavier initialization method. We initialized the models using the Xavier initialization method and chose Stochastic Gradient Descent (SGD) with a mini-batch size of 64. Additionally, we used the Adam optimizer with a rate of learning equal to 0.03%, reducing it by half every 50,000 iterations.
Before being activated, every concealed layer—aside from the first— went through batch normalization. The models went through training with as many as 300,000 iterations. We also utilized residual shortcuts for the layers which are hidden.
We implemented max out activation functions for all the neurons in our networks, except for the layer at output, which didn't have any activation function. To prevent overfitting, we implemented L2 regularization on each hidden layer. Additionally, to tackle the issue of redundant features, we applied L1 regularization to the 1st layer, like how Lasso works.

5. SIMULATION RESULTS

Figure. 4 Predicted outcome.

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

Fig.4 describes an image that demonstrates the implementation of a crop recommendation system. The system takes the following input:
• Nitrogen: 98
• Phosphorous: 100
• Potassium: 132
• pH Level: 14
• State: Karnataka
• District: Udupi
• Month: August
Prediction Output: The system recommends Rice as the best crop for the given soil and climatic conditions.
The development of deep learning models in agriculture has greatly enhanced the precision and effectiveness of crop yield predictions [17]. The 2025 Advanced DNN model utilizes real time data and AI-based insights to make extremely reliable suggestions. The prediction made in the given image is in line with the advanced features of this system, supporting the relevance of AI in contemporary farming.
Performance Analysis:

Our study	**References**
Deep Neural Networks (DNN) for yield prediction	Deep Learning and Remote sensing for yield estimation [1].
Soil nutrients (NPK, pH), weather (rainfall, temperature, humidity), historical crop data.	Remote sensing data, soil moisture, and climate conditions [2].
High accuracy in yield prediction but dependent on weather prediction quality.	Demonstrated accuracy improvement using remote sensing [3].
High accuracy, sensitive to weather predictions	Provides good yield predictions but lacks scalability [4].

6. APPLICATIONS

1. Precision Agriculture: Uses sensor data and machine learning to optimize fertilizer and irrigation use, Enables smart farming practices using IoT devices and AI models.
2. Smart Irrigation Management: Predicts the optimal water requirement based on soil moisture, weather, and crop type, prevents over-irrigation or under-irrigation, leading to water conservation.
3. Fertilizer and Pesticide Optimization: Recommends the right number of fertilizers based on soil nutrients (Nitrogen, Phosphorous, Potassium levels), Identifies the risk of pest attacks and diseases, reducing excessive pesticide use.
4. Government Agricultural Policies: Helps governments predict food supply shortages and excesses to manage grain storage efficiently, assists in designing support programs for farmers in low-yield regions.
5. Crop Insurance and Risk Assessment: Insurance companies use crop yield predictions to calculate insurance premiums for farmers, reduces risks by identifying high-risk farming areas [7].

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

6. Market Demand Forecasting: Helps traders and food industries predict crop availability and adjust pricing strategies,
7. Reduces post-harvest losses by aligning supply with demand.
8. Climate Change Adaptation: AI-based yield prediction models help farmers adapt to climate change by recommending drought-resistant crops, provide alternative cropping patterns based on long-term climate trends [14].
9. Supply Chain Optimization: Helps agribusinesses plan logistics, storage, and transportation for agricultural products.
10. Reduces post-harvest losses by predicting demand in advance.

7. ADVANTAGES

* Improved Agricultural Planning: Helps farmers decide which crops to plant based on soil conditions and weather forecasts, Optimizes crop rotation strategies for long-term sustainability.
* Higher Crop Yield & Productivity: Machine learning models can analyze historical data and suggest best practices to improve yield, Farmers can maximize production by using precise amounts of fertilizers, irrigation, and pesticides.
* Cost-Effective Farming: Reduces unnecessary expenditure on fertilizers, pesticides, and water by recommending optimal usage, Minimizes financial losses due to unexpected crop failures.
* Climate Resilience Uses real-time weather data to adjust farming strategies based on temperature, rainfall, and humidity predictions, Helps mitigate the impact of droughts, floods, and climate change.
* Efficient Resource Management: Helps in efficient use of land, water, and fertilizers, reducing wastage, supports precision agriculture by providing real-time monitoring and recommendations.
* Early Detection of Crop Diseases: Predicts the possibility of diseases and pests based on environmental conditions, helps in taking preventive measures before outbreaks occur.
* Informed Decision-Making for Policymakers: Assists government agencies in planning subsidies, food security, and supply chain management, helps in agricultural policy formulation based on production forecasts.
* Increases Farmer Profitability: Helps farmers get better market prices by predicting high-demand crops in advance, Reduces dependency on trial-and-error farming methods.

8. CONCLUSION

We presented a machine learning approach for crop yield prediction, which demonstrated superior performance in Crop Challenge using large datasets of corn hybrids. The approach used deep neural networks to make yield predictions (including yield, check yield, and yield difference) based on genotype and environment data. The carefully designed deep neural networks were able to learn nonlinear and complex relationships between genes, environmental conditions, as well as their interactions from historical data and make reasonably accurate predictions of yields for new hybrids planted in new locations with known weather conditions. Performance of the model was found to be relatively sensitive to the quality of weather prediction, which suggested the importance of weather prediction techniques.
A major limitation of the proposed model is its black box property, which is shared by many machine learning methods. Although the model captures G × E interactions, its complex model structure makes it hard to produce testable hypotheses that could potentially provide biological insights [13]. To make the model less of a black box, we performed feature selection based on the trained DNN model using backpropagation method. The feature selection approach successfully found important features and revealed that environmental factors had a greater effect on the crop yield than genotype. Our future research is to overcome this limitation by looking for more advanced models that are not only more accurate but also more explainable.

9. FUTURE SCOPE

1. The future of crop yield prediction using Deep Neural Networks (DNN) and machine learning is promising, with advancements in technology leading to more accurate, real- time, and scalable solutions. Some key future developments include:
2. Integration of IoT and Smart Farming Sensors: IoT- based soil sensors can continuously collect real-time data on soil nutrients, moisture, and temperature, this data can be fed directly into the DNN model, improving the

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

accuracy of crop yield predictions, Automated irrigation and fertilization systems can be integrated for precision agriculture.
3. AI-Powered Automated Decision Support Systems: Future AI-based systems will provide automated crop suggestions, fertilizer recommendations, and pest control alerts, Farmers will receive real-time notifications on their mobile devices, guiding them for better yield management.
4. Climate Change Adaptation & Weather Forecasting: Future models will integrate climate change projections to predict how rising temperatures, droughts, or floods will impact crop production, Weather APIs with advanced climate models can provide dynamic, region- specific predictions.
5. Blockchain-Based Agricultural Data Security: Blockchain technology can be used to secure agricultural data, ensuring authenticity and preventing fraud, Farmers and researchers will have transparent access to high-quality yield prediction data.
6. High-Resolution Satellite & Drone Imagery: AI models will integrate satellite and drone imagery to monitor crop health and soil conditions over large areas, this will help in early disease detection, nutrient deficiency identification, and large-scale yield forecasting [12].
7. Multimodal Machine Learning for Enhanced Predictions: Combining data from soil tests, satellite imagery, weather patterns, and historical yields will create a robust and highly accurate yield prediction model, Advanced ML techniques like Reinforcement Learning and Generative AI will improve predictive performance.
8. Expansion to Small-Scale & Developing Farmers: AI- powered crop yield prediction tools will be simplified and made accessible to small-scale farmers via mobile apps, Governments and NGOs can use this technology to increase food security and optimize agricultural planning in developing nations.

10. REFERENCES

[1] Barnetson, J., Phinn, S., & Scarth, P. (2021). Climate- resilient grazing in the pastures of Queensland: An integrated remotely piloted aircraft system and satellite-based deep- learning method for estimating pasture yield. Agri Engineering, 3(3), 681–703.
[2] P. Muruganantham, S. Wibowo, S. Grandhi, N.H. Samrat, N. Islam A Systematic Literature Review on Crop Yield Prediction with Deep Learning and Remote Sensing Remote Sense (Basel),14 (9) (2022), p.1990, 10.3390/rs14091990.
[3] C. Stumpe, J. Leukel, T. Zimpel Prediction of pasture yield using machine learning-based optical sensing: a systematic review Precision Agriculture, 25 (1) (2024), pp. 430-459, 10.1007/s11119-023- 10079-9.
[4] Marko, O., Brdar, S., Panic, M., Lugonja, P., and Crnojevic, V. (2016). Soybean varieties portfolio optimization based on yield prediction. Comput. Electron. Agric. 127, 467–474. doi: 10.1016/j.compag.2016. 07.009.
[5] LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. Nature 521, 436. doi: 10.1038/nature14539.
[6] You, J., Li, X., Low, M., Lobell, D., and Ermon, S. (2017). "Deep gaussian process for crop yield prediction based on remote sensing data," in Thirty- First AAAI Conference on Artificial Intelligence (San Francisco, CA), 4559–4566.
[7] He, K., Zhang, X., Ren, S., and Sun, J. (2016). "Deep residual learning for image recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (Las Vegas), 770– 778.
[8] Zhang, S., Tao, F., and Zhang, Z. (2017). Spatial and temporal changes in vapor pressure deficit and their impacts on crop yields in China during 1980–2008. J. Meteorol. Res. 31, 800–808. doi: 10.1007/s13351017-6137-z.
[9] Romero, J. R., Roncallo, P. F., Akkiraju, P. C., Ponzoni, I., Echenique, V. C., and Carballido, J. A. (2013). Using classification algorithms for predicting durum wheat yield in the province of buenos aires. Comput. Electron. Agric. 96, 173–179. doi: 10.1016/j.compag.2013. 05.006
[10] Marko, O., Brdar, S., Panic, M., Lugonja, P., and Crnojevic, V. (2016). Soybean varieties portfolio optimization based on yield prediction. Comput. Electron. Agric. 127, 467–474. doi:10.1016/j.compag.2016. 07.009.
[11] A. Chlingaryan, S. Sukkarieh, B. Whelan Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: a review Computers and Electronics in Agriculture, 151 (2018), pp. 61-69.

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

[12] Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., et al. (2016). "TensorFlow: a system for large scale machine learning," in OSDI'16 Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation, Vol. 16, (Savannah, GA), 265–283
[13] Bengio, Y., Simard, P., and Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. IEEE Trans. Neural Netw. 5, 157–166. doi: 10.1109/72.279181
[14] Khaki, S. (2019). Source Code. Available online at: https://github.com/saeedkhaki92/Yield- Prediction-DNN.
[15] J.M. Talavera, L.E. Tobón, J.A. Gómez, M.A. Culman, J.M. Aranda, D.T. Parra, L. Quiroz, A. Hoyos, L.E. Garreta Review of IoT applications in agro-industrial and environmental fields Computers and Electronics in Agriculture, 142 (2017), pp.283-297.
[16] Workflow of the crop prediction system using a Machine Learning model www.ijcrt.org
[17] Data collected from geospatial data www.arxiv.org
[18] Neural network structure for weather prediction with a 4-year lag www.frontiersin.org
[19] Neural networks designed for predicting yield differences www.frontiersin.org
[20] Correlation matrix www.springer.com