IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 2, November 2024

Crop Yield Prediction Using Machine Learning

Mr. Laxmikant Malphedwar[1],Dr.Alpana Adsul[2],Mr. Sanket Nagare[3], Mr. Yash Nimse[4], Mrs. Sanchiti Nimble[5], Mrs. Sakshi Pakhle[6]

Assistant Professor, Department of Computer Engineering [1]
Associate Professor, Department of Computer Engineering[2]
Students, Department of Computer Engineering[3,4,5,6]
Dr. D. Y. Patil College of Engineering & Innovation Talegaon, Pune, India

Abstract*: Machine learning plays a key role in helping farmers and decision-makers predict crop yields, assisting them in choosing which crops to grow and how to manage them throughout the growing season. To better understand the effectiveness of different machine learning methods, we conducted a systematic literature review. From six electronic databases, we initially gathered 567 relevant studies and, after applying our selection criteria, narrowed it down to 50 studies for in-depth analysis. We focused on identifying the commonly used algorithms and features in these studies. The analysis revealed that temperature, rainfall, and soil type are the most frequently used features for crop prediction, while Artificial Neural Networks (ANN) are the most applied machine learning models. In addition, we expanded our investigation to include studies specifically using deep learning techniques. We reviewed 30 papers in this category and found that Convolutional Neural Networks (CNN) are the most popular choice. Other widely used deep learning algorithms include Long Short-Term Memory (LSTM) networks and Deep Neural Networks (DNN). Overall, our findings offer insights into the prevalent approaches and highlight opportunities for future research in crop yield prediction using both machine learning and deep learning techniques.*

Keywords: Machine Learning, Crop Yield Prediction, Agricultural Analytics, Data Preprocessing, Feature Selection, Meteorological Data, Soil Characteristics.

I. INTRODUCTION

Machine learning (ML) has been utilized in various domains, from analyzing customer behavior in retail to predicting phone usage patterns. In agriculture, ML has been applied for years, particularly in crop yield prediction, which is a challenging aspect of precision farming. Crop yield prediction involves complex factors like climate, soil, fertilizer use, and seed type, requiring multiple datasets for accurate modeling. While current prediction models provide reasonable estimates, there is still room for improved accuracy. As a branch of Artificial Intelligence (AI), ML aims to create models that learn from data, identifying patterns and relationships to make informed predictions. These models are trained using historical data, which helps to determine the model parameters, and then tested on unseen data to evaluate their performance. Depending on the problem, ML models can be descriptive, explaining past outcomes, or predictive, forecasting future events.Developing an effective ML model involves selecting suitable algorithms and ensuring that the system can handle the data's volume and complexity. To understand the current landscape of ML in crop yield prediction, we conducted a systematic literature review (SLR). An SLR identifies research gaps, synthesizes existing knowledge, and offers a structured overview to guide future research in the field.

II. LITERATURE REVIEW / DISCUSSION

Recent advancements in machine learning (ML) have significantly transformed agricultural practices, particularly in predicting crop yields. For instance, Gajula et al. (2021)[1] employed the KNearest Neighbors (KNN) algorithm to enhance crop yield predictions by analyzing soil quality and climatic conditions. Sharma et al. (2022)[2] highlighted the importance of Random Forest and Decision Tree models, achieving an impressive accuracy of 89 in crop yield forecasting based on various environmental factors. Additionally, Sai Teja et al.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 2, November 2024

(2022) [3]utilized Support Vector Machines (SVM) to recommend profitable crops while integrating historical weather data to maximize productivity. These studies underscore the effectiveness of ML techniques in addressing the challenges posed by unpredictable climatic patterns, offering practical solutions to improve agricultural efficiency. Furthermore, Kuriakose and Singh (2022)[4] introduced Long Short- Term Memory (LSTM) networks for dynamic crop yield predictions based on soil fertility and weather patterns, emphasizing the critical role of advanced analytics in modern agriculture.
Collectively, these works illustrate a growing trend towards leveraging data-driven approaches to enhance crop management and decisionmaking in farming.

III. MOTIVATION AND OBJECTIVES

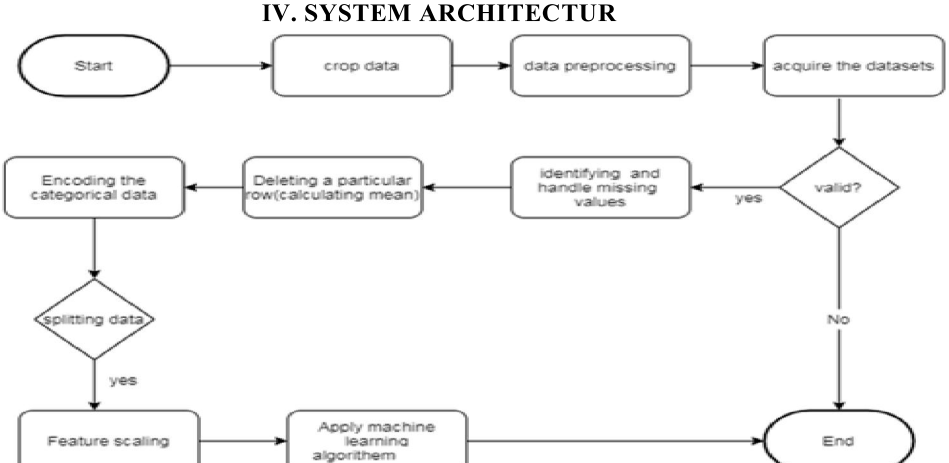
Motivation

The increasing global population is driving a significant demand for higher agricultural productivity to ensure food security. However, unpredictable weather patterns and climate change pose challenges to traditional yield prediction methods, making it essential for farmers to adapt. Machine learning technologies offer valuable solutions by optimizing the use of resources such as water and fertilizers, which is crucial for sustainable farming practices. By leveraging data-driven decision-making, predictive models provide actionable insights that enhance crop management and improve yields. This integration of advanced technologies not only helps farmers navigate the complexities of climate variability but also supports the development of sustainable agricultural practices. Accurate yield predictions can contribute significantly to maintaining food security while minimizing environmental impacts. In a world where agricultural challenges are becoming increasingly complex, adopting these innovative approaches is vital for building resilient and efficient farming systems that can meet the demands of a growing population. Ultimately, combining technology with sustainable practices holds the key to ensuring that agriculture can thrive in the face of global challenges, paving the way for a more secure and sustainable food future.

Objective

1. Develop Predictive Models: Utilize machine learning algorithms to forecast crop yields.
2. Analyze Environmental Factors: Incorporate weather data such as temperature, humidity, and precipitation.
3. Utilize Geographic and Soil Data: Include location-based and soil characteristics to improve prediction accuracy.
4. Enhance Decision-Making: Provide farmers with actionable insights for resource optimization.
5. Improve Agricultural Productivity: Help boost crop yields and efficiency by leveraging datadriven predictions

IV. SYSTEM ARCHITECTUR



V. PROJECT FEASIBILITY AND SCOPE

Project Feasibility

The target market for this initiative includes farmers, agricultural cooperatives, and government agencies. There is a notable trend toward the growing integration of technology in agriculture, which underscores the demand for advanced predictive tools. From a technical feasibility standpoint, essential technologies will include machine learning frameworks such as Scikit-learn and TensorFlow, data analytics tools, and a user-friendly interface. The implementation plan will focus on data collection, model training, and application development. In terms of economic feasibility, development costs are estimated to range from 50,000to70,000, with revenue potential through a subscription model and partnerships that could achieve breakeven within 2-3 years. Legal considerations will require compliance with data privacy laws and obtaining consent from farmers for data collection. Operationally, the system can be integrated into existing agricultural organizations, although training will be necessary for effective usage. The project timeline is anticipated to span approximately 6-12 months, with key milestones including data analysis and system launch. Lastly, potential risks such as resistance to technology adoption and data variability can be mitigated through strategies like training and pilot projects, ensuring a smoother transition to the new system.

Scope

1. Integration of More Data Sources: Incorporate satellite imagery, remote sensing data, and IoT sensor data for real-time monitoring. Use social and economic factors to enhance predictions.
2. Collaboration with Agricultural Institutions: Partner with universities and research institutions for ongoing research and development. Conduct field trials to validate predictions and improve model accuracy.
3. Regional Customization: Tailor predictions to specific crop types and geographic regions for more localized insights. Collaborate with agricultural experts to validate models and recommendations.

VI. CONCLUSION

This study revealed that the selected publications employ a diverse range of features based on the research scope and data availability. While all papers focus on yield prediction using machine learning, they vary in the specific features analyzed. Differences in scale, geographical location, and crop type further distinguish the studies. The choice of features often hinges on dataset availability and research objectives, and findings indicate that models with more features do not always yield the best performance. To determine the most effective model, both extensive and minimal feature sets should be evaluated. Various algorithms have been utilized across studies, yet no definitive conclusion emerges regarding the best model. However, it is evident that certain machine learning models, such as random forest, neural networks, linear regression, and gradient boosting trees, are more commonly applied. As neural networks are frequently used, we also examined the extent to which deep learning algorithms contribute to crop yield prediction. An analysis of 30 papers that employed deep learning identified CNN, LSTM, and DNN as the most favored algorithms, alongside others. This article aims to stimulate further research in crop yield prediction, with our future work focusing on developing a deep learningbased model.

REFERENCES

[1] A. K. Gajula, J. Singamsetty, V. C. Dodda, and L. Kuruguntla, “Prediction of crop and yield in agriculture using machine learning technique,” in 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kharagpur, India, Jul. 2021, pp. 1-5. doi: 10.1109/ICCCNT51525.2021.9579843.

[2] A. Sharma, A. Tamrakar, S. Dewasi, and N. S. Naik, “Early Prediction of Crop Yield in India using Machine Learning,” in 2022 IEEE Region 10 Symposium (TENSYMP), Mumbai, India, Jul. 2022, pp. 1-6. doi: 10.1109/TENSYMP54529.2022.9864490.

[3]M. S. Teja, T. S. Preetham, L. Sujihelen, C. Jancy, and M. P. Selvan, “Crop Recommendation and Yield Production using SVM Algorithm,” in 2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, May 2022, pp. 1-5. doi: 10.1109/ICICCS53718.2022.9788274.

[4]N. Suresh, N. V. K. Ramesh, S. Inthiyaz, P. P. Priya, K. N. Sowmika, K. V. N. Harish Kumar, M. Shaik, and B. N. K. Reddy, "Crop Yield Prediction Using Random Forest Algorithm," in 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, Mar. 2021, pp. 1-6. doi: 10.1109/ICACCS51430.2021.9441871.

[5]Jain A., "Analysis of growth and instability in the area, production, yield, and price of rice in India", Journal of Social Change and Development, vol. 2, pp. 46-66, N/A, 2018.

[6] Wolfert S, Ge L, Verdouw C, Bogaardt MJ, "Big data in smart farming– a review. Agricultural Systems", vol. 153, pp. 6980, May 2017.

[7] Sangeeta, Shruthi G. "Design and implemen tation of crop yield prediction model in agriculture." International Journal of Engineering Research Technology (IJERT), vol. 9, no. 4, pp. 305310, Apr. 2020.

[8] Squirrel search method for deep learningbased anomaly identification in videos L Malphedwar, TR Kumar Bulletin of Electrical Engineering and Informatics13 (2), 1197-1206,April 2024