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: 7.301
Volume 3, Issue 2, December 2023

Forecasting Crop Yield For Sustainable Agriculture

Sakshi Mundhe[1], Jitu Sanap[2], Pooja Jadhav[3], Vaishnavi Kalsadkar[4], Prof. Chiranjit Das[5]
Students, Department of Data Science[1,2,3,4]
Faculty, Department of Data Science[5]
G. H. Raisoni College of Engineering and Management, Wagholi, Maharashtra, India

Abstract: *Forecasting crop yield prediction is an very important decision support tool. Crop yield prediction is the process of using data and technology to estimate how much agricultural produce, such as crops, will be harvested in a specific area for a given period. This involves analyzing various factors like weather patterns, soil quality, historical data, and farming practices to make educated guesses about the future harvest. It helps farmers, policymakers, and food supply chain stakeholders make informed decisions regarding planting, resource allocation, and food security. Essentially, it's like forecasting the future of crop production to ensure efficient and sustainable agriculture.*

Forecasting crop yield is a complex and vital endeavor in modern agriculture, aiming to foresee the quantity of crops that will be harvested in a specific area during a particular growing season. This process relies on a synergy of data-driven analysis, technological advancements, and the integration of various factors that influence crop growth and productivity. In essence, it involves the art and science of anticipating nature's bounty, allowing us to make informed decisions, enhance food security, and optimize resource allocation.

Keywords: Forecasting crop yield, Decision support system, Systematic literature review, Machine learning

I. INTRODUCTION

Agriculture stands as the bedrock of human civilization, offering sustenance, livelihoods, and economic stability to communities across the globe. However, this vital sector confronts unprecedented challenges in the 21st century. The inexorably expanding global population, the uncertainties stemming from climate change, and the compelling need to conserve our precious natural resources all coalesce, demanding innovative solutions that safeguard both food security and environmental sustainability. In this context, our research on the paramount issue of forecasting crop yields. Crop yield prediction assumes a pivotal role in addressing these challenges, as it equips farmers, policymakers, and stakeholders with the knowledge necessary for judicious decision-making. Precision in predictions facilitates timely resource allocation, risk mitigation, and, most crucially, the capacity to fulfill the nutritional requirements of a burgeoning populace while minimizing the ecological impact of agriculture. Our research endeavors to delve into cutting-edge methodologies and technologies for prognosticating crop yields, recognizing the multifaceted factors that exert influence on crop growth. These factors encompass intricate meteorological patterns, soil vitality, the dynamics of pests and diseases, and the ever-evolving landscape of agricultural practices. By harnessing the prowess of data science, machine learning, and remote sensing technologies, our aim is to construct a resilient and adaptable framework capable of forecasting crop yields with a remarkable degree of precision and granularity.

At its core, crop yield prediction begins with understanding the fundamental elements that shape the agricultural landscape. Factors such as soil quality, moisture levels, and nutrient content form the foundation upon which predictions are built. Soil samples and laboratory analysis help assess these parameters, ensuring a comprehensive understanding of the land's potential.

Weather patterns and climatic conditions are pivotal components in yield prediction. Monitoring temperature, precipitation, sunlight, and humidity provides critical data points to anticipate how environmental variables will impact crop growth. Advanced meteorological models and historical weather data contribute to the precision of these predictions.

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: 7.301

Volume 3, Issue 2, December 2023

Choosing the right machine learning algorithm is pivotal. Various algorithms such as regression, decision trees, random forests, support vector machines, and neural networks can be considered.

Key Aspects:

1. The significance of crop yield forecasting in achieving food security and sustainable agriculture.
2. An overview of the challenges posed by climate change and population growth.
3. The role of technology and data-driven approaches in improving yield predictions.
4. Case studies and real-world applications of crop yield forecasting models.
5. The potential socio-economic and environmental benefits of accurate crop yield forecasts.

By advancing our understanding of crop yield forecasting, we aspire to contribute to the overarching goals of sustainable agriculture—balancing food production with ecological conservation and ensuring that future generations inherit a resilient and flourishing planet. This research seeks to bridge the gap between science and practice, fostering a more sustainable and prosperous future for agriculture and society as a whole

II. LITERATURE REVIEW

[1] crop yield prediction using machine learning, International Journal of Science and Research (IJSR). In this paper it uses machine learning, random forest algorithm and climate analysis like whether, climate etc for predicting crop yield.

[2] crop prediction using machine learning approaches, International Journal of Engineering Research. This paper describes about the recommendation of crop for particular land considering the factors of soil, rainfall, humidity, and pH. The system takes required inputs from farmer and analyze it using machine learning, SVM and decision tree.

[3] smart farming crop yield prediction using machine learning, International Research Journal of Engineering and Technology (IRJET). This paper represent the concept of smart farming using machine learning algorithms like support vector machine, random forest for predicting the crop yield considering the parameters of soil type, temperature, atmospheric pressure, humidity and crop type.

[4] smart crop prediction using IoT and machine learning, K.J. Somaiya of Engineering Mumbai This smart crop management system uses IoT based sensor (for temperature, humidity, soil type) and machine learning algorithms for predict the most suitable crop in that environment

[5] crop yield prediction using machine learning and deep learning techniques. In this paper it uses both machine learning and deep learning together for crop prediction. They have used random forest, SVM, gradient decent, long term-short memory and lasso regression techniques. In this paper they said random forest performed well as compared to others

[6] crop prediction using IoT and machine learning, International Research Journal of Engineering and Technology (IRJET). This paper describes about crop prediction based on machine learning and IoT techniques. It uses sensors like rainfall sensor this tells how much rain is falling, humidity sensor, which determines amount of moisture and temperature in the air, LDR, and digital temperature and humidity for accurate crop prediction.

III. PROBLEM STATEMENT

The objective of this project is to develop a comprehensive and reliable system for forecasting crop yields in a sustainable agriculture context. The system should leverage advanced data analytics, machine learning, and remote sensing technologies to provide accurate and timely predictions of crop yields for various crops and regions.

3.1 Data Collection and Integration:

Gather historical and real-time data on various factors affecting crop yields, including weather data, soil quality, crop types, pest and disease incidence, and agricultural practices.

Integrate data from multiple sources, such as weather stations, satellite images government reports, and farmer surveys, into a centralized database.

3.2Data Preprocessing and Feature Engineering:

Clean and preprocess the collected data to remove outliers, handle missing values, and ensure data consistency.

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: 7.301
Volume 3, Issue 2, December 2023

Conduct feature engineering to extract relevant features and variables that impact crop yields, such as temperature, precipitation, soil moisture, and land use.

3.3 Validation and Performance Evaluation:
Validate the developed models using historical crop yield data and real-time observations.
Implement cross-validation techniques to assess model performance and generalizability.
Evaluate the models' accuracy, precision, recall, F1-score, and other relevant metrics.

3.4 Stakeholder Engagement:
Collaborate with local farmers, agricultural experts, and government agencies to gather domain-specific knowledge and feedback for model improvement.
Conduct workshops and training sessions to educate stakeholders on how to use the forecasting system effectively.

3.5 User-Friendly Interface:
Design a user-friendly interface or dashboard that allows farmers, agricultural agencies, and policymakers to access crop yield forecasts easily.
Provide interactive visualizations and customizable options for users to explore forecasted data.
The project should culminate in a crop yield forecasting system that provides accurate predictions, is user-friendly, and can be readily adopted by farmers, agricultural organizations, and policymakers to support sustainable agriculture practices.

IV. METHODOLOGY

In a crop yield prediction project using machine learning, we commence by gathering historical data encompassing crop production records, meteorological data, soil quality information, and agricultural practices. Subsequently, the collected data undergoes thorough preprocessing, including handling missing values and data normalization. Feature selection and engineering steps follow, aimed at identifying the most influential variables. After partitioning the dataset into training and testing subsets, we explore diverse machine learning algorithms, ultimately selecting one for training and fine-tuning using the training set. Model performance evaluation, conducted with appropriate metrics, guides the final model selection. Post-validation, the model is deployed, possibly through a web application or API, to facilitate real-time yield predictions. Continuous monitoring, feedback integration, and ethical considerations remain pivotal throughout the project's life cycle to ensure its accuracy, utility, and ethical compliance in supporting farmers with reliable yield forecasts
The first stage is planning the review. In this, research question are identification addition to research questions, initial search strings, publication selection criteria are also defined.
The second stage is conducting the review, when conducting the review, the publications were selected by going through all the databases. the data was extracted, it means that their information regarding authors, year of publication, type of publication and more information regarding to the research questions were stored.

V. HOW DOES MACHINE LEARNING WORKS

Machine learning works by enabling computer systems to learn from data and make predictions or decisions based on that learning. Here's a high-level overview of how machine learning works:
5.1 Data Collection:
The process begins with the collection of relevant data. This data can come from various sources, such as databases, text documents, images, or user interactions. Data is essential because it serves as the foundation for training and evaluating machine learning models.

5.2 Data Preprocessing:
Raw data often contains noise, inconsistencies, and missing values. Data preprocessing involves cleaning and transforming the data into a format suitable for training machine learning models. This may include tasks like data

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: 7.301
Volume 3, Issue 2, December 2023

cleaning, feature scaling, and feature engineering.

5.3 Model Selection:
Machine learning involves choosing an appropriate algorithm or model architecture that is well-suited to the specific task or problem. Different algorithms are used for different types of tasks, such as classification, regression, clustering, or reinforcement learning.

5.4 Model Training:
The selected machine learning model is trained using the training dataset. During training, the model learns the underlying patterns and relationships in the data. It adjusts its internal parameters to minimize the difference between its predictions and the actual target values (in supervised learning) or to find the optimal structure (in unsupervised learning).

5.5 Model Evaluation:
After training, the model's performance is evaluated using the testing dataset. This evaluation assesses how well the model generalizes to new, unseen data. Common evaluation metrics depend on the type of problem and can include accuracy, mean squared error, F1-score, and more.

5.6 Model Deployment:
If the model performs well during evaluation, it can be deployed in a real-world application to make predictions or decisions. This could involve integrating the model into a software application, a website, or an autonomous system, depending on the use case.
Machine learning encompasses a wide range of algorithms and techniques, including supervised learning, unsupervised learning, reinforcement learning, and deep learning, among others. The choice of algorithm and the quality of data play crucial roles in the success of a machine learning project. Additionally, domain expertise and careful consideration of ethical and fairness aspects are essential for responsible machine learning applications.

VI. RELATED WORK

- Crop Modeling: Developing crop models based on collected data helps in predicting yields. Models like the Crop Growth Simulation Models (e.g., DSSAT, APSIM) simulate crop growth and yield under different scenarios.
- Weather Forecasting: Accurate weather forecasts are crucial for crop yield predictions. Collaborating with meteorological agencies and using weather data and forecasts can improve forecasting accuracy.
- Soil Health Assessment: Soil testing and assessment can provide insights into soil quality and nutrient levels, which are critical for crop growth. Soil health assessments can inform nutrient management practices.
- Remote Sensing: Satellite imagery and remote sensing technologies can monitor crop health, detect pests and diseases, and assess vegetation conditions. These data sources are valuable for yield forecasting.

VII. CONCLUSION

In Conclusion, forecasting crop yield is not only essential for the economic well-being of farmers and the food security of nations but also a critical tool in promoting sustainable agriculture. With the integration of modern technologies and a commitment to data-driven decision-making, we can make agriculture more efficient, resilient, and environmentally friendly. By leveraging machine learning models, historical weather data, soil information, and crop-specific factors, we were able to forecast accurate predictions. This technology has the potential to bring the new technology in agriculture by assisting farmers in making informed decisions, optimizing resource allocation, and ultimately increasing food production efficiency. However, it's important to continue refining the models, incorporating real-time data, and ensuring accessibility for farmers in diverse regions to maximize the project's impact.

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: 7.301

Volume 3, Issue 2, December 2023

REFERENCES

[1] Ahamed, A.T.M.S., Mahmood, N.T., Hossain, N., Kabir, M.T., Das, K., Rahman, F., Rahman, R.M., 2015. Applying data mining techniques to predict annual yield of major crops and recommend planting different crops in different districts in Bangladesh. In: 2015 IEEE/ACIS 16th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, SNPD 2015 - Proceedings, https://doi.org/10.1109/SNPD.2015.7176185

[2] Ahmad, I., Saeed, U., Fahad, M., Ullah, A., Habib-ur-Rahman, M., Ahmad, A., Judge, J., 2018. Yield forecasting of spring maize using remote sensing and crop modeling in Faisalabad-Punjab Pakistan. J. Indian Soc. Remote Sens. 46 (10), 1701–1711. https://doi.org/10.1007/s12524-018-0825-8..

[3] Ayodele, T.O., 2010. Introduction to Machine Learning.Baldi,P.,2012.Autoencoders, unsupervised learning, and deep architectures. In: Proceedings of ICML workshop on unsupervised and transfer learning, pp. 37–49. Baral, S., Kumar Tripathy, A., Bijayasingh, P., 2011.Yield Prediction Using Artificial Neural Networks, pp. 315–317. https://doi.org/10.1007/978-3-642-19542-6_57

[4] Bargoti, S., Underwood, J.P., 2017. Image segmentation for fruit detection and yield estimation in apple orchards. J. Field Rob. 34 (6), 1039–1060. https://doi.org/10. 1002/rob.21699

[5] Beulah, R., 2019. A survey on different data mining techniques for crop yield prediction. Int. J. Comput.Sci.Eng.7(1),738–744. https://doi.org/10.26438/ijcse/v7i1.738744.

Bhojani, S.H., Bhatt, N., 2020. Wheat crop yield prediction using new activation functions in neural network. Neural Comput. Appl. 1–11.

[6] Cheng, H., Damerow, L., Sun, Y., Blanke, M., 2017. Early yield prediction using image analysis of apple fruit and tree canopy features with neural networks. J. Imag. 3 (1), 6. https://doi.org/10.3390/jimaging3010006.

[7] Crane-Droesch, A., 2018. Machine learning methods for crop yield prediction and climate change impact assessment in agriculture. Environ. Res. Lett. 13 (11), 114003. https://doi.org/10.1088/1748-9326/aae159

[8] Everingham, Y., Sexton, J., Skocaj, D., Inman-Bamber, G., 2016. Accurate prediction of sugarcane yield using a random forest algorithm. Agron. Sustainable Dev. 36 (2). https://doi.org/10.1007/s13593-016-0364-z.

[9] Fernandes, J.L., Ebecken, N.F.F., Esquerdo, J.C.D.M., 2017. Sugarcane yield prediction in Brazil using NDVI time series and neural networks ensemble. Int. J. Remote Sens. 38 (16), 4631–4644. https://doi.org/10.1080/01431161.2017.1325531.

[10] Filippi, P., Jones, E.J., Wimalathunge, N.S., Somarathna, P.D.S.N., Pozza, L.E., Ugbaje, S.U., Bishop, T.F.A., 2019a. An approach to forecast grain crop yield using multilayered, multi-farm data sets and machine learning. Precis. Agric. 1–15. https://doi. org/10.1007/s11119-018-09628-4.

[11] Gandhi, N., Armstrong, L., 2016. Applying data mining techniques to predict yield of rice in humid subtropical climatic zone of India. In: Proceedings of the 10th INDIACom; 2016 3rd International Conference on Computing for Sustainable Global Development, INDIACom 2016, 1901–1906. Retrieved from https://ieeexplore.ieee.org/abstract/document/7724597/.

[12] Girish, L., Gangadhar, S., Bharath, T., Balaji, K., n.d. Crop Yield and Rainfall Prediction in Tumakuru District using Machine Learning. Ijream.Org. Retrieved from https:// www.ijream.org/papers/NCTFRD2018015.pdf

[13] Kashif Mehboob, Junaid Arshad and Muhammad Mubashir Khan January 2018 Secure digital voting system based on Blockchain technology Int. J. of Electronic Government Research 14(1) pp 53-62. https://core.ac.uk/download/pdf/155779036.pdf

[14] Noor Mohammedali and Ali Al-Sherbaz 2019 Election system based on Blockchain technology Int. J. of Computer Science & Information Technology (IJCSIT) 11, No 5 pp 13-31. https://aircconline.com/ijcsit/V11N5/11519ijcsit02.pdf

[15] Archit Pandey, Mohit Bhasi and K Chandrasekaran 2019 VoteChain: A Blockchain based e-voting system Global Conf. for Advancement in Technology (GCAT) (Bangaluru: India) pp 1-4. https://ieeexplore.ieee.org/document/8978295

[16] Clement Chan Zheng Wei and Chuah Chai Wen 2018 Blockchain-based electronic voting protocol Int. J. on Informatics Visualization 2 pp 336- 341. http://doi.org/10.30630/joiv.2.4-2.174

[17] Haibo Yi 2019 Securing e-voting based on blockchain in P2P network J. on Wireless Communications and Networking (EURASIP). https://doi.org/10.1186/s13638-019-1473-6

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: 7.301
Volume 3, Issue 2, December 2023

[18] Freya Sheer Hardwick, Apostolos Gioulis, Raja Naeem Akram and Konstantinos Markantonakis 2018 E-voting with Blockchain: An e-voting protocol with decentralization and voter privacy Int. Conf. on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData) (Ghaziabad: India). https://ieeexplore.ieee.org/document/8777471
[19] Ahmed Ben Ayed 2017 A conceptual secure Blockchain-based electronic voting system Int. J. of Network Security & Its Applications (IJNSA) 9, No.3.
[20] Shah, A., Dubey, A., Hemnani, V., Gala, D., Kalbande, D.R., 2018. Smart Farming System: Crop Yield Prediction Using Regression Techniques. Springer, Singapore, pp. 49–56. https://doi.org/10.1007/978-981-10-8339-6_6.
[21] Shekoofa, A., Emam, Y., Shekoufa, N., Ebrahimi, M., Ebrahimie, E., 2014. Determining the most important physiological and agronomic traits contributing to maize grain yield through machine learning algorithms: a new avenue in intelligent agriculture. PLoS ONE 9 (5), e97288. https://doi.org/10.1371/journal.pone.0097288.
[22] Shidnal, S., Latte, M.V., Kapoor, A., 2019. Crop yield prediction: two-tiered machine learning model approach. Int. J. Inf. Technol. 1–9.
[23] Šmite, D., Wohlin, C., Gorschek, T., Feldt, R., 2010. Empirical evidence in global software engineering: a systematic review. Empirical Softw. Eng. 15 (1), 91–118. https://doi. org/10.1007/s10664-009-9123-y.
[24] Somvanshi, P., Mishra, B.N., 2015. Machine learning techniques in plant biology. In: PlantOmics: The Omics of Plant Science. Springer India, New Delhi, pp. 731–754. https://doi.org/10.1007/978-81-322-2172-2_26.
[25] Su, Y.X., Xu, H., Yan, L.J., 2017. Support vector machine-based open crop model (SBOCM): case of rice production in China. Saudi J. Biol. Sci. 24 (3), 537–547.
[26] Ahamed, A. M. S., Mahmood, N. T., Hossain, N., Kabir, M. T., Das, K., Rahman, F., & Rahman, R. M. (2015, June). Applying data mining techniques to predict annual yield of major crops and recommend planting different crops in different districts in Bangladesh. In 2015 IEEE/ACIS 16th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD) (pp. 1-6). IEEE.
[27] Shastry, A., Sanjay, H. A., & Hegde, M. (2015, June). A parameter based ANFIS model for crop yield prediction. In 2015 IEEE International Advance Computing Conference (IACC) (pp. 253-257). IEEE
[28] Veenadhari, S., Misra, B., & Singh, C. D. (2014, January). Machine learning approach for forecasting crop yield based on climatic parameters. In 2014 International Conference on Computer Communication and Informatics (pp. 1-5). IEEE.
[29] Molden, D.; Vithanage, M.; de Fraiture, C.; Faures, J.M.; Gordon, L.; Molle, F.; Peden, D. Water Availability and Its Use in Agriculture. Treatise Water Sci. 2011, 4, 707–732. [CrossRef].
[30] EastAfrican. 3000 Rwandan Families Face Hunger due to Drought—Rwanda|ReliefWeb, (n.d.). Available online: https://reliefweb.int/report/rwanda/3000-rwandanfamilies-face-hunger-due-drought (accessed on 3 February 2022).
[31] Kironde, E.G. Rwanda State of Environment and Outlook Report. In REMA; 2016; 1, pp. 93–115.Available online: https://www.rema.gov.rw/soe/chap9.php (accessed
[32] Chakraborty, D.; Saha, S.; Sethy, B.K.; Singh, H.D.; Singh, N.; Sharma, R.; Chanu, A.N.; Walling, I.; Anal, P.R.; Chowdhury, S.; et al. Usability of the Weather Forecast for Tackling Climatic Variability and Its Effect on Maize Crop Yield in Northeastern Hill Region of India. Agronomy 2022, 12, 2529. [CrossRef
[33] The Internet of Things Applications for Challenges and Related Future Technologies & Development Mritunjay Kumar1, Km Annoo2, Raman Kumar Mandal3 https://scholar.google.com/citations?view_op=view_citation &hl=en&user=IRSdLtMAAAAJ&citation_for _view=IRSdLtMAAAAJ:u5HHmVD_uO8
[34] Characteristics of Research Process and Methods for Web-based Research Support Systems Mritunjay Kumar, Km Annoo, Harish Chandra Maurya https://scholar.google.com/citations?view_op=view _citation&hl=en&user=IRSdLtMAAAAJ&citation _for_view=IRSdLtMAAAAJ:d1gkVwhDpl0C
[35] Security of Internet of Things of Application, Challenges and Related Future Technologies Prabhanjan Kumar, Sneha Mishra, Mritunjay Kumar, Arjit Tomar, Jayati Mukherjee hΣps://scholar.google.com/citaΘons?view_ op=view_citaΘon&hl=en&user=IRSdLtMAAAAJ&citaΘon _for_view=IRSdLtMAAAAJ:738O_yMBCRsC