IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Analysis On Predicting Crop Yields With Machine Learning

[1] Mahesha L S, Lecturer, Department of Computer Science & Engineering, L V Government Polytechnic, Hassan, India.

ABSTRACT: Crop prediction is a critical component of modern agriculture, enabling farmers to make informed decisions about crop selection and cultivation practices. Machine learning techniques have revolutionized this process by leveraging historical data, weather patterns, and various environmental factors to predict the most suitable crops for a given region. The proposed system begins with the collection of comprehensive data, including historical crop yields, weather conditions, soil properties, and geographic information for a specific area Machine learning algorithms such as SVM, Random Forests and KNN are indispensable tools in the realm of agriculture. By analyzing vast datasets, these models contribute to informed decision-making, helping farmers choose the most suitable crop varieties for future cultivation based on a comprehensive understanding of the underlying patterns and relationships in the data.

KEYWORDS: Crop prediction, machine learning

I.INTRODUCTION

Agriculture is the backbone of many economies around the world, providing food, fiber, and raw materials for various industries. As the global population continues to grow, the demand for agricultural products is increasing, making efficient and sustainable farming practices more crucial than ever. Crop prediction, the process of forecasting which crops are best suited for a specific region and when to plant them, plays a pivotal role in ensuring the success and sustainability of agriculture. Machine learning, a subset of artificial intelligence, has emerged as a powerful tool in this domain by leveraging data-driven techniques to make accurate predictions and informed decisions for crop selection and cultivation practices.

Traditional methods of crop selection and planting have typically relied on the knowledge and experience of farmers, local climate patterns, and historical practices. While these approaches have served well in the past, they may not be sufficiently equipped to address the challenges of modern agriculture, including climate change, shifting weather patterns, and the need for optimized resource management. This is where machine learning steps into revolutionize the way we approach crop prediction.

Crop prediction using machine learning involves the utilization of historical data, including past crop yields, weather information, soil properties, geographic factors, and various environmental parameters. This data is used to train machine learning models, enabling them to identify intricate patterns, relationships, and dependencies among different variables. Once these models are trained, they can make predictions about which crop varieties are best suited for a particular region and when they should be planted to maximize yield and quality.

The power of machine learning in crop prediction is not limited to historical data alone. Real-time data sources, such as current weather conditions, satellite imagery, and IoT-based sensor networks, can be integrated into the models to provide up-to-date and accurate information for decision-making. This dynamic approach ensures that farmers and agricultural stakeholders have access to timely recommendations, enabling them to adapt to changing conditions and optimize their farming practices. The proposed agricultural decision support system incorporates a range of machine learning algorithms, including Support Vector Machines(SVM), Random Forests, k-Nearest Neighbors(KNN), and neural networks. This system aims to enhance farming practices by utilizing diverse datasets encompassing soil composition, climate conditions, and historical crop performance. Through careful data preprocessing, feature engineering, and ensemble modeling, the system provides precise recommendations for optimal crop varieties based on localized conditions. The user-friendly interface allows farmers to easily access and interprets the recommendations, fostering informed decision-making.

II.PROPOSEDMETHODOLOGY

1. **Data Collection:**
   a. Collect historical data on crop yields, weather conditions, and soil properties for the target region.
   b. Integrate real-time data sources such as weather forecasts, satellite imagery, and IoT sensor data.
   c. Utilize geographic information system (GIS) data to capture geographical factors.
2. **Data Preprocessing:**
   a. Clean and preprocess the collected data, addressing missing values, outliers, and inconsistencies.
   b. Standardize or normalize numerical features to ensure uniform scales.
   c. Encode categorical variables using techniques like one-hot encoding.
3. **Feature Engineering:**
   a. Extract relevant features from the data, including seasonality, temperature indices, precipitation patterns, and soil quality indicators.
   b. Create additional features if needed, such as growth stage indicators and pest and disease risk scores.
4. **Data Splitting:**
   a. Divide the data set into training, validation, and test sets to evaluate model performance.
   b. Consider techniques like cross-validation to optimize hyper parameters and avoid over fitting.
5. **Machine Learning Model Selection:**
   a. Choose appropriate machine learning algorithms based on the nature of the problem. Common choices include:
      i. Random Forests
      ii. Support Vector Machines
      iii. Gradient Boosting
      iv. Neural Networks
      v. Decision Trees
6. **Model Training:**
   a. Train the selected machine learning models on the training data set using historical data and features.
   b. Tune hyper parameters and validate model performance on the validation set.
7. **Model Evaluation:**
   a. Evaluate model performance using various metrics, such as mean absolute error (MAE), root mean square error (RMSE), or R-squared.
   b. Consider using domain-specific evaluation criteria, such as crop-specific yield metrics.
8. **Crop Recommendation and Decision Support:**
   a. Usethetrainedmodeltomakecroprecommendationsbasedonreal-timeandhistoricaldata.
   b. Provide farmers with actionable insights on optimal crop selection, planting times, and resource management.

Linear Regression:

- Ordinary Least Squares (OLS) method is used to estimate the coefficients (ββ) that minimize the sum of squared residuals (εε).
- Assumptions such as linearity, homoscedasticity, and normality of residuals are checked to ensure the validity of the model.

Polynomial Regression:

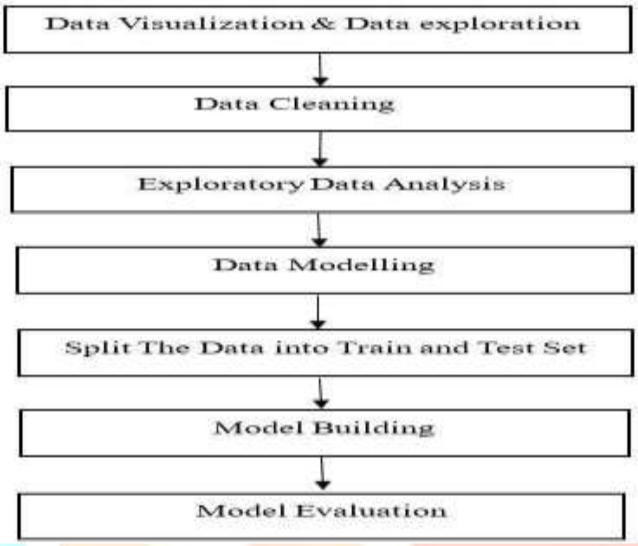
- Higher-order polynomial functions are fitted to capture non-linear relationships between predictors and crop yield.
- Model complexity is controlled using techniques like regularization (e.g.,Ridge, Lasso) to prevent over fitting.

Decision Trees:

- Recursive partitioning algorithm (e.g.,CART,ID3) is used to construct decision trees by selecting splits that maximize information gain or minimize impurity.
- Pruning techniques are applied to prevent over fitting and improve generalization performance.

Random Forests:

- Bootstrap aggregation (bagging) is used to train multiple decision trees on random subsets of the data, reducing variance and improving robustness.
- Feature importance measures (e.g., Gini importance) are computed to assess the contribution of predictors to the model.



III.CONCLUSION

In this paper, we have undertaken a comprehensive exploration of the field of crop prediction using machine learning, highlighting the significant progress, challenges, and promising directions that have emerged in recent years. The importance of crop prediction in modern agriculture cannot be overstated, as it offers a data-driven approach to address the dynamic and complex challenges faced by farmers and agricultural stakeholders.

The reviewed literature demonstrates the vast potential of machine learning in crop prediction. Researchers and practitioners have leveraged a wide array of machine learning algorithms and data sources to build accurate predictive models. These models take into account historical crop yields, weather conditions, soil properties, geographic information. By doing so, they provide farmers with valuable insights for informed decision-making, optimizing crop selection, planting schedules, and resource management.

One of the notable trends in this field is the integration of deep learning techniques, particularly convolutional neural networks (CNNs), which have shown great promise in processing remote sensing data and images for crop mapping, disease detection, and yield prediction. The adoption of these advanced deep learning models marks a significant shift in the capabilities of crop prediction systems, offering higher accuracy and greater adaptability to changing environmental conditions.
Data availability, data quality, and the interpretability of machine learning models remain important considerations.

IV.FUTURE SCOPE

The future of crop yield prediction using machine learning algorithms is brimming with possibilities and avenues for exploration:

1. **Enhanced Model Interpretability:** Develop interpretable machine learning models and visualization techniques to elucidate the underlying factors influencing crop yield predictions and foster trust among end-users.
2. **Integration of Multi-Modal Data:** Explore the fusion of heterogeneous data sources, including remote sensing imagery, IoT sensor data, genetic information, and socio-economic indicators, to capture complex interactions and improve prediction accuracy.
3. **Dynamic Model Adaptation:** Develop adaptive and self-learning models capable of continuously updating and refining predictions in response to changing environmental conditions, management practices, and emerging threats.
4. **Predictive Analytics for Precision Agriculture:** Expand the application of crop yield prediction models to support precision agriculture practices, including variable rate technology, site-specific management, and optimization of inputs for improved resource efficiency.
5. **Decision Support Systems:** Integrate crop yield prediction models with decision support systems and mobile applications to provide real-time recommendations, alerts, and risk assessments to farmers and stakeholders.
6. **Climate Change Resilience:** Incorporate climate change scenarios, resilience strategies, and adaptive management practices into crop yield prediction models to enhance resilience and sustainability in agricultural systems.
7. **Global Collaboration and Data Sharing:** Foster collaboration among researchers, institutions, and governments to share data, best practices, and methodologies for crop yield prediction across different regions and climates.

REFERENCES

[1] Dahikar S and Rode S V 2014 Agricultural crop yield prediction using artificial neural network approach International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering vol 2 Issue 1 pp 683-6.

[2] Suresh A, Ganesh P and Ramalatha M 2018 Prediction of major crop yields of Tamilnadu using K-means and Modified KNN 2018 3rd International Conference on Communication and Electronics Systems (ICCES) pp 88-93 doi: 10.1109/CESYS.2018.8723956.

[3] Medar R, Rajpurohit V S and Shweta S 2019 Crop yield prediction using machine learning techniques IEEE 5th International Conference for Convergence in Technology (I2CT) pp 1-5 doi: 10.1109/I2CT45611.2019.9033611.

[4] Nishant P S, Venkat P S, Avinash B L and Jabber B 2020 Crop yield prediction based on Indian agriculture using machine learning 2020 International Conference for Emerging Technology (INCET) pp 1-4 doi: 10.1109/INCET49848.2020.9154036.

[5] Kalimuthu M, Vaishnavi P and Kishore M 2020 Crop prediction using machine learning 2020 Third International ConferenceonSmartSystemsandInventiveTechnology(ICSSIT)pp926-32doi: 0.1109/ICSSIT48917.2020.9214190.

[6] Geetha V, Punitha A, Abarna M, Akshaya M, Illakiya S and Janani A P 2020 An effective crop prediction using random forest algorithm 2020 International Conference on System, Computation, Automation and Networking (ICSCAN) pp 1-5 doi: 10.1109/ICSCAN49426.2020.9262311

[7] Pande S M, Ramesh P K, Anmol A, Aishwaraya B R, Rohilla K and Shaurya K 2021 Crop recommender system usingmachinelearningapproach20215thInternationalConferenceonComputingMethodologiesandCommunication (ICCMC) pp 1066-71 doi: 10.1109/ICCMC51019.2021.9418351.

[8] Sellam V, and Poovammal E 2016 Prediction of crop yield using regression analysis Indian Models AICECS 2021 Journal of Physics: Conference Series 2161 (2022) 012033 IOP Publishing doi:10.1088/1742-6596/2161/1/012033 Journal of Science and Technology vol 9(38) pp 1-5.

[9] Bharath S, Yeshwanth S, Yashas B L and Vidyaranya R Javalagi 2020 Comparative Analysis of Machine Learning Algorithms in The Study of Crop and Crop yield Prediction International Journal of Engineering Research & Technology (IJERT) NCETESFT – 2020 vol 8 Issue 14.

[10] Mahendra N, Vishwakarma D, Nischitha K, Ashwini and Manjuraju M. R 2020 Crop prediction using

machine learning approaches, International Journal of Engineering Research & Technology (IJERT) vol 9 Issue 8 (August2020).
[11] Gulati P and Jha S K 2020 Efficient crop yield prediction in India using machine learning techniques International Journal of Engineering Research & Technology (IJERT) ENCADEMS – 2020 vol 8 Issue 10.
[12] GuptaA,NagdaD,NikhareP,SandbhorA,2021,SmartcroppredictionusingIoTandmachinelearning International Journal of Engineering Research & Technology (IJERT) NTASU – 2020 vol 9 Issue 3.
[13] Anna Chlingaryan, Machine Learning Approaches for Crop Yield Prediction and Nitrogen Status Estimation inPrecision Agriculture Elseveir – 2018 vol151.
[14] PriyabartaKaramkar , AReviewofRemoteSensing-Based Crop Mappingand Classification, Elseveir – Jan2024 vol33
[15] Minaee, Shervin, Nal Kalchbrenner, Erik Cambria, Narjes Nikzad, Meysam Chenaghlu, and Jianfeng Gao. (2021) "Deep Learning--Based Text Classification: A Comprehensive Review." ACM Computing Surveys 54 (3): 1–40. doi:10.1145/3439726.
[16] Belgiu, Mariana, and Lucian Drăguţ. (2016) "Random Forest in Remote Sensing: A Review of Applications and Future Directions." ISPRS Journal of Photogrammetry and Remote Sensing 114 (April): 24–31. doi:10.1016/j.isprsjprs.2016.01.011.
[17] Suthaharan, Shan. (2016) "Support Vector Machine." In Machine Learning Models and Algorithms for Big Data Classification: Thinking with Examples for Effective Learning, edited by Shan Suthaharan, 207–235. Integrated Series in Information Systems. Boston, MA: Springer US. doi:10.1007/978-1-4899-7641-3_9.
[18] Hochreiter, Sepp, A. Steven Younger, and Peter R. Conwell. (2001) "Learning to Learn Using Gradient Descent." In Artificial Neural Networks — ICANN 2001, edited by Georg Dorffner, Horst Bischof, and Kurt Hornik, 87–94. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer. doi:10.1007/3-540-44668-0_13.
[19] Sherstinsky, Alex. (2020) "Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) Network." Physica D: Nonlinear Phenomena 404 (March): 132306. doi:10.1016/j.physd.2019.132306.
[20] Yu,Yong,XiaoshengSi,ChanghuaHu,andJianxunZhang.(2019)"AReview ofRecurrentNeuralNetworks: LSTM Cells and Network Architectures." Neural Computation 31 (7): 1235–1270. doi:10.1162/neco_a_01199.
[21]Ranstam,J,andJACook.(2018)"LASSORegression."BritishJournalofSurgery105(10):1348. doi:10.1002/bjs.10895.
[22] https://data.world/thatzprem/agriculture-india
[23]https://agriculture.rajasthan.gov.in/content/agriculture/en/Agriculture-Department-dep/agriculturestatistics.html
[24] https://www.indiawaterportal.org/met_data[25]https://water.rajasthan.gov.in/content/water/en/waterresourcesdepartment/WaterManagement/IWRM/annua lrainfall.html#
[26]https://agriculture.rajasthan.gov.in/content/agriculture/en/Agriculture-Department-dep/agriculturestatistics.html
[27]https://agriculture.rajasthan.gov.in/content/agriculture/en/AgricultureDepartmentdep/DepartmentalIntroduction/Agro-Climatic-Zones.