e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:11/November-2024 Impact Factor- 8.187 www.irjmets.com

CROP YIELD AND PRICE PREDICTION USING MULTIFACTORIAL ANALYSIS

Tanmay Chaudhari[*1], Umesh Sake[*2], Vinanti Shinde[*3], Srushti Ghise[*4], Nikita Girase[*5], Dr. Sulochana Sonkamble[*6]

[*1,2,3,4,5]Department Of Computer Engineering, JSPM Narhe Technical Campus, Pune, Maharashtra, India.

[*6]Prof., Department Of Computer Engineering, JSPM Narhe Technical Campus, Pune, Maharashtra, India.

DOI : https://www.doi.org/10.56726/IRJMETS63856

ABSTRACT

This system presents a novel approach for Crop yield prediction by integrating convolutional long short-term memory (Conv LSTM), three-dimensional convolutional neural network(CNN), and vision transformers (ViT). Leveraging multispectral remote sensing data, the model utilizes CNNs to capture spatial hierarchies, ConvLSTM for temporal sequencing, and ViT for global context analysis, enabling the extraction of complex patterns in agricultural datasets. According on experimental findings, the suggested model performs noticeably better than current techniques, evidenced by lower root mean square error and higher correlation coefficients. The CNN component effectively extracts spatial features, while ConvLSTM captures crop growth dynamics over time, and ViT refines these features using self-attention mechanisms. This approach enhances crop yield prediction accuracy, assisting farmers in resource optimization, irrigation scheduling, and fertilizer application, ultimately promoting sustainable agricultural practices. The model's robustness across diverse conditions further demonstrates its applicability to various crops and geographic regions, marking a valuable contribution to agricultural remote sensing for large-scale multispectral data analysis.

Keywords: Agricultural Data Processing, Crop Yield Forecasting, Deep Learning, Model Generalization.

I. INTRODUCTION

Precision agriculture has become a vital area in recent years for tackling the problems of sustainability and global food security. Yield prediction is a crucial component of precision agriculture that aids farmers and other agricultural stakeholders in making well-informed choices on crop management, resource allocation, and strategic planning. To increase crop health and productivity, precise yield prediction enables timely interventions, focused fertilizer application, and optimal irrigation scheduling. Conventional yield estimation techniques rely on historical data and oversimplified statistical models, which frequently fall short in capturing the intricacies of crop growth influenced by management techniques, soil conditions, and environmental factors.

Advancements in machine learning have provided new avenues for improving yield prediction accuracy. Remote sensing data, particularly multispectral imagery, captures critical spatial and spectral information about crops, offering insights into plant health, growth stages, and environmental stress factors. Machine learning models, especially deep learning architectures, have shown great promise in leveraging this data to reveal complex patterns that traditional approaches may overlook.

However, existing deep learning models typically address spatial or temporal dimensions independently, which limits their ability to capture the multifaceted dynamics of crop growth over time. Experimental results demonstrate that the proposed model outperforms traditional yield prediction methods, achieving lower root mean square error and higher correlation coefficients.

II. LITERETURE REVIEW

This work endeavors to summarize recent key technologies and applications of smart agriculture, delineate the prevalent challenges it faces, highlight its publicly available datasets for adoption, and offer some policy guidelines for stakeholders, assisting them in making informed decisions regarding technology adoption and investment. We conclude that smart agriculture can potentially revolutionize the agricultural sector, provided we overcome the challenges by ensuring effective collaboration among stakeholders, a strong infrastructure, digital literacy, adoption incentives, data privacy, interoperability, favorable policy frameworks, and continuous research and development [1].

Authors provide a comprehensive investigation of deep learning (DL) models for the categorization of paddy illnesses in this survey work. In addition to exposing several paddy diseases and the symptoms that go along with them, our paper explores the reason behind this research project and dissects the distinct deep-learning models used for disease detection. We have also covered the methods that researchers employ to enhance the performance of DL models, as well as modifications made for certain application scenarios. Furthermore, we present pertinent research results, examine datasets used in this field, and evaluate data augmentation strategies. We highlight current research gaps, difficulties, and unresolved concerns through a thorough analysis, and we wrap up with a discussion of potential directions for further research [2].

By putting out a strong classification model that can accurately identify Diabrotica speciosa and caterpillar infestations, we tackle the issue of soybean leaf infestation. On unseen testing data, we attain balanced accuracies between 93.71% and 94.16% using our transfer-learning based model, which classifies soybean leaves using a VGG19 convolutional neural network. This surpasses earlier work using the same dataset and establishes a new standard. There are theoretical and practical ramifications to our work. In the agriculture sector, the soybean is essential. Soybean infestation causes significant environmental and financial consequences. Our methodology, which is shown above, enables early and precise identification to stop the spread of plant pests, minimizing ecological and financial harm [3].

We provide a methodical methodology to gather visual words, context, and taxonomy—analogs of natural language in vision. Our concept covers elements including textures, forms, and lines and is based on Marr's computational theory of vision. We demonstrate the suitability and usefulness of those analogs for elucidating self-supervised representations. Our main conclusions show that, regardless of the data modality, the relationship between language and vision may be a useful and understandable tool for comprehending how machine learning models operate. Our methodology creates a wide range of research opportunities for more understandable and transparent AI [5].

In this study, we do a systematic assessment of a substantial amount of scholarly literature on the use of machine learning techniques in agriculture that was published between 2000 and 2022. We list and talk about some of the most important data problems, including high dimensionality, data sparsity, and class imbalance. We also investigate how these data problems affect different machine learning techniques in the agricultural context. Lastly, we point out a few typical mistakes made in machine learning and agricultural research, such as the improper use of machine learning assessment methods. In order to do this, this survey offers a comprehensive assessment of the current situation in the area of machine learning and agriculture and suggests some appropriate mitigating techniques to deal with these issues [6].

High-quality agricultural products and efficient production can result from the application of machine learning (ML) techniques in this industry. Nonlinear correlations between data were not detectable by conventional predictive machine models. With the development of machine learning, prediction systems have undergone a recent revolution that can lead to highly accurate decision-making networks. Agricultural goods have been evaluated using a variety of techniques up to this point, including CNN-LSTM, ConvLSTM, and DeepYield. Preferable forecast accuracy is necessary, nevertheless. Two architectures have been proposed in this study. The first model incorporates LSTM-Attentions, skip connections, and 2D-CNN. ConvLSTM Attention, skip connections, and 3D-CNN make up the second model. The input data from Land-Cover and other MODIS packages .Surface-Temperature and MODIS-Land-surface data for 1800 counties in the United States, where soybeans are primarily grown, from 2003 to 2018. The suggested techniques have been contrasted with the latest models. The outcomes then demonstrated that the second suggested approach performed noticeably better than the previous approaches. For MAE, the second suggested approach yielded results of 4.3, 6.003, 6.05, 6.3, and 7.002, respectively, for Deep Yield, ConvLSTM, 3DCNN, and CNN-LSTM [7].

The primary goal was to examine how Sentinel-2 vegetation indices and Sentinel-1 interferometric coherence data may be used in tandem to estimate rice grain yield using Gaussian kernel regression. Using in situ measured yield data gathered across Xinghua county in Jiangsu Province, China, in 2019 and 2020, the prediction accuracy was evaluated. In every instance, Bayesian linear inference and probabilistic Gaussian regression were surpassed by Gaussian kernel regression. The optical red edge difference vegetation index (RDVI1) (r2 = 0.65, RMSE = 0.61 t/ha) outperformed the interferometric coherence (r2 = 0.52) and cal indices for crop yield mapping with Gaussian kernel regression in terms of prediction accuracy using the separately used optical and SAR data [8].

III. METHODOLOGY

The methodology for predicting crop yields and prices using machine learning involves a systematic approach encompassing data collection, preprocessing, model development, validation, and deployment. Below is an outline of the key steps involved in the methodology:

Data Collection:

Gather historical data on various factors influencing crop yields and prices, including weather patterns, soil quality, agricultural practices, market trends, and economic indicators. Utilize diverse sources such as government databases, satellite imagery, agricultural research institutions, and market databases to obtain comprehensive datasets covering relevant variables.

Data Preprocessing:

Cleanse the collected data to remove outliers, missing values and inconsistencies, ensuring data quality and integrity. Normalize and standardize the data to facilitate model training and improve predictive performance. Perform feature engineering to extract relevant features and create meaningful input variables for the machine learning models.

Model Development for Yield Prediction:

Create machine learning models to forecast crop yields based on past weather data, soil moisture, temperature, and other factors, as well as agricultural practices.To capture the intricate interactions between input variables and crop yields, use ensemble, regression, and classification algorithms. Part of the dataset should be set aside for model validation, while the rest should be divided into training and testing sets.

Train the machine learning models using historical yield data and relevant features, optimizing model parameters through techniques such as cross-validation and hyperparameter tuning. Evaluate the performance of the models using metrics such as RMS error, mean absolute error, and coefficient of determination to assess predictive accuracy and reliability.

Regression analysis:

This statistical technique examines a link between two or more variables so that the data on the other variables can be used to predict or explain one of the variables.

Random Forest Algorithm

One well-known machine learning method that is categorized as supervised learning is Random Forest. In machine learning, it can be applied to both regression and classification issues. Its foundation is the idea of ensemble learning, which combines multiple classifiers to tackle a challenging problem and enhance the model's performance.

The Random Forest Working Order Algorithm:

1. Select K data points at random from the training set.
2. Create the decision trees that are connected to the selected data points (subsets).
3. Choose the number N that you want to use to build the decision trees. Complete Steps 1 and 2
4. Ascertain each decision tree's predictions for the additional data points, then.

Convolutional Neural Network (CNN):

Jobs involving image processing and recognition are particularly well-suited for this type of deep learning system. which is made up of multiple layers, including convolutional, pooling, and fully linked layers. CNNs' architecture, which is based on how the human brain processes visual information, makes them perfect for spotting spatial correlations and hierarchical patterns in pictures.

IV. MODELING AND ANALYSIS

The model for website follows Figure 1. The figure represents the working and handling of requests where multiple process multiple requests that Model receives.

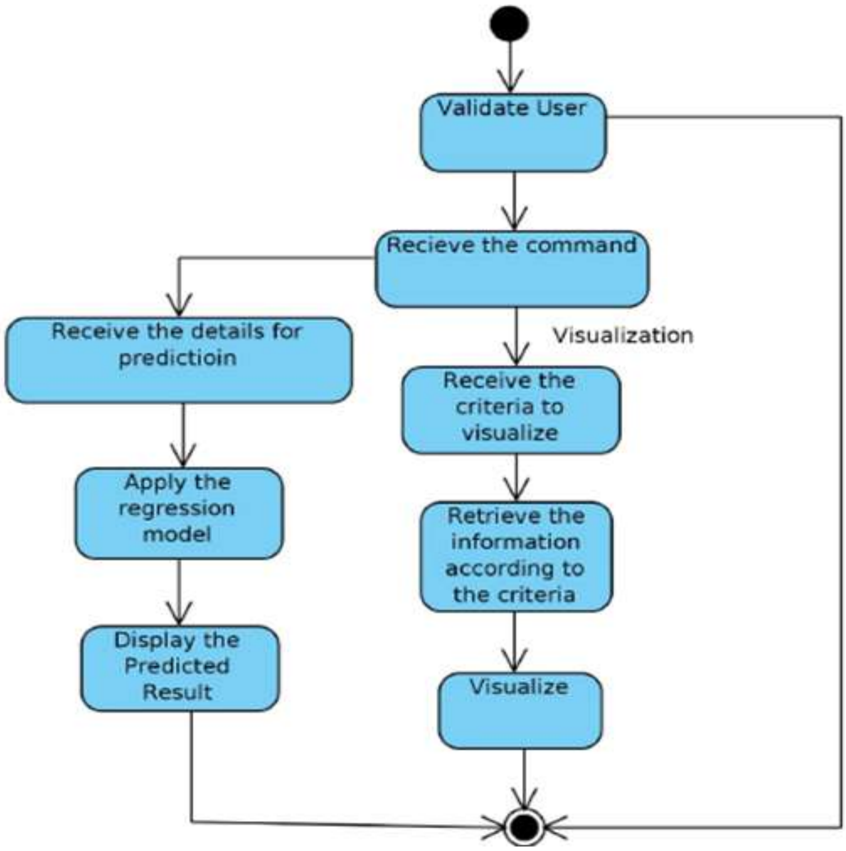


Figure 1: Interaction of user with model.

The next figure shows The phases of Developing a Model.

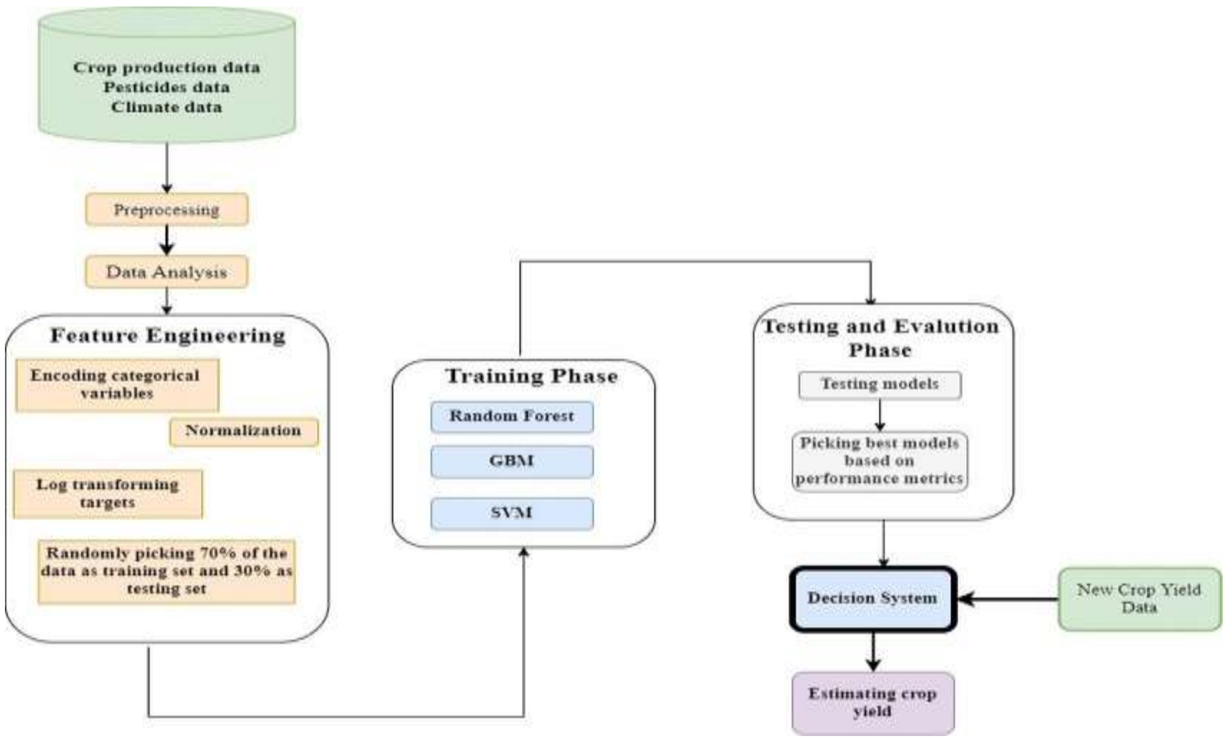
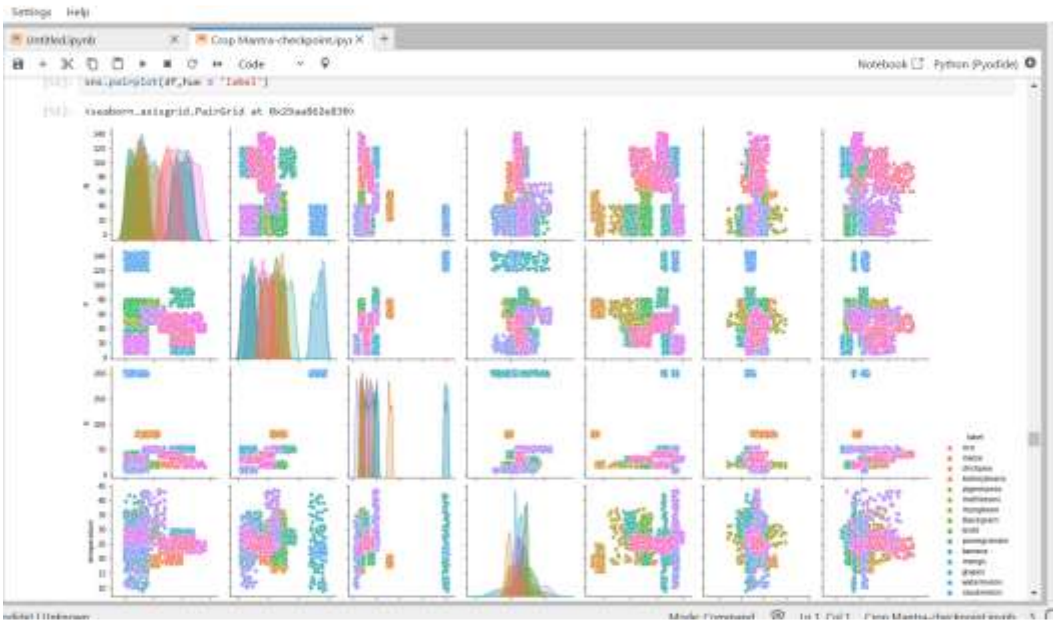


Figure 2: Phases of Developing Model

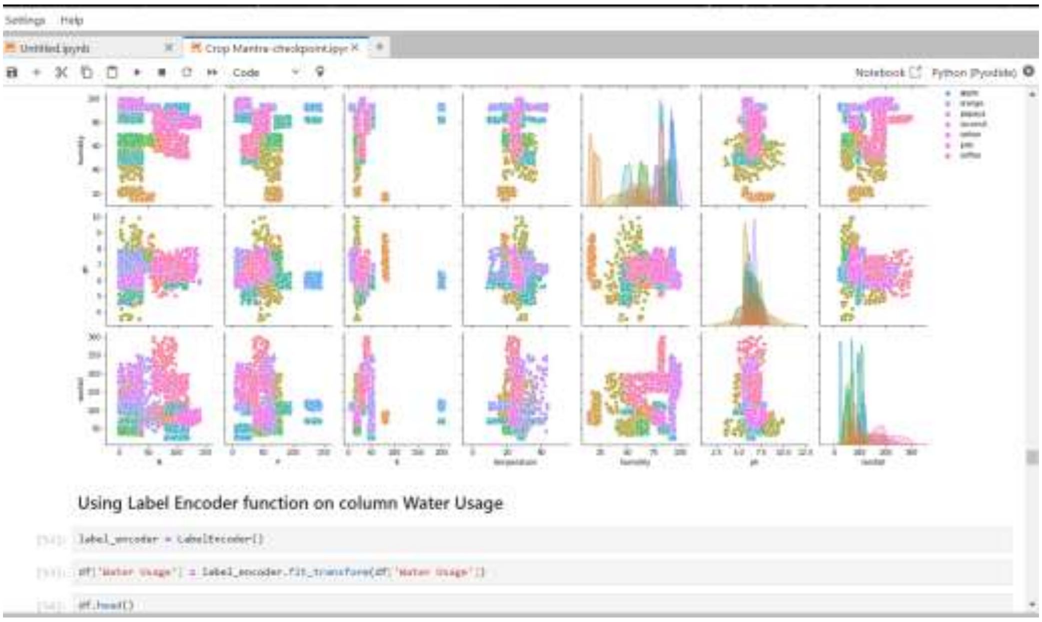
e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:11/November-2024 Impact Factor- 8.187 www.irjmets.com

V. RESULTS AND DISCUSSION

During this we discussed the working and functionality of different predictions and recommendation Systems. We also discussed on how we can implement these systems .Authors have collected dataset and performed various operations as following.



a



b

Figure 3 a-b: Represent the pairwise relationship on Multiple Factors used for Model Training

VI. CONCLUSION

This project presents a robust and innovative approach to crop yield prediction by integrating convolutional long short-term memory (ConvLSTM), convolutional neural network(CNN), and vision transformers (ViT) to address key limitations in existing models. By utilizing multispectral remote sensing data, the proposed model effectively captures spatial, temporal, and global contextual features, resulting in significantly improved accuracy and robustness over traditional methods. Experimental results demonstrate that the model achieves lower root mean square error and higher correlation coefficients, highlighting its superior predictive performance.

IRJMETS

e-ISSN: 2582-5208

International Research Journal of Modernization in Engineering Technology and Science

(Peer-Reviewed, Open Access, Fully Refereed International Journal)

Volume:06/Issue:11/November-2024 Impact Factor- 8.187 www.irjmets.com

This research introduces a novel approach to crop yield prediction, combining the strengths of ConvLSTM, CNN, and ViT to analyze multispectral remote sensing data. By capturing spatial, temporal, and global contextual features, the proposed model significantly outperforms traditional methods in terms of accuracy and robustness. Experimental results validate the model's effectiveness in providing precise yield estimates, enabling informed decision-making in agriculture, such as optimized resource allocation and improved crop management practices. This work contributes to advancing the field of precision agriculture by offering a reliable and efficient tool for large-scale crop yield prediction.

VII. REFERENCES

[1] Bilal Ahmed, Hasnat Shabbir, Syed Rameez Naqvi and Lu Peng, "Smart Agriculture: Current State, Opportunities, and Challenges", IEEE 2024.

[2] Mahrin Tasfe, Akm Nivrito, Fadi Al Machot, Mogib Ullah And Habib Ullah, "Deep Learning Based Models for Paddy Disease Identification and Classification: A Systematic Survey", IEEE 2024.

[3] Niklas Farah, Nicolas Drack, Hannah Dawel And Ricardo Buettner, "A Deep Learning-Based Approach for the Detection of Infested Soybean Leaves", IEEE 2023.

[4] Xiaochuan Guo, Wenbo Xie And Xin Li, "Spatial-Temporal Correlation Neural Network for Long Short-Term Demand Forecasting During COVID-19", IEEE 2023.

[5] Witold Oleszkiewicz, Dominka Basaj, Igor Sieradzki, Michal Gorszczak, Barbara Rychalska, Koryana Lewandowska, Tomasz Trzcinski and Bartosz Zielinski, "Visual Probing: Cognitive Framework for Explaining Self-Supervised Image Representations", IEEE 2023.

[6] Sarah Condran, Michael Bewong, MD Zahidul Islam, Lancelot Maphosa and Lihong Zheng, "Machine Learning in Precision Agriculture: A Survey on Trends, Applications and Evaluations Over Two Decades", IEEE 2022.

[7] Seyed Mahdi Mirhoseini Nejad,Dariush Abbasi-Moghadam,Alireza Sharif,Nizom, "Multispectral Crop Yield Prediction Using 3D-Convolutional Neural Networks and Attention Convolutional LSTM Approaches", IEEE 2023.

[8] Yeshanbele Alebele, Wenhui Wang, Weiguo Yu, Xue Zhang, Xia Yao, Yongchao Tian, Yan Zhu, Weixing Cao, and Tao Cheng, "Estimation of Crop Yield From Combined Optical and SAR Imagery Using Gaussian Kernel Regression", IEEE 2021.

[9] Uferah Shafi, Rafia Mumtaz, Zahid Anwar, Muhammad Muzyyab Ajmal, Muhammad Ajmal Khan, Zahid Mahmood, Maqsood Qamar And Hafiz Muhammad Jhanab, "Tackling Food Insecurity Using Remote Sensing and Machine Learning-Based Crop Yield Prediction", IEEE 2023.