Synopsis – Crop Yield Prediction Driven By Machine Learning And Artificial Intelligence

Name of Team Members: Nikhil Subhash Shelke (44)

Manas Sharad Torane (62)

Rohan Vishnu Pawar (24)

Pranay Sanjay Satpute (39)

Class: BE

Div: B

Domain: Machine Learning

Topic and Sponsorship: Crop Yield Prediction

Abstract:

Agricultural yield prediction is a critical task for ensuring food security, optimizing resource allocation, and supporting policy-making in the agricultural sector. However, predicting crop yield is a complex problem influenced by numerous dynamic factors such as weather conditions, soil characteristics, crop type, and farming practices. This project explores the use of machine learning techniques to accurately forecast agricultural yield based on historical and environmental data. A key challenge in this domain is dealing with missing, noisy, and regionally skewed data, which can affect model performance. This study applies various supervised learning algorithms including Linear Regression, Random Forest, Support Vector Machines, and advanced deep learning models like LSTM and XGBoost to enhance predictive accuracy. Feature engineering and data balancing methods are also investigated to address dataset inconsistencies. The proposed system has practical applications in precision agriculture, smart farming, policymaking, and supply chain management. While the techniques explored are applicable across domains such as finance and healthcare, this work is focused on the agricultural context, where the challenges of real-world data and model generalization remain significant.

Keywords: Crop Yield Prediction, Agriculture, Machine Learning, Precision Farming, Climate Data, Soil Analysis, Data Mining, Supervised Learning, Remote Sensing, Forecasting Models

Challenges identified :

- Agricultural datasets often suffer from missing values, noise, and inconsistency due to diverse data sources (e.g., satellite, sensors, and surveys), making preprocessing a critical challenge.

- Crop yield is highly influenced by regional factors like soil type, climate, and farming practices, which vary significantly across locations and require models to generalize well.

- While deep learning models like LSTM or XGBoost offer high accuracy, their black-box nature can be a hurdle in gaining trust from domain experts like farmers and agronomists.

- Yield prediction involves complex, non-linear interactions among variables (e.g., rainfall and temperature effects), making it difficult for traditional models to capture all dependencies.

Novelty or Industrial Application: -

- Useful for precision agriculture platforms, agri-tech startups, and government policy planning
- Applicable to crop insurance, supply chain optimization, and farm resource management
- Deployable as a real-time dashboard, mobile app, or API for yield forecasting and advisory services

Base IEEE/ Springer / Equivalent publication (paper URL):

https://ieeexplore.ieee.org/document/10735786

List of References:

- "Crop Yield Prediction using Machine Learning: A Systematic Literature Review", IEEE, 2020 "Predicting Personality from Social Media using NLP", Springer, 2021
- "Smart Farming: Crop Yield Prediction Using Supervised Machine Learning", Springer, 2021 "Neural Approaches for MBTI Classification from Text", Elsevier, 2022
- "Crop Yield Forecasting Using Artificial Neural Networks and Satellite Data", ACM, 2019