Crop Yield Prediction Using Regression & Ensemble Techniques

---

1. Introduction

Agriculture plays a major role in India. Accurate crop yield prediction helps farmers plan irrigation, fertilizer usage and harvesting. This project uses Machine Learning algorithms—Linear Regression and Bagging Regressor—to predict crop yield based on environmental factors.

2. Problem Statement

Crop yield depends on rainfall, soil condition, temperature and fertilizer. Manual prediction is inaccurate.
This project aims to develop a machine learning model that predicts yield and compares different regression models.

3. Objectives

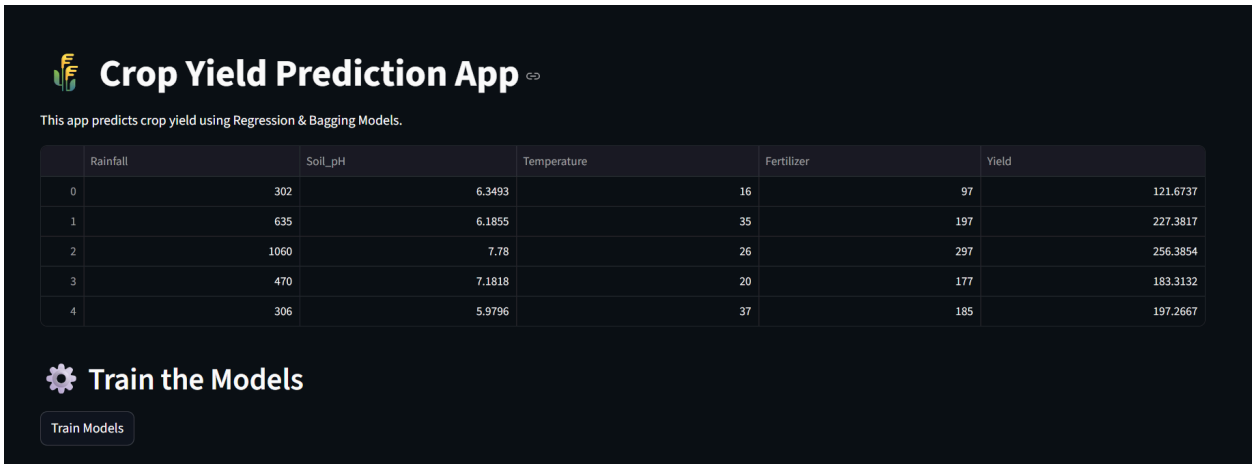
- Prepare dataset with agricultural features
- Train different regression models
- Compare model performance (MSE, R^2)
- Build a Streamlit web app for prediction

4. Dataset Description

100 synthetic data samples were created with:

Feature	Description
Rainfall	mm
Soil_pH	Soil acidity
Temperature	°C
Fertilizer	kg
Yield	Output

Dataset Display



5. Methodology

1 Data Generation

Random values were generated for rainfall, pH, temperature and fertilizer.
Yield = Weighted formula of all + random noise.

2 Model Training

Models used:

- Linear Regression
- Bagging Regressor

3 Model Evaluation

Metrics used:

- Mean Squared Error (MSE)
- R^2 Score

6. Model Results

Two models were compared:

Linear Regression

- Simple
- Moderately accurate

Bagging Regressor

- More stable
- Handles noise better

- Higher accuracy

Model Performance Output

7. Streamlit Application

The app allows users to input:

- Rainfall
- Soil pH
- Temperature
- Fertilizer

And predicts:

- Linear Regression Yield

- Bagging Regressor Yield

---

8. Applications

- Smart farming
- Government crop forecasting
- Fertilizer planning
- Research in agriculture analytics

---

9. Conclusion

This project shows that Machine Learning can improve the accuracy of crop yield predictions.
Bagging Regressor outperforms Linear Regression for noisy datasets.
The Streamlit interface makes the system accessible and user-friendly.

10. Future Enhancements

- Random Forest / XGBoost models
- Real-world datasets
- Mobile app for farmers
- Weather forecast integration