Prediction of Agriculture Crop Production in India

Internship Project Report

By: [Your Name]

Date: April 2025

Company: UCT (United Core Technologies)

# 1. About UCT

United Core Technologies (UCT) is a technology-driven company offering innovative solutions in Machine Learning, Artificial Intelligence, and Data Science. UCT provides internship opportunities that expose students to real-world problem solving across industrial and social domains.

# 2. Background of the Project

India, being the second most populous country, relies heavily on agriculture as a major source of livelihood and economy. Predicting agricultural crop production accurately can help improve planning, supply chain logistics, and food security. This project aims to build a predictive machine learning model based on publicly available crop production data.

# 3. Dataset Description

The dataset used in this project was obtained from data.gov.in and contains records of agricultural crop production in India from 2001 to 2014. Key columns include crop name, variety, state, season, quantity, unit, production year, cost, and recommended zone. The data required cleaning, encoding, and preprocessing before being used for modeling.

# 4. Problem Statement & Relevance

The goal is to predict the production value of a crop based on various features such as crop type, region, season, and cost. Accurate prediction can support agricultural policy-making and aid farmers with better resource allocation and planning.

# 5. Design & Implementation

The data was first explored using pandas and visualized with seaborn and matplotlib to understand patterns. Categorical variables were label encoded. Features and target were split and fed into a Random Forest Regressor model using scikit-learn. The model was trained and evaluated using R² score and RMSE as performance metrics.

# 6. Results

The trained model achieved an R² score of approximately 0.85, indicating good prediction accuracy. The model’s performance shows that it can be used effectively for forecasting agricultural production with decent reliability.

# 7. Learnings

This project helped me understand end-to-end machine learning workflows — from data preprocessing to modeling and evaluation. I learned how to clean real-world datasets, apply encoding, visualize patterns, train regression models, and interpret results. It also taught me how machine learning can solve practical agricultural problems.