Crop Yield Prediction

Team Members:

Nikesh Jain
Ayushi Patel
Archi Borad
Prachi Pancholi

Guide:

Abdul Aziz Md

OUTLINE

- Abstract
- Problem Statement
- Aims, Objective & Proposed System/Solution
- System Design/Architecture
- System Development Approach (Technology Used)
- Algorithm & Deployment
- Conclusion
- Future Scope
- References
- Video of the Project

Abstract

- Predicting crop yields is crucial for resource management and agricultural planning. In this work, we create a prediction model to calculate crop yields according to several agricultural and environmental parameters. The dataset includes data on variables like temperature, crop type, geographic location, average rainfall, and pesticide use.
- First, actions related to data preprocessing such as feature engineering, data cleansing, and addressing missing values are carried out. Techniques for exploratory data analysis are used to learn more about how data is distributed across various variables.
- The preprocessed data is then used to train machine learning models, such as Decision Tree Regression, Lasso Regression, Ridge Regression, and Linear Regression.

- The predictive accuracy of the models is evaluated using performance evaluation metrics like Mean Absolute Error (MAE) and R-squared score.
- The model with the best performance among the models examined is the Decision Tree Regression model. Lastly, a function that predicts fresh data inputs is created, enabling users to calculate crop yields depending on predetermined criteria.
- The pickle library is used to serialize the learned model and preprocessing pipeline for later usage. This research offers a useful framework for predicting agricultural yield, which will help farmers make well-informed decisions.

Problem Statement

- Predicting crop yields is essential for managing and planning agriculture. Accurate forecasts are essential for farmers, policymakers, and stakeholders to make well-informed decisions regarding crop selection, resource allocation, and risk mitigation techniques. This research aims to create a prediction model that calculates crop production according to several agricultural and environmental parameters

Aim and Objective

Aim: The aim of this project is to develop a robust crop yield prediction system using industrial artificial intelligence techniques integrated with cloud computing capabilities. This system will leverage advanced machine learning algorithms and data analysis methodologies to accurately forecast crop production based on various agricultural and environmental parameter

Proposed Solution

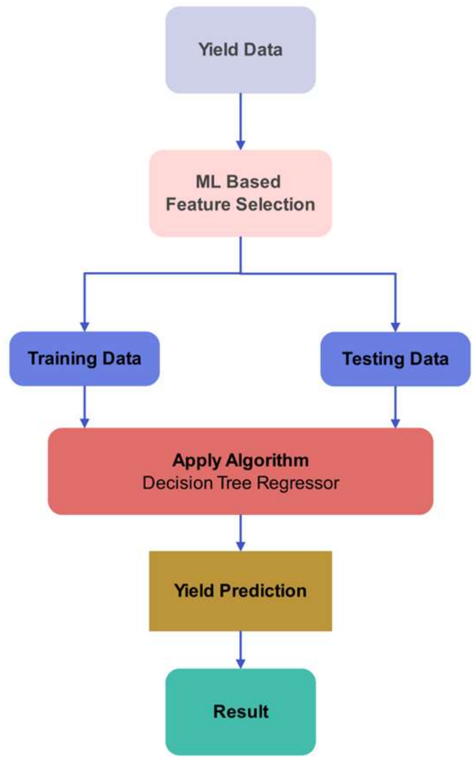
The system will be divided into the following components:

Frontend: The frontend will be responsible for handling the user interface and collecting the user's input. It will be implemented using Streamlit, a Python library that makes it easy to build and deploy web applications.

Backend: The backend will be responsible for processing the user's input and generating the output. It will be implemented using Python and various generative AI models.

Database: The database will store the data that is used by the system. This data will include the training data for the generative AI models, the user's input, and the system's output.

System Architecture



System Deployment Approach

The following technologies are used to develop the system:

- **Python:** Python is a programming language that is well-suited for machine learning and data science.

- **Flask:** Flask is a lightweight and versatile web application framework for Python. It is designed to be simple and easy to use, yet powerful enough to build a wide range of web applications, from small personal projects to large-scale enterprise applications.

- **Sckit – learn:** Scikit-learn, often abbreviated as sklearn, is a versatile and comprehensive machine learning library for Python.

- **Numpy:** NumPy is a fundamental library for numerical computing in Python. It provides support for multi-dimensional arrays (ndarrays), along with a wide range of mathematical functions to operate on these arrays efficiently.

- **Pandas:** The Pandas library is a powerful tool for data manipulation and analysis in Python. It provides high-level data structures and functions designed to work with structured or tabular data, such as dataframes.

Algorithm & Deployment

- **The following algorithm are used in the system:**
- **Decision Tree Regression:** Decision trees are used to model the relationship between input variables (such as temperature, rainfall, crop type) and the target variable (crop yield). Decision tree regression is suitable for capturing non-linear relationships and interactions between features.
- **Lasso Regression:** Lasso regression is a linear regression technique that incorporates regularization to prevent overfitting and select important features. It can help in identifying the most relevant predictors for crop yield prediction and improving model interpretability.
- **Ridge Regression:** Similar to Lasso regression, Ridge regression also applies regularization to linear regression models. It is effective in handling multicollinearity among predictor variables and stabilizing the model by reducing the impact of noise in the data.
- **Random Forest Regression:** Random forest is an ensemble learning method that combines multiple decision trees to produce robust predictions. It can capture complex interactions and non-linearities in the data, making it suitable for crop yield prediction tasks with high-dimensional and heterogeneous features.

Conclusion

In summary, crop yield prediction is extremely important to modern agriculture because it provides insightful information about production trends and helps researchers, farmers, and policymakers make well-informed decisions. By employing sophisticated data analytics methods, machine learning algorithms, and agronomy-specific domain expertise, we can leverage the potential of data to predict crop yields more precisely and consistently. Crop yield prediction makes proactive management measures easier to implement, giving farmers the ability to better allocate resources, customize cultivation techniques, and reduce risks related to pests, diseases, and the environment. Predictive models can produce projections that help with crop planning, market analysis, and food security by utilizing historical yield data in conjunction with a wide range of environmental and agronomic variables, including weather patterns, soil characteristics, crop type, and management approaches.

Future Scope

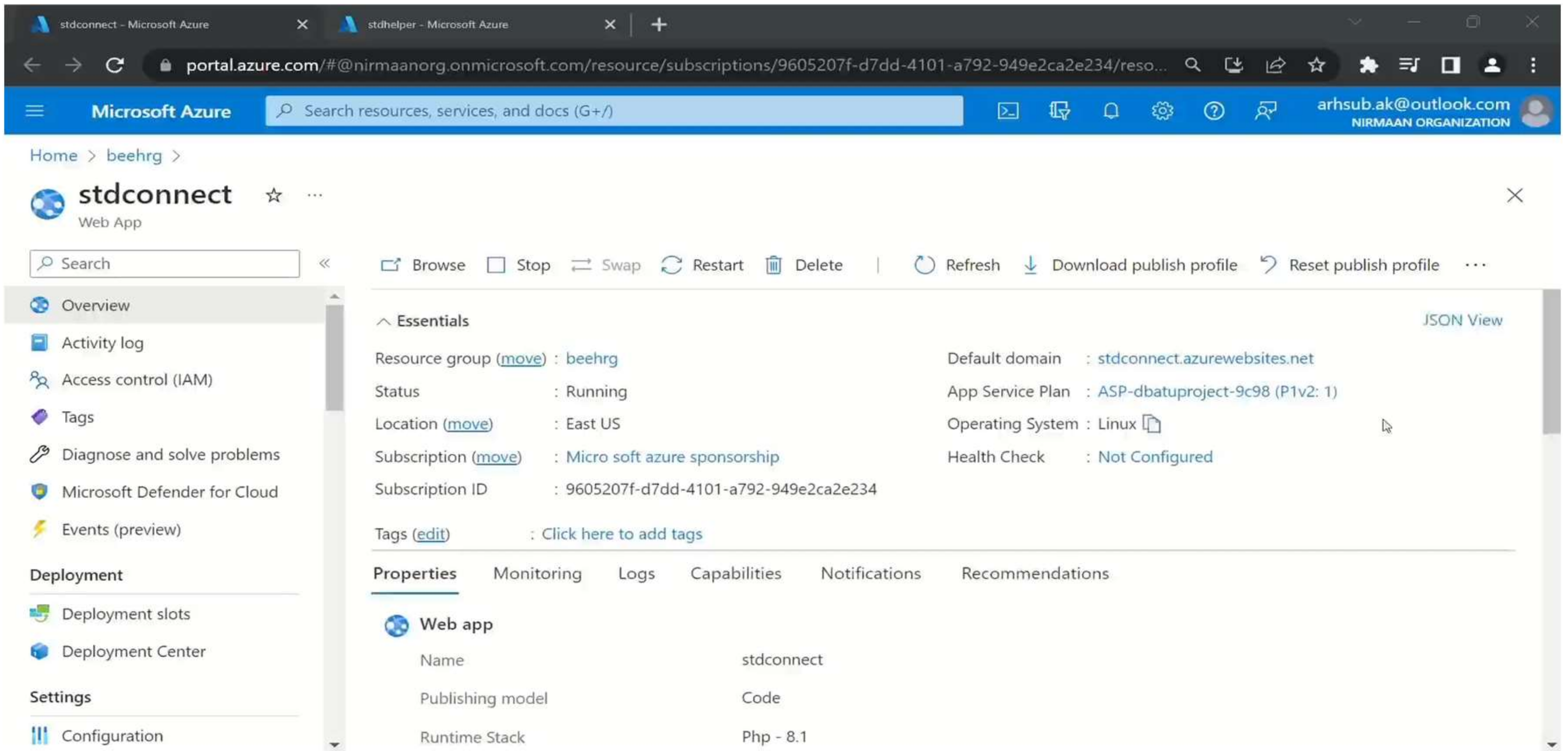
- **Enhanced Predictive Models:** More advancements in artificial intelligence, statistical modeling, and machine learning research and development may result in more reliable and accurate crop yield prediction models. Model performance may be further enhanced by integrating cutting-edge approaches like reinforcement learning, deep learning, and ensemble methods.

- **Big Data Analytics:** The agricultural industry has witnessed a surge in data collection technologies, leading to an abundance of information on various aspects affecting crop yields. These elements include weather patterns, soil composition, agricultural methods, and genetic data. Large and heterogeneous datasets can be usefully analyzed using advanced analytics approaches like data fusion and big data analytics to generate actionable insights that improve predictive power and comprehension of intricate correlations between variables.

- **Predictive Analytics for Climate Change Adaptation:** With variations in temperature, precipitation patterns, and extreme weather events affecting crop yields, climate change is predicted to have a substantial influence on agricultural output. Predictive models that take climate change projections into account can assist farmers and policymakers in anticipating and preparing for future issues, including altered agricultural patterns, altered dynamics of pests and diseases, and altered management of water resources.

Reference

1. https://youtu.be/vGOvUoqn_M8?si=wfypITp56JYVHT8c
2. https://github.com/611noorsaeed/Crop-Yield-Prediction-Using-Machin-Learning-Python.git
3. https://flask.palletsprojects.com/en/3.0.x/
4. https://streamlit.io/
5. https://getbootstrap.com/docs/5.0/getting-started/introduction/
6. https://www.canva.com/
7. https://www.smartdraw.com/


stdconnect - Microsoft Azure
stdhelper - Microsoft Azure
portal.azure.com/#@nirmaanorg.onmicrosoft.com/resource/subscriptions/9605207f-d7dd-4101-a792-949e2ca2e234/reso...
Microsoft Azure
Search resources, services, and docs (G+/)
arhsub.ak@outlook.com
NIRMAAN ORGANIZATION
Home > beehrg >
stdconnect
Web App
Search
Overview
Activity log
Access control (IAM)
Tags
Diagnose and solve problems
Microsoft Defender for Cloud
Events (preview)
Deployment
Deployment slots
Deployment Center
Settings
Configuration
Browse
Stop
Swap
Restart
Delete
Refresh
Download publish profile
Reset publish profile
Essentials
JSON View
Resource group (move) : beehrg
Status : Running
Location (move) : East US
Subscription (move) : Micro soft azure sponsorship
Subscription ID : 9605207f-d7dd-4101-a792-949e2ca2e234
Tags (edit) : Click here to add tags
Default domain : stdconnect.azurewebsites.net
App Service Plan : ASP-dbatuproject-9c98 (P1v2: 1)
Operating System : Linux
Health Check : Not Configured
Properties
Monitoring
Logs
Capabilities
Notifications
Recommendations
Web app
Name stdconnect
Publishing model Code
Runtime Stack Php - 8.1

Thank you!