edunet foundation

FarmYield Predictor: Crop Recommendation Using Environmental Data

By:
G.Dhavamani
R.Varshini
N.Srimathi
AS.Suvathi

Content

- Abstract
- Problem Statement
- Objective
- Data Collection and Preparation
- Proposed Solution (Methodology)
- Model Performance Evaluation
- Screenshots / Demonstration (video)
- Future Scope
- Conclusion

Abstract

- The FarmYield Predictor is a machine learning-based system that analyzes environmental data, including temperature, humidity, soil moisture, and precipitation, to provide accurate crop recommendations tailored to specific locations and seasons. By processing historical data and forecasting environmental conditions, the system predicts which crops are most likely to thrive under the given environmental factors, optimizing farming practices.

- The goal of the project is to empower farmers with data-driven insights, enabling them to make informed decisions about crop selection. By improving crop yield, reducing resource wastage, and minimizing the impact of climate variability, the system contributes to more sustainable agricultural practices and enhances the efficiency of food production in various regions.

Problem Statement

- Farmers often face challenges in selecting the right crops for their land due to unpredictable environmental conditions such as varying soil types, weather patterns, and climate changes.
- Traditional methods of crop selection, based on historical practices or local knowledge, may not be accurate in today's rapidly changing environmental landscape
- This results in suboptimal yields, inefficient resource use, and increased costs
- There is a pressing need for a data-driven approach that utilizes environmental data to predict the most suitable crops for specific regions.
- Such a system would help farmers make informed decisions, optimize crop production, and ensure sustainable farming practices.

Objective

- To collect and analyze environmental data (temperature, humidity, soil moisture, precipitation) that influences crop growth and yield.
- To develop a machine learning model that predicts the most suitable crops based on real-time and historical environmental data for specific regions and seasons.
- To assist farmers in making informed crop selection decisions, optimizing yield and resource usage, and reducing the environmental impact.
- To create a user-friendly interface where farmers can input environmental data and receive personalized crop recommendations.
- To promote sustainable farming practices by improving crop yield, reducing resource waste, and adapting to changing climate conditions.

Data Collection and Preparation

- Data collection for the Farm Yield Predictor involves gathering environmental data such as soil properties (e.g., pH, nutrient levels, texture), weather conditions (temperature, rainfall, humidity), and historical crop yield information. Sensors, remote sensing technologies, and local agricultural databases provide real-time and historical data.

- The collected data is cleaned by removing inconsistencies, handling missing values, and standardizing variables for uniformity.

- Feature selection is performed to identify the most relevant factors influencing crop yield, such as soil fertility and climatic conditions.

- The data is then split into training and testing sets, ensuring accurate model training and validation for crop recommendations.

Proposed Solution (Methodology)

- Data Collection: Collect environmental data (temperature, humidity, soil moisture, precipitation) from weather stations or Kaggle datasets.
- Data Preprocessing: Clean the data by handling missing values, removing outliers, and scaling numerical features for consistency.
- Model Selection: Use machine learning models (e.g., Random Forest, Decision Trees) to learn the relationship between environmental conditions and crop suitability.
- Model Training: Train the model with historical data to predict which crops will thrive in specific environmental conditions.
- Prediction and Recommendation: The trained model provides crop recommendations based on new environmental data, helping farmers make better decisions.

Model Performance Evaluation

- Accuracy: Measure the accuracy of the model by comparing the predicted crop recommendations to the actual crop yields from the test data. A higher accuracy indicates better performance.
- Precision: Evaluate how many of the recommended crops were correct (i.e., true positives).
- Recall: Measure how many of the actual correct crops were identified by the model (i.e., true positives out of all possible correct crops).
- F1-Score: Calculate the F1-Score, which balances precision and recall, providing a single metric to evaluate model performance.
- Cross-Validation: Use k-fold cross-validation to check the model's performance across multiple data subsets. This helps ensure the model generalizes well to unseen data.
- Confusion Matrix: Use a confusion matrix to evaluate how well the model performs in classifying the correct crop types. It shows the number of true positives, true negatives, false positives, and false negatives.

Code

```
import pandas as pd
From sklearn.model_selection
import train_test_split
from sklearn.ensemble
Import RandomForestClassifier
df=pd.read_csv("Crop_recommendation.csv")
X = df.drop("label", axis=1)y = df["label"]X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
random_state=42)
model=RandomForestClassifier(n_estimators=100, random_state=42)model.fit(X_train, y_train)
print("Enter the following details:")
n = float(input("Nitrogen (N): "))
p = float(input("Phosphorous (P): "))
k = float(input("Potassium (K): "))
temperature = float(input("Temperature (°C): "))
humidity = float(input("Humidity (%): "))
ph = float(input("pH value: "))
rainfall = float(input("Rainfall (mm): "))
input_data = [[n, p, k, temperature, humidity, ph, rainfall]]prediction =
model.predict(input_data)print(f"\nRecommended Crop: {prediction[0]}")
```

Screenshots / Demonstration (video)

```
print(f"\nRecommended Crop: {prediction[0]}")

Enter the following details:
Nitrogen (N): 90
Phosphorous (P): 40
Potassium (K): 40
Temperature (°C): 26
Humidity (%): 80
pH value: 6.5
Rainfall (mm): 100

Recommended Crop: jute
/usr/local/lib/python3.11/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names, but RandomForestClassifier was fitted with feature names
  warnings.warn(
```

Future Scope

- The future scope of the Farm Yield Predictor includes integrating more advanced data sources, such as satellite imagery, IoT-based sensor data, and climate change models, to enhance the accuracy of crop recommendations.
- Incorporating real-time predictive analytics and incorporating more variables like pest outbreaks or water availability could further optimize crop selection.
- The system could evolve to offer personalized recommendations for different farm scales and regions, adapting to specific farmer needs.
- Additionally, the integration of blockchain technology for traceability and supply chain management and the expansion to global agricultural markets could significantly improve sustainability and global food security.

Conclusion

- The Farm Yield Predictor: Crop Recommendation Using Environmental Data offers a promising solution to optimize agricultural productivity through data-driven insights.
- By analyzing environmental factors such as soil conditions, climate, and historical yields, the system provides accurate crop recommendations that help farmers maximize yield, reduce resource wastage, and promote sustainable farming practices.
- With the integration of advanced machine learning algorithms and real-time data, the system empowers farmers to make informed, efficient decisions.
- As the system evolves, it has the potential to enhance global agricultural sustainability, addressing challenges like climate change, food security, and resource management for future generations.