DSA210 PROJECT FINAL REPORT

CLIMATE CHANGE IMPACT ON AGRICULTURAL PRODUCTIVITY

CEREN TEKİN

32492

SUPERVISOR: BORA ÇELİKÖRS

# IN THIS REPORT

This report investigates the effects of various climate-related variables on crop productivity. Using a dataset of climate indicators and agricultural outputs, the study employs supervised machine learning models to predict crop yield and uncover patterns. Exploratory analysis highlights the roles of temperature, precipitation, CO₂ emissions, and other environmental factors in agricultural performance. The results, though limited in predictive power, provide meaningful insights into how climate change affects food production.

With global climate change becoming increasingly evident, understanding its impact on crop yield is essential for agricultural policy and food security. This project explores the relationship between environmental factors and agricultural productivity through machine learning and data analysis, excluding linear regression methods and utilizing hyperparameter tuning.

# DATASET & PARAMETERS

The dataset contains 1031 entries and includes the following parameters:  
- Average Temperature (°C)  
- Total Precipitation (mm)  
- CO₂ Emissions (MT)  
- Crop Yield (MT/HA)  
- Extreme Weather Events  
- Irrigation Access (%)  
- Soil Health Index  
- Crop Type and Region

# METHODOLOGY

Data cleaning involved standardizing decimal notation and removing missing entries. Exploratory data analysis was performed using histograms, scatter plots, boxplots, and correlation heatmaps. Three machine learning models were applied:  
- Random Forest Regressor  
- Decision Tree Regressor  
- K-Nearest Neighbors Regressor  
Each model was tuned using GridSearchCV and evaluated using R² Score, Mean Squared Error (MSE), and Mean Absolute Error (MAE).

# RESULTS & ANALYSIS

## Univariate Analysis

Crop yield and temperature distributions showed moderate right-skew, while CO₂ emissions and precipitation were normally distributed. Histogram and violin plots indicated yield variations across crop types and regions.

diyagram, çizgi, öykü gelişim çizgisi; kumpas; grafiğini çıkarma, paralel içeren bir resim

Açıklama otomatik olarak oluşturuldu

metin, çizgi, öykü gelişim çizgisi; kumpas; grafiğini çıkarma, ekran görüntüsü içeren bir resim

Açıklama otomatik olarak oluşturuldu

metin, öykü gelişim çizgisi; kumpas; grafiğini çıkarma, diyagram, çizgi içeren bir resim

Açıklama otomatik olarak oluşturuldu

## Bivariate Analysis

Scatter plots revealed a weak negative trend between CO₂ emissions and crop yield. Boxplots showed that more extreme weather events are associated with lower yields. Irrigation and soil health had strong positive associations with yield.

metin, çizgi, ekran görüntüsü, öykü gelişim çizgisi; kumpas; grafiğini çıkarma içeren bir resim

Açıklama otomatik olarak oluşturuldu

metin, çizgi, ekran görüntüsü, paralel içeren bir resim

Açıklama otomatik olarak oluşturuldu

## Multivariate Analysis

A correlation heatmap showed:  
- Strong positive correlation: Irrigation Access & Soil Health → Crop Yield  
- Negative correlation: Extreme Weather & CO₂ Emissions → Crop Yield

metin, ekran görüntüsü içeren bir resim

Açıklama otomatik olarak oluşturuldu

## Regression Performance

metin, çizgi, öykü gelişim çizgisi; kumpas; grafiğini çıkarma, sayı, numara içeren bir resim

Açıklama otomatik olarak oluşturuldu

Model performances were as follows:

|  |  |  |  |
| --- | --- | --- | --- |
| Model | R² Score | MSE | MAE |
| Random Forest | -0.978 | 29045.67 | 102.95 |
| Decision Tree | -2.782 | 55545.78 | 86.21 |
| K-Nearest Neighbors | -0.397 | 20521.98 | 77.88 |

# FINDINGS

- Higher CO₂ emissions and extreme weather events tend to reduce yield.  
- Soil health and irrigation are the strongest positive predictors.  
- Multivariate influences suggest yield is the result of complex climate interactions.

# CONCLUSION

Despite poor R² scores, the project revealed meaningful relationships between environmental variables and crop yield. Findings support the need for climate adaptation in agriculture. Future work should explore time-series forecasting, more granular regional data, and domain-specific crop modeling for actionable policy recommendations.