AI-Driven Food Crisis Prediction Using Satellite and Climate Data

Sıla Akkaya   Ahmet Bekir Arslanalp

1. Introduction

This project aims to build an early warning system to predict global food crises using satellite imagery, climate data, deep learning, NLP, and time-series models. The system focuses on identifying drought-induced agricultural stress to support policymakers, farmers, and humanitarian organizations.

2. Technologies Used

- **Programming:** Python
- **Development:** Google Colab, VSCode
- **Libraries/Tools:** Google Earth Engine, Flask
- **Web Stack:** HTML, CSS, JavaScript

3. Data and Preprocessing

- **Satellite Data (Sentinel-2):** Applied cloud masking and divided images into 128×128 pixel tiles.
- **Climate Data (ERA5):** Temperature converted from Kelvin to Celsius; precipitation normalized and drought score computed as $1 - \text{NormalizedPrecipitation}$.
- **Temporal Features:** Month and week extracted from date; region one-hot encoded.
- **News Data (NewsAPI):** Title, description, and content combined, lowercased for consistent text processing.

4. NDVI Explained

NDVI (Normalized Difference Vegetation Index) is a measure of vegetation health, ranging from -1 to +1.

- 0.6 – 1.0: Dense, healthy vegetation

- 0.2 – 0.5: Sparse to moderate vegetation
- -1 – 0.1: Water, urban areas, or bare soil

5. Classification Model (ResNet-18)

Table 1: ResNet-18 Classification Metrics

Class	**Precision**	**Recall**	**F1-Score**
No Drought (0)	0.9981	0.9991	0.9986
Drought (1)	0.9849	0.9703	0.9775
Accuracy		0.9973	
Macro Avg	0.9717	0.9415	0.9881
Weighted Avg	0.9585	0.9127	0.9973

6. Time Series Models (LSTM)

A. NDVI Forecasting

Table 2: LSTM Model Performance – NDVI Prediction

Metric	**Value**
MAE	0.0647
MSE	0.0069
R^2	0.8453

B. Drought Score Forecasting (ERA5)

Table 3: LSTM Model Performance – Drought Score Prediction

Metric	**Train**	**Val**
MAE	0.069	0.0729
MSE	0.0088	0.0104
R^2	0.7194	0.7265

7. NLP-Based Media Analysis

- **Source:** NewsAPI
- **Analysis:** Sentiment analysis using TextBlob's PatternAnalyzer
- **Status:** Prototype stage – plans to improve with custom or deep learning models

8. Challenges and Future Work

- Large raw satellite data limited by local GPU capacity
- Difficulty integrating multiple models
- Hardware limits for processing wide areas
- Future improvements: better NLP and scalability for large regions

9. CNN Confusion Matrix

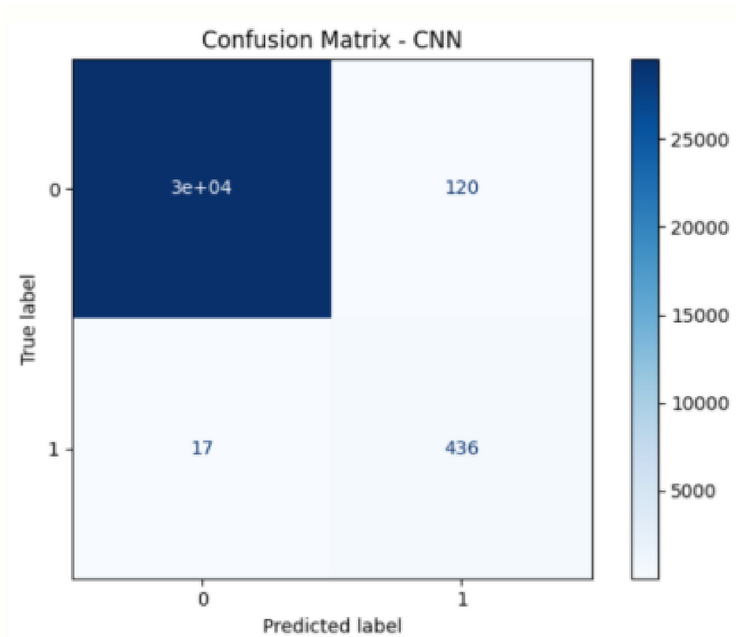

Figure 1: Confusion Matrix - CNN

Interpretation:

- **True Positive:** 436, **False Positive:** 120
- **True Negative:** 30,000+, **False Negative:** 17
- Low false negatives are crucial – the model rarely misses true drought events.

10. Spatial Comparison: NDVI vs. Drought Heatmap

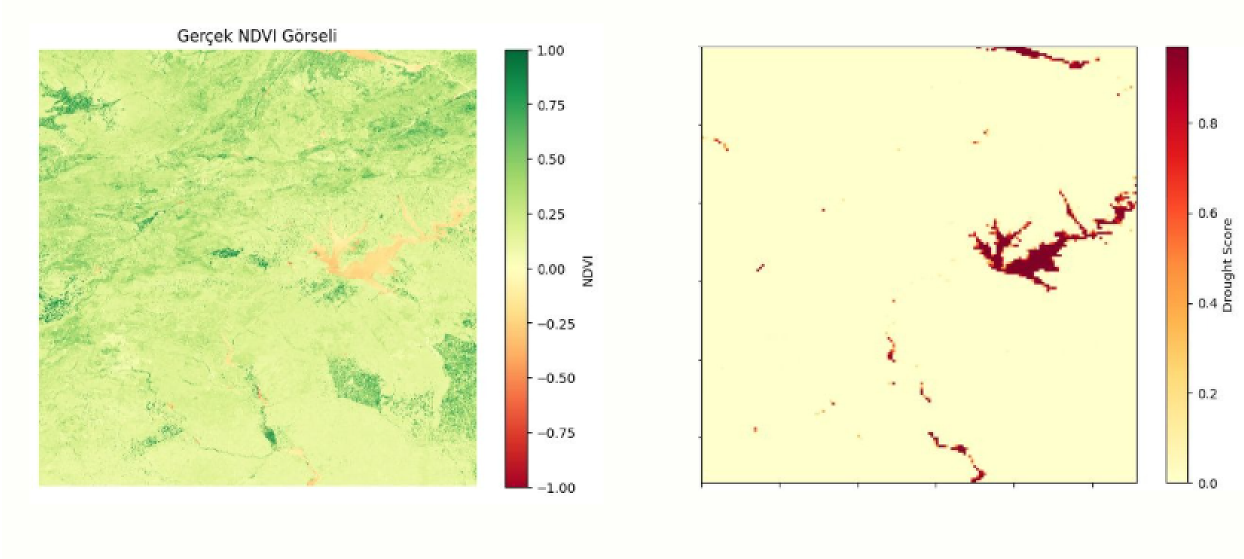

Figure 2: Spatial Comparison: NDVI vs. Drought Heatmap

Interpretation:

- The NDVI image (left) shows real vegetation health; red areas indicate weak vegetation.
- The drought score heatmap (right) highlights areas with high drought risk.
- High spatial overlap validates the relationship between vegetation stress and drought indicators.

11. Conclusion

This project successfully integrates satellite, climate, and textual data to develop a robust early warning system for food crises. It demonstrates high accuracy in classification, effective time-series forecasting, and strong visual correlation between NDVI and drought scores, paving the way for real-world applications.