Global Temperature Anomaly Analysis & Prediction Project

This project analyzes historical global temperature anomaly data and builds predictive models to forecast future temperature trends. Let me break down what it accomplishes and the data it works with:

The Data

The dataset contains global land-ocean temperature anomalies from 1880 to the present, sourced from NASA's GISTEMP dataset: https://data.giss.nasa.gov/gistemp/tabledata_v4/GLB.Ts+dSST.csv

Key characteristics:

- Temperature anomalies represent deviations from the 1951-1980 average temperatures
- The dataset includes monthly, seasonal, and annual temperature measurements
- Values are in degrees Celsius
- The time series spans approximately 145 years of climate data

What the Project Does

1. Data Analysis

- Loads and cleans the temperature dataset
- Converts string values to numeric data
- Visualizes historical temperature trends through various plots
- Shows the acceleration of warming, particularly since the 1970s

2. Model Development

The project implements and compares four different prediction models:

- **Linear Regression**: A simple model assuming a linear relationship between year and temperature
- **Decision Tree**: A non-linear model that makes predictions based on decision rules
- **Random Forest**: An ensemble of decision trees that improves prediction accuracy
- **Custom Gradient Descent**: A manually implemented linear regression using gradient descent optimization

3. Model Evaluation

- Splits data into training (80%) and testing (20%) sets
- Evaluates models using RMSE (Root Mean Square Error) and R^2 metrics

- Identifies Random Forest as the best-performing model with the lowest RMSE (0.0971) and highest R^2 (0.9422)

4. Prediction System

- Saves the best model (Random Forest) to a pickle file ("best_model.pkl")
- Creates a reusable prediction function in "prediction_function.py"
- Enables temperature anomaly predictions for any future year
- For 2025, predicts a temperature anomaly of approximately 0.4887°C

5. Visualization

- Displays historical temperature trends
- Shows monthly temperature patterns
- Compares model predictions with actual values
- Projects future temperature anomalies to around 2050

Key Findings

- Clear upward trend in global temperatures, especially since the mid-20th century
- The warming trend is accelerating in recent decades
- Random Forest model provides the most accurate predictions
- Future projections suggest continued warming, with temperature anomalies potentially reaching 0.9°C by 2050

The project delivers both analytical insights into past climate changes and a practical tool for estimating future temperature anomalies, which could be valuable for climate research, policy planning, and environmental impact assessments.