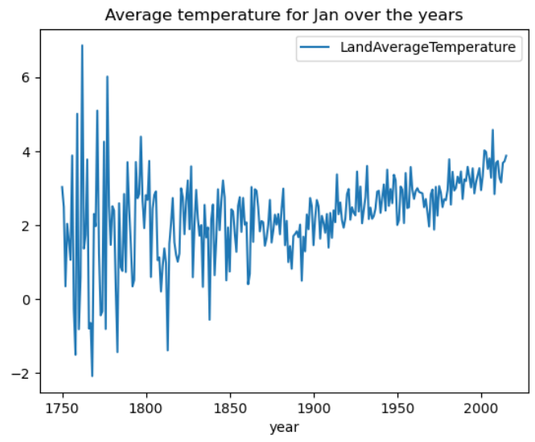
Climate change: Surface temperature prediction and visualization

Abhishek Pandey, Akhona Msibi, Kevin Pietruszka, Timothy Cody, Trevor McCrary

Motivation and Introduction

What is the problem:

- Climate change is a serious problem for future ecosystems.
- The trend, if continues can cause irreversible damage.
- There are not many tools that are easily accessible to the general public that effectively communicate the current situation.



Why is the Problem Important:

- Providing a tool that effectively communicates the reality of climate change to the general public can lead to increased community awareness
- Our model enables users to see the direct impact of climate change by seeing predicted temperature graphs for many cities around the world

Data Collection

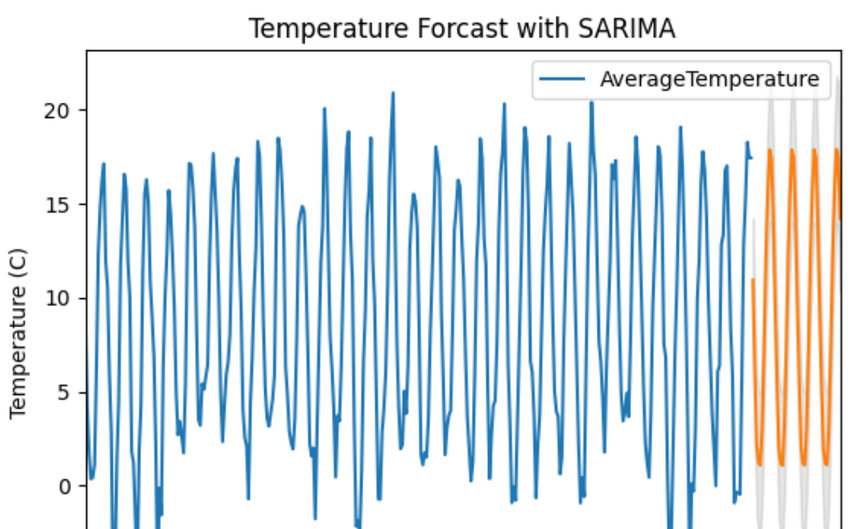
How the Data was Collected:

- Obtained from public datasets on kaggle.
- Parsed and cleaned using python scripts into an appropriate format for model training and visualization.

Characteristics of the Data:

Raw data: 8.6M rows, 500MB on disk.
Processed data: 9M rows, 574MB

Data Fields: Latitude, Longitude, City, Country, Recorded temperature



Approach

Our Approach and How it Works:

- Implemented several models using classical and machine learning techniques to predict future global temperatures.
- Used the first 80% of our dataset (1900-1988) to train our models and the last 20% (1989-2010) to test its predictive accuracy.
- d3.js for building visualization on top of model results.

How Our Approach Solves the Problem:

- Complex models for temperature prediction and time series forecasting. ~SOTA performance.
- Modern and interactive UI to visualize results.

What is Novel in Our Approach:

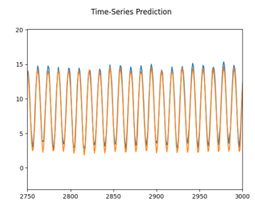
- Most modern methods use classical models for time series predictions of temperature.
- We explore multiple methods that use deep learning methods.
- Interactive and Intuitive UI for the general population.

Experiments and Results

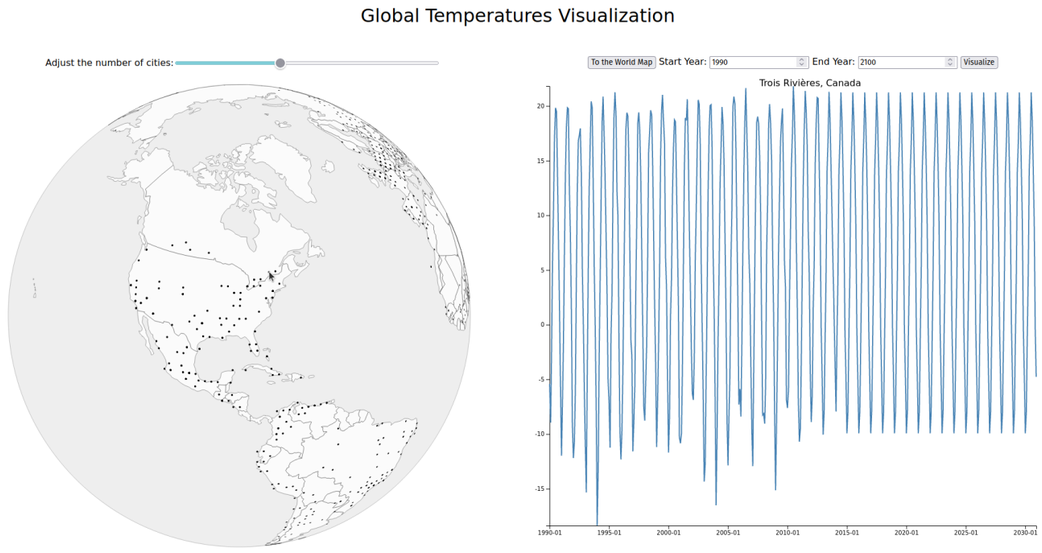
How We Evaluate Our Project:

- We evaluated our model's performance based on mean squared error.
- Our random forest model was chosen as the final model for our visualization based on its superior performance.

Model	Validation MSE
Random Forest	0.311
XGBoost	2.67
Decision Tree	0.412
Fuzzy time series	N/A
S-ARIMA	2.93
LSTM	0.843



Results of Our Project:



Comparison with other methods

- Our model agrees with modern research that predicts a gradual increase in temperature over time.