Project Proposal

Title: Machine Learning Forecasts of Agrofood CO_2 and Temperature Outcomes under Policy Scenarios

Course: Advanced Machine Learning (AML)

Team (Group 13): Abby Peck (ap69393), Abhiroop Kumar (ak56448), Grace Lin (cl47623), Shruthi Chembu (sc72339), Vyshnavi Maringanti (vm27288)

Due: Friday - October 24, 2025 at 11:59pm

Background and Motivation

Agriculture

Agriculture is one of the most important aspects of human civilization; however, compared to other sectors such as energy or transportation, it remains slower to transform and adopt new technologies. Unlike other rapidly industrialized industries, agriculture is deeply intertwined with natural ecosystems, biological variability, and land-dependent production processes, causing technological adoption to be limited. This stagnation in agricultural innovation shows the urgent need for data-driven predictive tools that can help the agricultural sector better align with global climate targets.

The Paris Agreement and its Temperature Benchmarks

The *Paris Agreement* is a legally binding international treaty on climate change set in 2015, which discussed two core temperature goals by the end of this century (2100):

1. Limit the increase in global average temperature to *well below 2°C* above pre-industrial levels
2. Pursue efforts to *restrict warming to 1.5°C* above pre-industrial levels

The primary drivers of global warming are energy production, transportation, industry, buildings, and agriculture. Among these fields, agriculture plays a dual role. It is both a significant emitter of greenhouse gases such as carbon dioxide, methane, and nitrous oxide, and a potential carbon absorption source through ways such as soil management, reforestation, and sustainable farming practices. Thus, addressing emissions from agriculture is a pivotal factor in determining if the world can meet the temperature targets.

Business Problem and Objectives

To address these challenges, this project will use machine learning techniques to forecast global temperature trajectories through 2100 based on agricultural indicators. Models will be trained on features representing emission sources and social-economic factors. The analysis that consists of three scenarios will then be developed to simulate the effects of different policy-driven outcomes:

- Baseline: A baseline, no-intervention scenario, reflecting the current *business-as-usual trajectory* if agricultural practices and emission patterns remain as is.
- Scenario A: A moderate emission reduction pathway targeting mild reductions in key emission related columns, including *Fertilizer manufacturing, manure management, etc.*
- Scenario B: An ambitious emission reduction pathway representing a stronger reduction in the same variables.

Instead of coming up with specific policies to trigger these changes, this project will focus on the quantitative outcomes of those policies by directly modifying the dataset's feature values. This design enables a data-driven assessment of how different levels of agricultural transformation influence long-term temperature outcomes, helping identify which agricultural changes are most effective for real-world policy intervention by aiming for minimum reductions with maximum results.

To do this accurately, the model will address the critical time-lag challenge, where temperature "stock" is driven by the accumulation of emission "flow." Therefore, the ML model will first predict these annual emission pathways and then translate them into final temperature trajectories using the FaIR simple climate model.

Dataset Overview

- Dataset: Agri-food CO_2 emission - Kaggle
- Size: ~7k rows and 31 columns (2.05 MB)
- Features: Country, Year, Crops, Fertilizer Use, Cultivated Area, Yield, CO_2 Emissions, etc.

Analytical Workflow

This project integrates regression and neural network modeling approaches to forecast global temperature outcomes. The analysis will connect agricultural emission sources with long-term climate outcomes by simulating "what-if" cases across different pathways (Scenarios A and B).

1. Data Preparation: Clean and standardize all variables, impute missing data, and normalize country-level indicators.
2. Exploratory Data Analysis: Correlation and PCA visualizations to uncover dominant emission clusters and feature importance.
3. Model Training & Scenario Simulation:
    - *Train:* Build models (Ridge, Lasso, MLP) to predict total CO_2 emissions from agricultural and socio-economic inputs.

    - *Simulate:* Project input features through 2100 to represent the three scenarios.
    - *Forecast:* Generate emission time-series for each scenario using trained models.
    - *Convert:* Translate emission pathways into temperature outcomes via a simplified climate model - FaIR (Finite Amplitude Impulse-Response).
4. Evaluation: Assess model accuracy and compare scenario trajectories with predefined benchmarks.

Evaluation Criteria

- R^2 / RMSE: Assess regression accuracy for CO_2 and temperature forecasts.
- Feature-Importance Coverage: Explainability of model output.
- Mean Absolute Temperature Error (MATE), Scenario Comparison Metric, etc.

Policy Recommendations

The two agricultural emission reduction pathways are designed to help model results under different levels of policy intervention. The intention is to connect machine learning forecasts with real-world policies under the Paris Agreement framework.

- Scenario A: The parameters of the required features will be adjusted moderately while still supporting the Paris Agreement objectives by guiding a transition toward sustainable intensification while balancing food security and emission control.
- Scenario B: The feature parameters will be reduced by a much larger percentage to obtain the extreme values of the required outcomes, consistent with higher-ambition climate strategies under the 1.5 °C Paris requirement.

Expected Impact

- Scenario A: Here, the model output will be able to offer suggestions for a balanced route to emission control and food security stability via sustainable agriculture.
- Scenario B: The model results will help provide inputs for high-ambition climate goals via transformative changes, using strategies such as technology adoption, carbon reduction, etc.

References

1. United Nations (2024). https://www.un.org/en/climatechange/net-zero-coalition.
2. UNFCCC (2024). National Inventory Submissions & Reporting Requirements.
3. Lobello, A. (2024). Agri-Food CO_2 Emission Dataset (Forecasting ML). Kaggle.