Vision Statement

Team Name: Flare
Project Name: LSTM (Long Short-Term Memory) For Wildfires
Project Sponsor: PwC
Members:

- Kelly Lin (Lead)
- Alvin Tu (Scribe)
- Nick Ong
- Shuya Shou
- Steven Chang

Project Summary

With the growing climate change predicament, wildfires are an increasingly dangerous problem, especially in wildfire-prone states like California. They pose a risk to local businesses by putting both structures and workers at risk. Various conditions increase the likelihood of wildfires, such as a large presence of flammable materials or dry climate. Our team hopes to use these conditions to accurately predict wildfires across the United States.

Problem Being Solved

Currently, many areas are in danger by wildfires, with very little means to detect and prevent them in the long term. As climate change worsens, we need to prepare for the increased occurrences of wildfires since it will not only help reduce business risks caused from wildfires, but also potentially save lives and protect the environment. Our team will tackle this problem by using LSTM neural network models to predict potential wildfires and use that information to help make crucial decisions.

Why is This Important?

Mitigating the risk and damages that occur from wildfires is urgent. Over the past ten years, the U.S. has averaged an annual 7.4 million acres burned[1], costing the country $2.4 billion a year[2]. Additionally, the world's largest corporations face up to $1 trillion in damages from climate change[3]. However, the most impacted group are individuals who have lost their homes, savings, and lives from fires. Investing in methodologies to predict wildfires would minimize the toll of these catastrophic events.

Current Existing Solutions

1. Statistical prediction model - costly; deviate from ground truth.
2. Satellite based solution - expensive; difficult for maintenance.
3. Climate forecasting/Extreme weather detection using GPU intensive methods.
4. Using the LSTM model for predicting locations of fire - current models have limited amounts of data used.

1. https://sgp.fas.org/crs/misc/IF10244.pdf
2. https://www.nifc.gov/fire-information/statistics/suppression-costs
3. https://www.cdp.net/en/articles/media/worlds-biggest-companies-face-1-trillion-in-climate-change-risks

Project Outcome

The goal of this project is to create a LSTM neural network to predict wildfire occurrences using time series of weather variables, as well as other data including land cover types and topography. Ideally, a wildfire prediction model will be developed for some states, and then tested on other areas, resulting in a strong, generalized model. Finally, the model can be improved on by ensuring robustness to outliers through data that predicts variances in the features we use to guarantee the model will not fail in the future.

Technologies

- Data Processing and Model Building: PyTorch, Tensorflow, Python, NumPy, Pandas
- Codebase: GitHub
- Data Visualization: Google APIs

Milestones

I. Survey other paper and learn about what has been done in the past
II. Understand LSTM models
    A. Study the architecture of the model
    B. Learn about the training and evaluation process
III. Feature selection
    A. Perform data analysis on given data
    B. Perform correlation analysis to evaluate feature importance
    C. Select a subset of features for training
    D. Prepare the data for training by performing feature normalization
IV. Train LSTM models
    A. Dataset should be split into training, validation, and test set
    B. Train multiple models to select the best hyperparameters
    C. One possible stretch goal is to improve the model by changing the network architecture
V. Evaluate our model
VI. Try other RNNs
    A. Compare the performance of other networks such as the Transformer to LSTM models
VII. Test our model on other areas
    A. To make sure our model generalize well