Spatio-Temporal Wildfire Forecasting

Balaram Behera, William Boynton, Sarah Cheah, Jerry Cheung, Shrey Patel, Hui Min Tee

Georgia Tech – CSE 6242 – Prof. Polo

Summary

Over the past 5 years, California has faced 34 thousand wildfires spanning over 10 million acres, killing dozens and destroying thousands of homes.

Increasing in number and size, these fires prompted research in **wildfire prediction, patterns, and prevention**, assisting state and federal institutions to reduce risk and improve mitigation strategy.

We thus propose **wildfire perimeter forecasting**: given a wildfire's current state, we find the forecasted geometry in 2, 4, and 6 days, applying recent deep learning innovations in spatio-temporal forecasting.

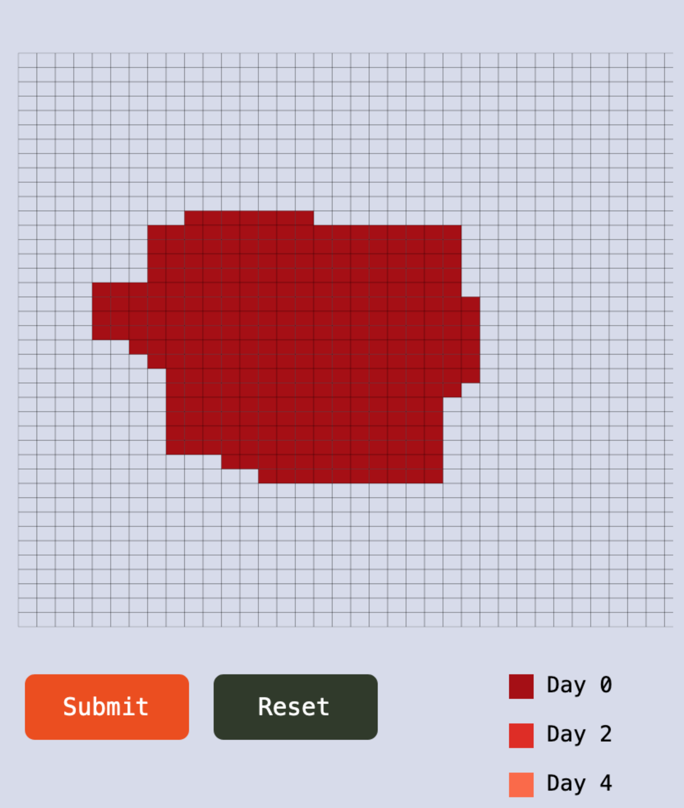
Data

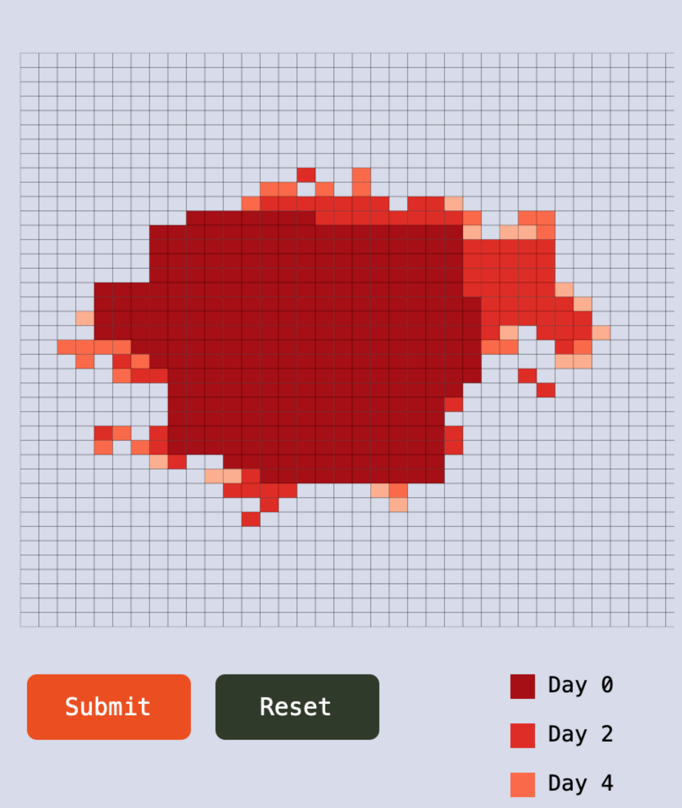
We downloaded our **spatio-temporal fire perimeter data** from the NIFC public server (**43 GB**) which includes all large wildfires since 2012 in California (**371 fires**).

This data required **immense cleaning** and filtering to incorporate different GIS formats and projections, corrupted data files, and missing data.

We also downloaded topological (elevation, vegetation, land cover, road networks), hydrological (water bodies), socioeconomic (population, income), and meteorological (climate, wind) data.

Due to difficulties in data integration and cleaning, we only used the fire dataset for prediction.





Model

We first split the data into 50% for training, 25% for validation, and 25% for testing to evaluate the models and tune hyperparameters effectively.

In previous research, many spatio-temporal models use Convolutional and LSTM units in their model; as our problem is **One-To-Many**, we first repeat the singular input, which is a widely known technique.

Our Conv+LSTM model has 2 **ConvLSTM2D** layers (from Keras) with ReLU activation and MaxPooling3D layers, and then following by 3 Dense layers ending with 40 x 40 x 3 units with Sigmoid activation.

We also created a simple **Feedforward Network** using 3 Dense Layers with ReLU activation and ending with a similar output layer with Sigmoid activation.

We optimized both models with the Keras **Adam** optimizer on **Mean Squared Error** loss.

Experiments and Results

We evaluated our model with **Binary Accuracy** (whether a cell is on fire or not), but since some data has many non-fire cells, metrics of **Precision** and **Recall** are more important in deciding a good model as both must be high.

As the simple **Dense model** trained much faster and performed better on test data over multiple trials, we choose that model over the Conv+LSTM.

Test Dataset Results

Model	Accuracy	Precision	Recall
Dense	98.27%	96.23%	93.05%
CNN+LSTM	98.19%	96.05%	92.73%

Visualization

We created a grid (with 40 x 40 km. cells) map of California's TigerLine as a **GeoJSON** using QGIS.

Utilizing **HTML, D3.js, and CSS**, we rendered the gridded map of California, and added the capability to select a square of the grid to procure a zoomed-in window of the square, panning the 40 x 40 km. area which represents a 1,600 km^2 area made of 1 km^2 squares.

Users "**draw a fire**" to start using their mouse to drag and draw the initial fire perimeter.

Then by pressing "Submit", the model predicts the **progression** of that fire in the next **2, 4 and 6 days** and displays them as extended layers of the initial perimeter.