Wildfire Prediction

a SENG474 project by:

Valeriya Savchenko, Tim Rolston

Predicting Wildfires

Predicting Wildfires can be a complex process that has many different approaches that each have their own pros and cons. Due to the size and lack access to data, in this project we will be exploring methods to build a small scale wildfire prediction system based off of satellite data gathered throughout Canada in the year 2021.

The reason for using data from 2021 is because in Canada there was a spike in the number of wildfires that occurred comparative to surrounding years. http://nfdp.ccfm.org/en/data/fires.php

Main Idea

Our design uses data gained from MODIS, a nasa satellite data collection technology, to gain insights on certain areas based off of ‘frp’ which is a number that shows the fire radiative power in Watts and ‘confidence’ which is a value between 1 and 100 (percent) showing the confidence that there was a wildfire in the area.

The website describing this particular dataset can be found here:
https://modis-fire.umd.edu/af.html

And the website the data was collected from:
https://firms.modaps.eosdis.nasa.gov/download/

Data Processing Methods

In order to make our data accurate and keep the training time short, we:

- Took out features with poor correlation with our target - confidence and frp
- Attempted different training methods to determine which one would be the fastest and most accurate for the target we focused on

Training Details

After trying a few methods we found that frp gave a much higher accuracy in training than confidence did and using a DTR method for training both managed to train the dataset very fast and also had a high accuracy.

Findings

Initially we tried using a SDG classifier as well as Logistic Regression both with confidence as the target, but we realized they were both not very accurate or efficient.

Using a Decision Tree classifier however managed to speed up the training process significantly, but we still had an issue of low accuracy - about 20%. So we finally decided to use frp as our target along with Decision Tree Regression which gave us 90% accuracy at very efficient training times (< 5s).

Resources

There was also a useful manual for this specific processed MODIS data, that we used in order to understand the complexities of the data. This was we could better judge the significance of the features and what they meant.

The manual can be found here:
https://modis-fire.umd.edu/files/MODIS_C6_C6.1_Fire_User_Guide_1.0.pdf

We also found another use case of someone else performing a similar process here:https://levelup.gitconnected.com/how-machine-learning-helps-us-to-predict-wildfires-fd4e5ba14578

Conclusion

In this notebook we have managed to show a method of using the collected MODIS data to be able to produce an accurate training model that can predict wildfires in Canada 90% of the time for the year of 2021. Our goal was to find an efficient way to process this data and an accurate method to train a model on in order to successfully predict wildfires. We found that, after removing a few features and training the data using Decision Tree Regression with frp as the target, we were able to train a fairly good model.