

# Wildfires



Noa Shadmon, Austin Chi  
Instructor: Professor James Sharpnack

# Goals

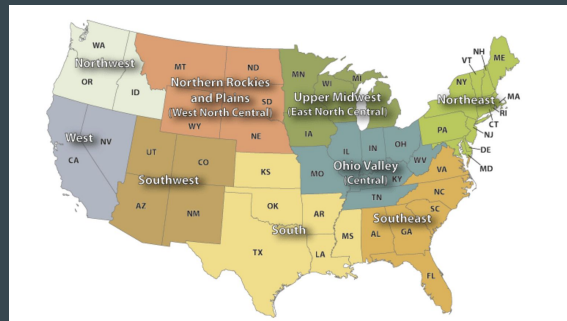
- Accurately predict wildfire ignition in an area with better time resolution than existing models.
- Predict wildfire ignitions more accurately than existing methods
- Develop a robust way to sample zeros when zeros are infinite and ones are fixed.

# Past Literature

- Modeling Wildfire Ignition Distribution and Making Prediction of Human-caused Wildfire by Weichen Ouyang , Chia Yung Han , and Susanna T.Y.Tong
  - Used Distance to Road, distance to wildland–urban interface, distance to development, development level, January mean temp, etc.
  - Accuracy: Logistic (0.718), Conditional Inference Tree (.755), Random Forest (.95)
- Wildfire ignition-distribution modelling: a comparative study in the Huron–Manistee National Forest, Michigan, USA by Avi Bar Massada, Alexandra D. Syphard, Susan I. Stewart, and Volker C. Radeloff
  - Used Landcover (% agriculture, %conifers etc.), Distance to nearest structure, road etc., and topographic (elevation, slope, etc)
  - Accuracy: Logistic (.664) Random Forest (.694) Maximum Entropy (.716)

# Data Collection

- Original dataset: <https://fsapps.nwcg.gov/afm/gisdata.php>
  - Consists of the 1s (fires)
- <https://darksky.net/>
  - Api to get features for 0s and 1s
- <https://www.google.com/maps>
  - Api to get elevation feature
  - Api to get state boundary
- <https://www.ncdc.noaa.gov/monitoring-references/maps/us-climate-regions.php>
  - To split points based on similar climate regions



# Problems with Sampling Zeros

- The way we sample our zeros affects our classification accuracy.
- Need to pick “difficult zeros”
- How to sample sample from an infinite number of zeros when we have a finite number of ones?
- Approximating a Counting Process using classification methods\*

# How we get our zeros

- Latitude, longitude, date, and time from the fire data
- Randomly select a '+' or '-' and use it to add/subtract epsilon from the latitude and longitude
- Uniformly select a date in 2016 as well as a time → convert to unix time
- Make API request to get weather features
- Make API request to get elevation

# Features

- Max Temperature
- Min Temperature
- Cloud Cover
- Dew Point
- Humidity
- Precipitation Intensity
- Precipitation Probability
- Elevation
- Pressure
- Visibility
- Wind Bearing
- Wind Speed
- Zone (location)
- Distance to Major Population Center\*
- Vegetation\*

# Analysis

Put 0s and 1s into one dataframe

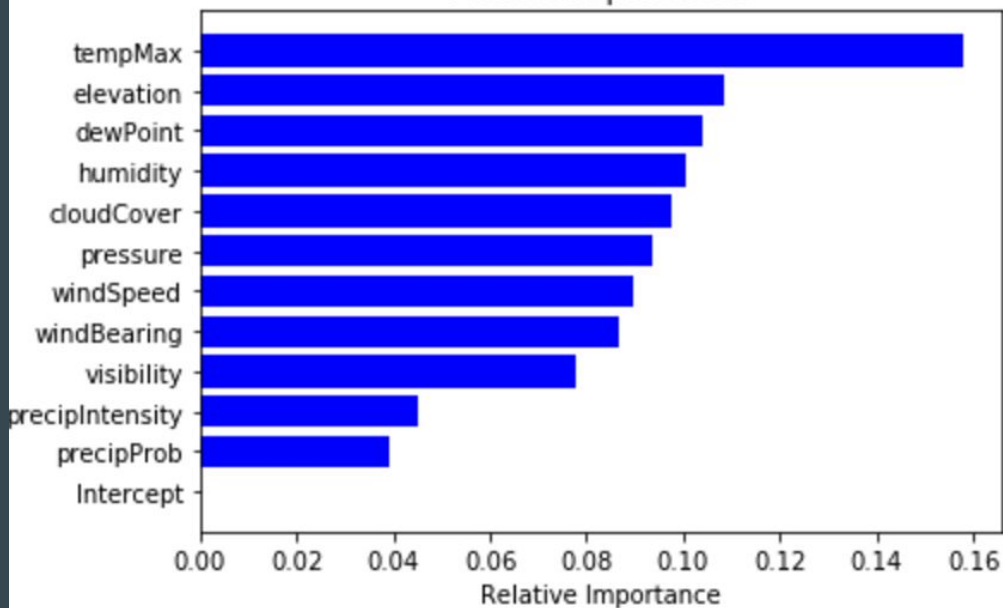
Our predictor = [0,1] → fire or not

Models - Logistic Regression, Random Forest, ANN

Accuracy - Logistic with 5-Fold validation (0.69), Random Forest (0.76), Feed Forward Neural Networks with 2 layers(.7377)



Feature Importances



# Next Steps

- More data
- Add distance to nearest population center as a feature
- Split data into ecoregions
- Apply our methods to data from prior papers