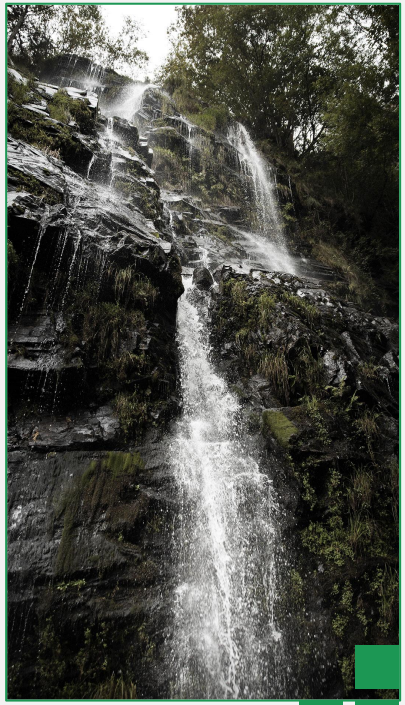
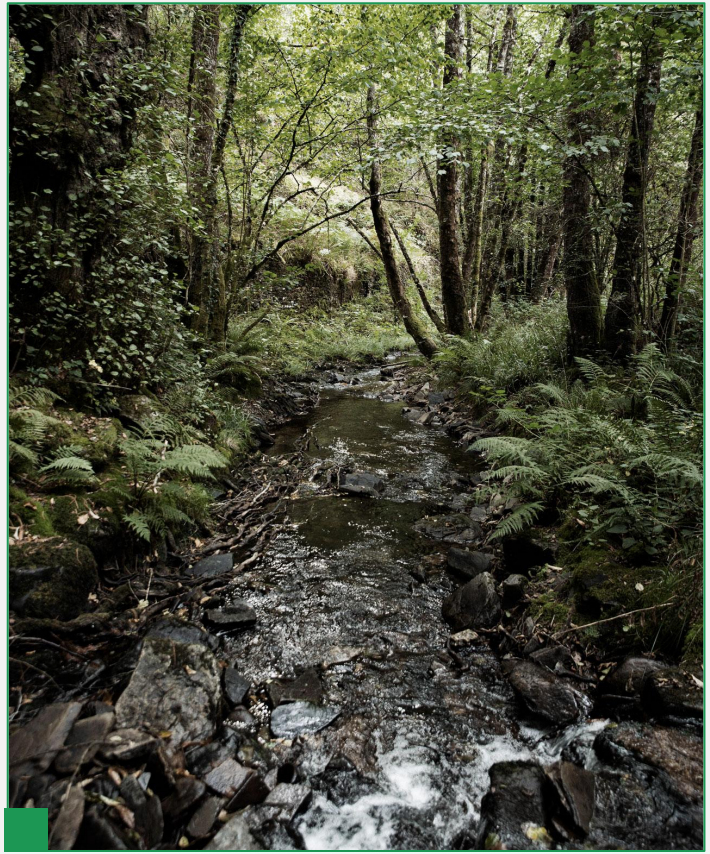
CSC 411

Wildfire Prediction

By James McSweeney
Supervised by Professor Aguiar-Pulido

Table of contents



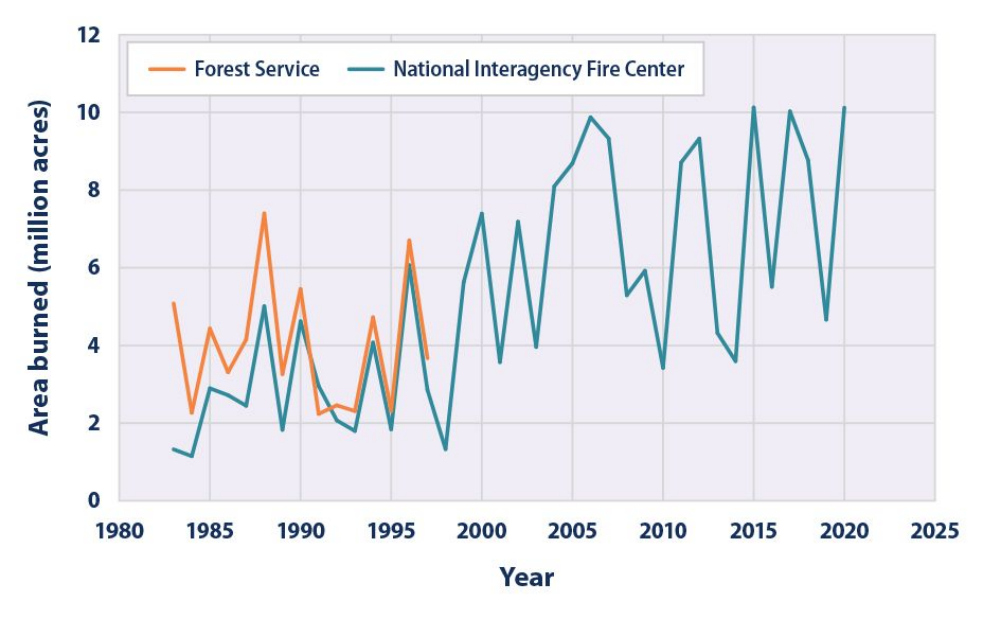


01.

Introduction

Why Wildfires

- Negative effect
  - Public health
  - Natural environment
  - Federal and State budgets

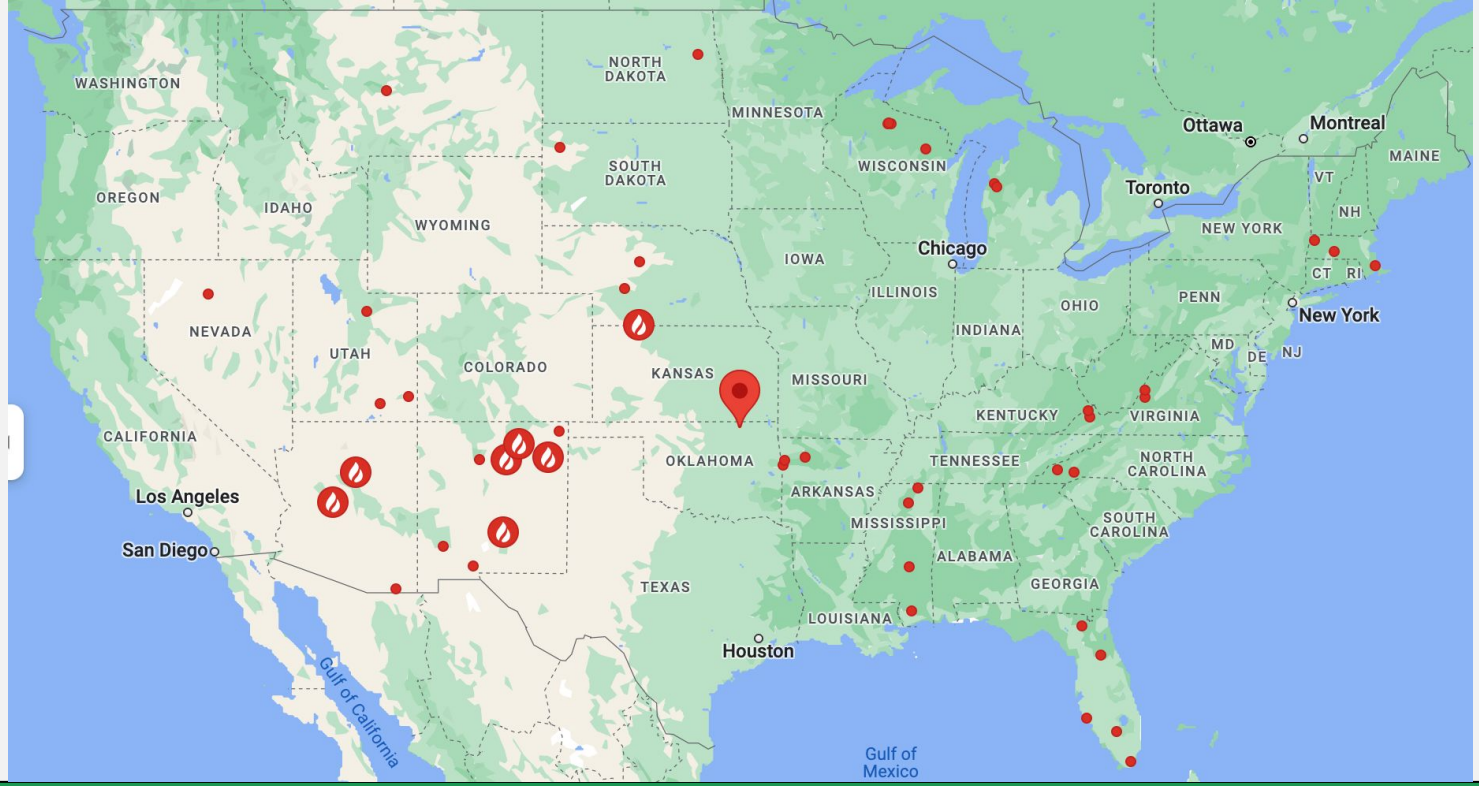


17,300

Fires since Jan 13th 2022

831,842

Acres burned since Jan 13th 2022



Current Wildfires on April 28nd, 2022

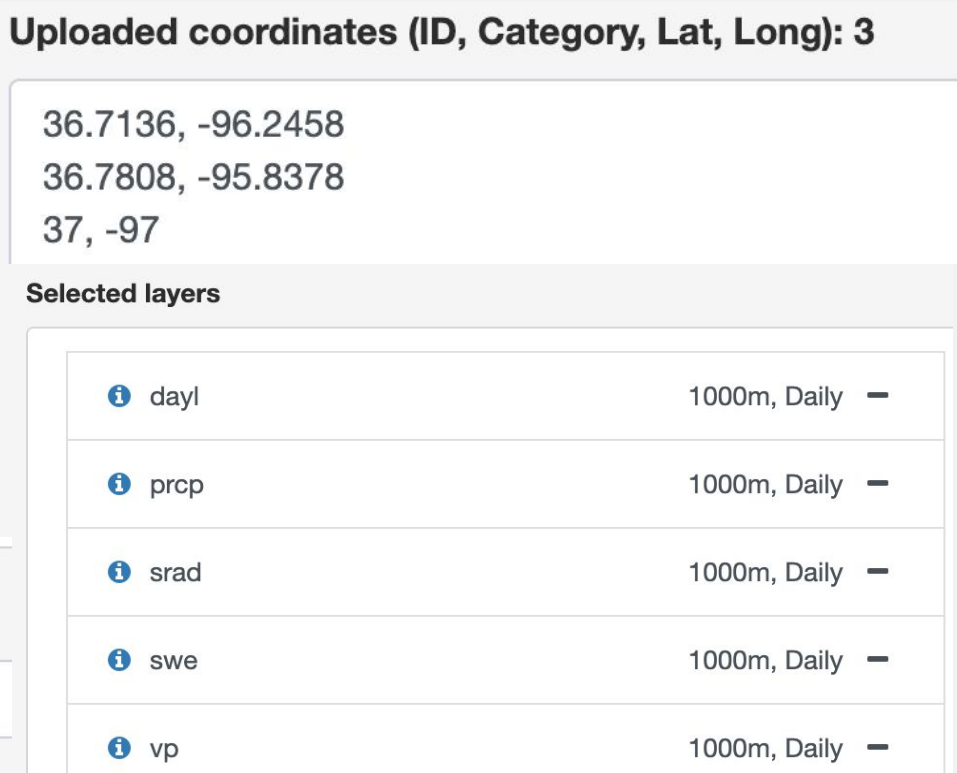
02. Data

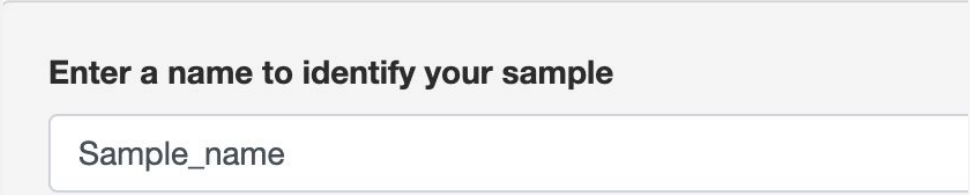
Data Source

Application for Extracting and Exploring Analysis Ready Samples (AppEEARS)

Requesting Data

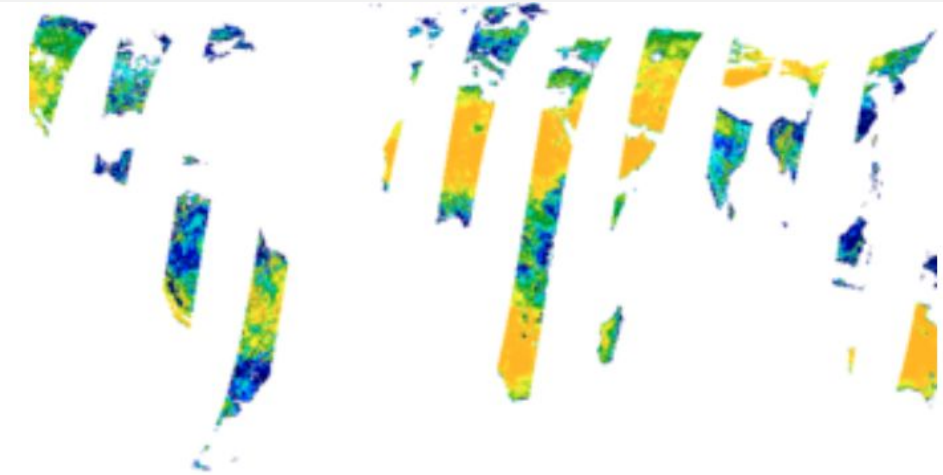
1. Input coordinates
2. Select the satellite
3. Select the variables you need
4. Input the date range
5. Submit





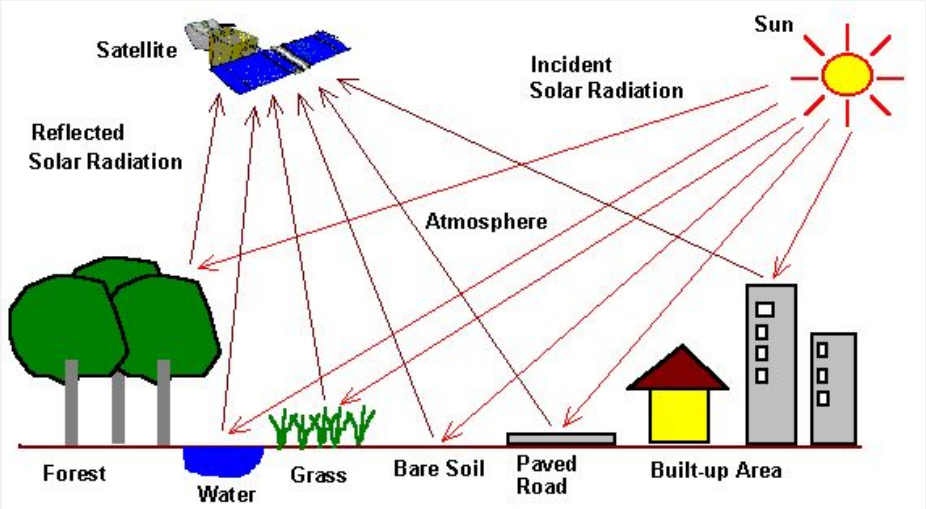
Data Capturing

- Two different Satellites
  - MODIS
  - Daymet
- Wildfire Points
  - Federal
  - State
  - Local



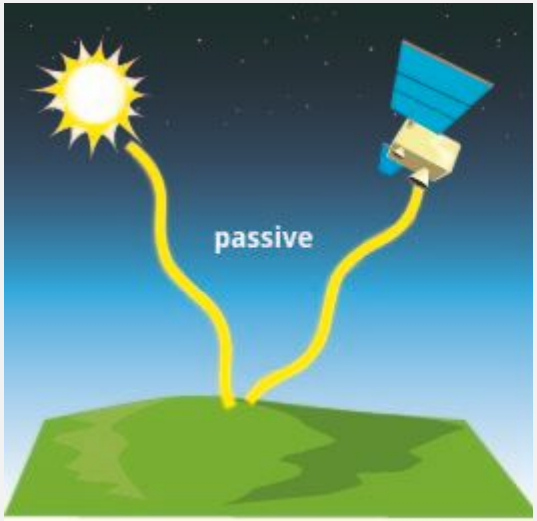
Remote Sensing Advantages

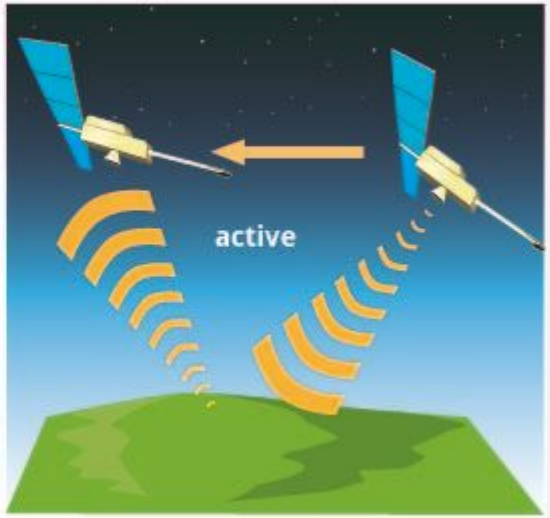
- Remote Areas
- Fast upload speeds
- Can cover a lot of land



Remote Sensing Disadvantages

- Spatial resolution
- Temporal resolution
- Spectral resolution





03.

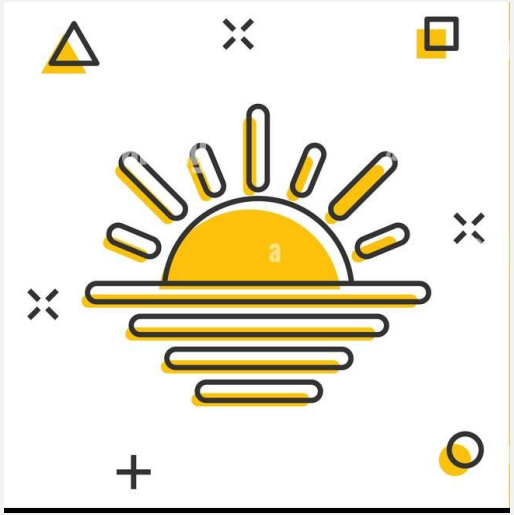
Variables

Feature Selection

- Longitude and Latitude
- Quality of measurement
- High, medium, and low classes
- Visible clouds
- SMAP

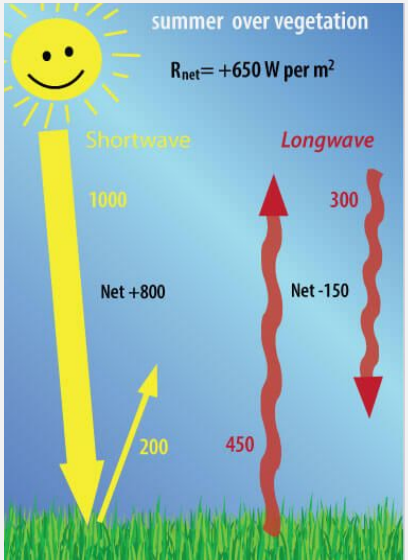
Daylight

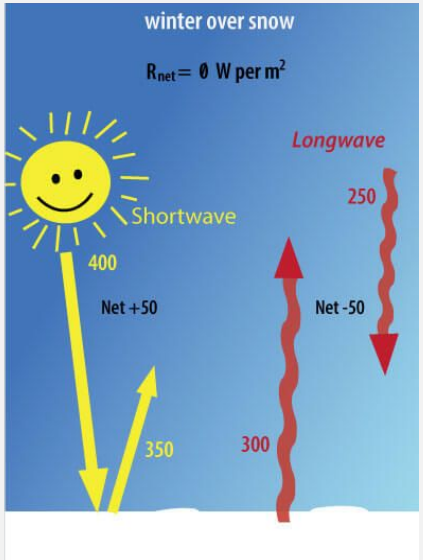
- Units: seconds per day of daylight
- Daylight can infer the season
- Considered very significant in the model



Short Wave Radiation

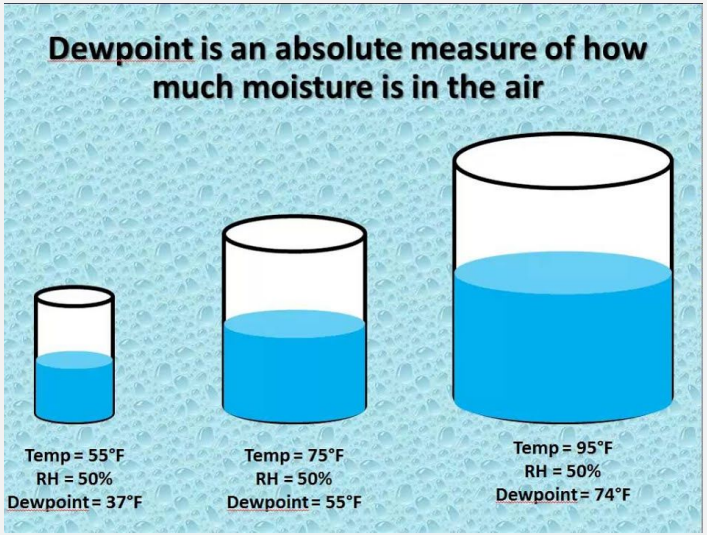
- Units:W/m^2
- Considered very significant in the model
- The intensity of the sunlight reflected off the earth





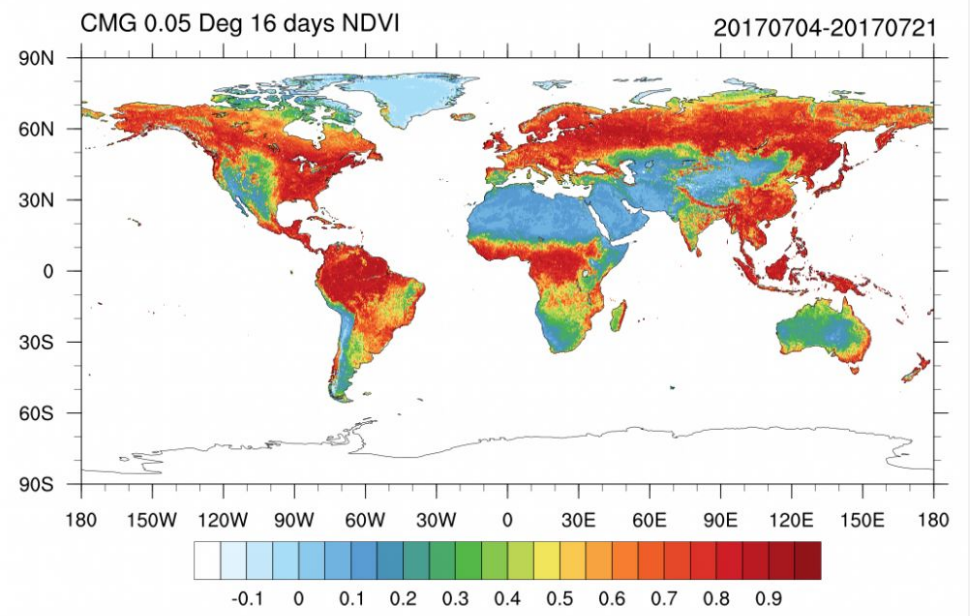
Vapor Pressure

- Units:Pa
- Measures humidity
- Considered very significant in the model
- Low pressure indicates active weather
  - High winds
  - Rain



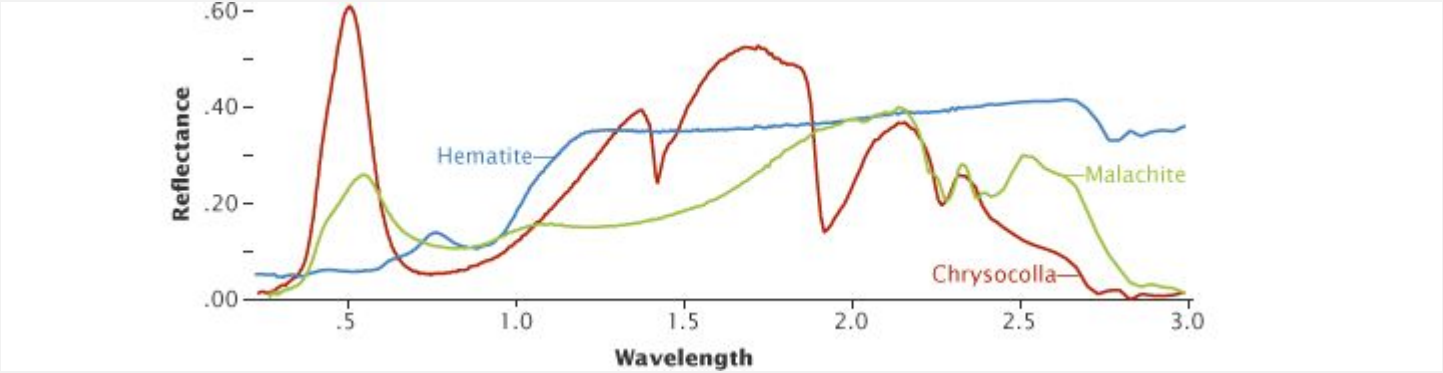
Vegative Indexes

- Enhanced vegetation index (EVI)
- EVI was considered very significant
- Normalized vegetation index (NDVI)
- NDVI was considered insignificant



Near Infrared Reflectance Spectroscopy

- Unit:nanometers
- Different organisms absorb different wavelengths of infrared waves
- Helps remotely classify different organisms
- Considered very significant in the model

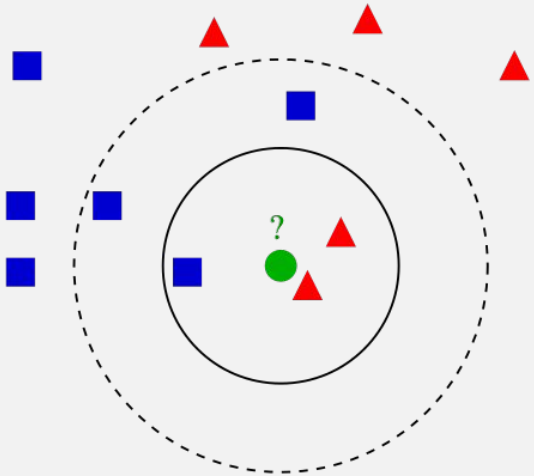


Classification Methods

04.

K-Nearest Neighbors

- Classifies based on its similarity to neighbors
- K is amount of neighbors to be considered for each new point
- Euclidean distance is used to find the closest K Neighbors
- Hyper parameter optimization found 5 as the optimal K value
- train(Formula, data,"KNN",CV,tunelength)



Truth

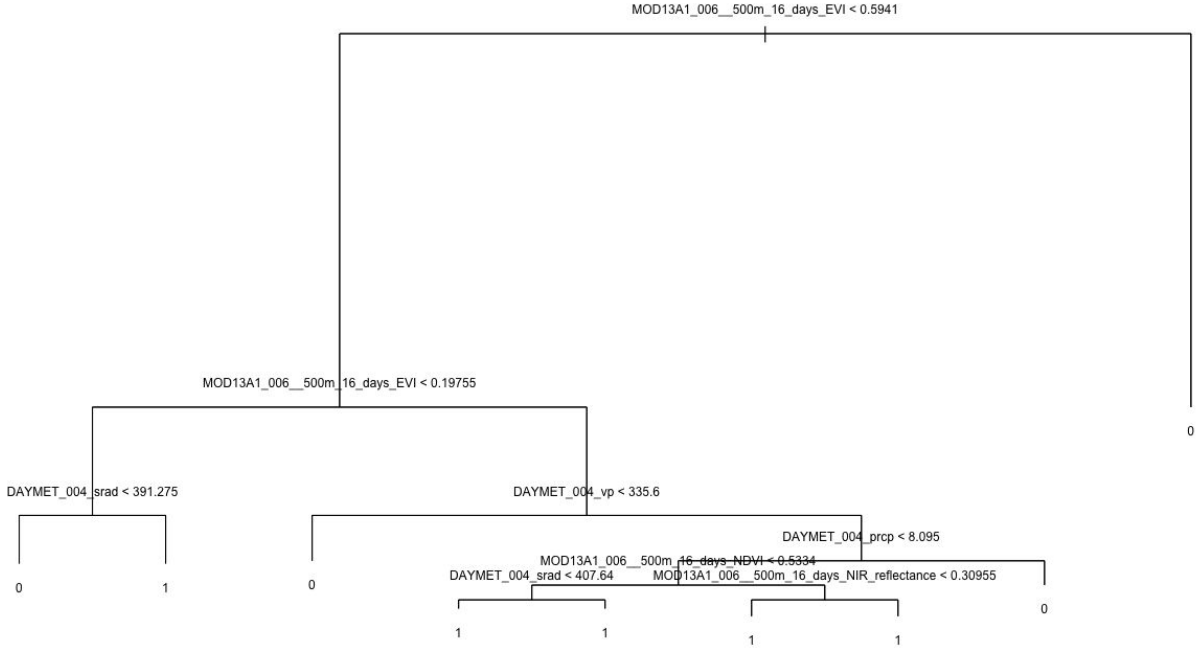
Prediction	Non-Fire	Fire
NonFire	50	4
Fire	0	43

Classification Tree

- A mapping of binary decisions that leads to a classification
- Best visually understandable classification method
- Splitting each decision into subpopulation based off binary decisions
- tree(Variable Formula,dataframe)

		Truth	
		Non-Fire	Fire
Prediction	Non-Fire	40	6
	Fire	1	50

Tree Visualization



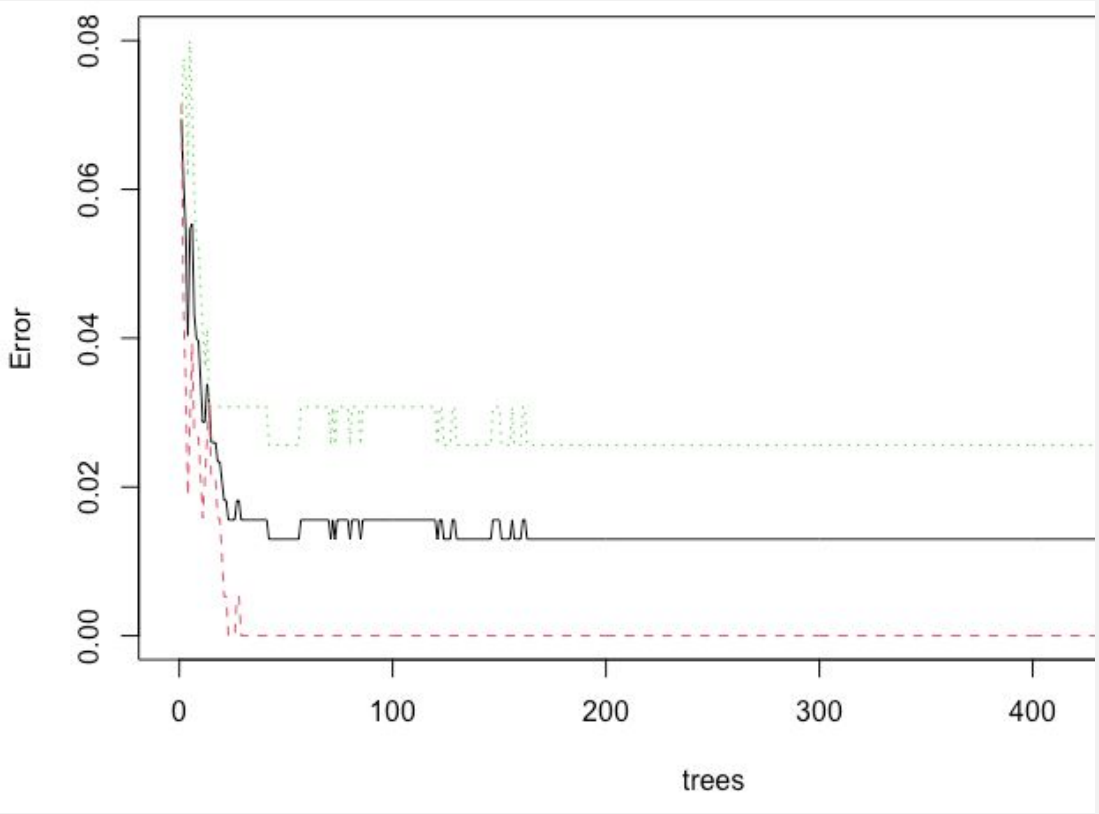
Random Forest

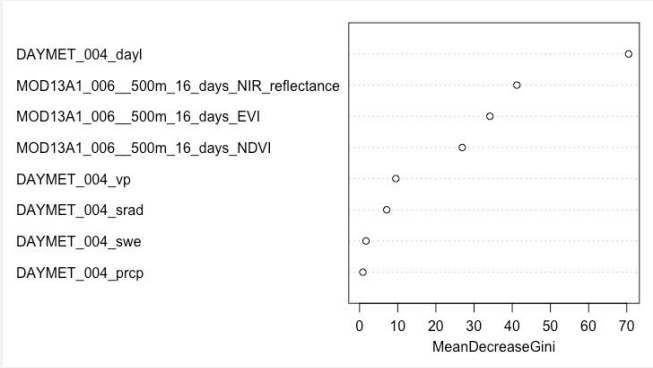
- Hyperparameter tuning found number of variable to be considered at each step to be 3
- train(Formula, data,subset, "RF", tuneGrid, trControl)

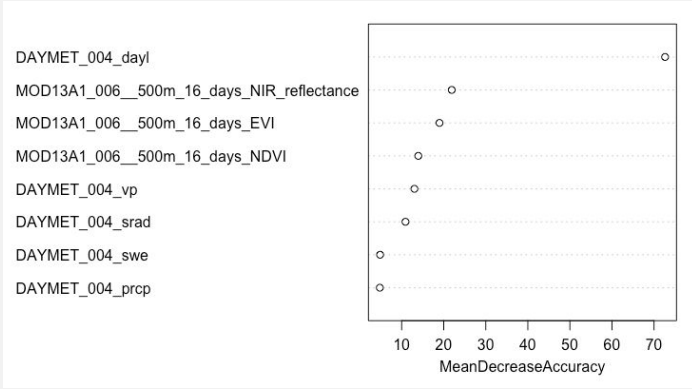
Truth

Prediction	Non-Fire	Fire
Non-Fire	50	1
Fire	0	46

Random Forest Visualization

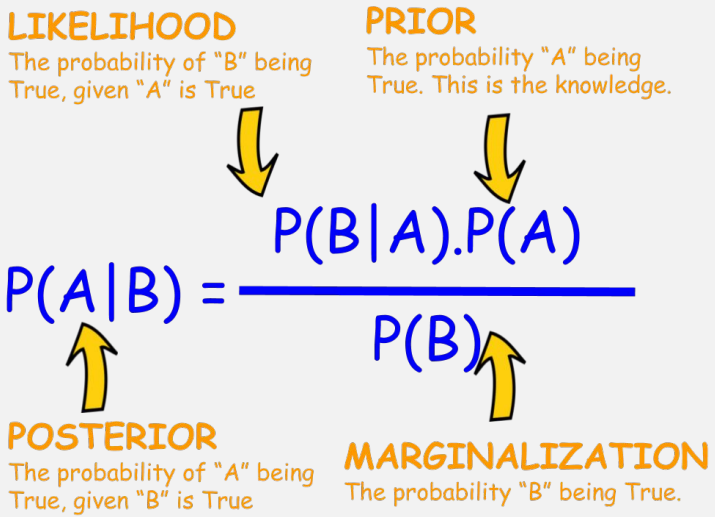






Naive Bayes

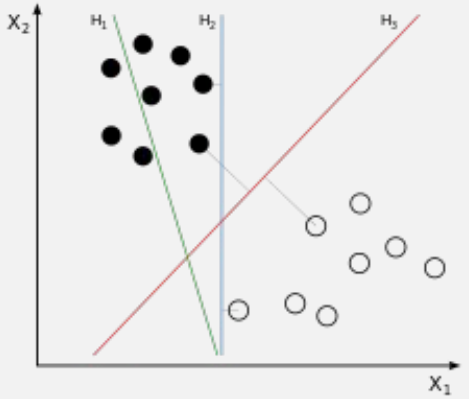
- Probabilistic classification
- Applies Bayes theorem with a strong assumption of independence between the variables
- naiveBayes(Formula, data,trControl)



	Truth	
Prediction	Non-Fire	Fire
Non-Fire	31	35
Fire	0	66

Support Vector Machines

- Supervised learning method
- Non-probabilistic binary linear classifier
- tune(svm,Formula, data,kernel,ranges,trControl)
- svm(Formula data, kernel, cost, scale)



Truth

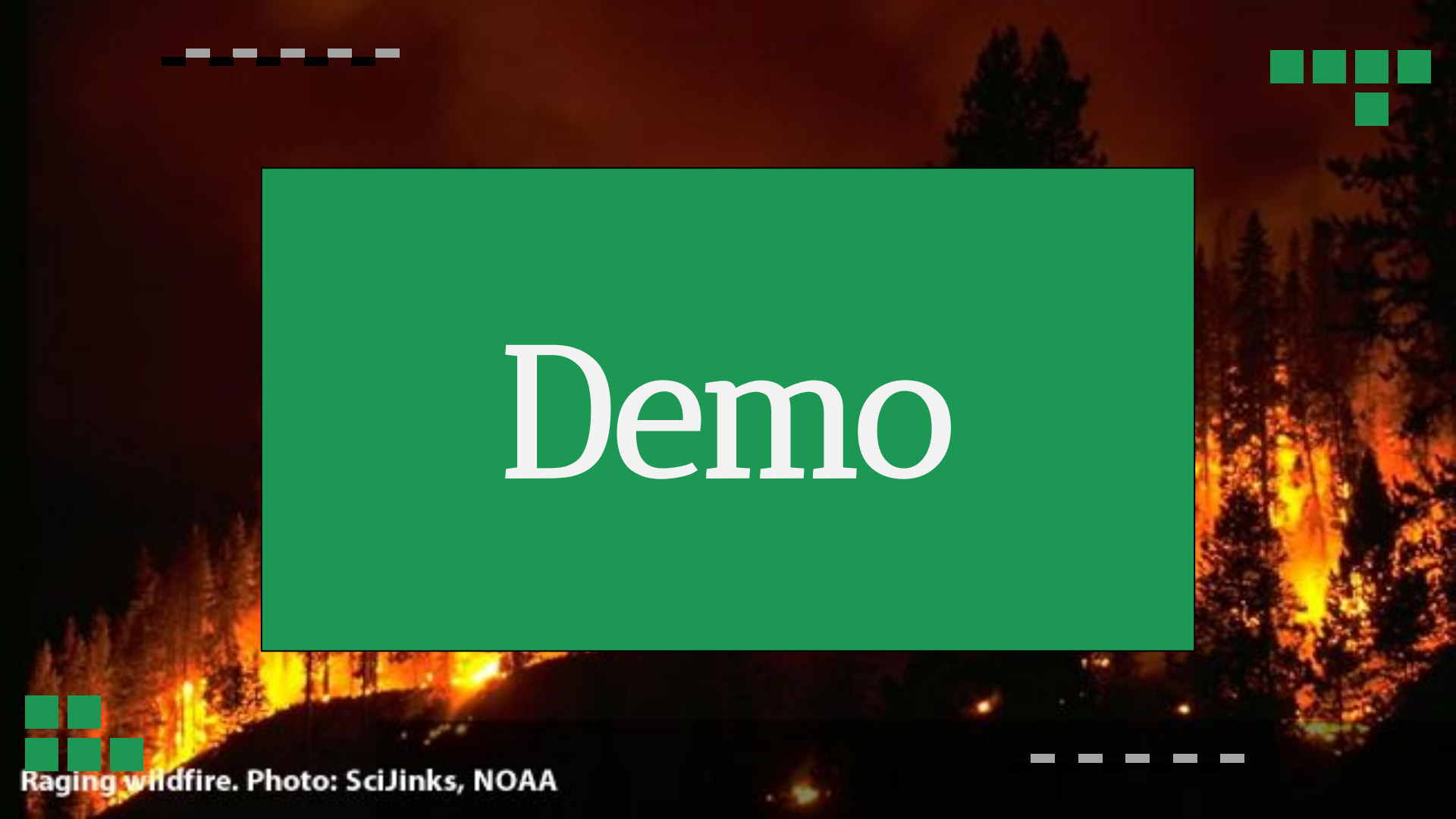
Prediction

	Non-Fire	Fire
Non-Fire	44	2
Fire	4	46

Results

Classification Method	Accuracy	Runtime in Seconds
KNN	95.8%	3.1
Classification Tree	92.7%	.7
Random Forest	98%	40
Naive Bayes	73.5%	.4
SVM	93.8%	1.09

Demo



Raging wildfire. Photo: SciJinks, NOAA

03.

Conclusion

Summary

- Classifies areas that are susceptible to wildfires
- Compared different classification models
- Preformed with NASA satellite data

Challenges

- Getting useable satellite data
- Time between data requests
- File organization

Future Work

- Results could be found in real time with a large server
- Coordinate with government models
- Add live weather data
- Factor humans into the model

Acknowledgements

The University of Miami

Dr Aguiar

Github

https://github.com/JamesMcSweeney/WildFires