Forest Fires in the U.S

By: Arturo Delgado

Thats a lot of damage.

Why care?


HEALTH
HOUSE
TRAVEL
LIFE
INSURANCE
CAR
BUSINESS
LIABILITY

My data

Kaggle

Total 1.88 million

Sample of 100k

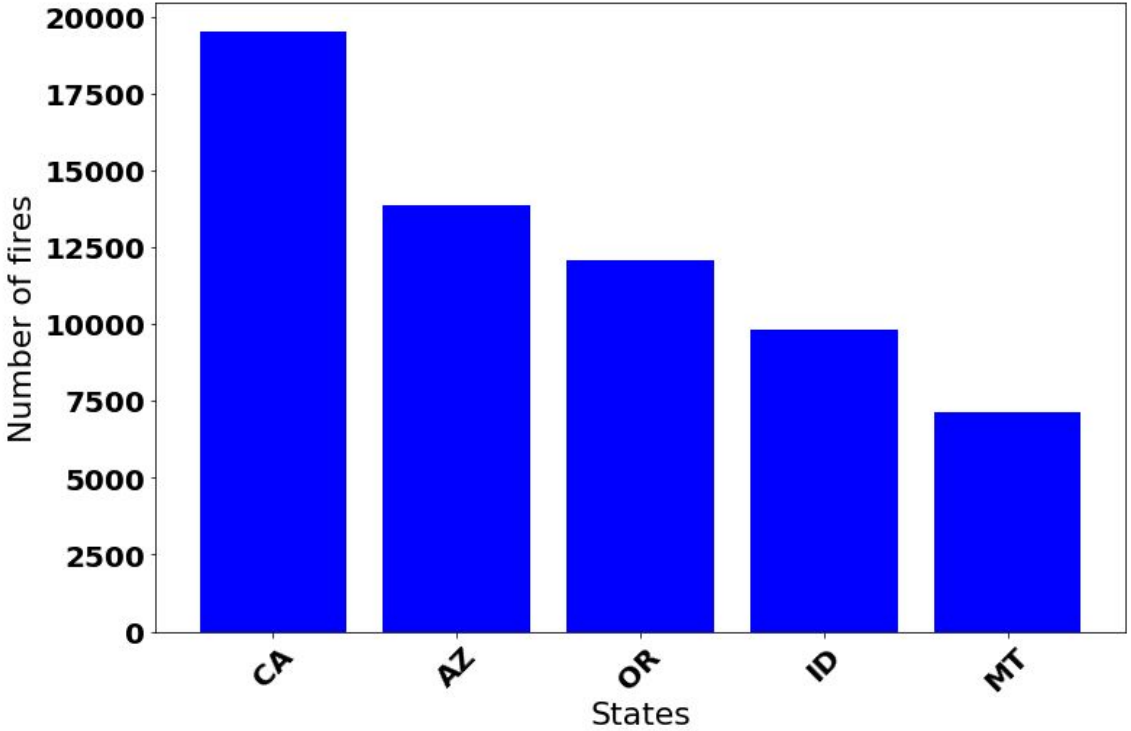
39 column

Collected: 1992-2015

Tools

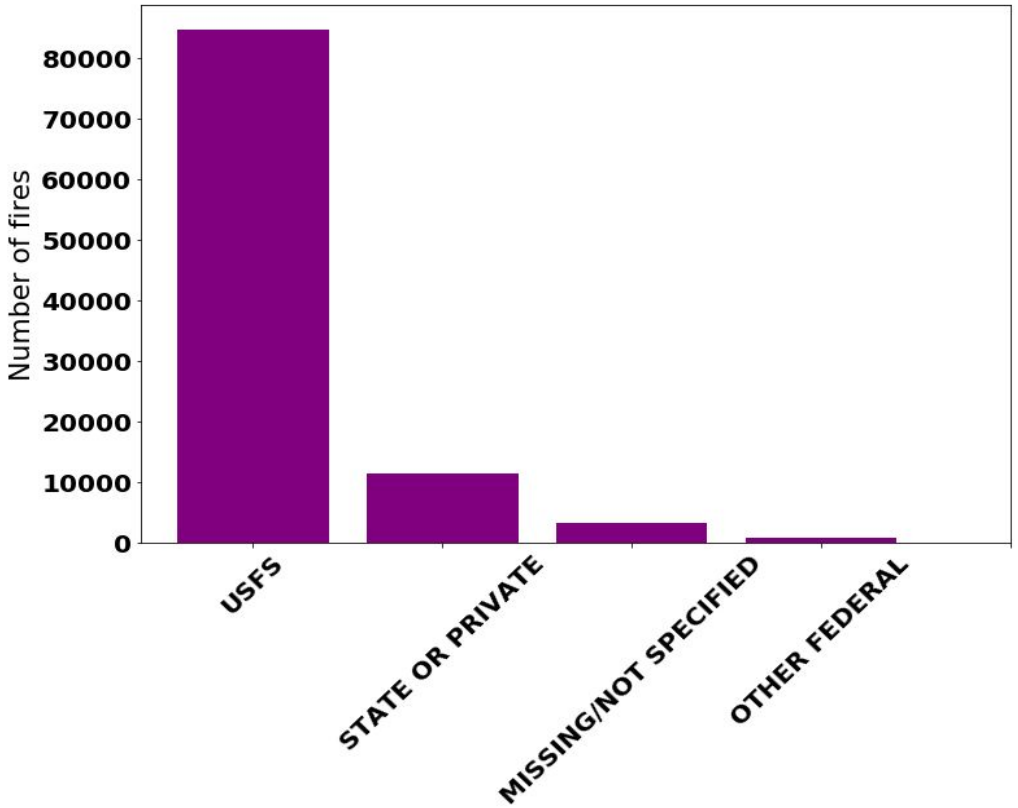
Libraries: pandas, sklearn, numpy, matplotlib
Python

States count of wildfires

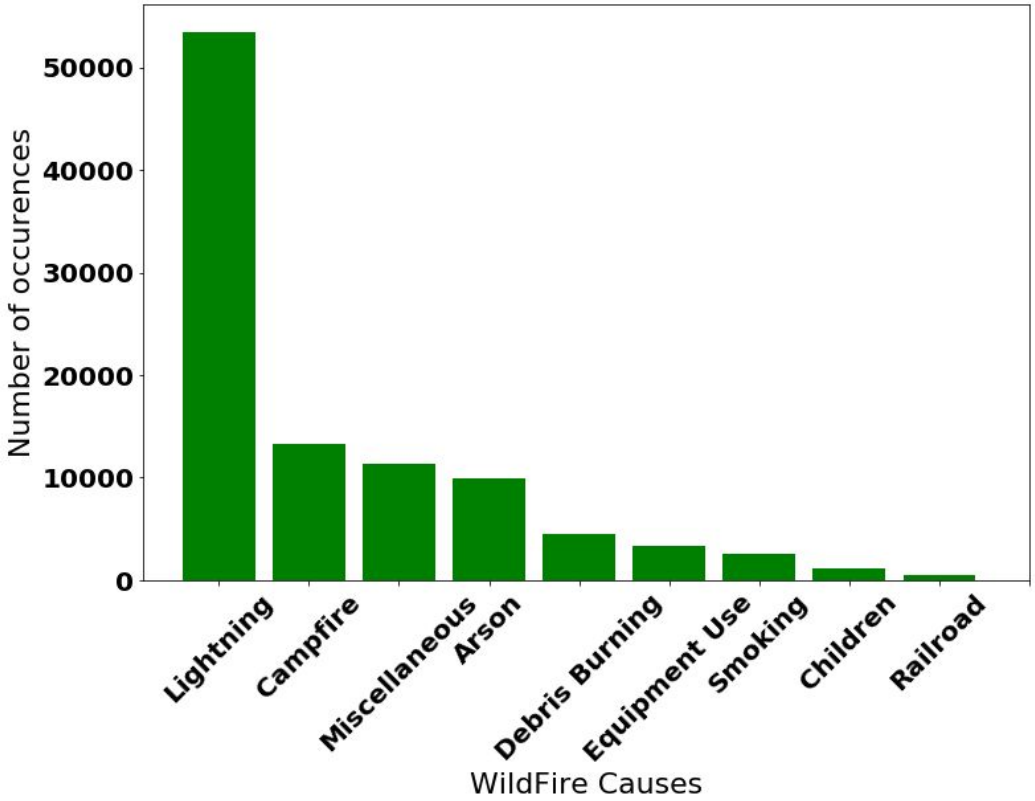


Owner of the land

USFS: United States Forest Service

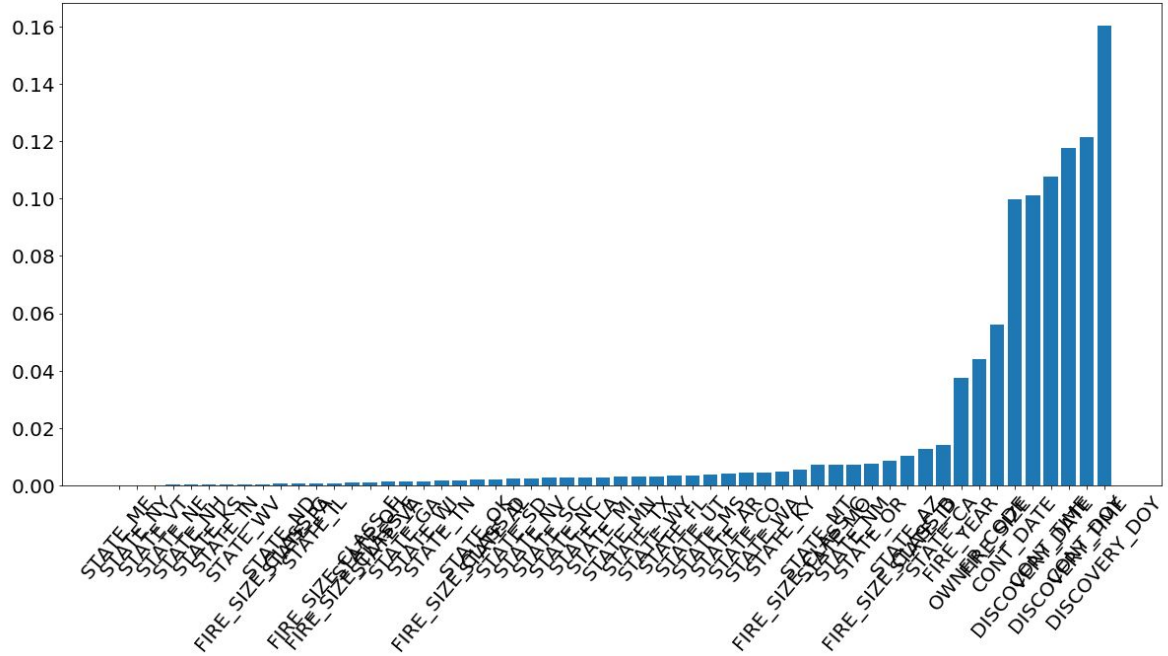


Causes of WildFire



Baseline Model

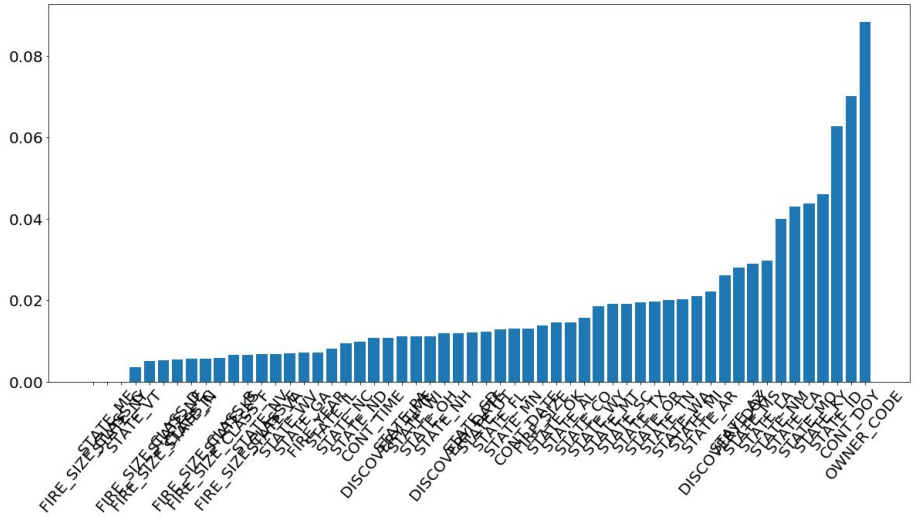
Random forest



XGBoost no params

Average scores

F1: .33~
Recall: .32~
Precision: .45~
Accuray: .68~



Next Steps

- Run a Nueral Network model to see if it can do a better job classifying
- Improve scores
- Do some feature engineering.
- Do Pca to reduce dimensionality

Thank You!

LinkedIn : Arturo Delgado

http://linkedin.com/in.arturo-delgado-ba94a1106