

Forest Fires in the U.S

By: Arturo Delgado



Thats a lot of damage.





The Problem



npr

Planet Money



Why care?



My data

Kaggle

Total 1.88 million

Sample of 100k

39 column

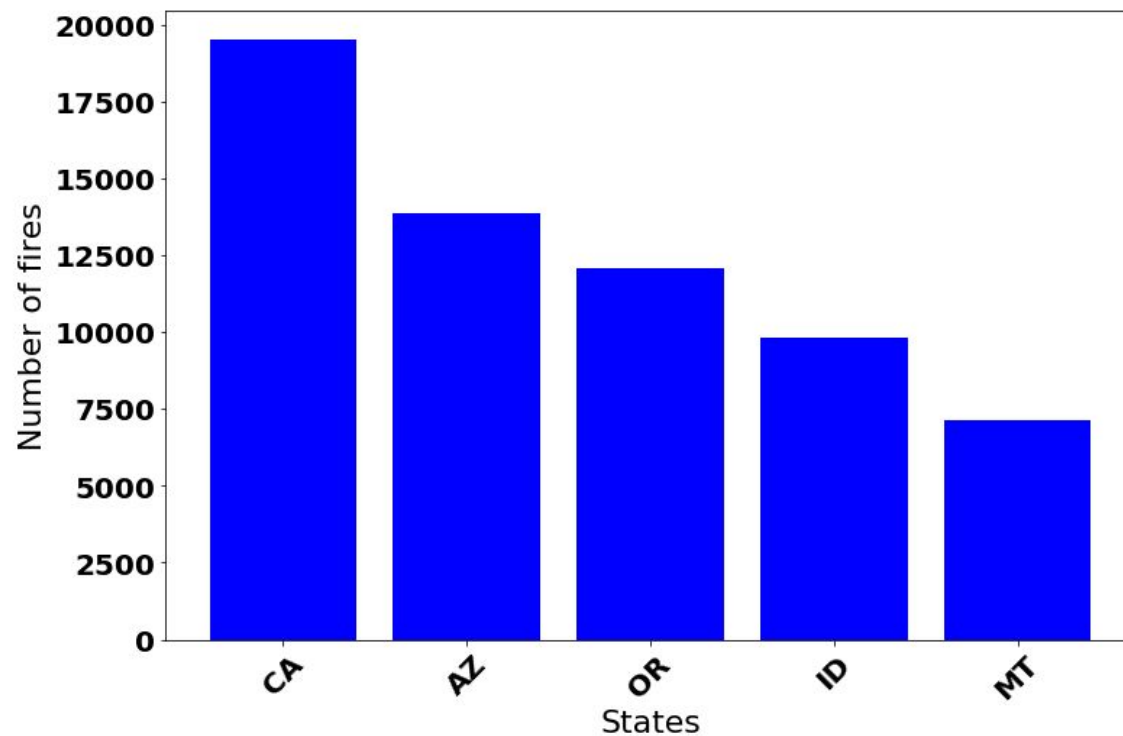
Collected: 1992-2015

Tools

Libraries: pandas, sklearn,
numpy, matplotlib
Python



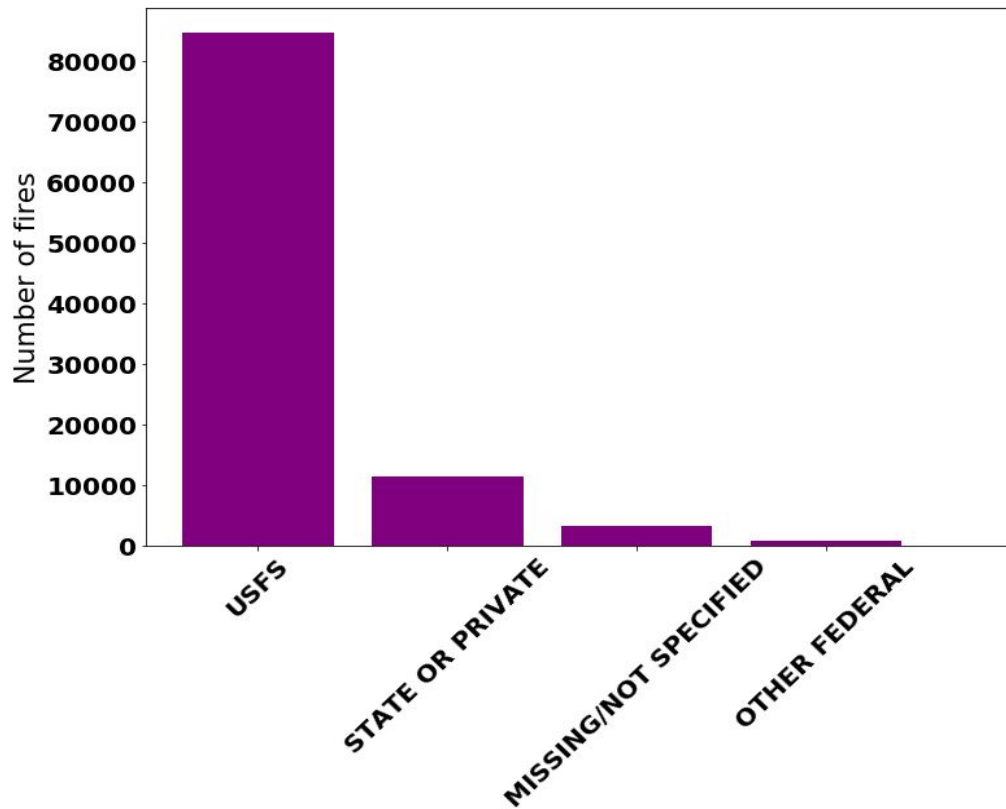
States count of wildfires



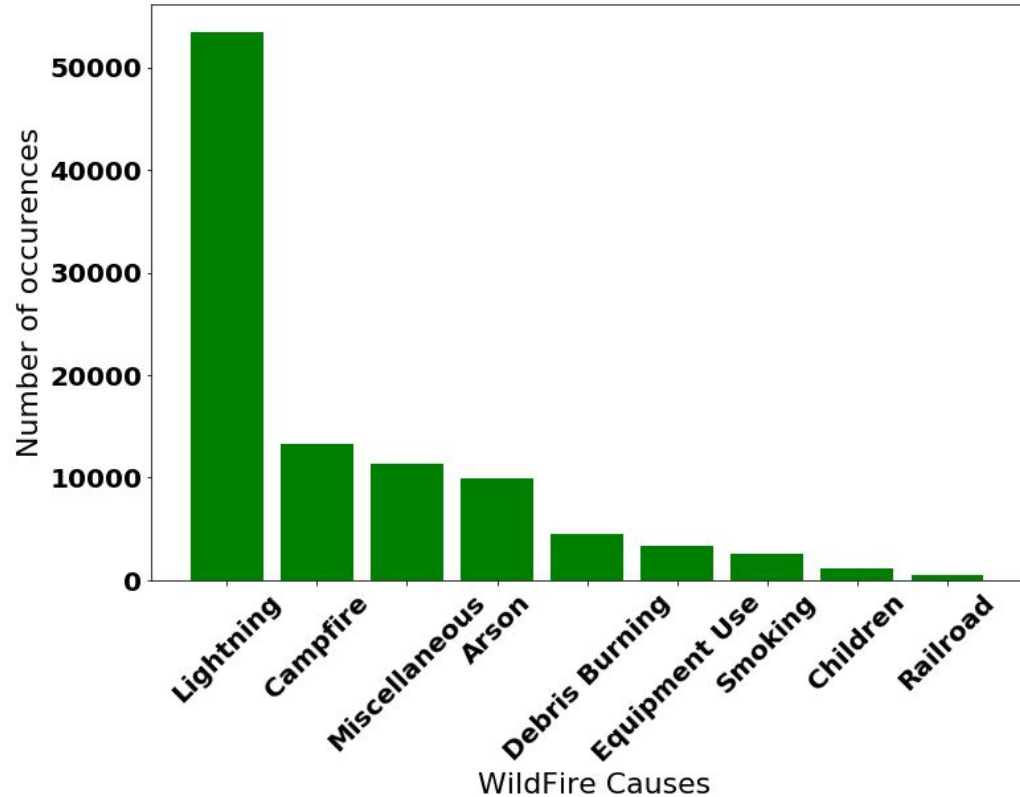


Owner of the land

USFS: United States Forest Service



Causes of WildFire





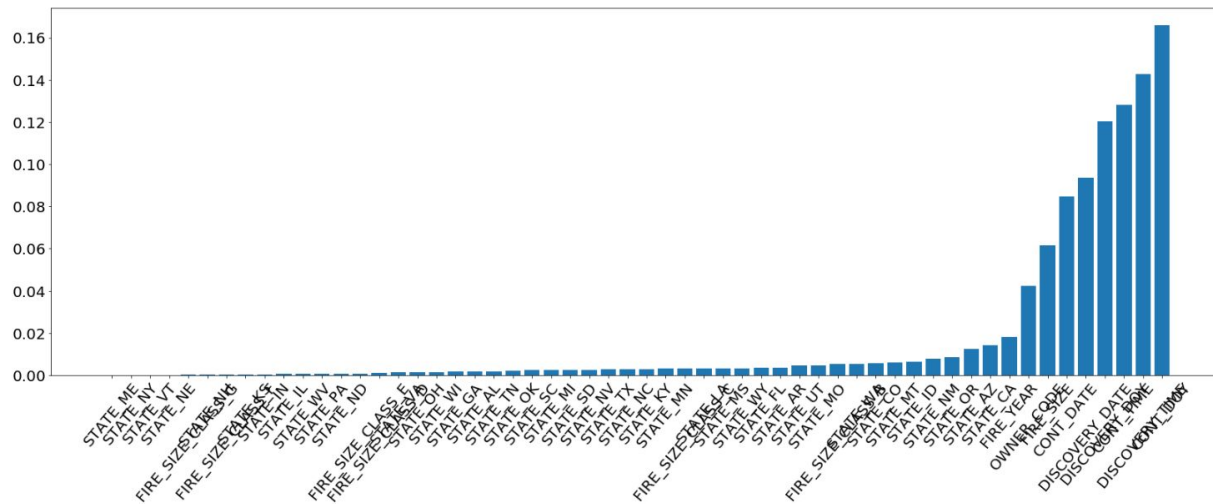
Decision Tree

Accuracy: 57.7%

Precision: 29.4%

Top 5 Feature Importances.

- Discovery time
- Cont Time
- Cont DOY
- Discovery DOY
- Cont Date





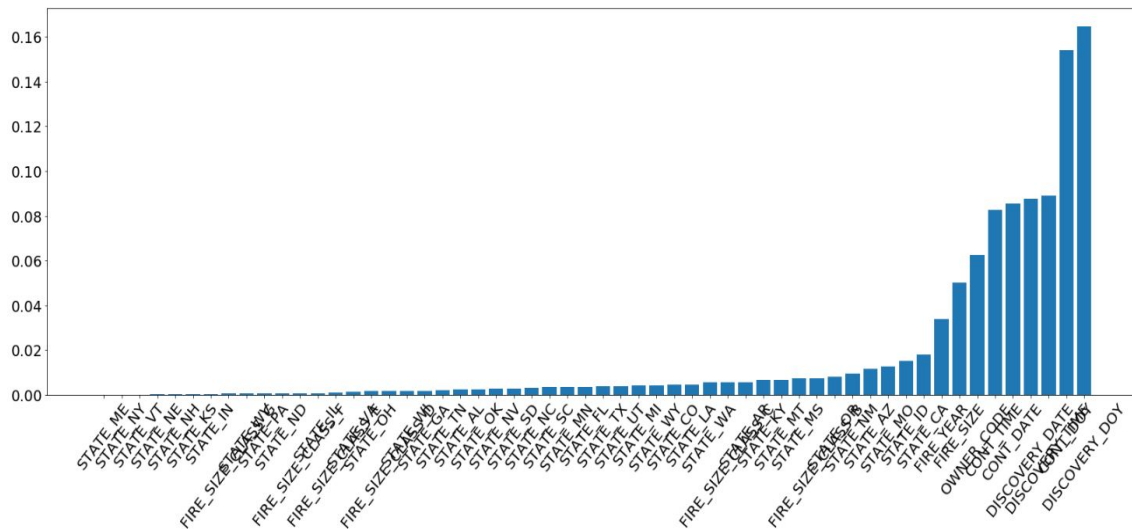
Random Forest hyperparameter tuned

Precision: .57~

- 28 percent increase

Top 5:

- Discovery DOY
- Cont DOY
- Discovery Time
- Discovery Date
- Cont Time



Next Steps

- Run a Neural Network model to see if it can do a better job classifying
- Build own grid search.
- Do some feature engineering.
- Do Pca to reduce dimensionality



Thank You!

LinkedIn : Arturo Delgado

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