

The background of the slide features a dramatic, high-contrast image of fire. On the left, bright orange and yellow flames are visible against a dark, almost black, background. To the right, wispy, orange-red plumes of smoke or embers drift across the frame. The overall effect is one of intense heat and danger.

FIRE RISK ANALYSIS

Smoke detection in satellite imagery with Xception.
Fire risk prediction on environmental conditions with
XGBoost.

ABOUT ME & MOTIVATION

CHELSEA ZALOUMIS

Data Science Student at Galvanize
BS in Applied Mathematics from Cal Poly

📍 Danville, CA
📍 Boulder, CO

- 2020 California fires are breaking records
 - Worst in 18 years
- Fires have damaged urban areas and uprooted tens of thousands from their homes.
- Smoke plumes billow across U.S. causing unhealthy air conditions.



FIRE RISK IN CALIFORNIA



Satellite images of northern and southern California.

Conditions based off regions of state.



DATA SOURCES



SATELLITE IMAGES

CONDITIONS DATA



NASA



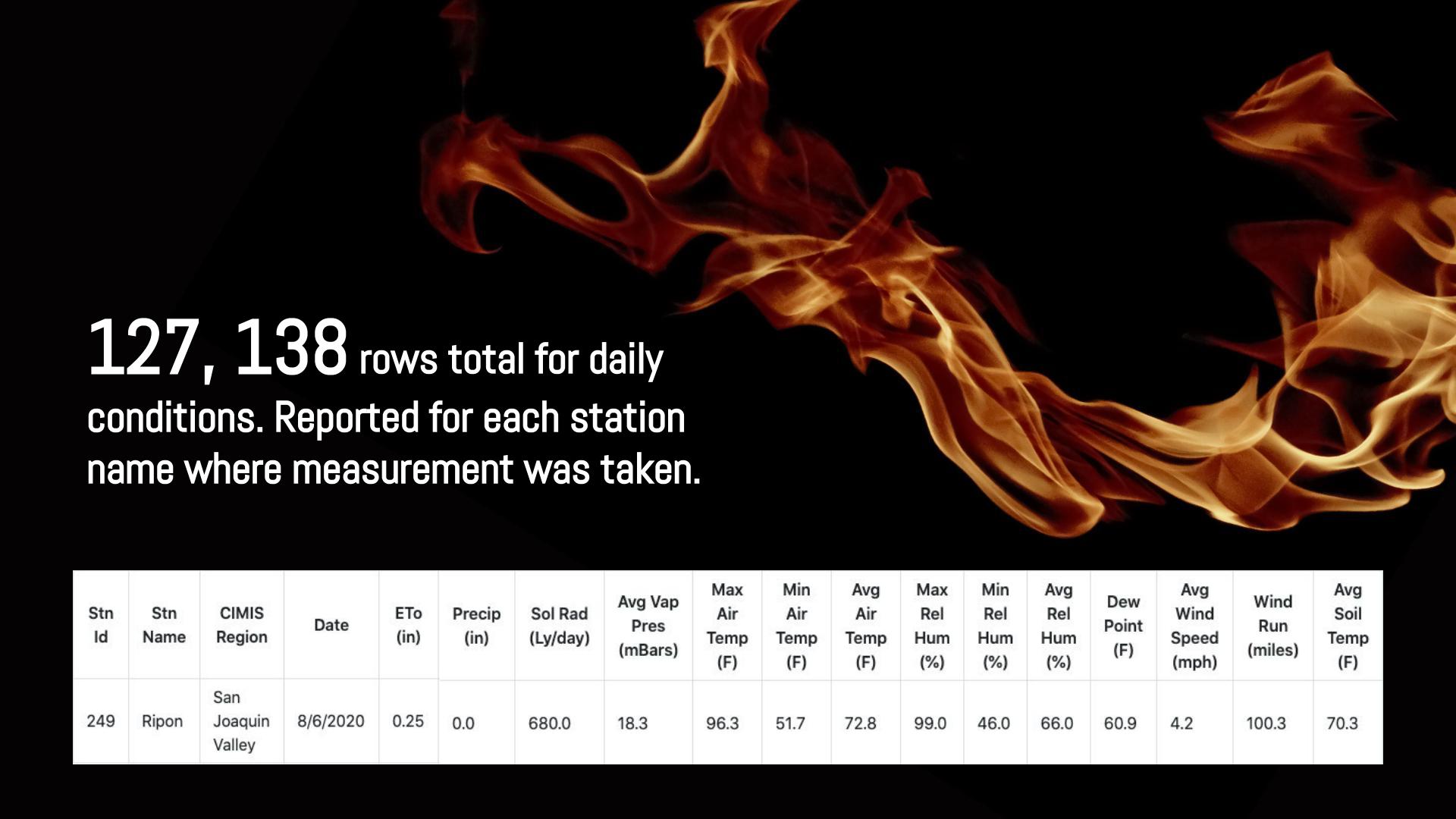
1,990 Images Total

SMOKE



FOG





127, 138 rows total for daily conditions. Reported for each station name where measurement was taken.

Stn Id	Stn Name	CIMIS Region	Date	ETo (in)	Precip (in)	Sol Rad (Ly/day)	Avg Vap Pres (mBars)	Max Air Temp (F)	Min Air Temp (F)	Avg Air Temp (F)	Max Rel Hum (%)	Min Rel Hum (%)	Avg Rel Hum (%)	Dew Point (F)	Avg Wind Speed (mph)	Wind Run (miles)	Avg Soil Temp (F)
249	Ripon	San Joaquin Valley	8/6/2020	0.25	0.0	680.0	18.3	96.3	51.7	72.8	99.0	46.0	66.0	60.9	4.2	100.3	70.3

HIGH OVERVIEW

I. DETECT SMOKE IN IMAGES



Transfer learning with a
pre-trained
convolutional neural
network, Xception.



HIGH OVERVIEW

I. DETECT SMOKE IN IMAGES



Transfer learning with a pre-trained convolutional neural network, Xception.

II. ASSESS FIRE RISK FROM CONDITIONS



XGBoost Classifier for fire risk.



HIGH OVERVIEW

I. DETECT SMOKE IN IMAGES



Transfer learning with a pre-trained convolutional neural network, Xception.

II. ASSESS FIRE RISK FROM CONDITIONS



XGBoost Classifier for fire risk.

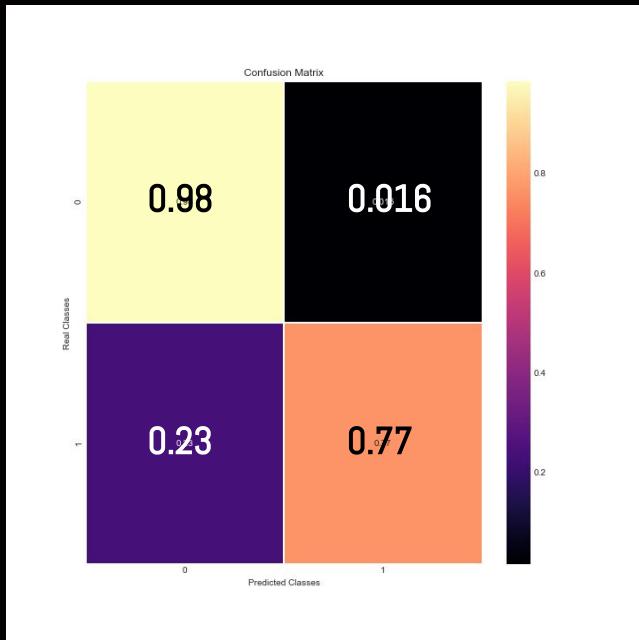
III. COMBINE MODELS & BUILD APP



Weighted model probabilities and deployed on Flask APP.

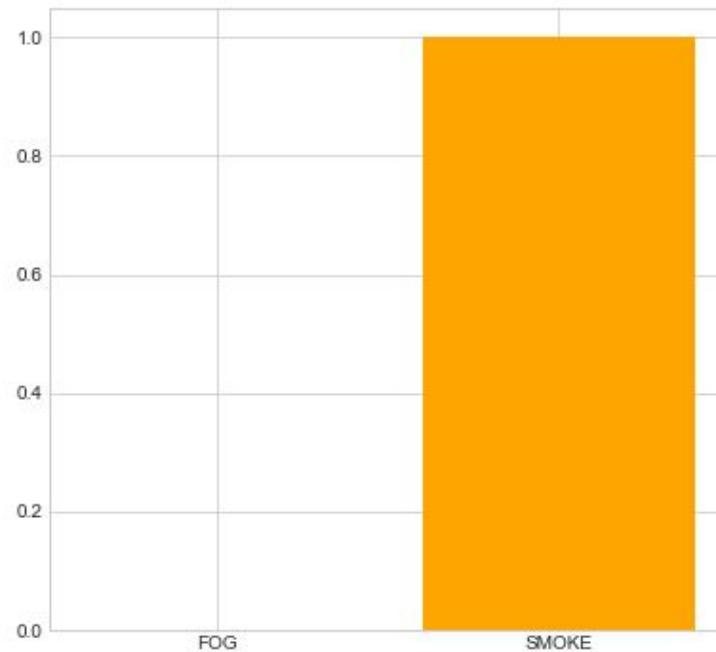
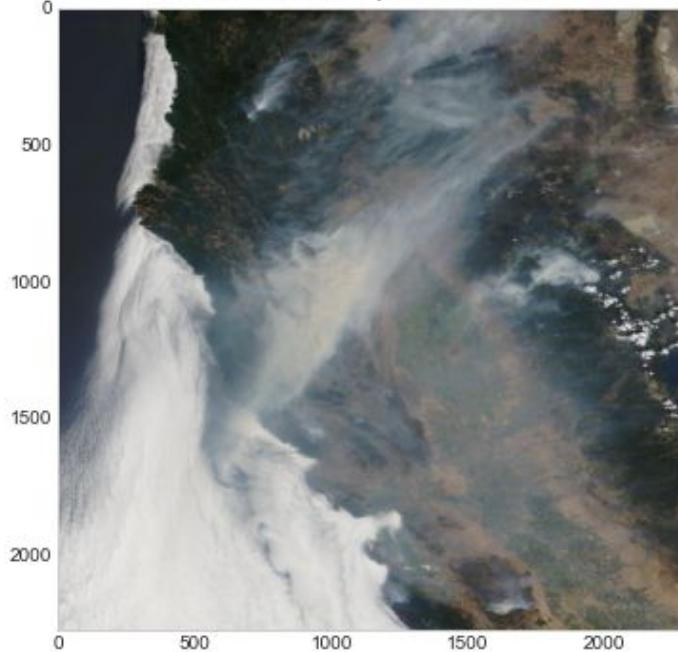


XCEPTION PERFORMANCE

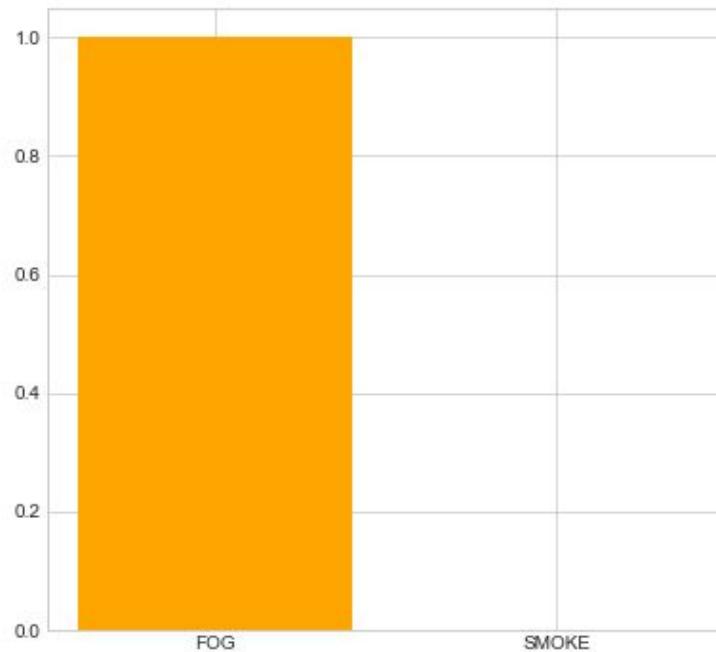
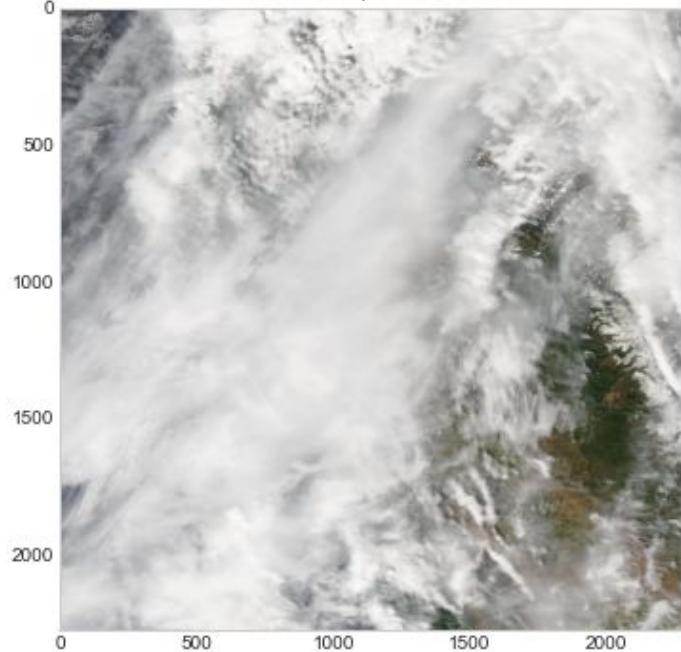


- Weighted imbalanced dataset.
 - Unfroze 4 layers for 10 epochs at a time.
 - Trained a total of 74 layers.
-
- **95% Accuracy**
 - **76% Recall**
 - **89% Precision**

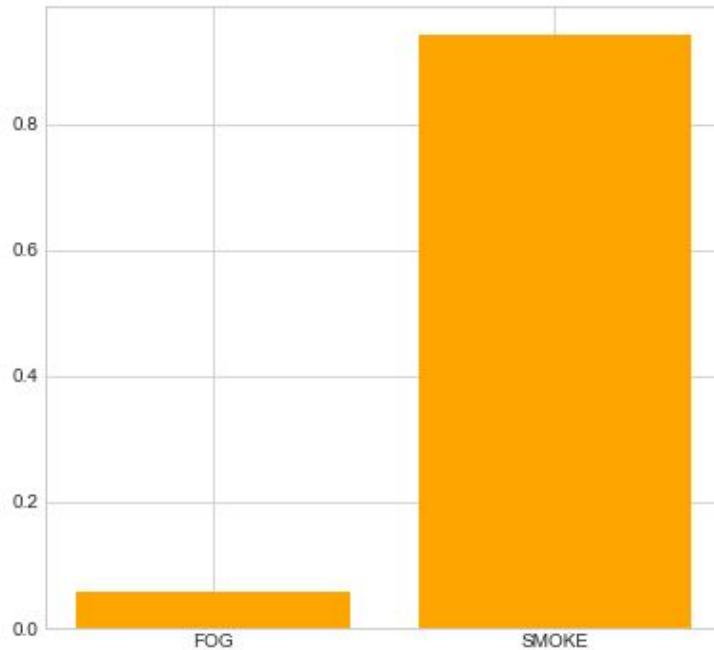
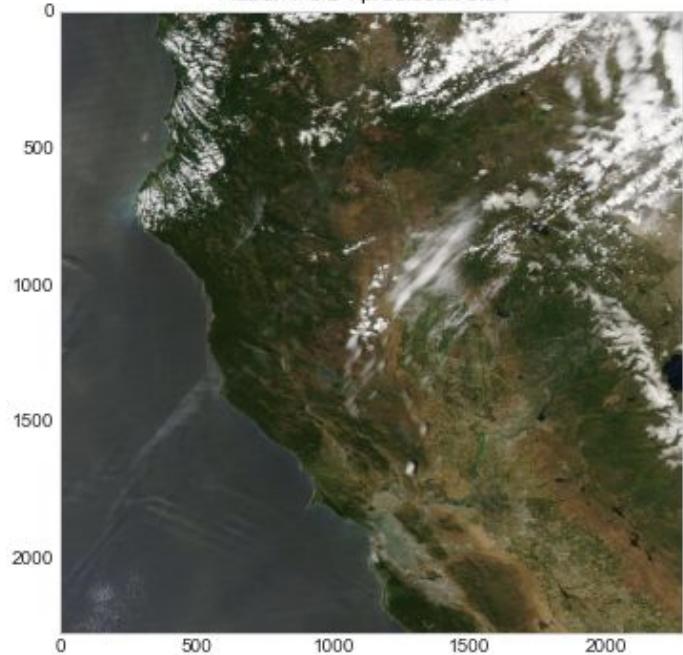
label SMOKE : prediction 1.00



label FOG : prediction 0.00

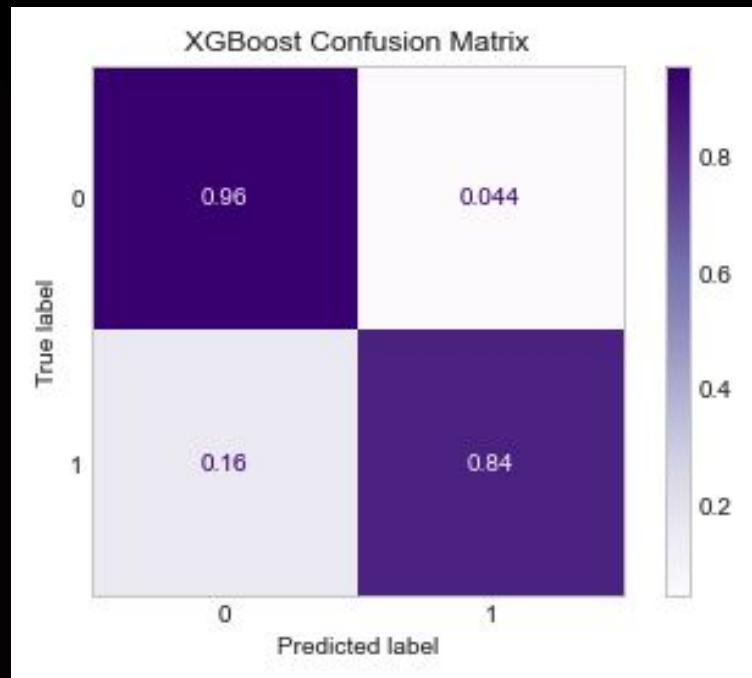


label FOG : prediction 0.94

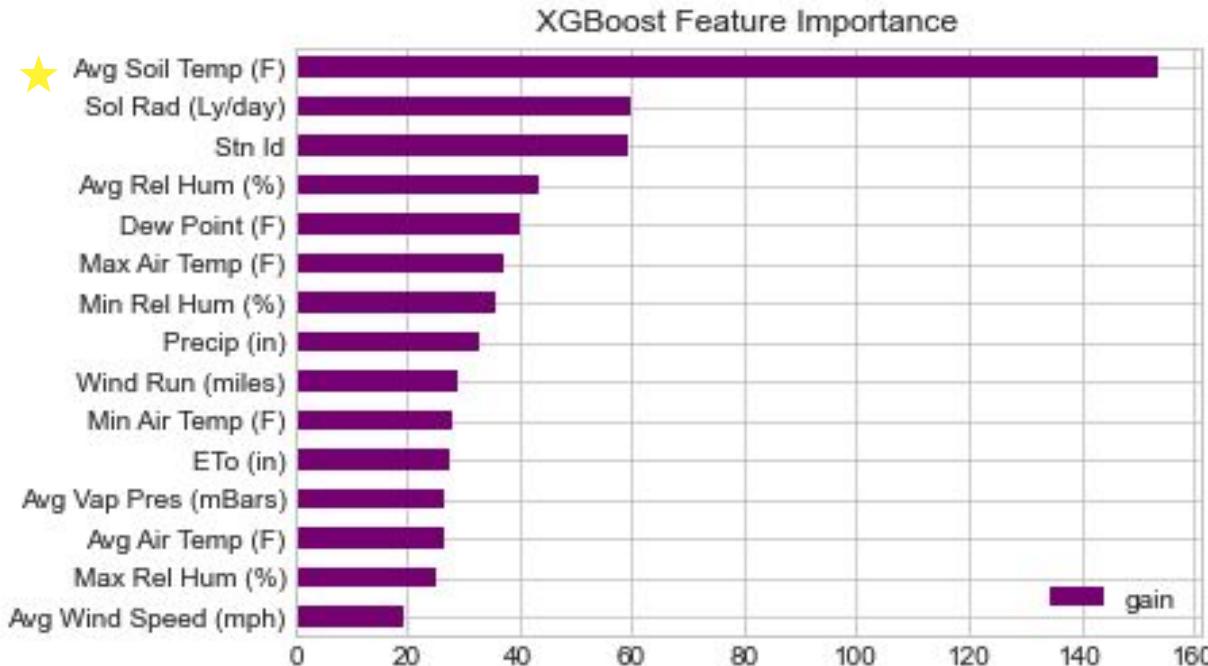


XGBOOST PERFORMANCE

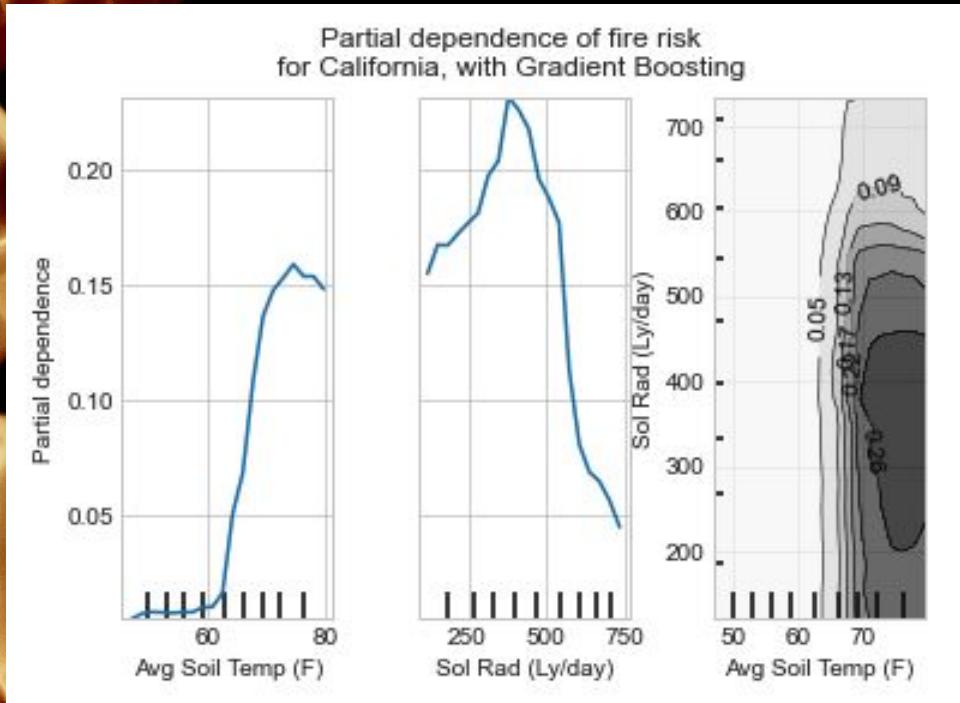
- Weighted imbalanced dataset.
- Hypertuned with gridsearch 9+ hours.
 - **95% Accuracy**
 - **78% Recall**
 - **42% Precision**



FEATURE IMPORTANCE FOR XGBOOST CLASSIFICATION



PARTIAL DEPENDENCE PLOTS



- The hotter, the greater risk for fire (until ~75 degrees)
- Solar radiation is linearly related to fire risk up until a peak of ~450 langley/day at which the risk begins decreasing drastically.
- For Avg Soil Temp < ~62 degrees, fire risk is independent of Solar Radiation.
- For Avg Soil Temp > 62, there is a strong dependence on solar radiation.

DETAILED STEPS



OBTAI
N DATA



DETAILED STEPS

Accurately
identify smoke in
images.

XCEPTION

OBTAIN DATA



DETAILED STEPS

XGBOOST

Assess risk for fire based on conditions.

XCEPTION

Accurately identify smoke in images.

OBTAIN DATA



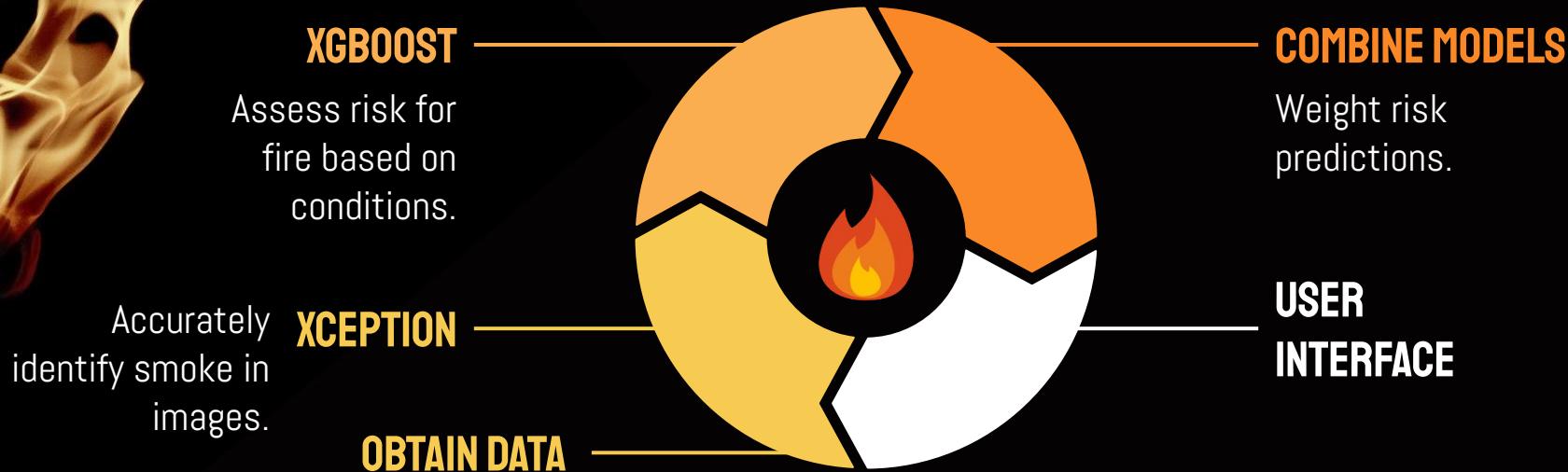


DETAILED STEPS





DETAILED STEPS





Fire Risk Analysis

Northern and Southern California

1/1/2018 - 9/13/2020

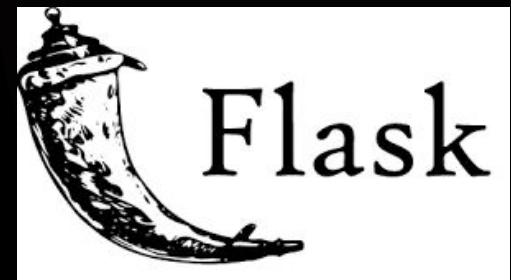
Select a region*

Enter a date between 1/1/2018 and 9/13/2020:

mm / dd / yyyy



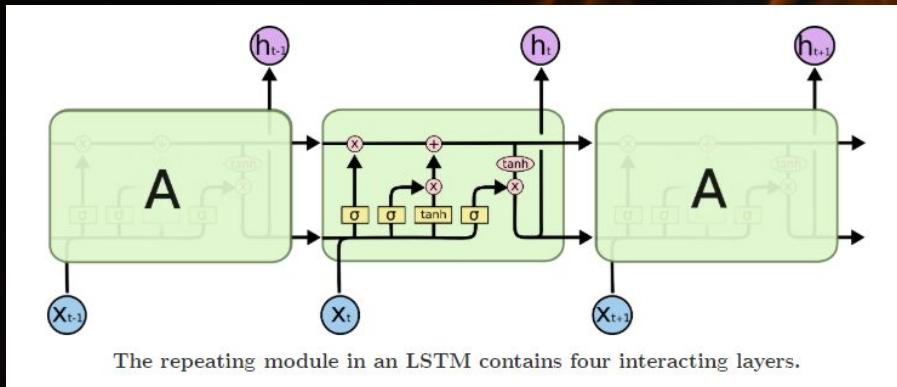
IN APPLICATION



FUTURE STEPS

- **LONG SHORT-TERM MEMORY MODEL**

Incorporate memory for conditions to predict CURRENT and future fire risk.



THANK YOU!

CHELSEA ZALOUMIS

Data Scientist

<https://github.com/czaloumi>

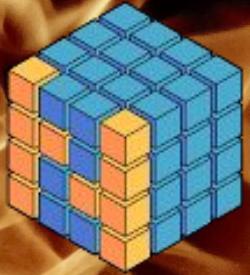


CREDITS: This presentation template was created by Slidesgo, including icons by Flaticon, and infographics & images by Freepik.



APPENDIX





NumPy

TECHNOLOGIES USED



127, 138 rows total for daily conditions. Reported for each station name where measurement was taken.

- More work to be done:
 - Drop max and min features since there are averages
 - Standardize feature values

Stn Id				ETo (in)	Precip (in)	Sol Rad (Ly/day)	Avg Vap Pres (mBars)	Max Air Temp (F)	Min Air Temp (F)	Avg Air Temp (F)	Max Rel Hum (%)	Min Rel Hum (%)	Avg Rel Hum (%)	Dew Point (F)	Avg Wind Speed (mph)	Avg Wind Run (miles)	Avg Soil Temp (F)
249	Ripon	San Joaquin Valley	8/6/2020	0.25	0.0	680.0	18.3	96.3	51.7	72.8	99.0	46.0	66.0	60.9	4.2	100.3	70.3

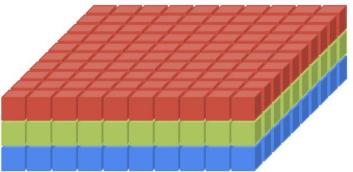


XCEPTION IDENTIFYING SMOKE IN SATELLITE IMAGES

TRADITIONALLY, FILTER SIMULTANEOUSLY CONSIDERS A SPATIAL DIMENSION (EACH 4X4 SQUARE) AND A CROSS-CHANNEL OR “DEPTH” DIMENSION (THE STACK OF THREE SQUARES). AT THE INPUT LAYER OF AN IMAGE, THIS IS EQUIVALENT TO A CONVOLUTIONAL FILTER LOOKING AT A 4X4 PATCH OF PIXELS ACROSS ALL THREE RGB CHANNELS.

EXTREME INCEPTION MAPS THE SPATIAL CORRELATIONS FOR *EACH OUTPUT CHANNEL SEPARATELY*, AND THEN PERFORMS A ONE X ONE DEPTHWISE CONVOLUTION TO CAPTURE CROSS-CHANNEL CORRELATION.

gif from codelabs.developers.google.com



$W_1[4, 4, 3]$

filter
size nb of filters
= input channels

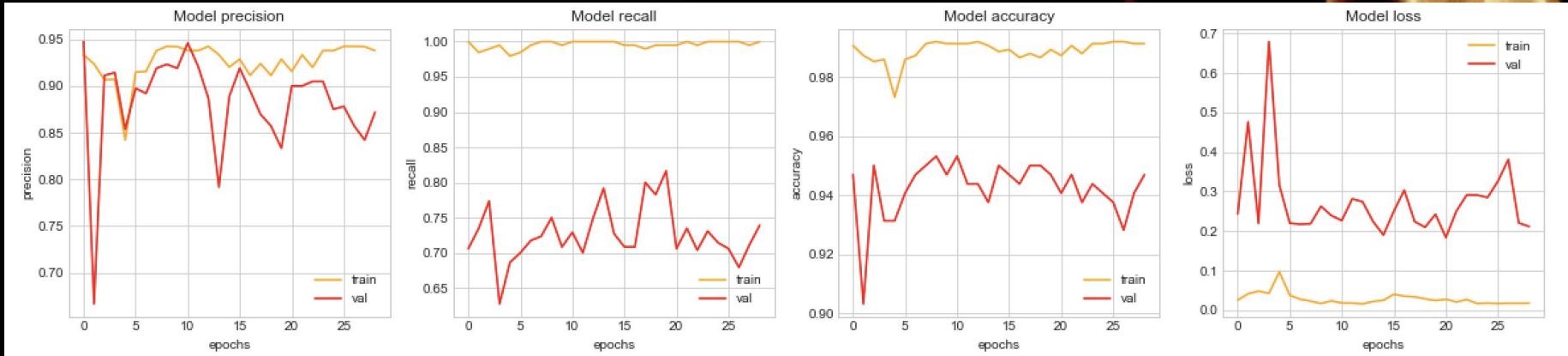
$W_2[3, 5]$

output channels

XCEPTION PERFORMANCE

Xception is pre-trained and has 126 layers with roughly 23 million parameters. Weighted imbalanced dataset. Used ImageNet weights.

- Unfroze 4 layers for 10 epochs at a time for a total of 74 trained layers.
- **95% Accuracy, 76% Recall, 89% Precision**



EXTREME GRADIENT BOOSTING

FIRE RISK BASED ON DAILY CONDITIONS

- Based off decision trees where “branches” define rules for making decision and split feature values along that decision.
- The next tree evaluates how correct tree prior was and aims to reduce error.
- Feature Importance: Gain implies the relative contribution of the corresponding feature to the model calculated by taking each feature's contribution for each tree in the model. A higher value of this metric when compared to another feature implies it is more important for generating a prediction.
- **95% Accuracy, 78% Recall, 42% Precision**

