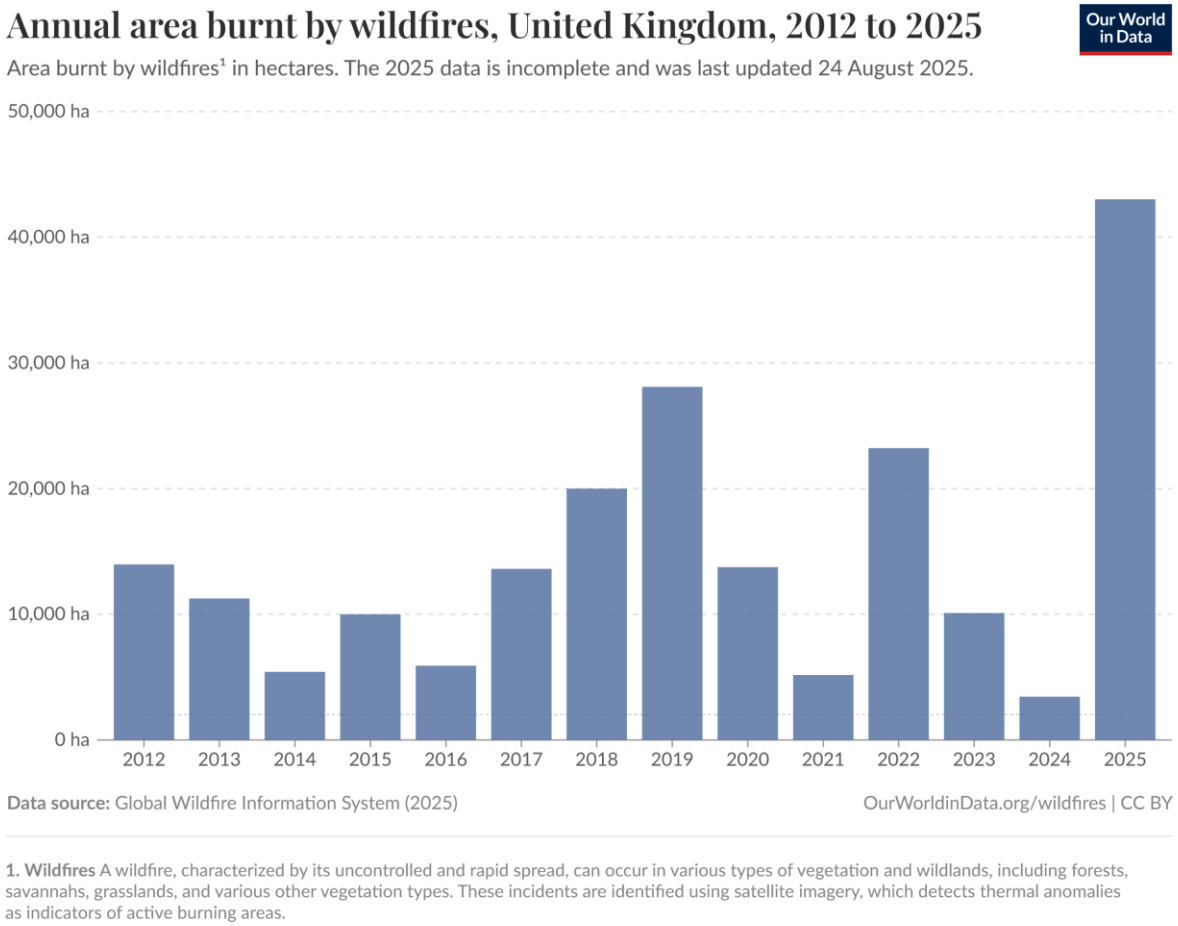
MACHINE LEARNING FOR PREDICTION OF FORREST FIRES

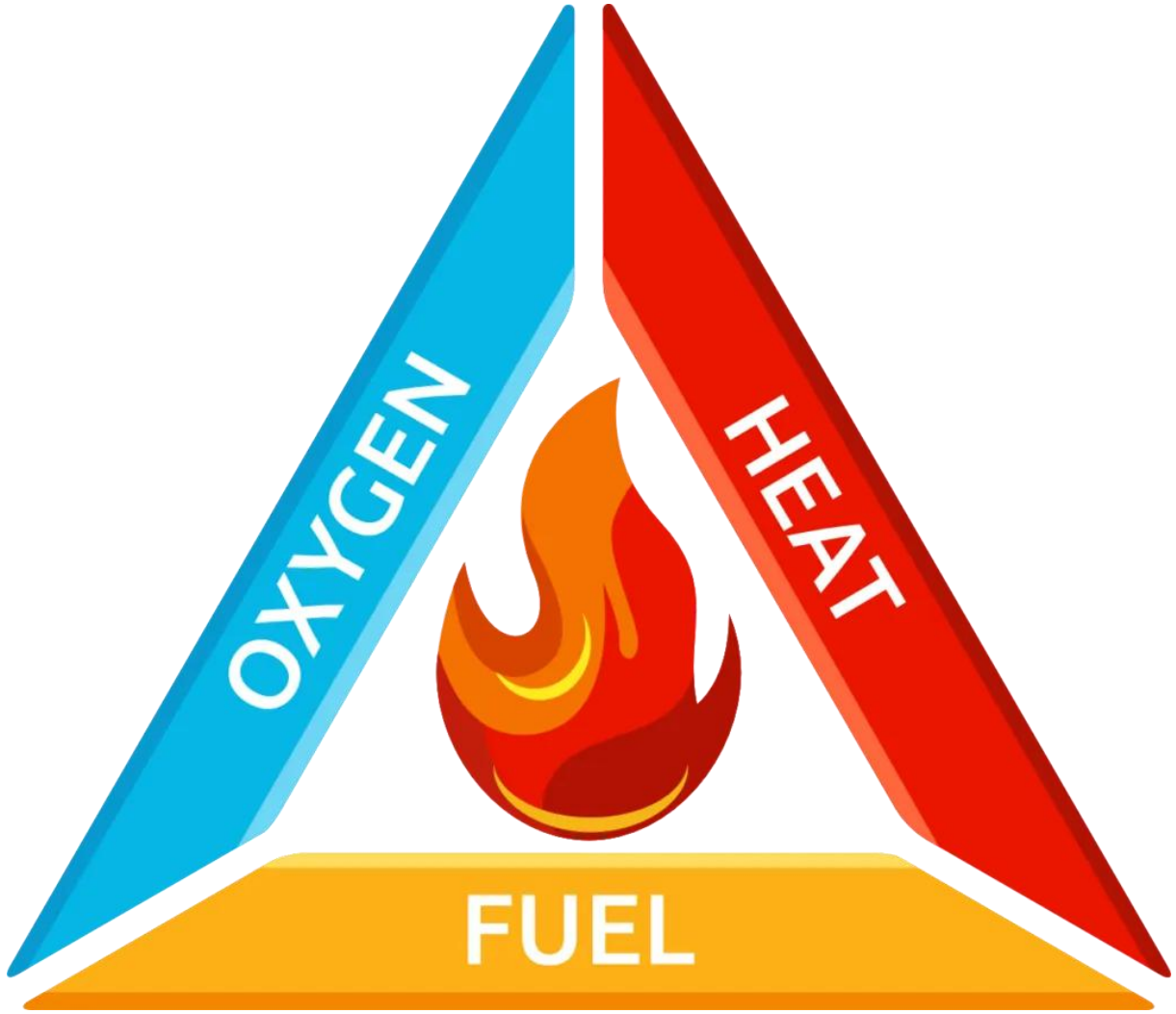
Michael Lin

WHY MODEL FIRES?

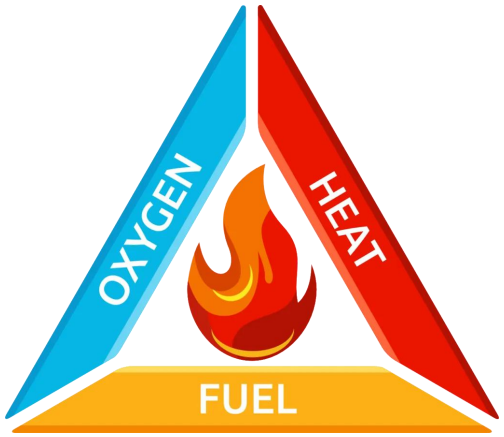
- Forest fires increasing in frequency
- Models help to mitigate effects
- Various contributing factors
  - eg Weather Whiplash



THE FIRE TRIANGLE



THE FIRE TRIANGLE



Fuel

- Amount available
  - Biomass: trees, shrubbery
  - Quantified by **Leaf Area Index**
- Flammability
  - Water content of plants

MOTIVATION FOR THE PROJECT

Geophysical Research Letters
RESEARCH LETTER
10.1029/2023GL107929

A Global Probability-Of-Fire (PoF) Forecast

J. R. McNorton[1], F. Di Giuseppe[1], E. Pinnington[1], M. Chantry[1], and C. Barnard[1]

Tree-based models (Random Forest / XGBoost)

Input Features

- daily mean precipitation
- 10m wind speed
- 2m dew point temperature
- 2m temperature
- *Fuel Characteristics*
  - Fuel load
  - Live Fuel Moisture Content (LFMC)
  - Dead Fuel Moisture Content (DFMC)

McNorton JR, Di Giuseppe F, Pinnington E, Chantry M, Barnard C. A Global Probability-of-Fire (PoF) Forecast. Geophys. Res. Lett. 2024

MOTIVATION FOR THE PROJECT

A global fuel characteristic model and dataset for wildfire prediction

Joe R. McNorton and Francesca Di Giuseppe

- Vegetation cover
  - Leaf Area Index
- Vegetation type
- *Soil moisture at 4 different levels*
- Surface pressure
- Skin temperature

McNorton JR & Di Giuseppe F. A global fuel characteristic model and dataset for wildfire prediction. Biogeosciences 2023

ANTECEDENT PROXY VARIABLES

Streamflow prediction using artificial neural networks and soil moisture proxies

Robert Edwin Rouse[1], Doran Khamis[2], Scott Hosking[3,4], Allan McRobie[1] and Emily Shuckburgh[5]

Proxy Variables

- Temperature
- Precipitation

Rolling Averages over…

- 30 days
- 90 days
- 180 days

Rouse RE, Khamis D, Hosking S, McRobie A, Shuckburgh E. Streamflow prediction using artificial neural networks and soil moisture proxies. *Environmental Data Science*. 2025

MACHINE LEARNING??

WHAT'S THAT?

MACHINE LEARNING: THE BASICS

What is it?

A computer program is said to **learn** from experience *E* with respect to some class of tasks *T* and performance measure *P*, if its performance at tasks in *T*, as measured by *P*, improves with experience *E*.

Aim

- Extremise an objective function

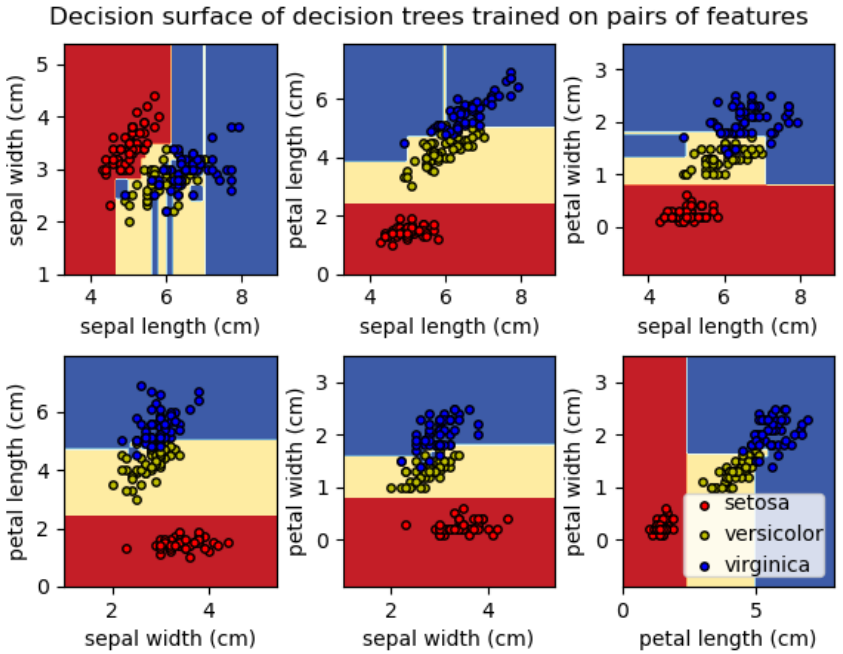
Example: Gini Coefficient

$$Gini(p) = 1 - \sum_{i=1}^{n} p_i^2$$

- Useful for classification problems
  - Ranges from 0 (pure) to 0.5 (completely mixed)

MACHINE LEARNING: DECISION TREES

- ‘like a flowchart’
- Grown using CART algorithm
  - Splits at each node to minimise Gini coefficient
- Fitting hyperplanes in n-dimensional feature space



https://sklearn.org/stable/modules/tree.html

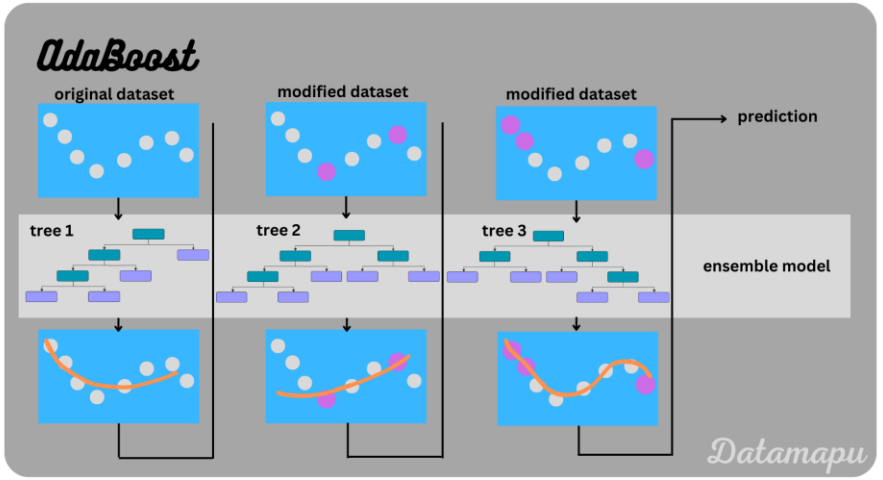
MACHINE LEARNING: ENSEMBLE METHODS

Random Forests

- Parallel
- Bootstrapped data
- Average result

Gradient Boosting

- Series
- Correcting loss fn
- Sum of results



THE MODEL

THE MODEL

- Data from Richardson Backcountry

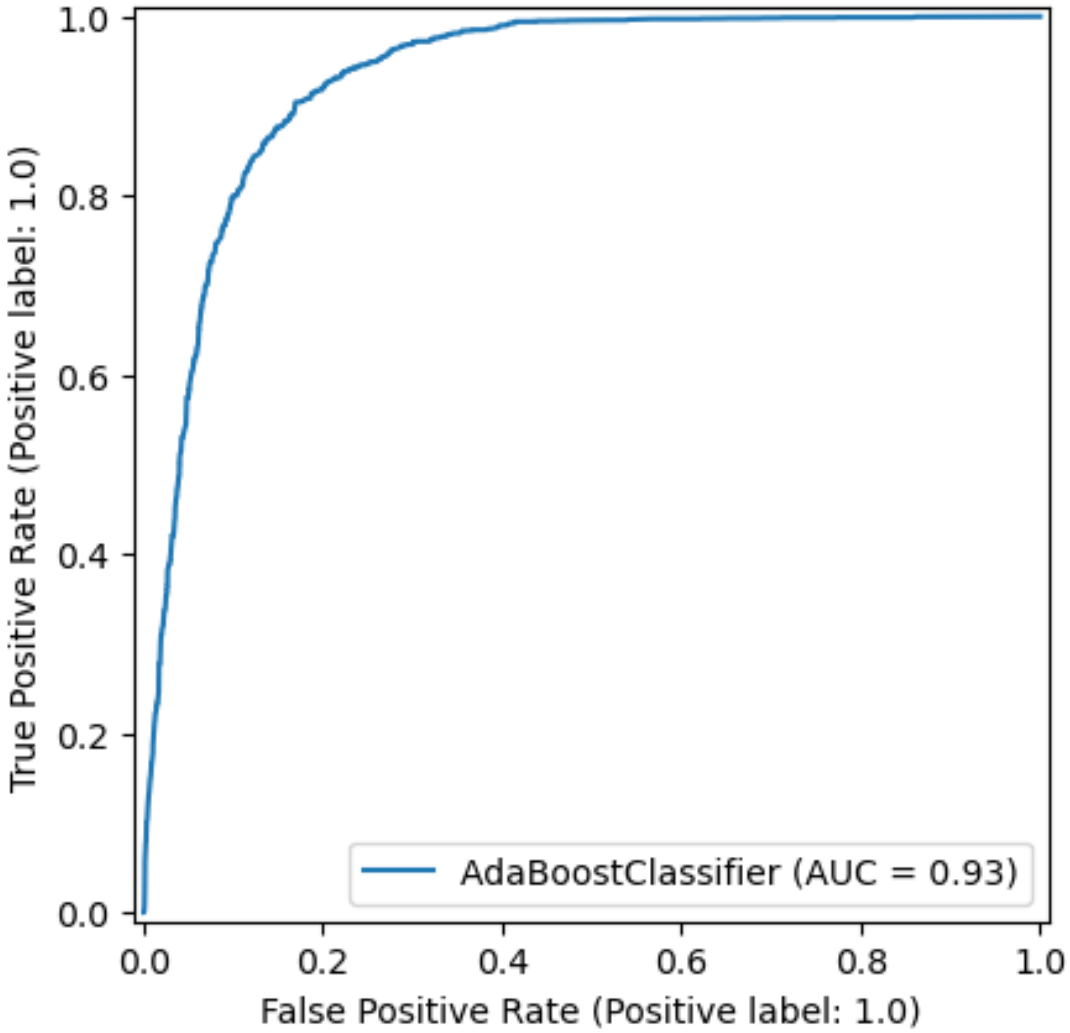
- Trained on 2010-2014
- Tested on 2015-2018

- Fire data from MODIS
  - >70% confidence

PERFORMANCE

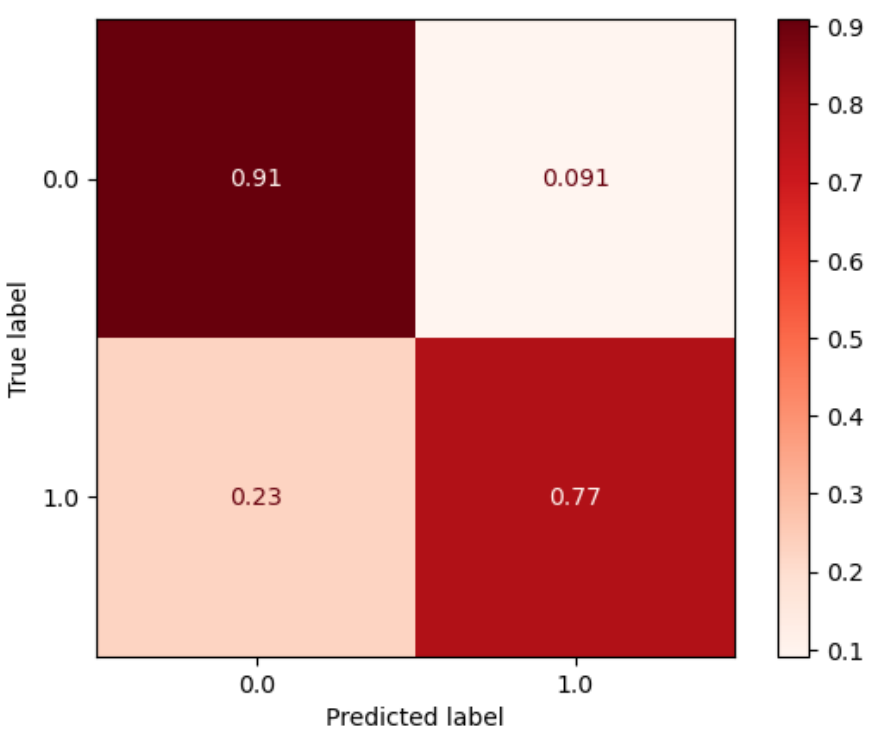
- AUC score of 0.93

	precision	recall	f1-score	support
0.0	1.00	0.91	0.95	58904573
1.0	0.00	0.77	0.00	2947
accuracy			0.91	58907520
macro avg	0.50	0.84	0.48	58907520
weighted avg	1.00	0.91	0.95	58907520

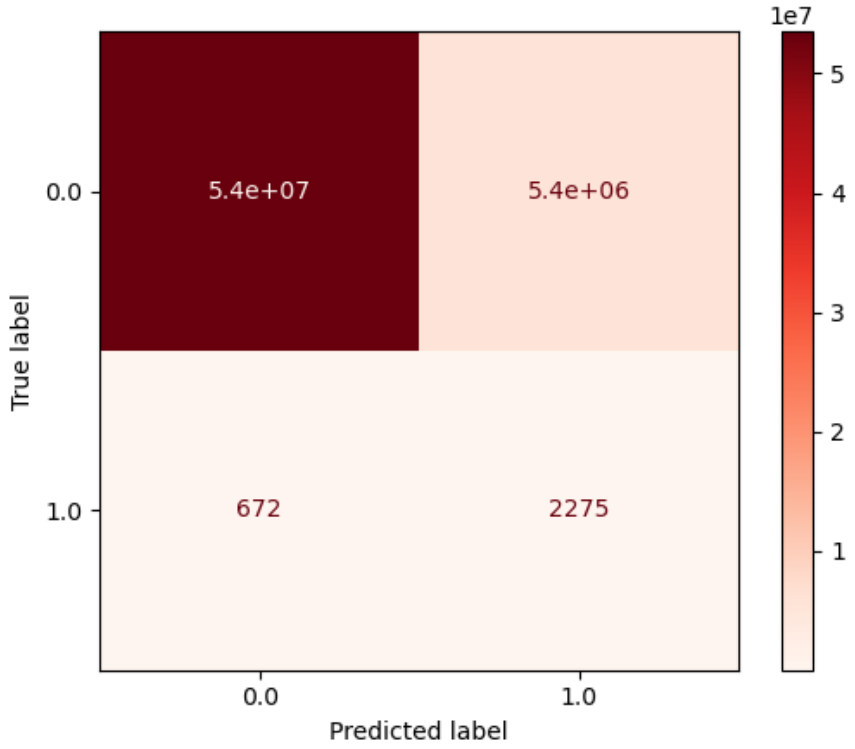


CONFUSION MATRIX

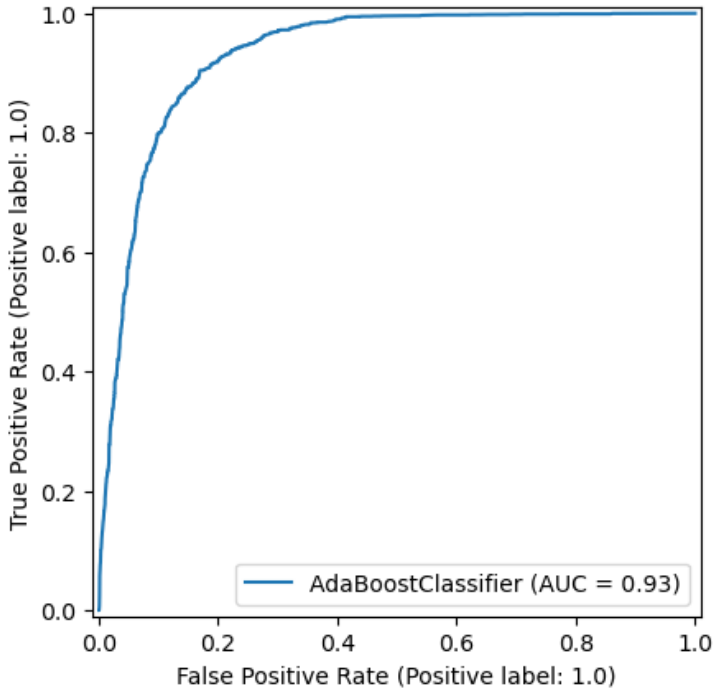
NORMALISED



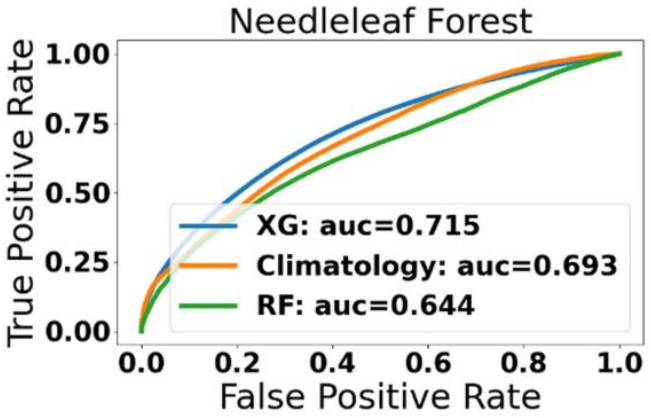
ORIGINAL

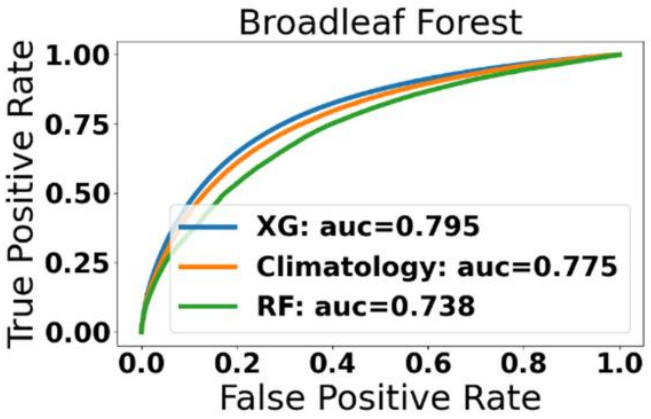


COMPARISON



RESULTS FROM PAPER





WHAT'S NEXT

- Daily data (for direct comparison)
- Generalise model
- Different / Additional proxies
  - Shorter-term proxies
  - Longer-term proxies
  - Stepped averages

THANKS FOR LISTENING

EFFECT OF DOWNSAMPLING

Results on original data

	precision	recall	f1-score	support
0.0	1.00	0.91	0.95	58904573
1.0	0.00	0.77	0.00	2947
accuracy			0.91	58907520
macro avg	0.50	0.84	0.48	58907520
weighted avg	1.00	0.91	0.95	58907520

Results on down-sampled data

	precision	recall	f1-score	support
0.0	0.98	0.91	0.94	29470
1.0	0.46	0.77	0.58	2947
accuracy			0.90	32417
macro avg	0.72	0.84	0.76	32417
weighted avg	0.93	0.90	0.91	32417

DATA VARIABLES

Variable Name	Short Name
2m Dewpoint Temperature	d2m
Leaf Area Index	lai_hv
	lai_lv
Proxy Variables	mu_tp_180
	mu_t2m_180
	mu_t2m_30
	mu_t2m_90
	mu_tp_90
	mu_tp_30
Surface Pressure	sp
2m Temperature	t2m
Total Precipitation	tp
10m Total Wind Speed	ws10