

Predictive Model for Forest Fire Preparedness

Determining the likelihood of forest fires by using weather projections



Why this project?

- Forest fires are becoming more commonplace and dangerous
- Weather is getting warmer by the year
- Anticipating forest fires may help communities be better equipped to handle the risks.



Where my data came
from:

Canada

Environment and natural resources

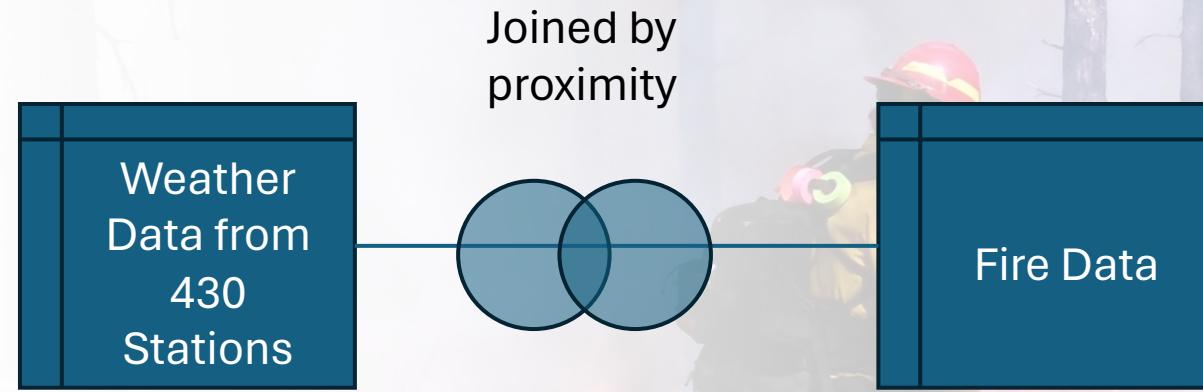
Forests and forestry



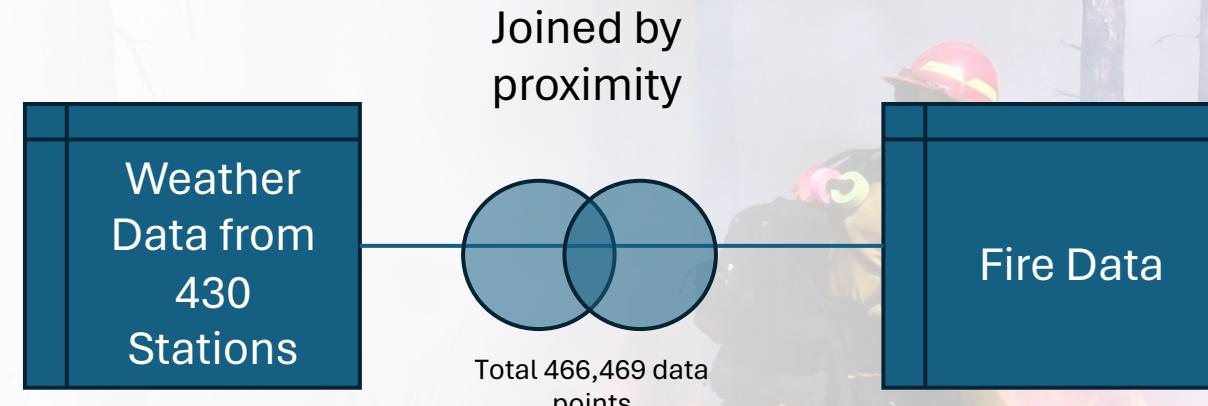
Merging DataFrames



Merging DataFrames



Merging DataFrames



Features used to predict whether a fire will occur:

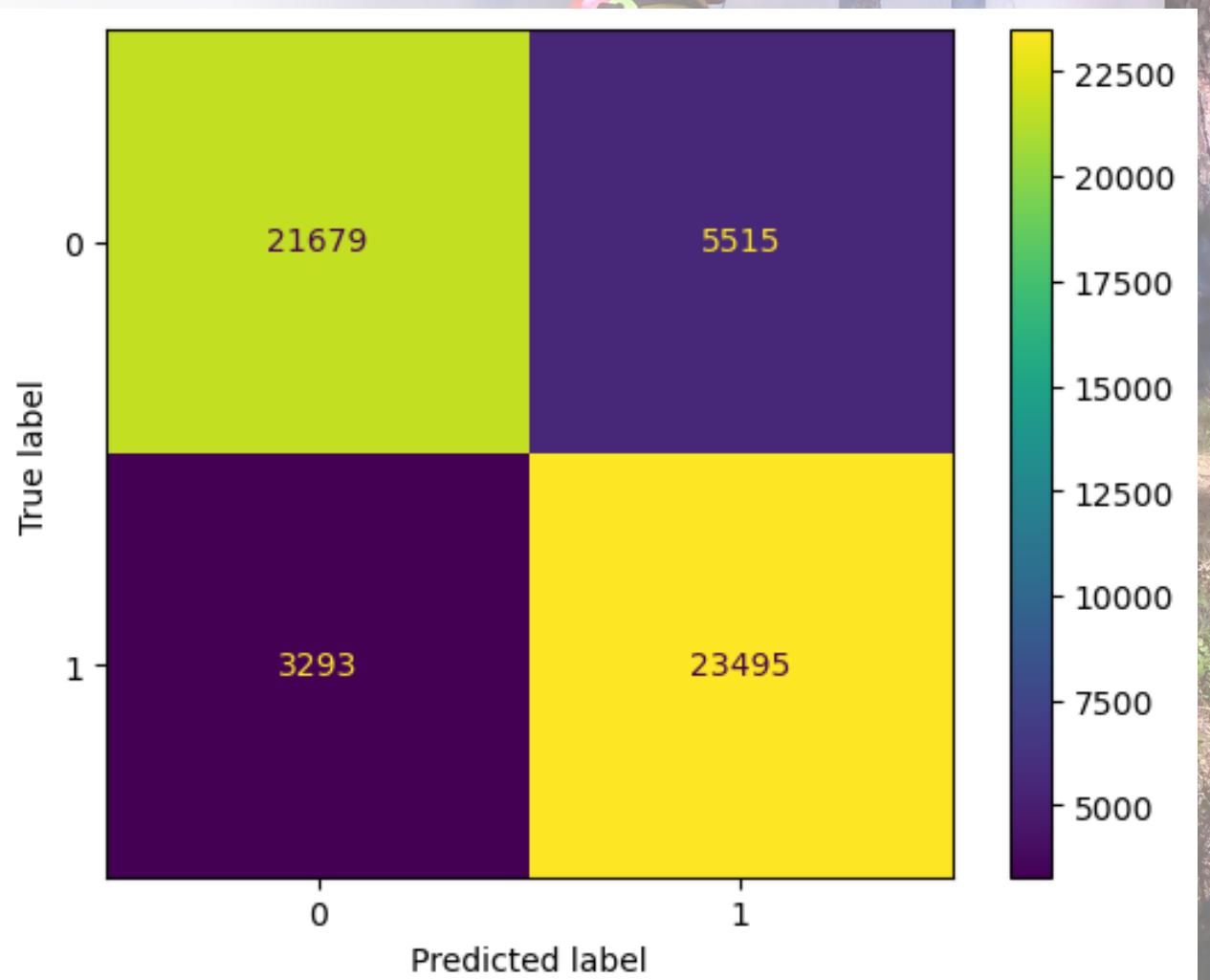
Month	Year	Longitude (x)	Latitude (y)	Mean Temp (°C)	Total Snow (cm)	Total Precipitation (mm)
-------	------	---------------	--------------	----------------	-----------------	--------------------------

MONTH has 0.0% missing data
YEAR has 0.0% missing data
Longitude (x) has 0.13% missing **data**
Latitude (y) has 0.13% missing data
Mean Temp (°C) has 11.35% missing data
Total Snow (cm) has 15.33% missing data
Total Precip (mm) has 9.7% missing data
Fire has 0.0% missing data

YEAR	MONTH	Longitude (x)	Latitude (y)	Mean Temp (°C)	Total Snow (cm)	Total Precip (mm)	MT 1 month prior
0.394077	0.011568	-0.626418	0.060766	1.14302	-0.128344	-0.425583	0.671183
TS 1 month prior	TP 1 month prior	MT 2 month prior	TS 2 month prior	TP 2 month prior	MT 3 month prior	TS 3 month prior	TP 3 month prior
0.049549	-0.048065	0.895108	0.01405	-0.201349	-0.310058	-0.012336	-0.420161

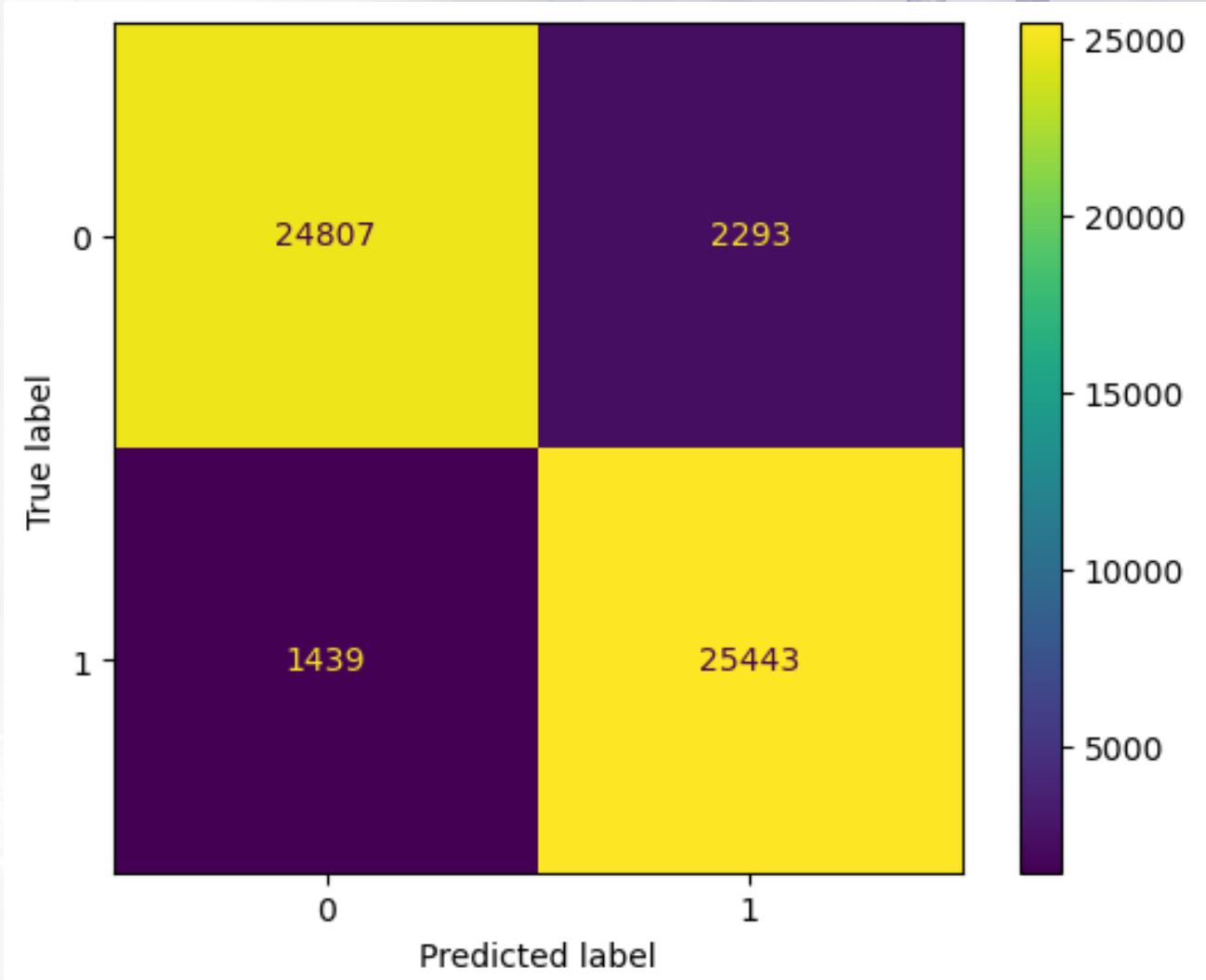
Logistic Regression

- Used backfill and front fill to take care of missing data
- Standard Scaler was used
- Train Score 83.47%
- Test Score 83.68%



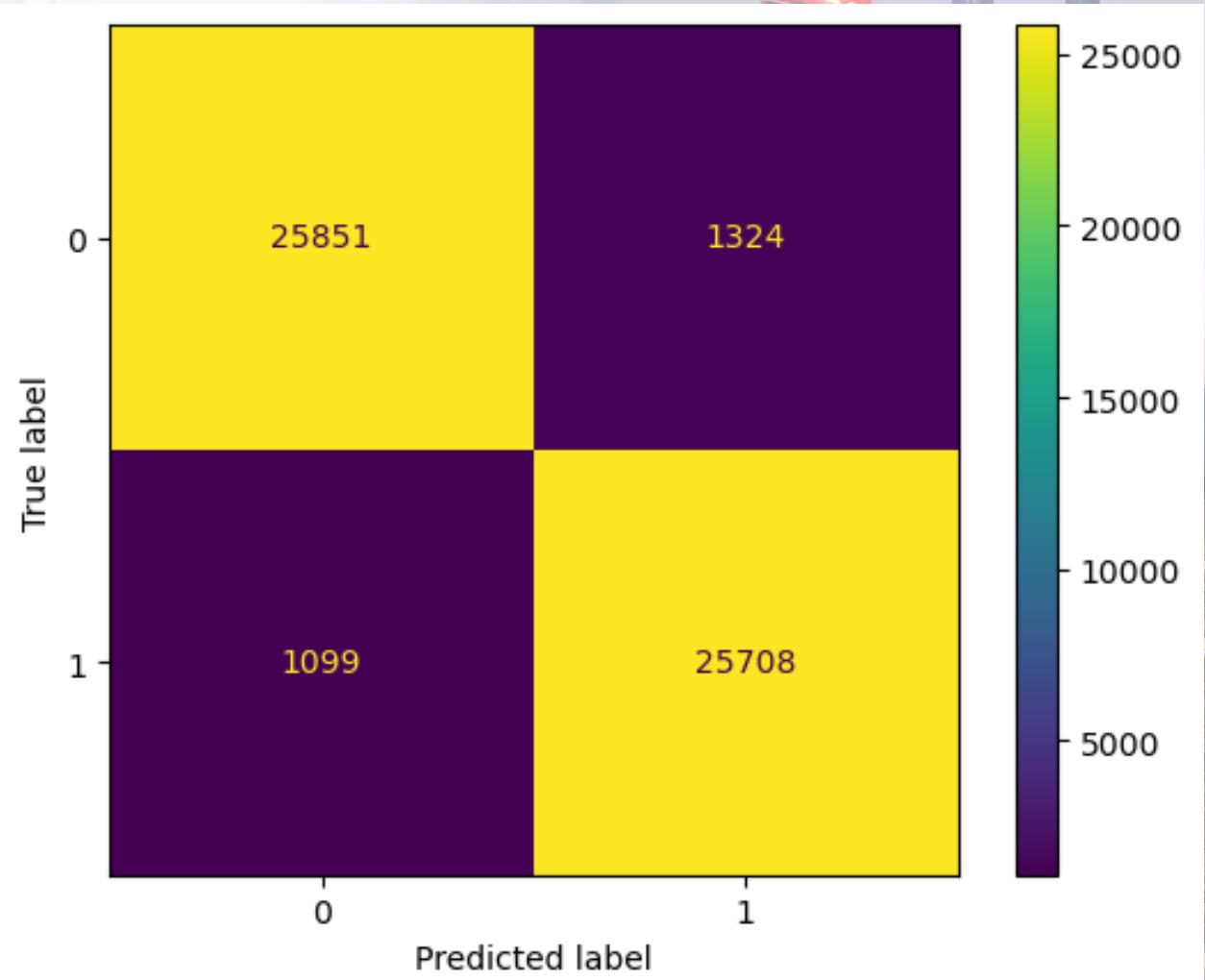
Decision Tree

- Used the same backfill and front fill data from before.
- Settled with 9 nodes deep
- Train score: 93.51%
- Test Score: 93.08%



Histogram-based Gradient Boosting Classification Tree

- Chose this model because it works well with Null data. No need to impute
- Used default parameters
- Train Score 95.64%
- Test Score 95.51%



Thank You

