California Wildfire Prediction

Jin YuHan Burgess

Outline

01. Background

02. Data and Preprocessing

03. Predictive Model

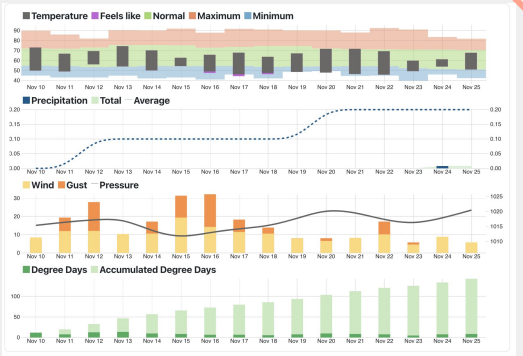
04. Comparisons

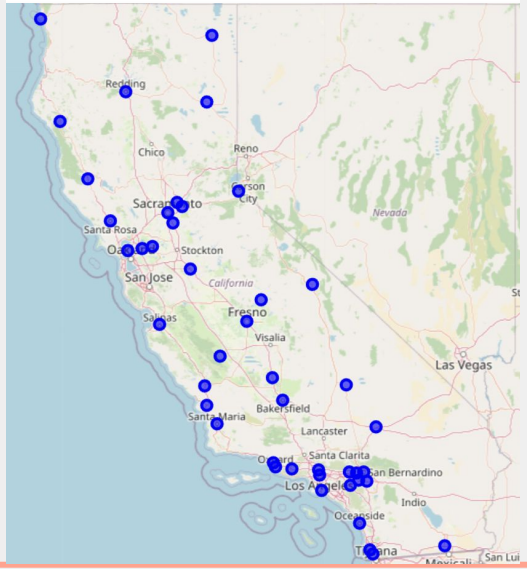
05. Questions

Historic Weather Data

Visual Crossing API
- 44 cities throughout the state

Known weather data: 128743





Parameters

Temperature	Dew Point	Precipitation
Precipitation Coverage	Wind Speed	Wind Direction
Wind Gust	Cloud Cover	Solar Radiation
Sea Level Pressure	Elevation	

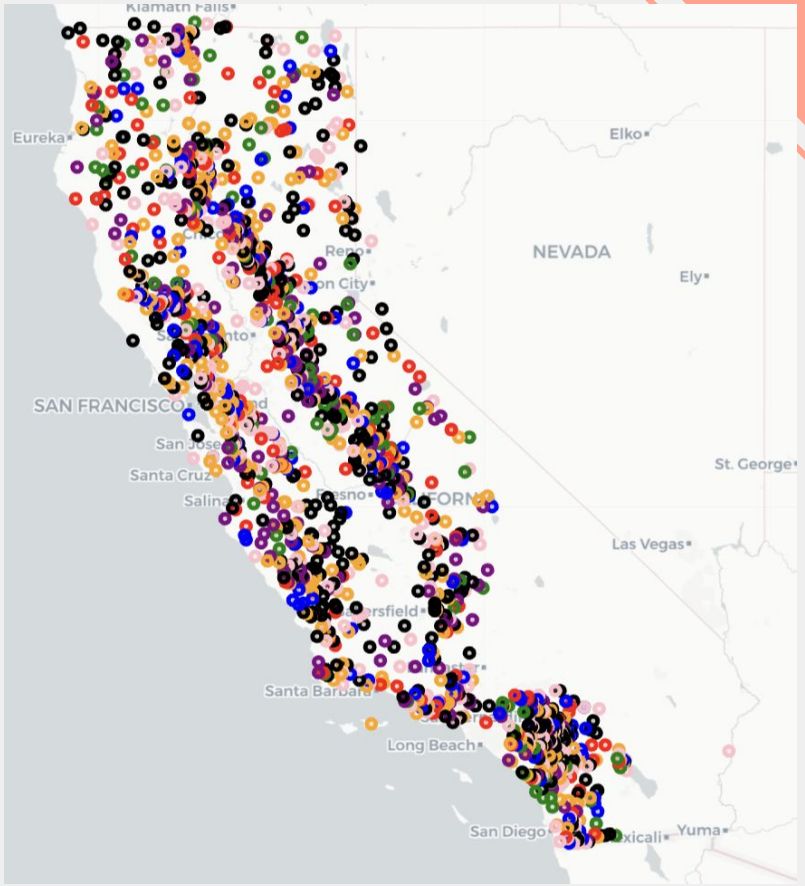

Training
Val
Test
2013 - 2017
2018
2019

Historic Wildfire Data

Kaggle: California WildFires (2013-2020)
Known wildfire labels: 1636

Parameters

Acres Burned	Fatalities	Injuries
Structures Damaged	Structures Destroyed	Air Tankers
Helicopters	Engines	Crews Involved
Dozers	Water Tenders	Major Incident
Start	Extinguished	



70448

I

Pre-train severity calculation model

Acres Burned as proxy label

II

Distribute Calculated Severity (Data Augmentation)

Distribute updated data of Acres Burned and Optimized Severity Score across duration of fire

III

Scale Optimization

Scale Optimized Severity Score using Log-Scale

IV

Use Distance of Fire to Assign to City

Wildfire Dataset is on a County Level. Use Coordinates to Assign it to the Closest City

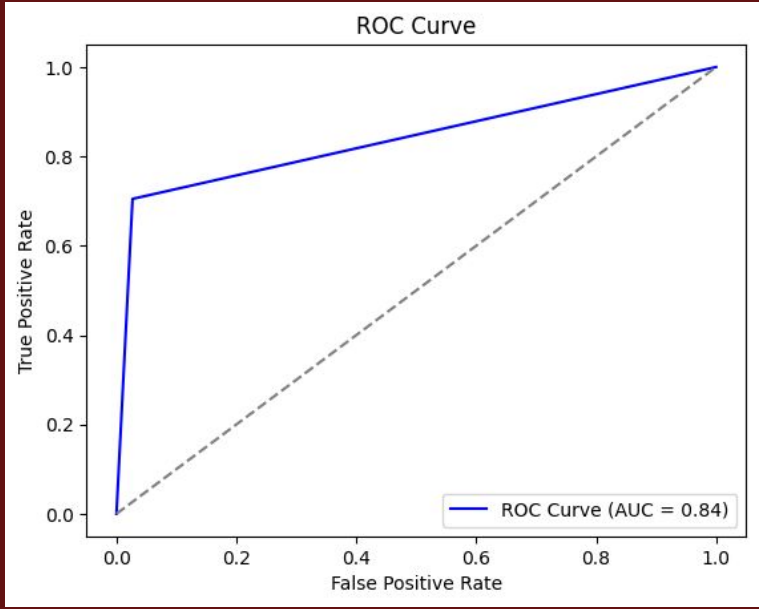
Data Distribution

Training	Validation	Test
Obs: 85170 Fire: 41183 No Fire: 43987	Obs: 32084 Fire: 3606 No Fire: 28478	Obs: 10389 Fire: 9381 No Fire: 1008

Prediction Models

Base Model - Random Forest Classifier - Random Forest Regressor
LSTM
AI Model

Base Model



Accuracy	70%
F1-Score	80%
MSE	1.67

Predicted

Actual		
	3012	74
	9111	20013

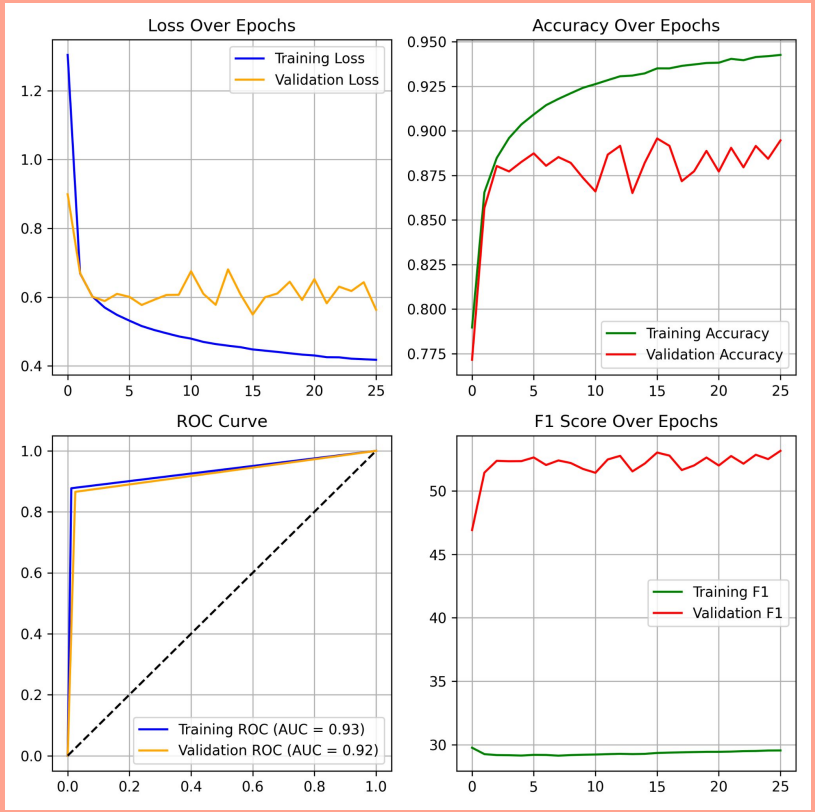
Parameter Fine Tuning

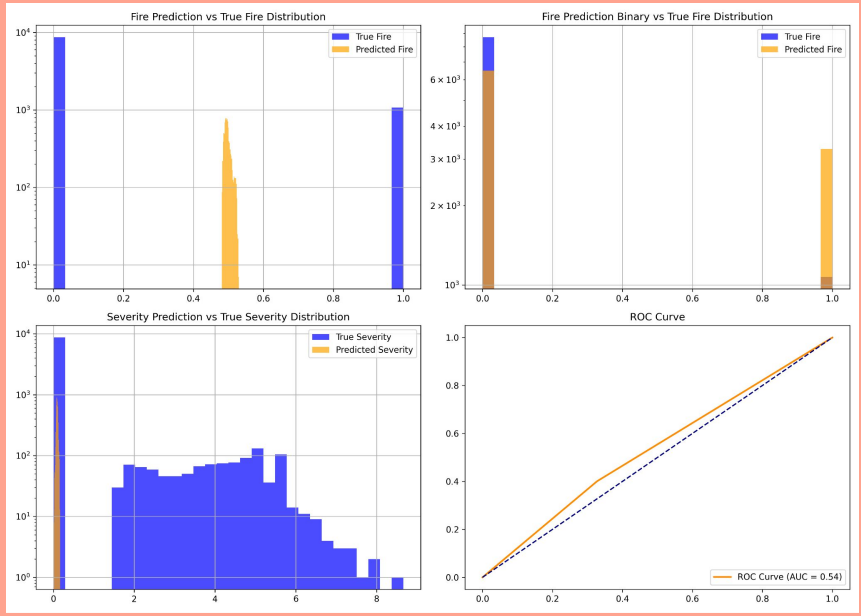
- Randomized Search
- Cross Validation: 3-5 fold

Hyper Parameters

Layers	32 64 128 256
Batch Norm	True, False
Dropout	0.2, 0.3, 0.4
L2 Regularization	0.001, 0.0005, 0.0001
Learning Rate	0.01, 0.001, 0.0001

LSTM Model





5848	2848
643	430

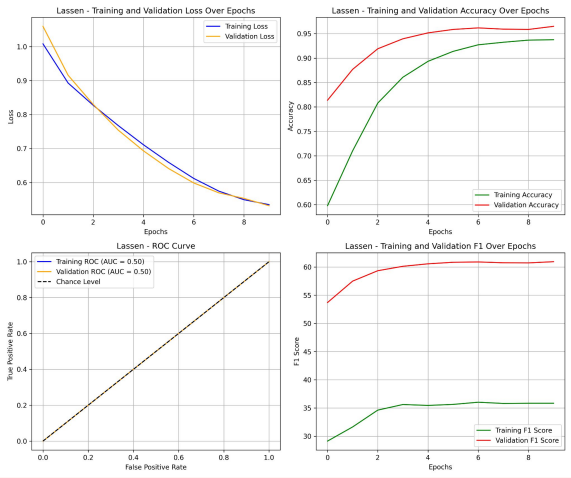
Data Structure

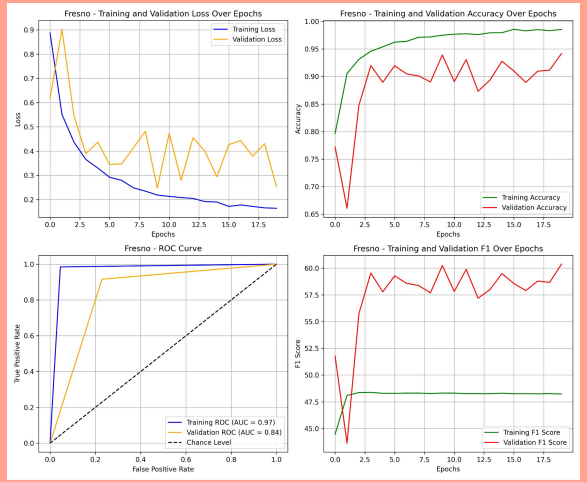
Different regions may have more extreme differences in parameters

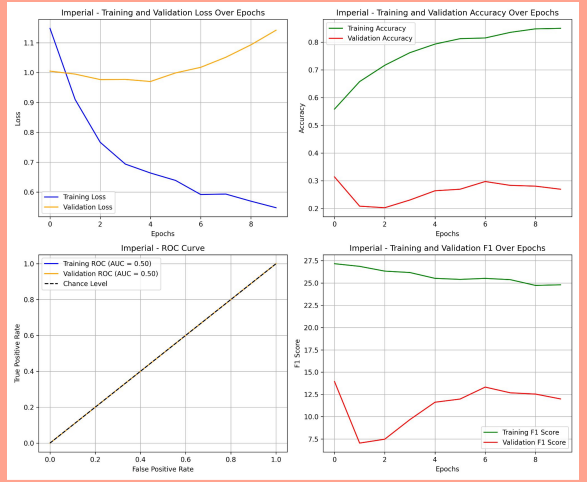
- Group by counties

Contra Costa	Del Norte	El Dorado	Fresno
Humboldt	Imperial	Inyo	Kern
Lassen	Los Angeles	Mendocino	Modoc
Monterey	Placer	Riverside	Sacramento
San Bernardino	San Diego	San Luis Obispo	Santa Barbara
Shasta	Sonoma	Stanislaus	Ventura

Model Regional Variation Handling







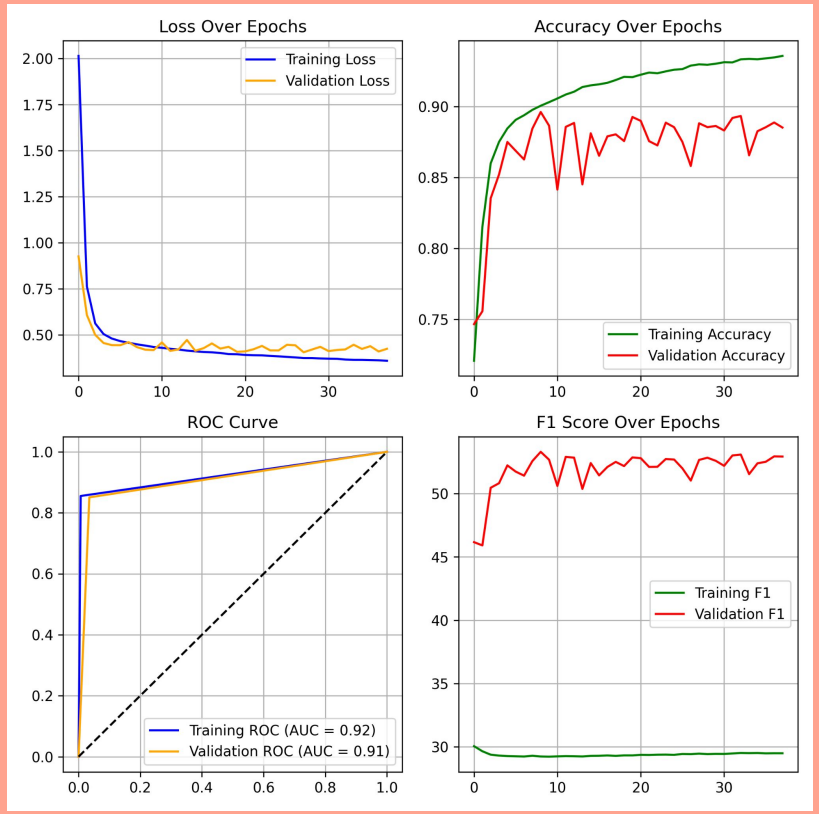
Imbalanced Data

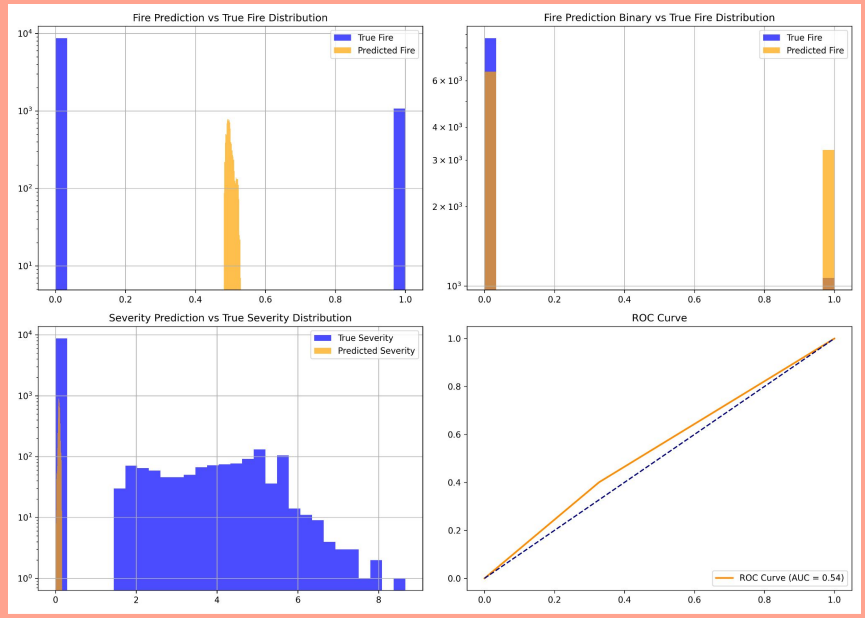
City	Fire Present	Fire Not Present
Los Angeles, CA	3	1217
Bakersfield, CA	3355	897
Coalinga, CA	4303	969

Undersampling Techniques

- Minority weighting
- Infrequent
- Sliding Window

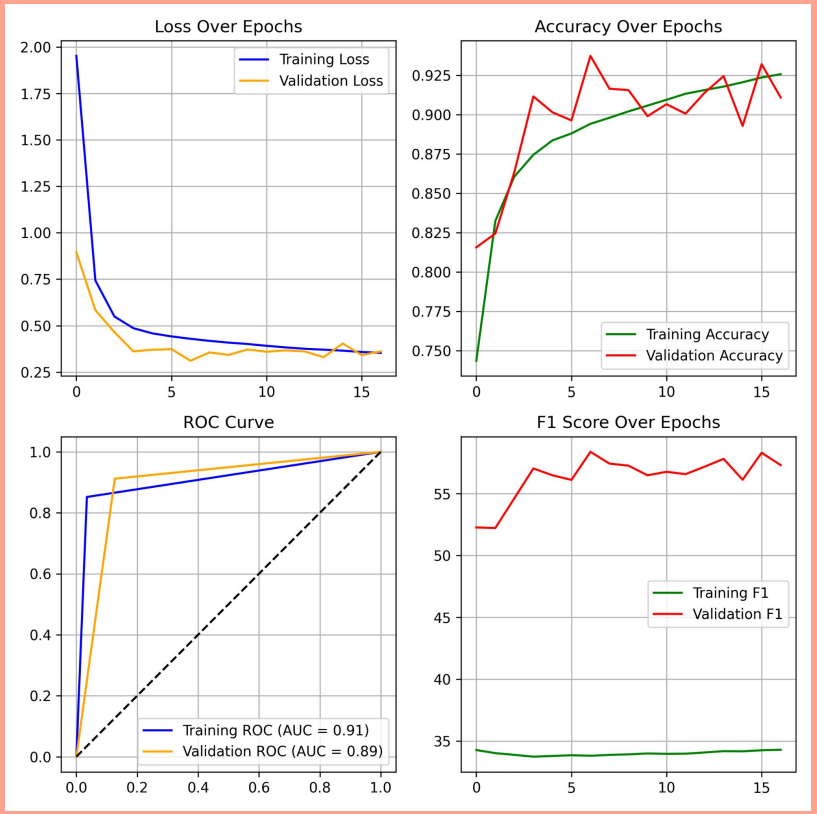
Model Imbalance Handling: Minority Weight

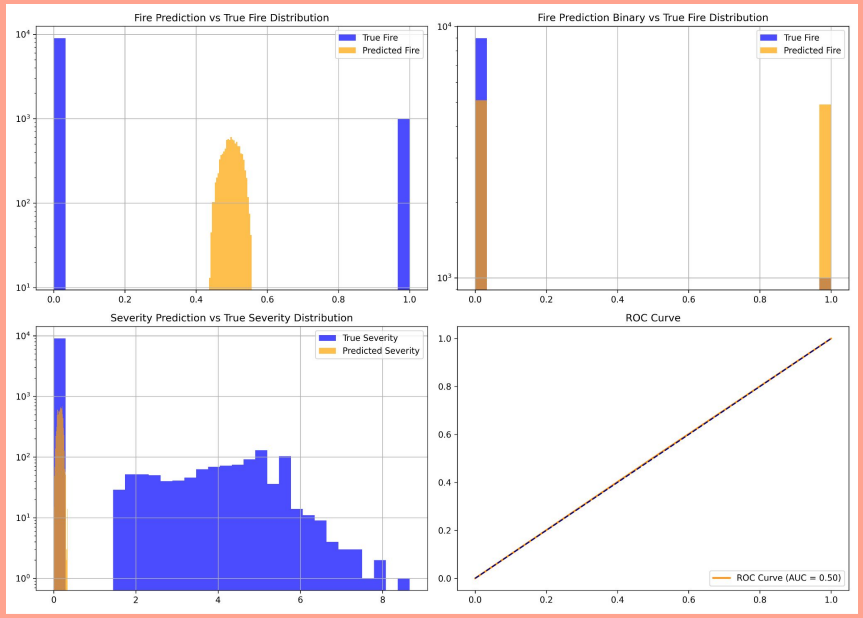




5848	2848
643	430

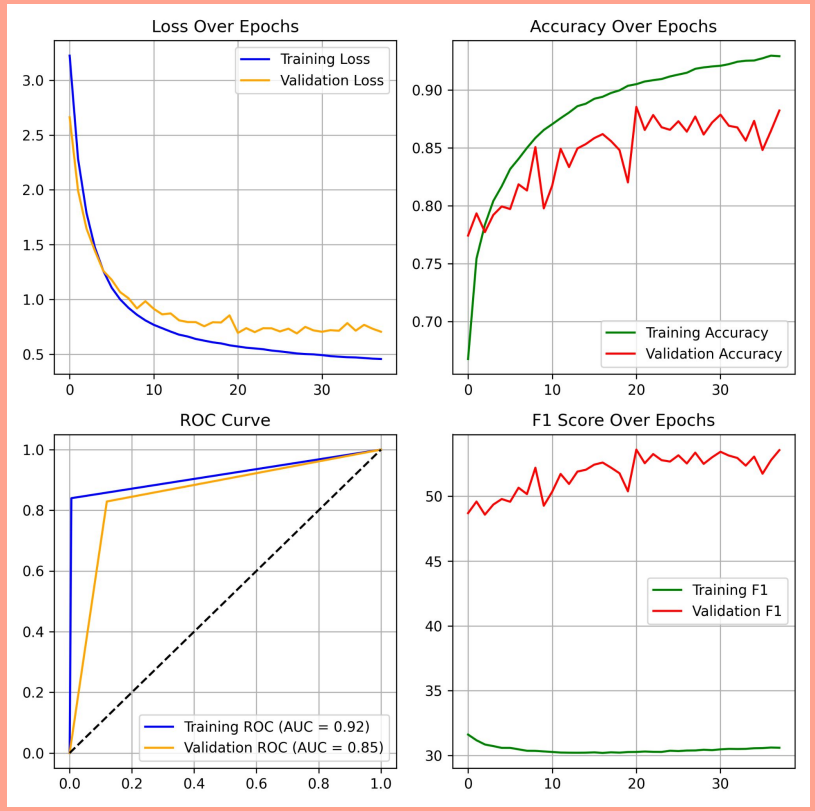
Model Imbalance Handling: Infrequent

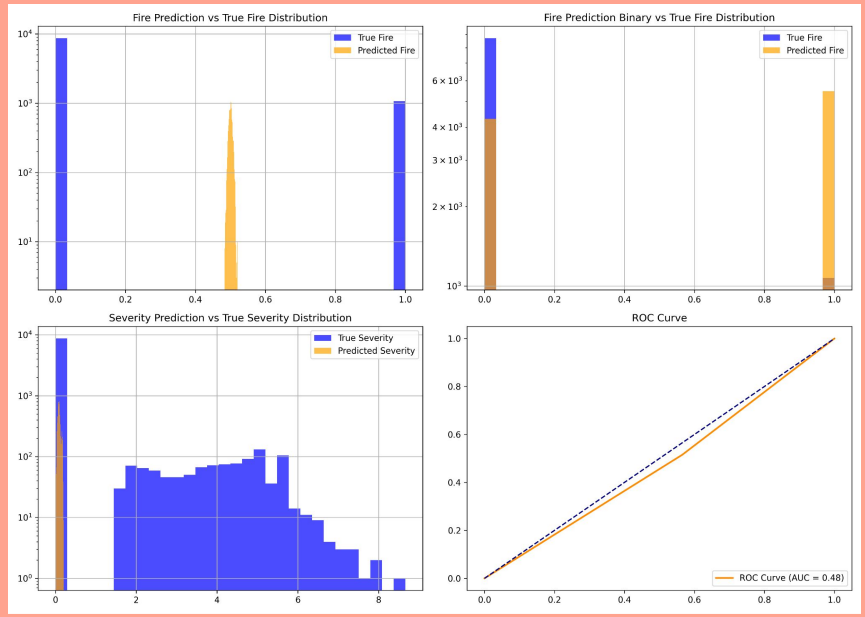




4585	4401
507	492

Model Imbalance Handling: Sliding Window





3926	5060
638	361

AI Setup

Parameters
Temperature
Dew Point
Windspeed
Sea Level Pressure
Solar Radiation
Elevation
*Fire Occurred

Role
You are an expert in wildfire predictions
Prompt
Analyze the wildfire risk for CITY on DATE: *Over the past 15 days:* - Average Parameters *Current weather:* - Current Parameters *Based on the past 15 days of weather and current conditions:* Step 1. Give a one word prediction the wildfire risk. Step 2. Give a one word prediction for the severity of a potential wildfire. Example of responses in Step 1 and Step 2: Low, Moderate, or High. There can be combinations of them such as Low-Moderate. Step 3. List the top 3 reasons why these predictions were chosen.

AI Model

City	Day	History	Current	Prompt	Label
Redding, CA	2019-05-25	Over the past 15 days: - Avg Temp: 17.7°C - Avg Dew: 8.6% - Avg Precip: 5.01mm - Avg Precip Cover: 17.78% - Avg Wind Speed: 30.5kph - Avg Sea Lvl Press: 1010.12mb - Avg Solar Rad: 224.77w/m^2 - Avg Elevation: 171.0m	Current weather: - Temp: 19.2°C - Dew: 7.5% - Precip: 8.686mm - Precip Cover: 12.5% - Wind Speed: 61.1kph - Sea Lvl Press: 1006.9mb - Solar Rad.: 316.7w/m^2 - Elevation: 171m	Step 1: Elevated Step 2: Moderate Step 3: 1. High wind speeds: 2. Low relative humidity given current and historic dew point 3. Recent precipitation: Insufficient to offset the dry conditions and fuel moisture levels	0

Comparison

Focus	AI	Main	Minority Weighted	Infreq	Sliding	Label
Wildfire Occurrence	Elevated	0.499	0.5	0.52	0.496	O
Severity	Moderate	0.203	0.079	0.16	0.0833	O

Questions?