

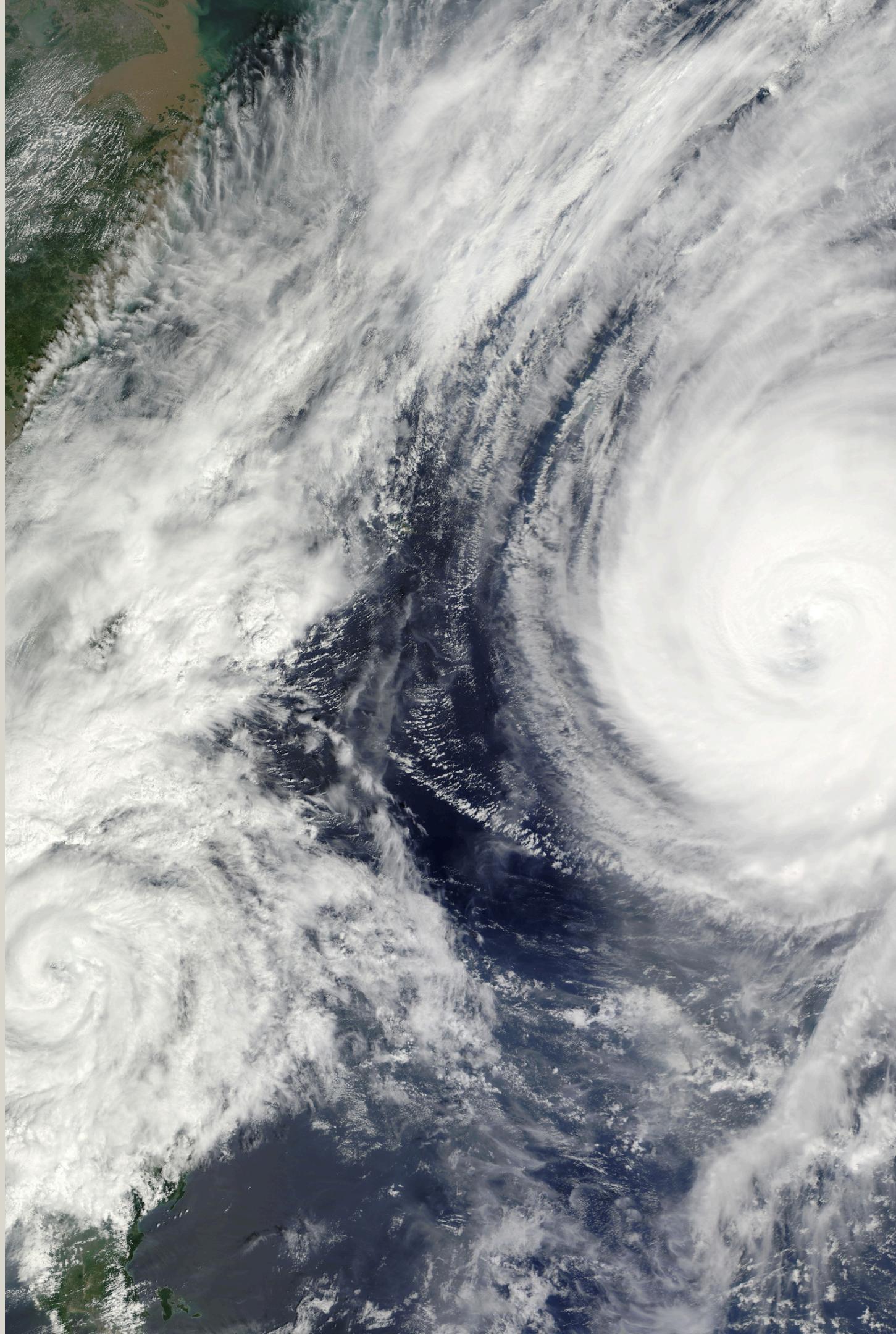
Predicting Heatwaves in Spain

A Machine Learning Approach

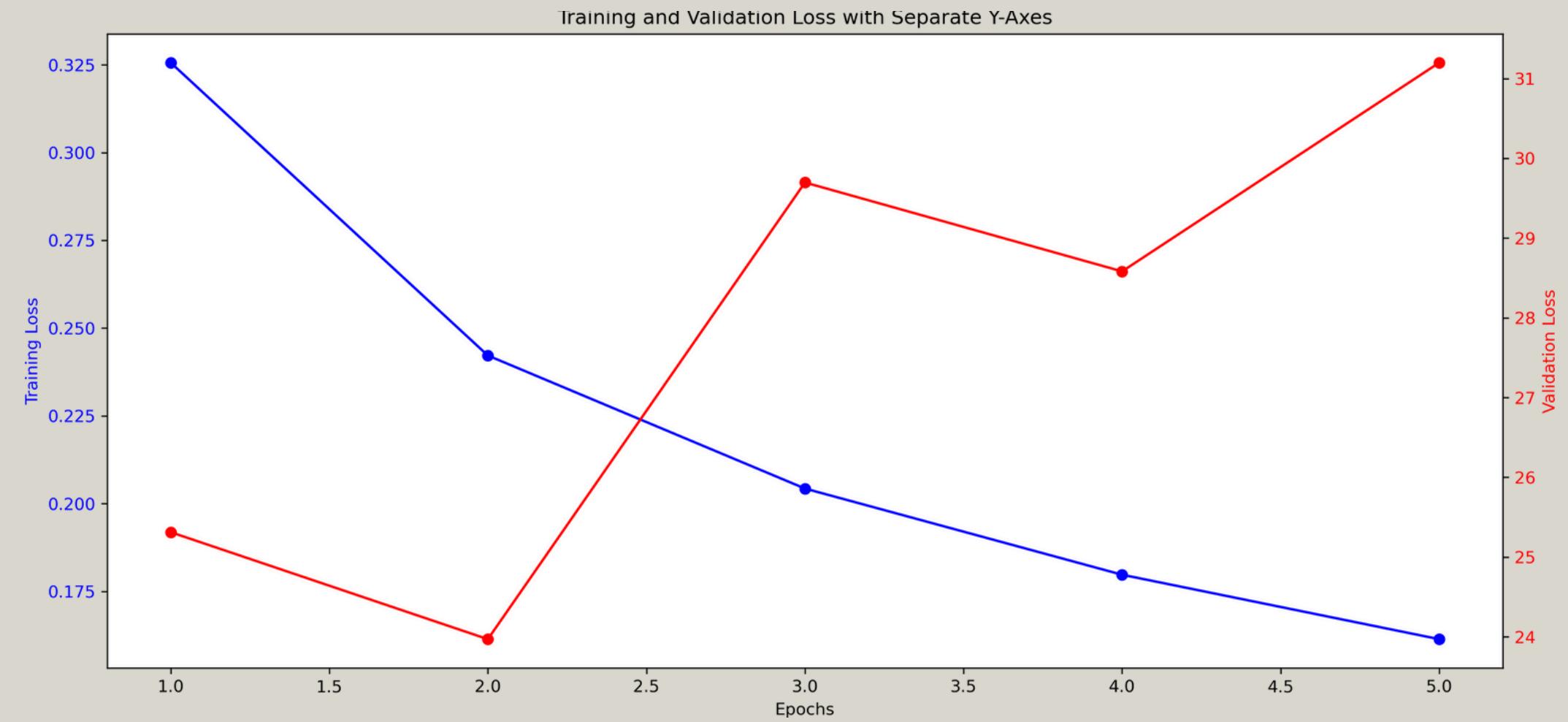
[GitHub link](#)

Data & Methodology

- Data pipeline: Copernicus CDS → ERAS dataset → Python preprocessing (xarray, Pygrib) → ML models.
- Key features: Temperature, wind, soil moisture, etc.



Model Challenges



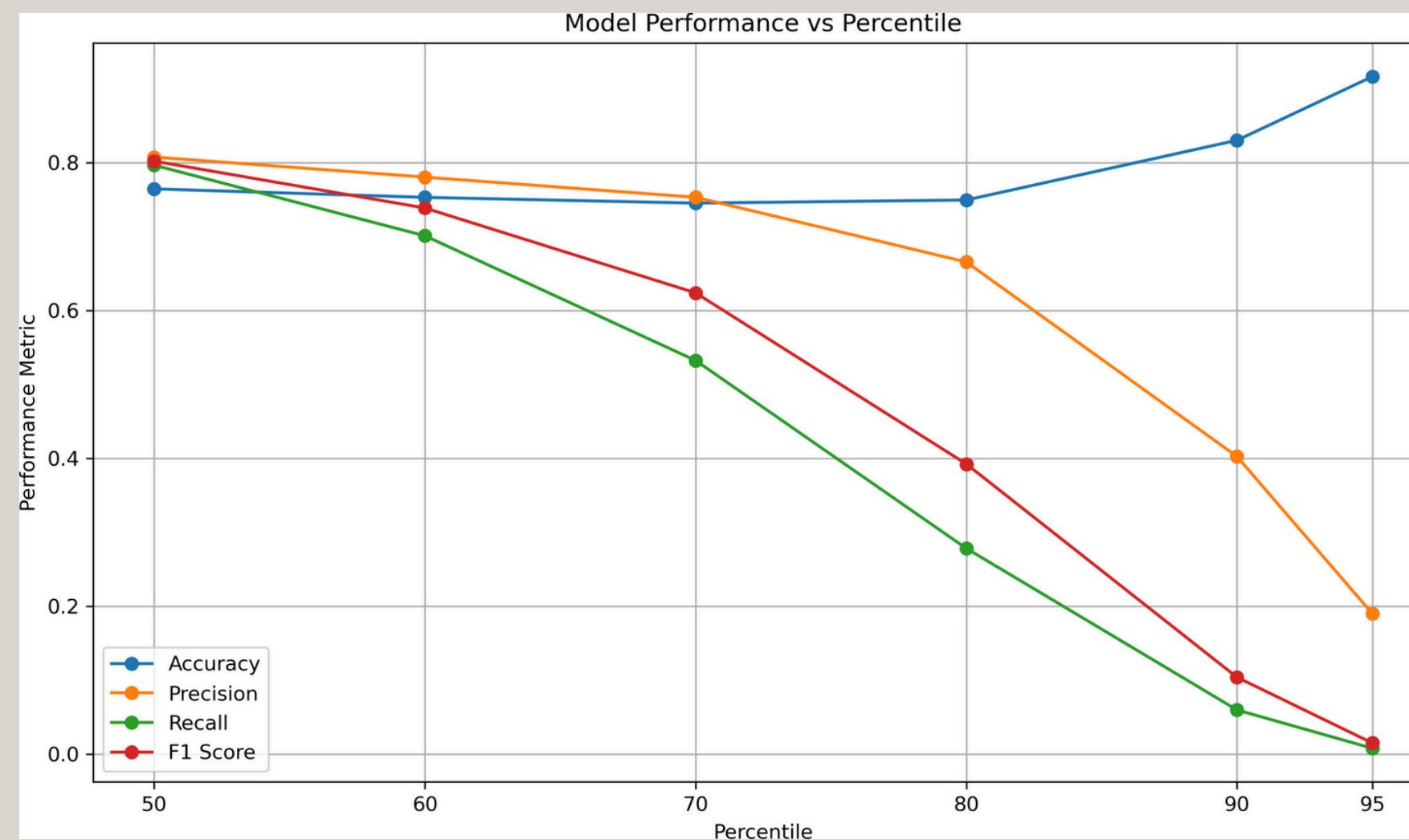
Accuracy	Precision	Recall	F1 Score
0.8951	0.1166	0.0436	0.0635

Breakthrough with SMOTE

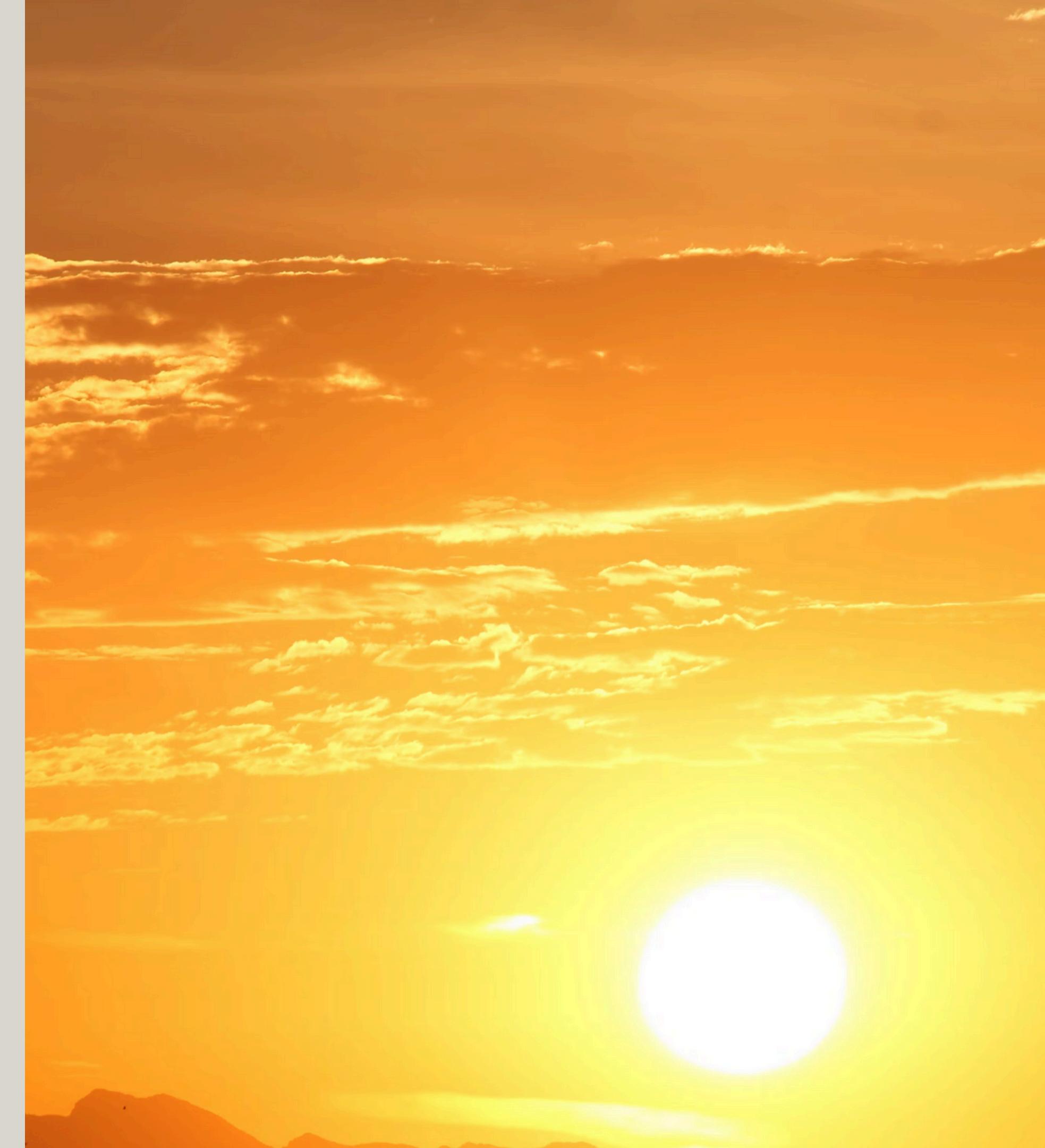
Model	Accuracy	Precision	Recall	F1 Score
Before SMOTE	0.9023	0.1472	0.0625	0.0885
Logistic Regression	0.7449	0.1586	0.8116	0.2653



Simplifying the Problem



Percentile	Case	Accuracy	Precision	Recall	F1 Score
70th	1	0.9469	0.9488	0.8836	0.9151
70th	2	0.8530	0.8131	0.6807	0.7411
70th	3	0.7452	0.7531	0.5323	0.6238





Future Directions

- Roadmap:
 - a. Incorporate regional weather data.
 - b. Cloud-based scaling (e.g., Google Colab).
 - c. Literature-driven model refinements.
- SMOTE + spatial data synergy.