An Evaluation of Data-Driven Interpretable Methods to Detect Tropical Cyclones

Marco Adriano Ferrero 977577

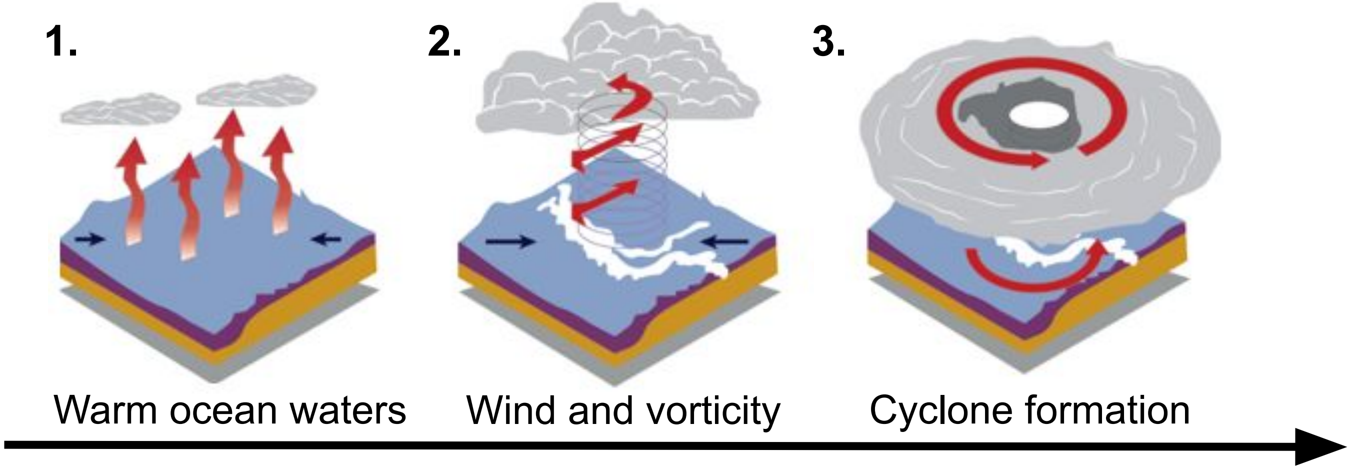
Academic Year: 2023-2024

Advisor: Francesco Amigoni

Co-Advisors: Federico Cerutti, Letizia Tanca, Davide Azzalini

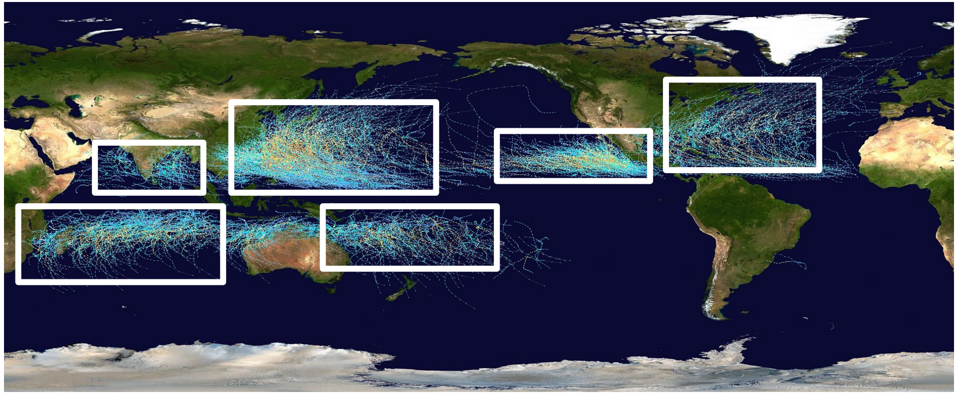
What is a Tropical Cyclone?

Tropical Cyclones Formation



[Palmen, 1948]

Tropical Cyclones Basins



Credits: NASA and IBTrACS

Tropical Cyclones Forecasting Problem

Data | **Models** | **Forecast** | **Decisions**

- Satellite data
- Oceanic data
- Reanalysis data

Genesis, intensity, tracking, extreme events

Coordinate humanitarian responses to extreme events

[Pradhan et al., 2017]
[Chen et al., 2020]

Purpose of the Thesis

- Apply Machine Learning based methods to detect the presence of tropical cyclones in the South-West Indian Ocean basin
- Provide explanations to support decision processes

Data

Global Drivers

El Nino Southern Oscillations
- Sea surface temperatures

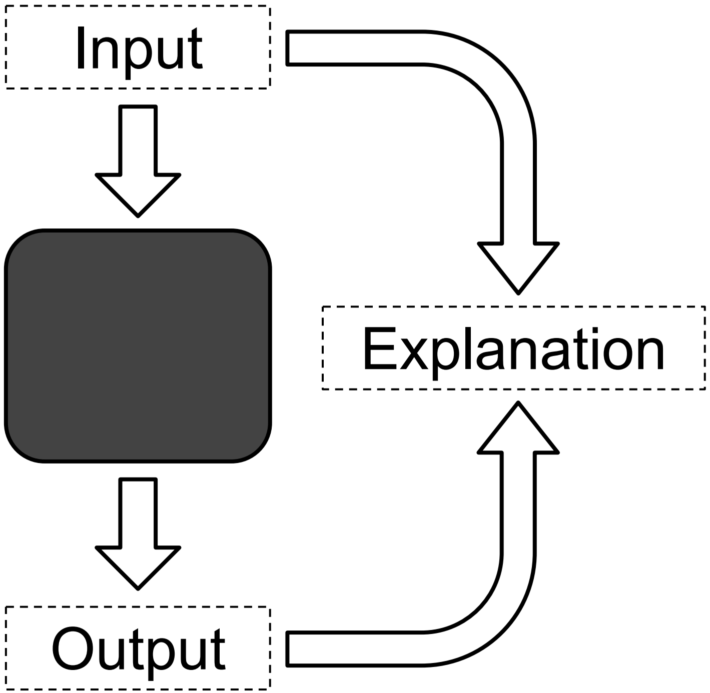
Madden-Julian Oscillation
- Real-Time Multivariate Index (RMM)

Local Drivers

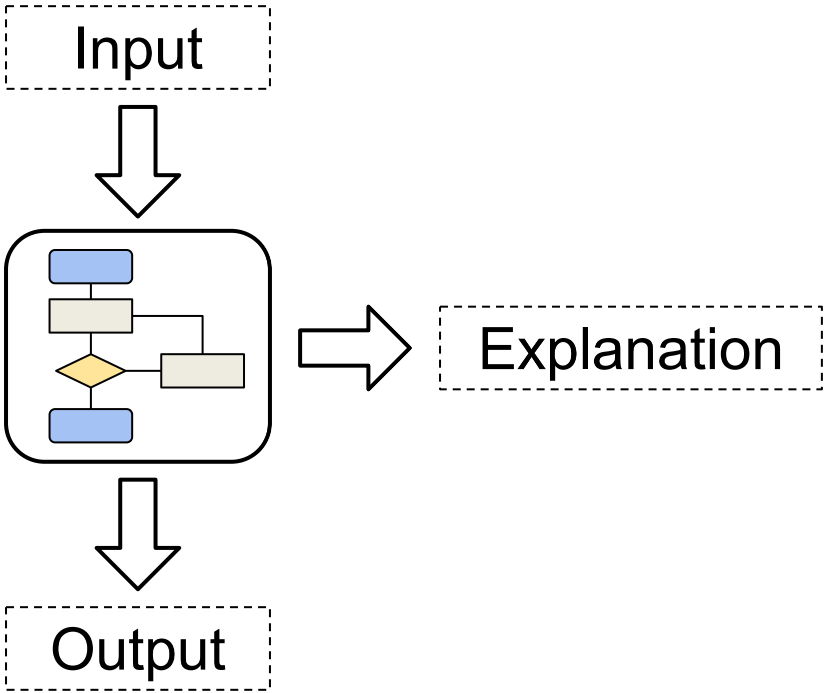
- Wind speeds
- Temperature
- Pressure
- Precipitation
- Relative vorticity
- Air density

ECMWF's 5th reanalysis system (ERA5)

Black-Box Model



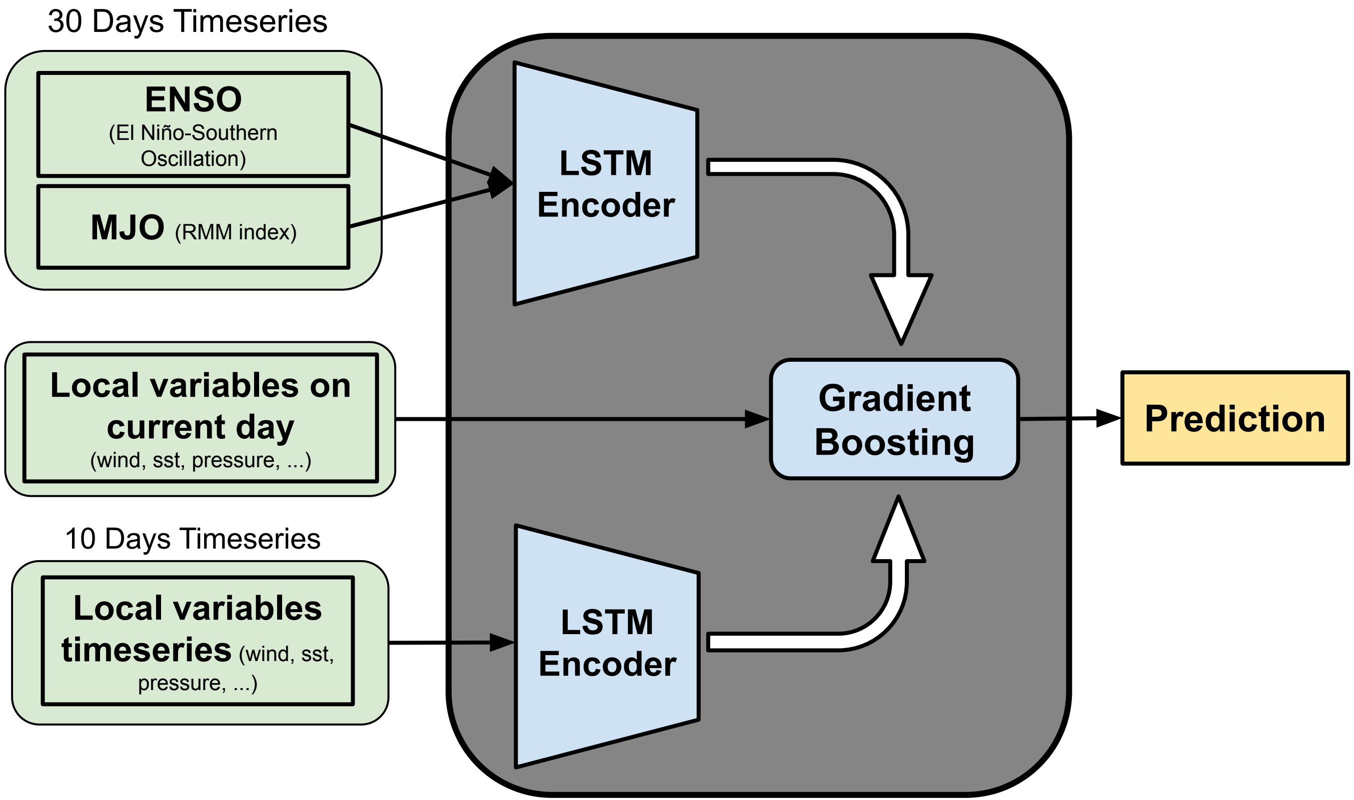
White-Box Model



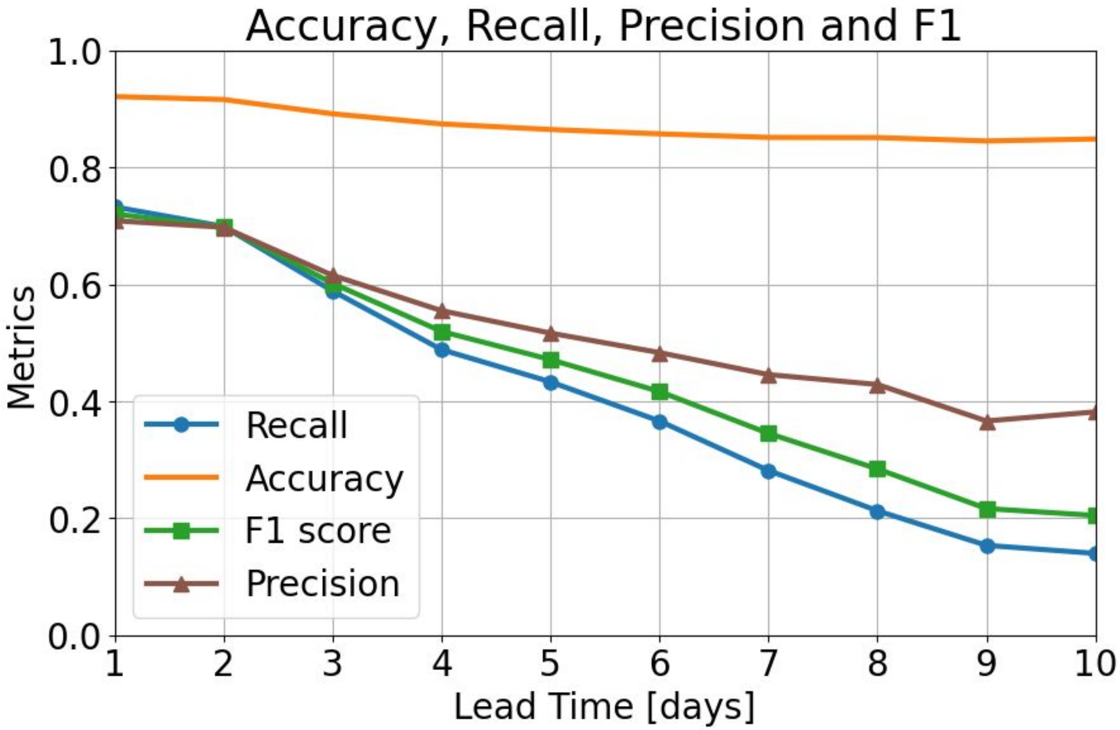
Black-Box - LSTM Encoders + Gradient Boosting

Black-boxes:

1. Gradient boosting decision trees (GBDTs)
2. LSTM networks
3. **LSTM Encoders + GBDTs**



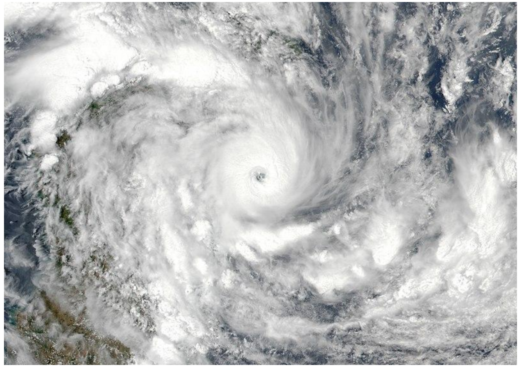
Black-Box - Results

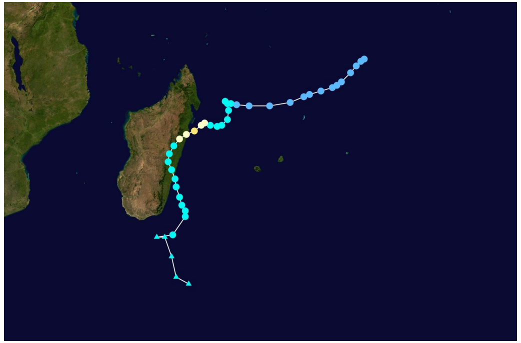


$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

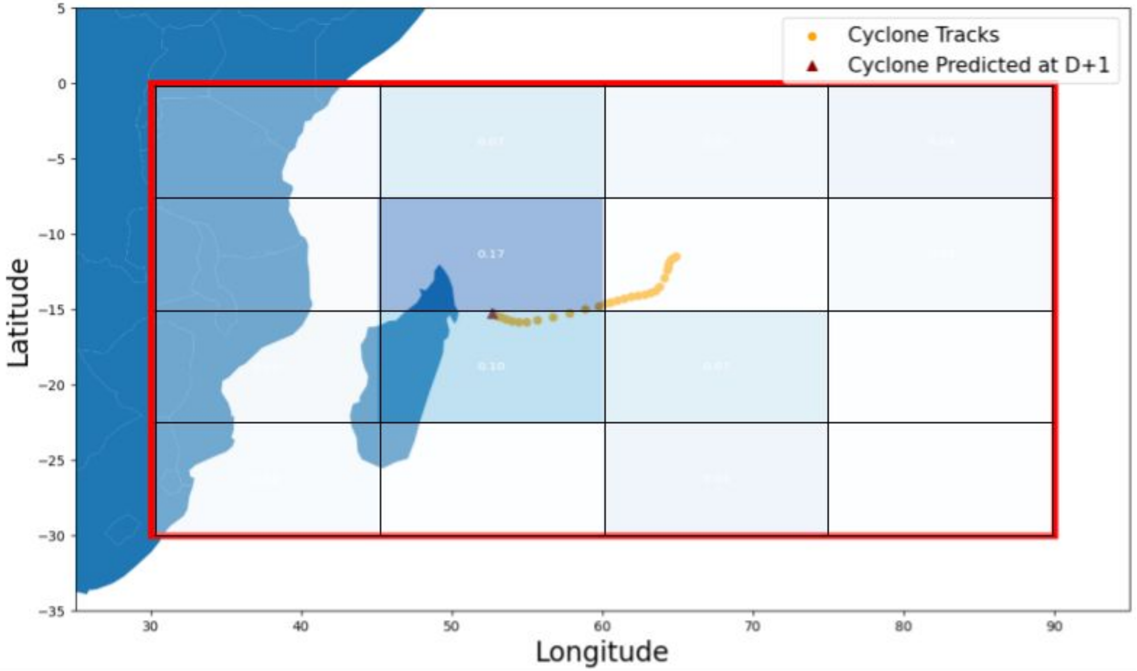
Black-Box - LIME



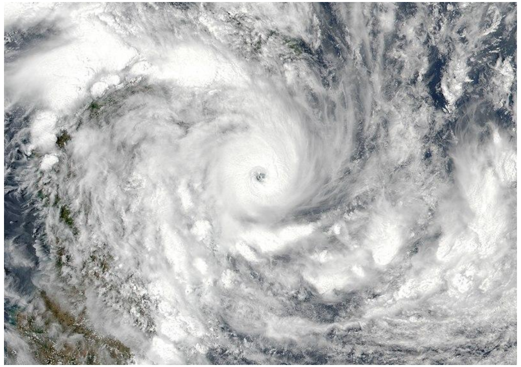


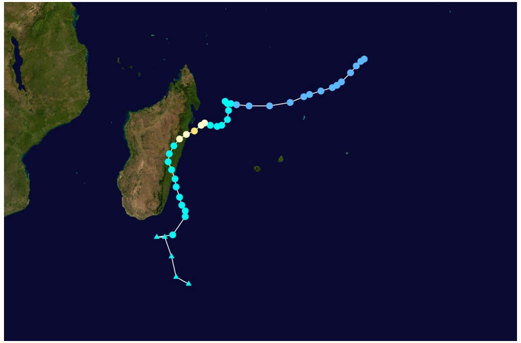
Credits: *earthobservatory.nasa.gov*

Cyclone Ava - 3 January 2018



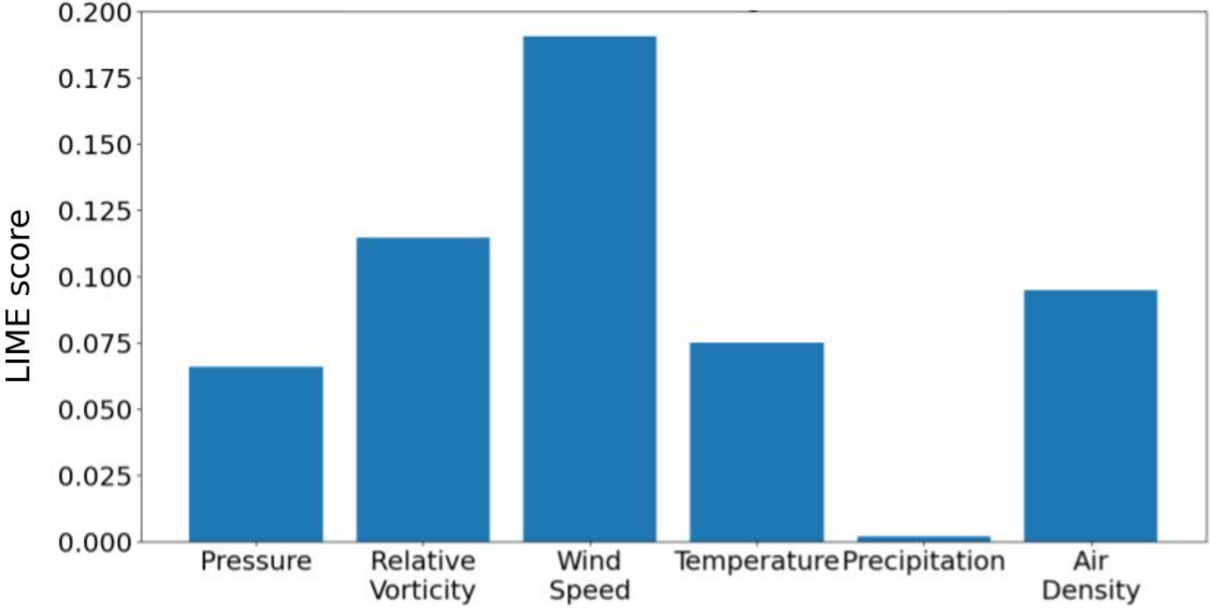
Black-Box - LIME



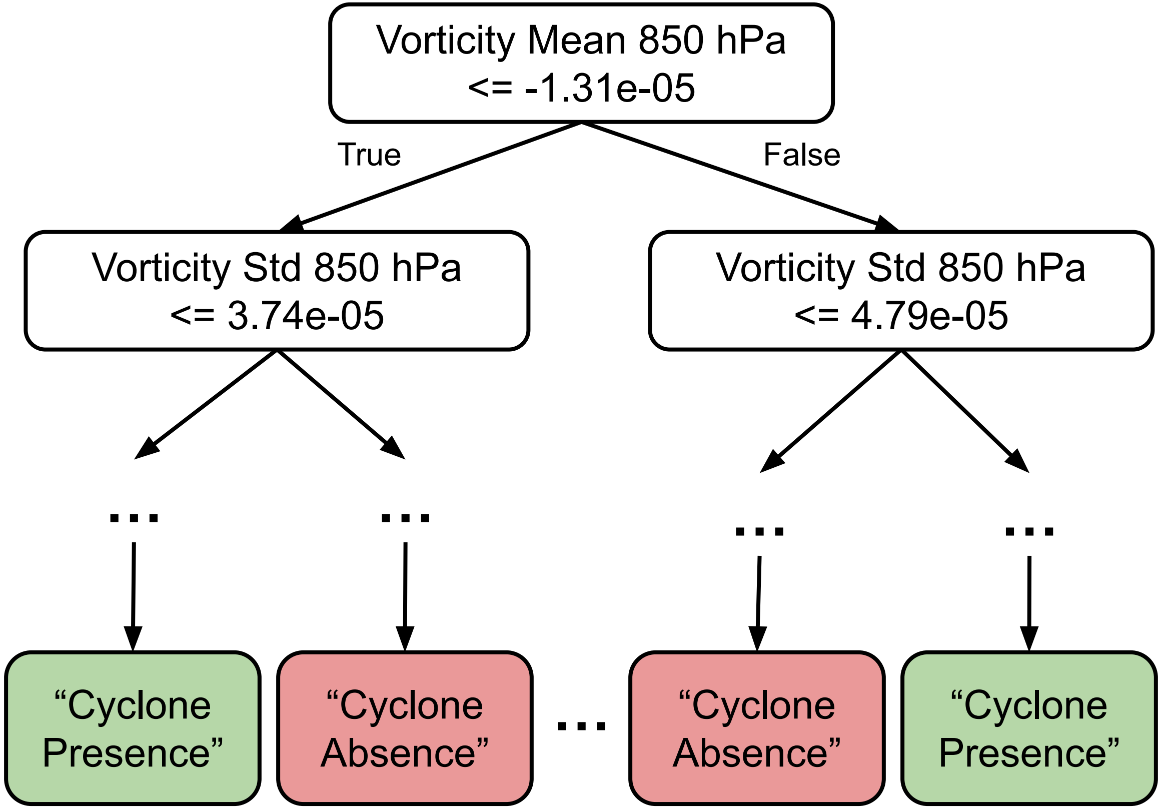


Credits: *earthobservatory.nasa.gov*

Cyclone Ava - 3 January 2018



White-Box - Decision Tree



Predictive Results - 24 h Horizon

Recall	82%
Precision	88%
False Alarm Rate	12%

Tree Structure

Max Depth	6
Number of Leaves	28

White-Box - Bayesian Rule Lists

Trained RuleListClassifier

```
===========================================================================================================
IF        P_Mean > 101230     and Wind_Gust_Std <= 2.11            THEN probability of class 1: 0.3%   (0.2%-0.5%)
ELSE IF   P_Mean > 101230     and Vor_850hPa_Std <= 3.26e-05       THEN probability of class 1: 0.6%   (0.2%-1.1%)
ELSE IF   P_Std <= 175        and Vor_850hPa_Mean >= -3.11e-06     THEN probability of class 1: 2.2%   (1.5%-3.0%)
ELSE IF   Vor_850hPa_Std <= 3.26e-05                               THEN probability of class 1: 10.7% (8.6%-13.1%)
ELSE IF   P_Mean > 101230     and T_1000hPa_Std > 0.92             THEN probability of class 1: 6.0%   (2.5%-11.0%)
ELSE IF   Vor_850hPa_Mean >= -3.11e-06                             THEN probability of class 1: 26.4% (15.6%-38.9%)
ELSE IF   P_Std <= 175                                             THEN probability of class 1: 29.4% (22.2%-37.1%)
ELSE IF   -1.32e-05 < Vor_850hPa_Mean <= -3.11e-06                 THEN probability of class 1: 55.8% (48.5%-62.9%)
ELSE IF   Wind_850hPa_Mean <= 11.5                                 THEN probability of class 1: 82.8% (75.6%-88.9%)
ELSE IF   P_Std > 196                                              THEN probability of class 1: 97.8% (94.8%-99.5%)
ELSE                                                                    probability of class 1: 33.3% (1.3%-84.2%)
===========================================================================================================
```

Predictive Results - 24 h Horizon

Recall	70%
Precision	78%
False Alarm Rate	22%

Conclusions

What has been done?

- Implementation of White-Box and Black-Box models to detect cyclones on multiple time horizons
- Global drivers have a weak influence on the overall forecasts, while local drivers better fit the problem
- Forecasting limitations due to poor quality and high dimensionality of meteorological data

Future Work:

- Evaluate interpretable data-driven methods to address track, intensity, or seasonality forecasting
- Apply interpretable techniques to additional state-of-the-art models (CNNs, transformers, …)

Thank you