



SPRINT 2 CAPSTONE

Presented by: Rodrigo Pelayo Ochoa





DETECTING ILLEGAL FISHING ACTIVITIES USING AIS DATA



The Problem area:

- Severe **environmental impacts** due to indiscriminate net fishing that catches all species.
- Estimated **\$34 billion annual economic loss** globally due to IUU fishing.
- Gives space to perform other **Illegal Activities**

The Solution:

- Develop a **predictive model to identify illegal fishing patterns** among events of vessels turning their AIS off.
- Focus on **analyzing vessel behaviour** to detect illegal activities.

THE DATA

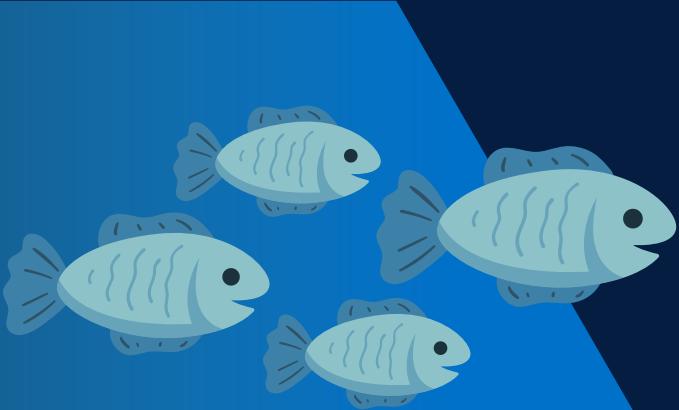
(55365, 22)

- Welch et - Unseen Fishing Hotspots
 - Data on **suspected AIS disabling events** highlights when vessels stop broadcasting, indicating potential illegal activities.
- Carrier Vessel Encounters
 - This dataset records fishing vessels transferring supplies or fish to carrier vessels. Indicating potential IUU.
- Fishing Effort
 - Adds information and features to Welch Dataset

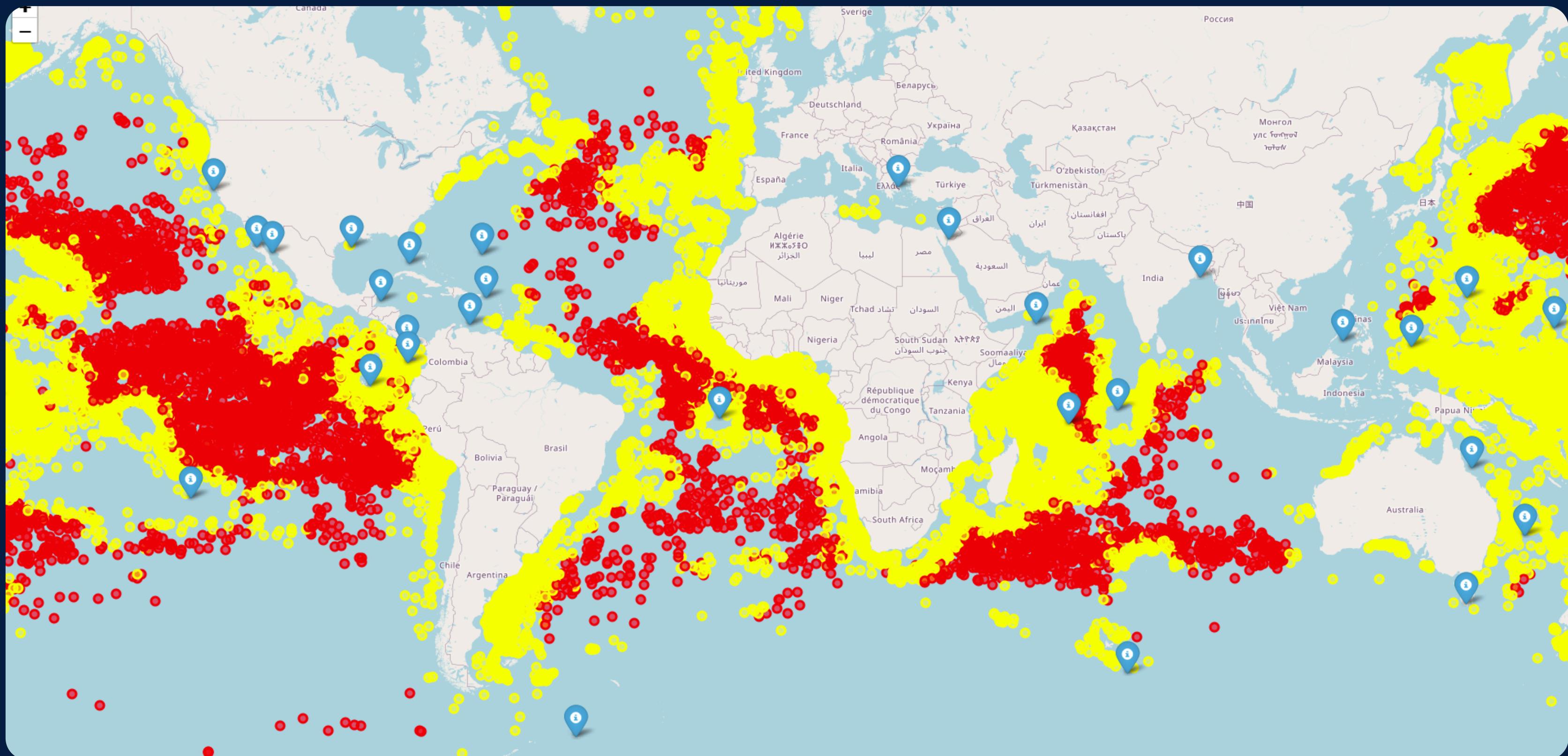
- Gap_id
- mmsi
- Vessel Class
- Flag
- Vessel Length m
- Vessel Tonnage Gt
- Gap Start Lat and Lon (also for end)
- Gap Hours
- Distance to a High IUU Risk Zone
- IUU Risk proximity
- Distance to a Marine Protected Area
- Closest Marine Protected Area
- Marine Protected Area Proximity Level
- Encountered Carrier
- IUU Score
- IUU Class ⚠

- Unsupervised Learning
 - Kmeans
 - Agglomerative Clustering
 - Gaussian Mixture
 - DBSCAN
 - Hierarchical Clustering
- Engineered a Scoring System
 - Vessel Class
 - Gap of Hours
 - High Risk Countries
 - Risky Distance to an MPA

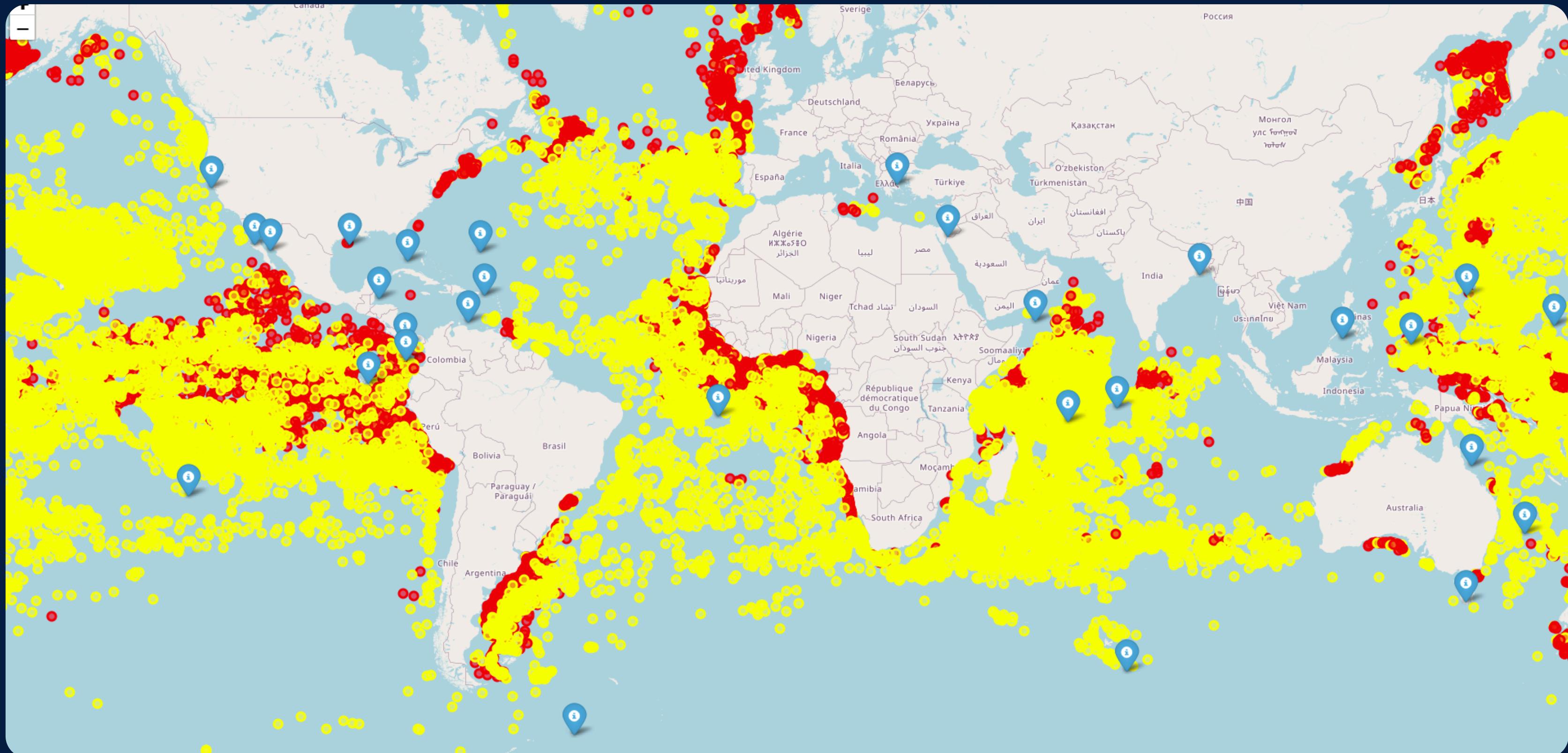
2017-2019



KMEANS LABELS

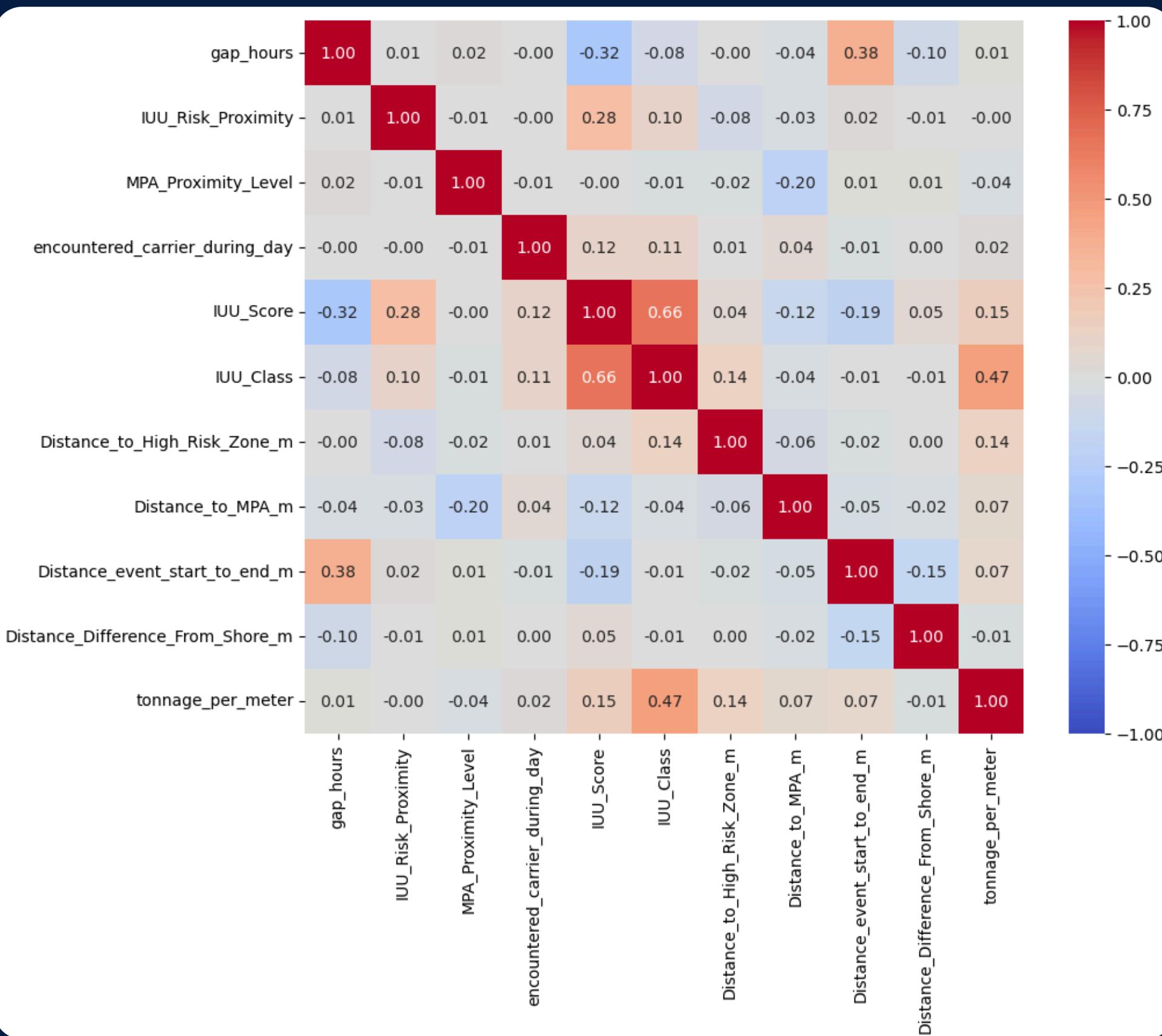


SCORE SYSTEM

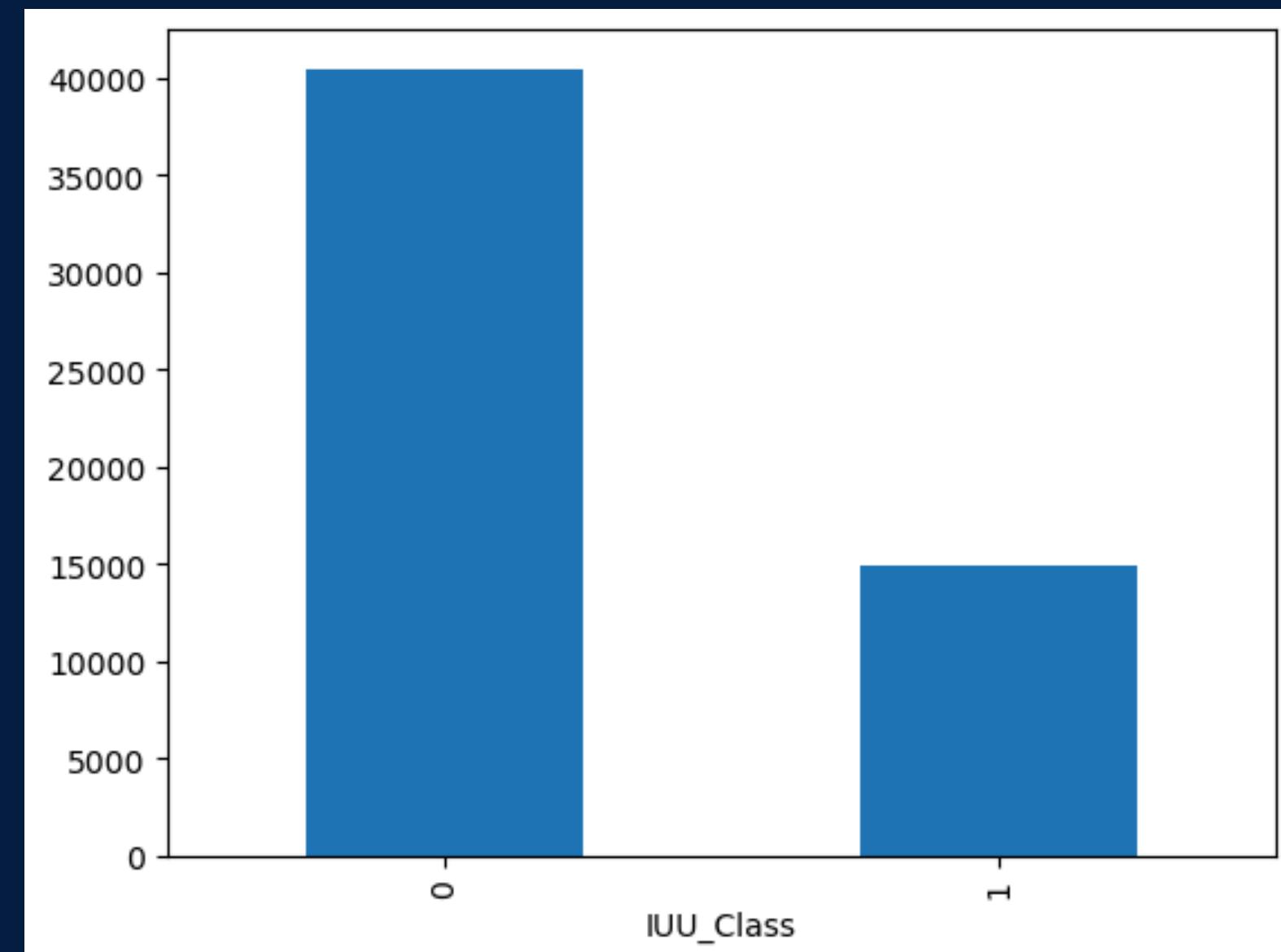


NOT FINAL

CORRELATION MATRIX

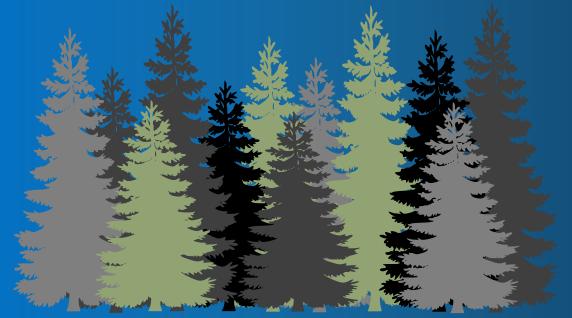


CLASS IMBALANCE TV



SMOTE FUNCTION

MODELS



- Logistic Regression
 - Simplicity
 - Good for Binary Classification
 - Quick Training
 - Precision:
 - Class 0 (Legal Fishing): 0.77
 - Class 1 (Illegal Fishing): 0.31
 - Recall:
 - Class 0: 0.55
 - Class 1: 0.55
 - F1-Score:
 - Class 0: 0.64
 - Class 1: 0.40
 - Accuracy: 55.20%
-
- Random Forest
 - Combines multiple decision trees
 - Resistance to Noise and Outliers
 - Scalability
 - Precision:
 - Class 0 (Legal Fishing): 1.00
 - Class 1 (Illegal Fishing): 0.99
 - Recall:
 - Class 0: 1.00
 - Class 1: 1.00
 - F1-Score:
 - Class 0: 1.00
 - Class 1: 0.99
 - Accuracy: 99.72%



THANK YOU

Presented by: Rodrigo Pelayo Ochoa

