Exploring Illegal, Unreported, and Unregulated Fishing Detection using AIS Data and Machine Learning

Samuel Brown, Danielle Katz, Dana Korotovskikh, and Stephen Kullman

UNIVERSITY OF VIRGINIA DATA SCIENCE INSTITUTE

What is the Issue with IUU Fishing?

- Environmental & economic impacts
- Tied to drug & human trafficking
- Estimated cost of $23 billion annually

Automatic Identification Systems (AIS)

- What are they?
  - Standard tracking devices on vessels used primarily for collision avoidance
- Commonly illegally exploited on fishing nets or buoys to protect large hauls
- These impersonated AIS signals are harmful towards legal vessels, leading to a critical need for improved IUU regulation.

Data Overview:

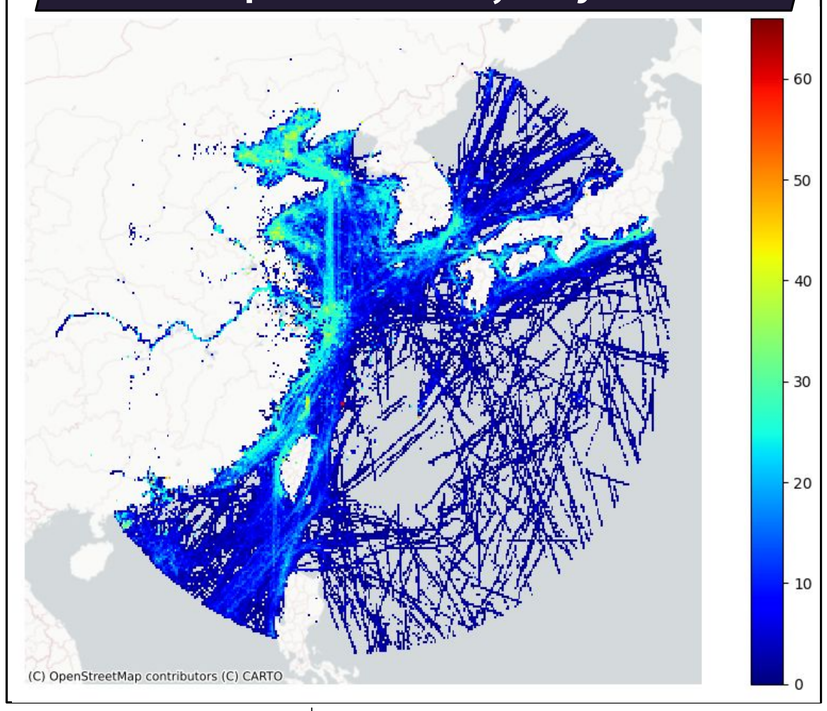
- Region of Interest: Southeast Asia
- Spatial, temporal, & user inputted data obtained from CCRi
- Pre-processing steps:
  - Aggregated by distinct device trips
  - *'Red flag'*: Indicators of non-standard naming conventions and movement characteristics (score from 0 - 4)
  - These indicators are then used for performance analysis or labelling

1. *net_name*	Names including a 'V', '%', 'buoy', or 'net'
2. mmsi_length	MMSI values not equal to 9 digits
3. spawn_offshore	Vessels whose first transmission is offshore (1 nautical mile off the coastline)
4. spoof	Devices with unreasonably high calculated speeds (>= 150 knots)

Regional Analysis

1. Overlay region with grid between 107° to 142° longitude and 14° to 44° latitude, with 1°x1° cells – approximately 36 mi²
2. Position each AIS signal within its corresponding grid cell
3. Calculate *hot_score*: total number of unique red flags, divided by the total count of unique vessels
4. Aggregate together each hours *hot_score* per grid cell and obtain a total
5. Assign a color gradient to each grid cell based on these scores

September 1st, Day 1



Goal:

Explore how AIS data be used to aid in IUU fishing net detection

Modeling Approaches ✎

⟶ UNSUPERVISED ⟵

- Treats dataset as truly unlabelled
- Performs iterative application of hierarchical density-based clustering
- Model develops 10 cluster groups of varying densities, with 3 distinct clusters as primarily illegal nets
- Proves: AIS positional features may serve as good indicators of IUU nets

⟶ SEMI-SUPERVISED ⟵

- Manually labels dataset based on created 'red flag' conditions
  - Illegal vessels chosen by having a bad naming convention and more than 3 red flags
- Trains on small sample of labelled AIS data
- Iteratively builds pseudo-labels on remaining observations and retrains on updated labelled dataset
- Employed with gradient boosted decision trees (XGBoost) and ANN

⟶ SUPERVISED ⟵

- Manually labels dataset based on created 'red flag' conditions
  - Illegal vessels chosen by having a bad naming convention and more than 3 red flags
- Uses pseudo-labeled training data once to build model
- Employed with fundamental ANN

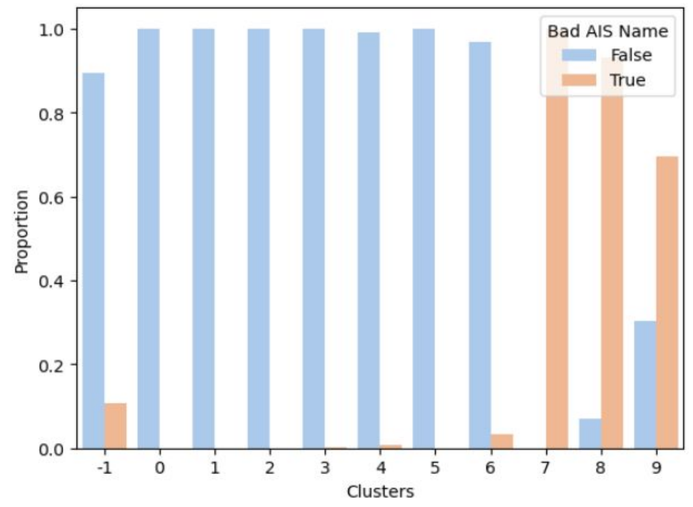
Results ✓

Model	Test Accuracy	TPR	TNR
Semi-supervised XGBoost	0.867	0.915	0.817
Semi-supervised ANN	0.853	0.899	0.805
ANN	0.844	0.920	0.764

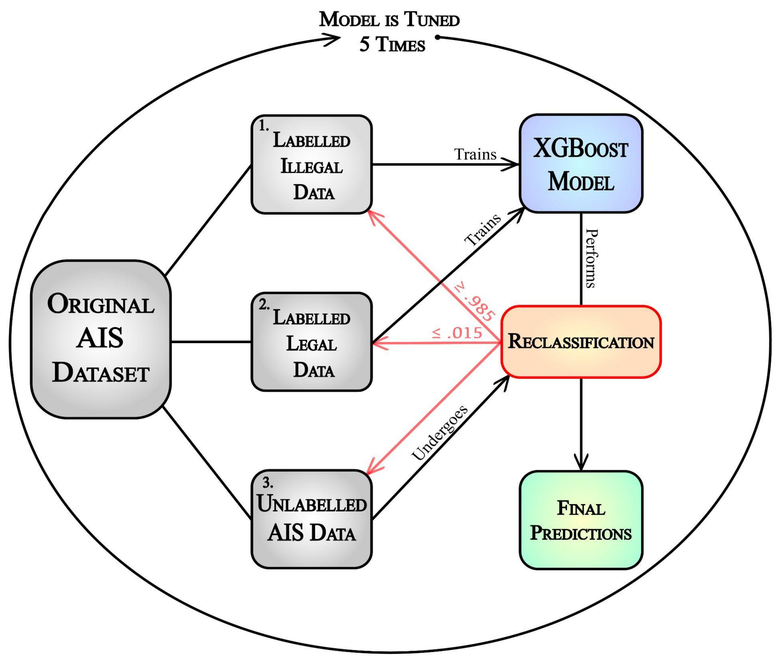
Future Work ↗

- Process a longer duration of AIS data
- Expand model applications to alternative, larger regions
- Explore a recurrent neural network (RNN) with AIS temporal data, then process with positional features

Clustering Analysis

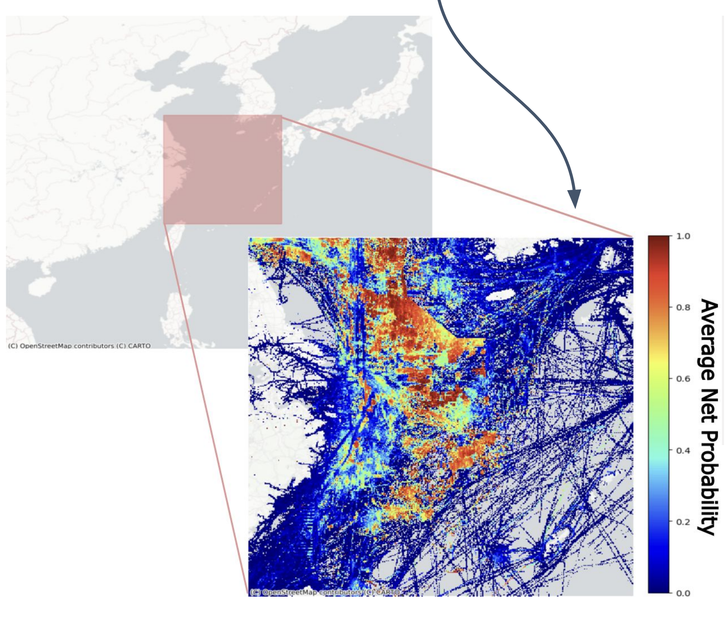


XGBoost Model Workflow



ANN Test Region Predictions

- Projecting 50 hours of data
- Output probability of 1 indicates illegal activity likely to have occurred in region
- Looking at restricted area



References

[1] AIS (Automatic Identification System) Overview. NATO Shipping Centre. (n.d.). https://shipping.nato.int/nsc/operations/news/2021/ais-automatic-identification-system-overview

[2] Brennan, J. (2020, March 16). Fast and easy gridding of point data with geopandas. Retrieved from james-brennan.github.io website: https://james-brennan.github.io/posts/fastg riddingg eopandas/

[3] MIS Marine. (2020). The Use of AIS Data to Identify Dark Activity for Marine Auditing Purposes. In Marine Assurance Technology Experts | MIS (pp. 0–25). Marine Information Solutions. Retrieved from Marine Information Solutions website: https://mismarine.com/wp-content/uploads/2020/02/The-Use-of-AIS-Data-to-Identify-Dark-Activity-for-Marine-Auditing-Purposes-v3.pdf

[4] National Intelligence Council. (2016, September 19). Global Implications of Illegal, Unreported, and Unregulated (IUU) Fishing. https://irp.fas.org/nic/fishing.pdf

[5] Nicholas Kalinowski, Jessica Kimbrell, Christopher Longchamp, and Anne Louise Seekford. (2023). Exploration of Illegal, Unreported, Unregulated Fishing and Vessel Relationships using AIS Data and Machine Learning. Zenodo. https://doi.org/10.5281/zenodo.8008295

[6] United Nations. (2021, June 5). International Day Against Illegal Fishing. United Nations. https://www.un.org/en/observances/end-illegal-fishing-day

UNIVERSITY of VIRGINIA | DATA SCIENCE