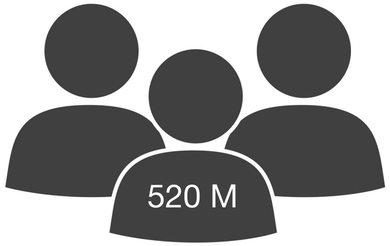
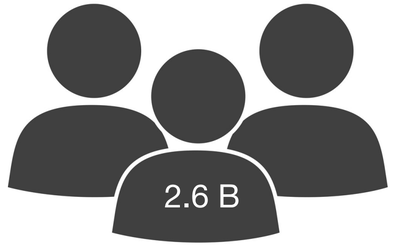
Predicting Fishing Activity

Context / Project Overview



Reliant on Fishing / Fishing Related Activities



Dependent on Fish as Important Part of Diet

- Illegal, unregulated, and unreported fishing activities threaten global food supply
- Vastness of oceans makes direct regulation impossible
- Regulators / policymakers need an effective way to monitor and identify vessels that are fishing
- Apply machine learning models to public ocean and vessel data to identify fishing activity

The Data

Overview of Datasets

Global Fishing Watch

- AIS vessel track data along with is_fishing labels to train machine learning models
- Latitude, Longitude, Vessel Type, Speed, Course, Distance From Port

+

National Centers for Environmental Information

- World Ocean Database - world's largest collection of uniformly formatted, quality controlled, publicly available ocean profile data
- Depth, Temperature, Salinity, pH, etc.

543,477
Total Data Entries

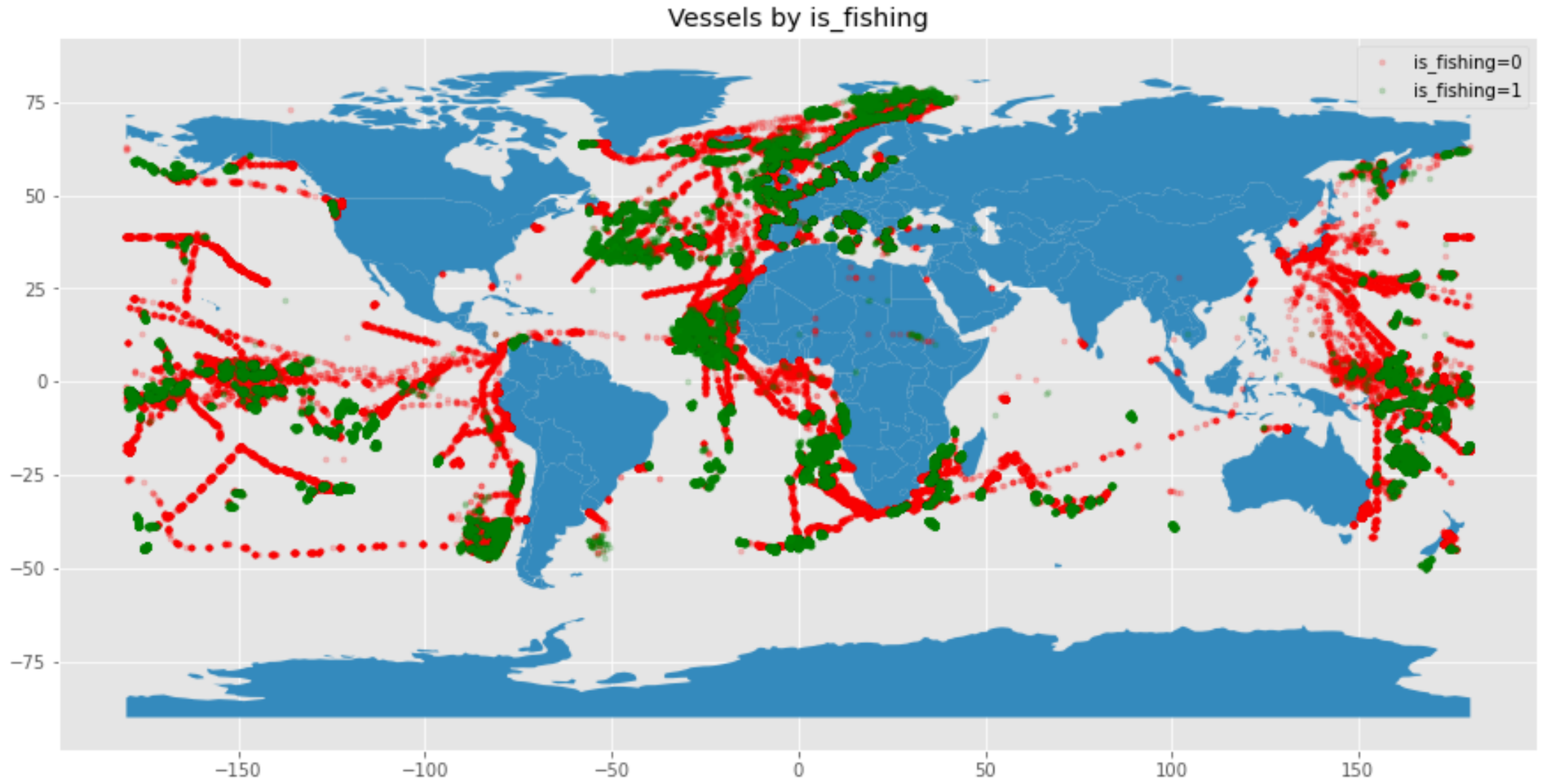
Over 22 Million
Unique Timestamps

Time Period:
2012 - 2016

7 Unique Vessel
Types

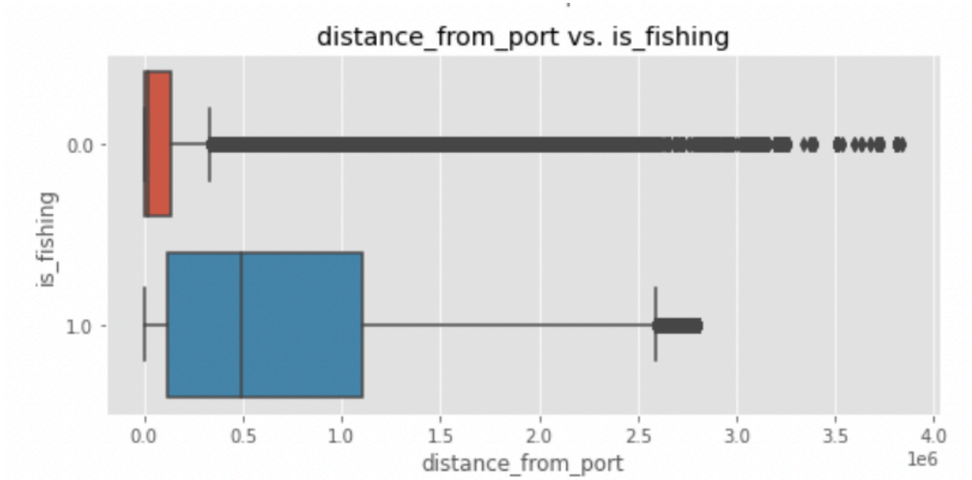
The Data

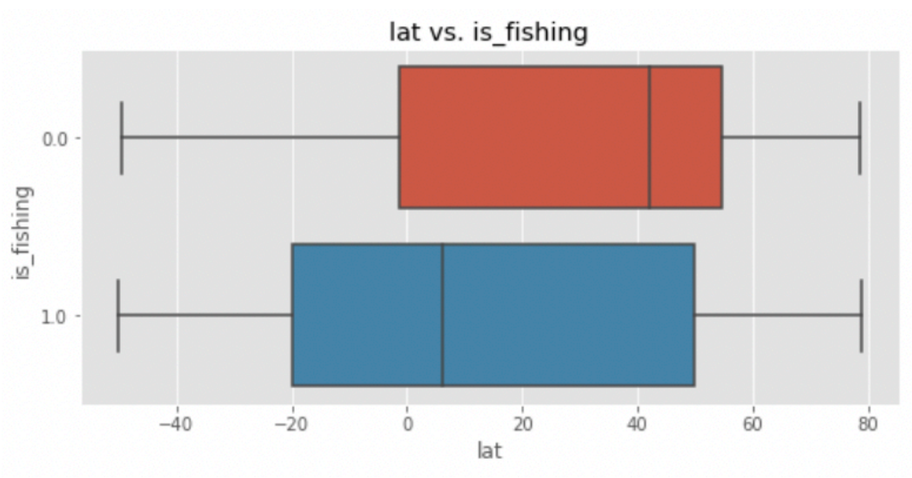
Breakdown of Target

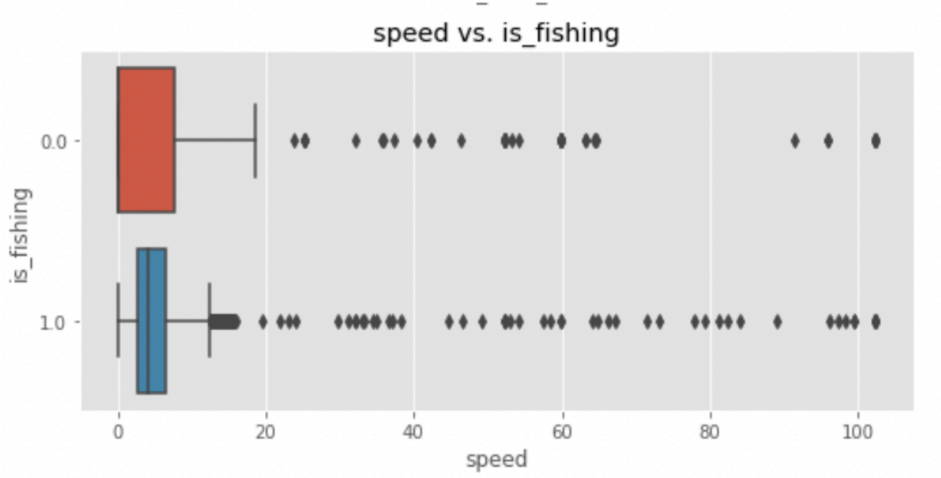


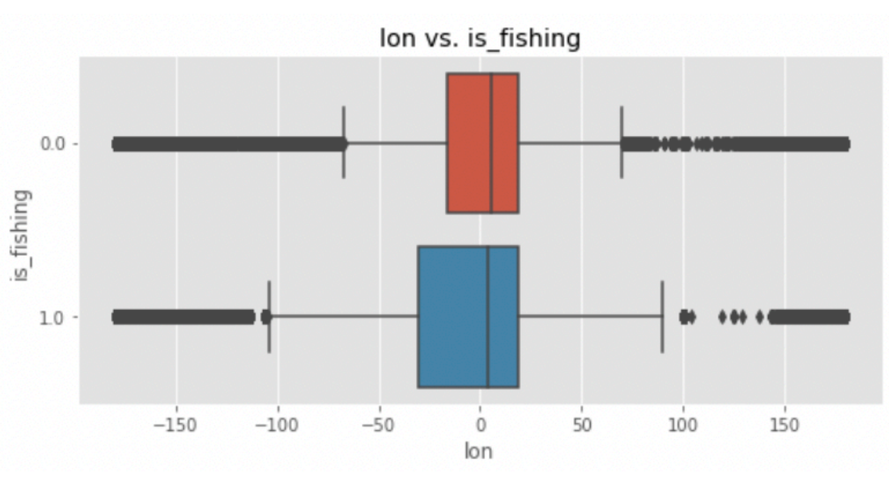
The Data

Select Global Fishing Watch Predictors vs. Target



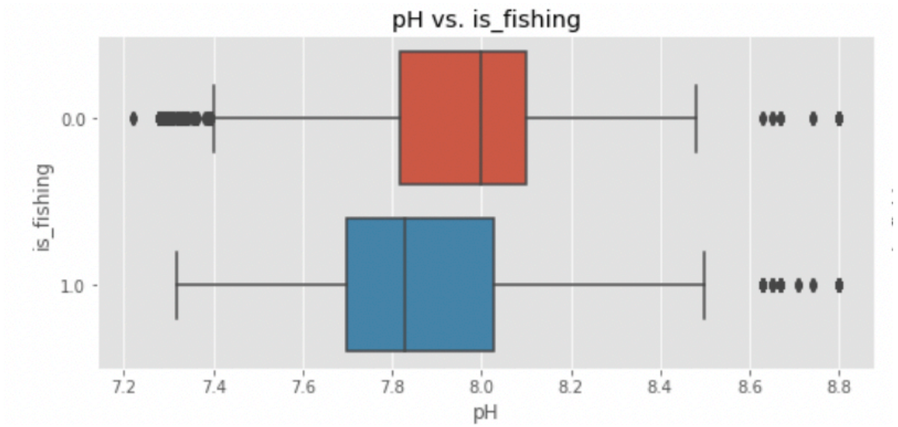


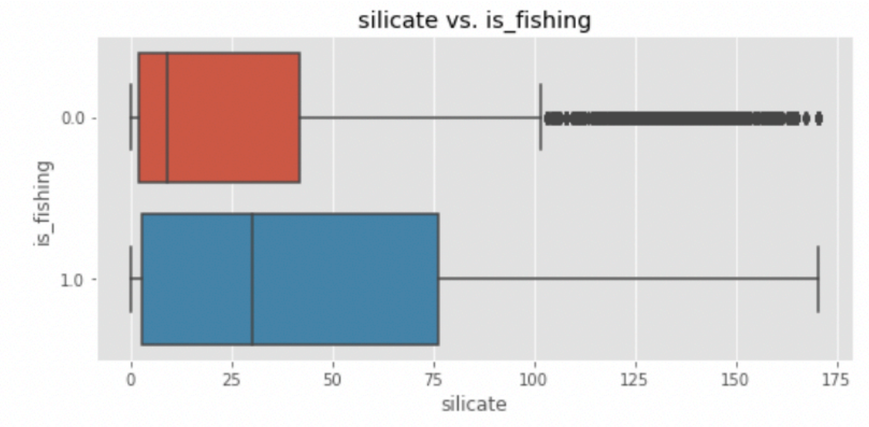


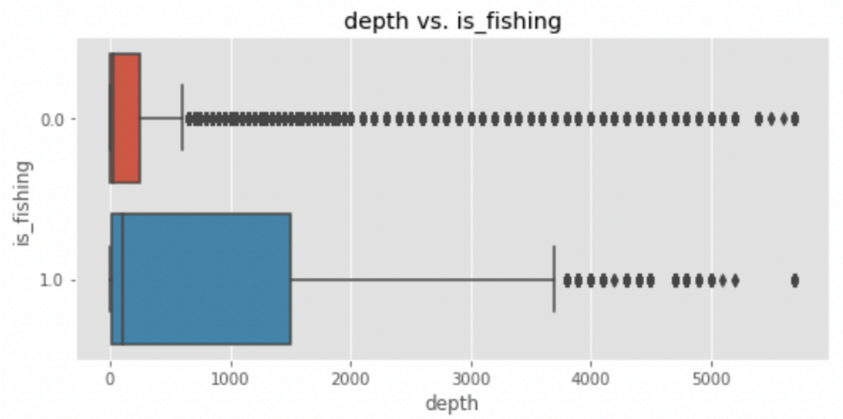


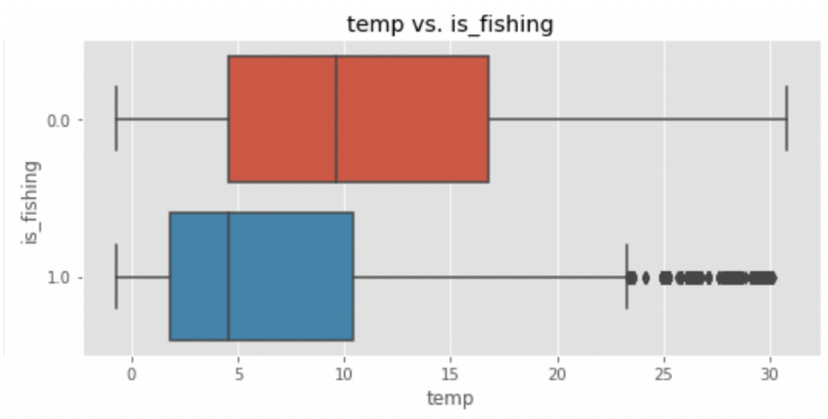
The Data

World Ocean Database Predictors vs. Target









Methodology

Overview of Process

1

- Source data from Global Fishing Watch and World Ocean Database
- Explore Data to Understand Patterns

2

- Prepare Data for Modeling
- Iteratively model using Decision Trees, Random Forest, XGBoost, and Neural Networks

3

- Evaluate performance of each model primarily using accuracy and recall (explained on next slide)
- Fit final model to training data and evaluate with test set

Results and Feature Importance

- Accuracy: 93%
- Model correctly classified 93% of all test data

- Recall: 96%
- Of all data points actually labeled as fishing, our model was able to correctly identify 96% of them