# Optimizing Global Freight: Machine Learning Predictions for Transit Times

By: Atziri Mendoza, Binod Gaihre, Sage Hanson

# What even is Logistics?

- getting the right things to the right place at the right time, using the most effective and cost-efficient methods.

# iPhone Journey

## Think about the journey these take to get to our hands!

Order your iPhone

The journey of the iPhone begins

Step 1. Assembling

Assembled in Zhengzhou China


Step 2. Shipped Across Oceans

iPhone is packages and sent on a freight ship to the US.

Step 3. The Final Mile

iPhones get transported from the port to the distribution warehouse.

# Shipment Steps

## This is how the magic happens!


Step 1:
Warehouse

Step 2:
Ground Transport


Step 3:
Port Arrival

Step 4:
Ocean Transport

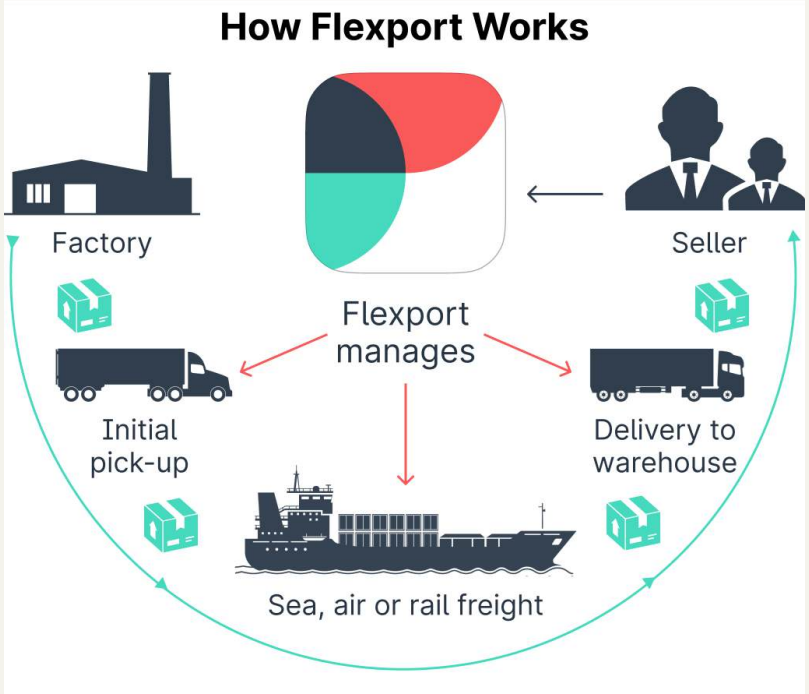

Step 5:
Repeat Steps 3-1

# Logistic Solutions for our client Flexport

# What Is Flexport?

**The problem we are solving:**

- **Freight Forwarding - It's really Freight Email Forwarding**
  - **Unstructured data**
  - **isolated data**
  - **poor visibility**
  - **poor data analytic capabilities**



# What Are Trade Lanes?

## Trans Pacific West Bound (TPEB)

## Transatlantic Westbound (TAWB)

## Far East Westbound (FEWB)

# Key Objectives


### Increase Client Trust

- Companies trust
- Give an accurate time on shipment
- impacts planning, customer satisfaction, cost

### Predict Transit Days for SE Asian Countries

- We want to give our clients a better time estimate- time is money afterall
- Focused Tradelanes: TPEB, FEWB, TAWB

### Make a Random Forest Regression

- This model should be able to predict the total days it takes for our shipment

### Implement our model in Flexport

- With this model Flexport will be able to give a more accurate time estimate than it was once doing before becoming more reliable

# Tariffs

The US is imposing tariffs on other countries goods! This will have an impact on the global transportation of goods that will end up hurting our pockets.

Having data on smaller SouthEast Asian countries will be important because of the tariffs imposed on China. We need higher accuracy.

# Random Forest Regression

# Data Source (Flexport)

## Data Engineering

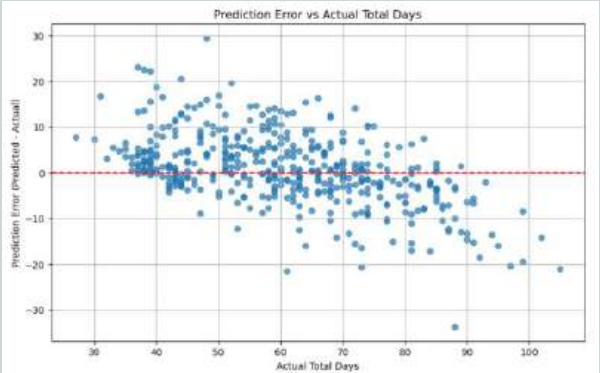
- Geo-DIstance
- Land-Distance
- Weight-to-Volume Ratio
- Outliers Remov

## Random Forest Regression Model

## Predictions

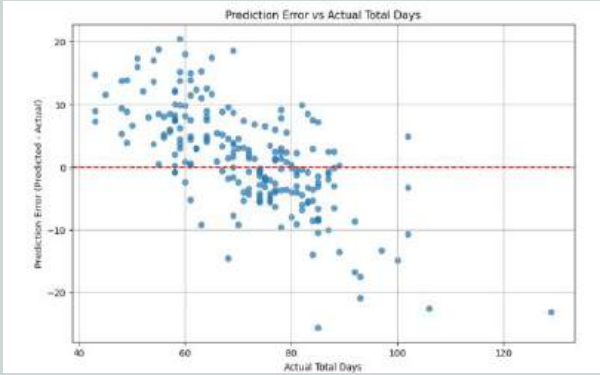
- TPEB
  Actual - 10 days
  Predicted - 5.72 days
- FEWB
  Actual - 13 days
  Predicted - 6.48 days
- TAWB
  Actual - 12.5 days
  Predicted - 8.32 days

# Results



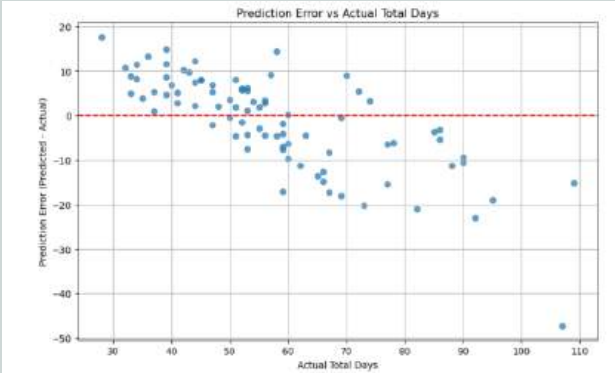

## TPEB Tradelane Results

- Present : 10 days
- MAE: 5.72 days
- R2: 0.7583




## FEWB Tradelane Results

- Present: 13 days
- MAE: 6.48 days
- R2: 0.6081

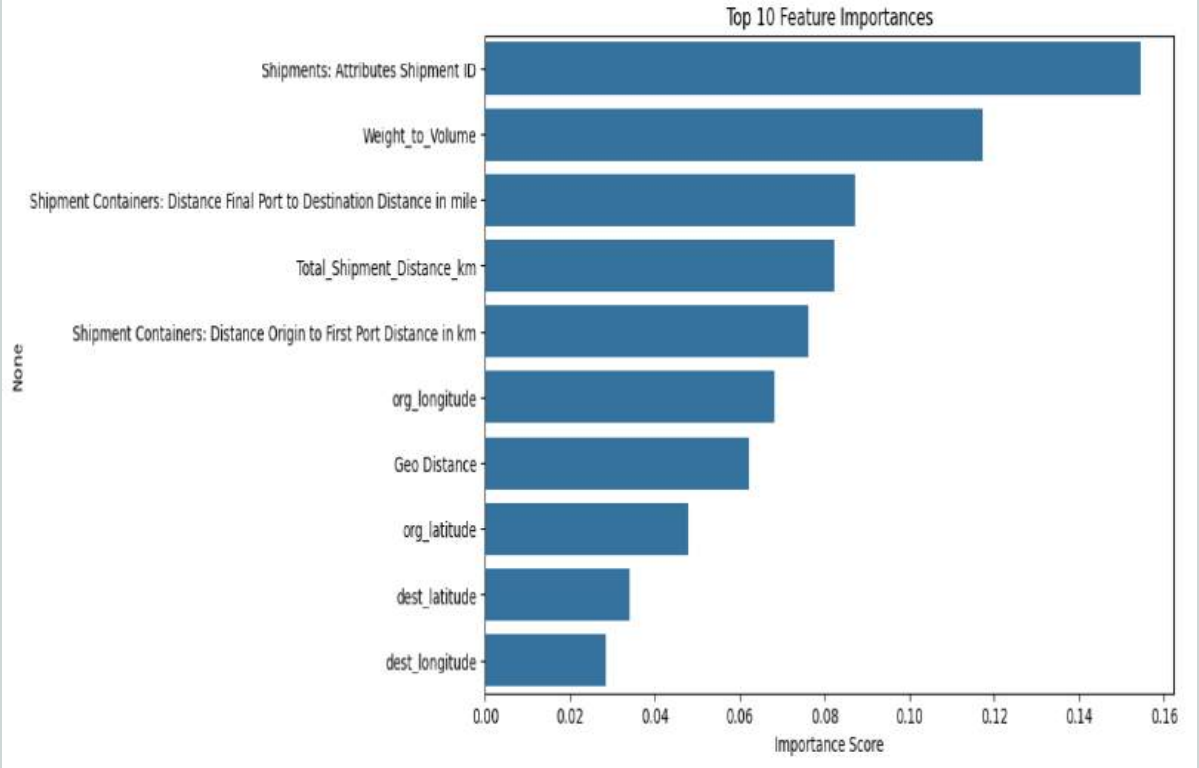



## TAWB Tradelane Results

- Present: 12.5 days
- MAE: 8.31
- R2: 0.6330

# Feature Analysis

- Process of evaluating the variables (features) used in a machine learning model
- **Goals:**
  - Identify most influential features
  - Detect irrelevant or redundant features
  - Understand relationships between features and target
  - Determine whether any features can be removed, transformed, or engineered to improve model performance



# LIMITATIONS

## Limited Data

- 13 weeks of Data
- Model may better predict w/ more data Reducing Mean Absolute Error

## Variables outside the Data

- Unaccounted variables that could cause potential issues in predictions
- Variables not included in our limited data

## Primary Key used as Feature

- Shipment ID is a primary key that was included in the features but we are doubtful that it's actually a predictor
- waiting on Flexport to verify

# Q&A