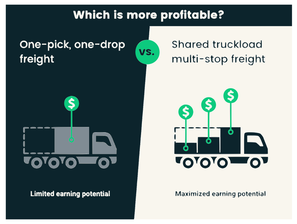
Offer Acceptance in the Freight Industry

Keagan Benson | ktbenson@ucsd.edu[1] Benson Duong | beduong@ucsd.edu[1] Radu Manea | rmanea@ucsd.edu [1] Nima Yazdani | niyazdan@ucsd.edu[1] Matthew Stirling | matt.stirling@flockfreight.com[2]

[1]University of California- San Diego [2]Flock Freight

Background and Problem Space

- **Flock Freight** (FF) deals with different carriers and shipments, acting as a marketplace for freight carriers to place bids on shipping orders needed to be fulfilled, which is otherwise known as a **freight broker**.
- There are two main styles of shipping loads across the country...

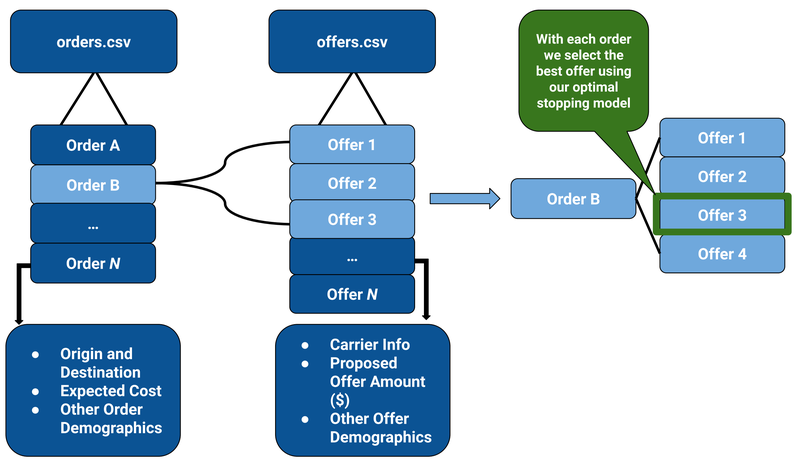


- The goal of Flock Freight is to connect shippers to carriers in order to eliminate inefficiency and waste through shared truckload shipping.

Optimal Stopping Algorithms

- The most famous example of an optimal stopping problem is the "**Interview Problem**". In this problem, a person must choose the best candidate for a position, the person must decide whether to accept a candidate or reject them, and once a candidate is rejected, they cannot be reconsidered.

Data Description

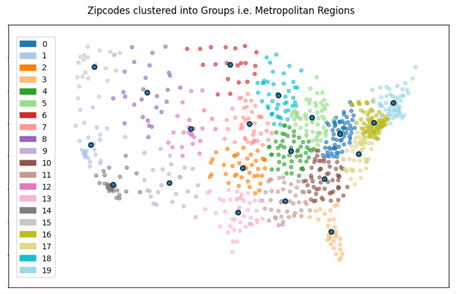


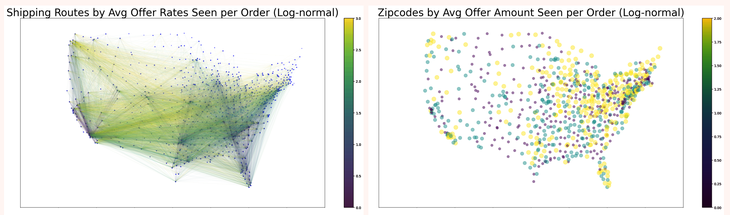
We start in the orders table. Each record is an order, with columns that include order information.	Then, in the offers table, each record is an offer made by a carrier. The columns include information about the offer. Multiple offers may be made on the same order.	We join the offers associated with an order and use our model to select the best offer.

Methods Overview

Geographic EDA: Geopandas was used to preprocess the given data, and extract potentially useful geospatial information for the models to be used.

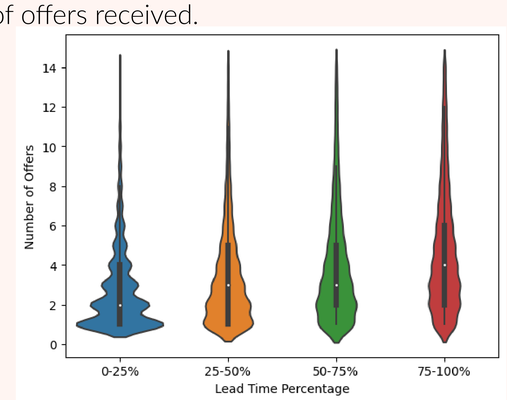
- Creating maps of zipcode nodes as "transportation hubs", and networks formed by shipping routes that connect said nodes.
- Applying **clustering** methods (K-means) on the zipcode nodes to group them apart into distinct "metropolitan regions"
- Collecting census and weather data of the counties that a shipping route crosses through, to approximate the shipping route's road conditions.
- The new geographic features- the clusters of regions, and the shipping route census data of the orders, ended up being used as features for the Avg. and StDev. Sub-models.





Lead Time EDA:

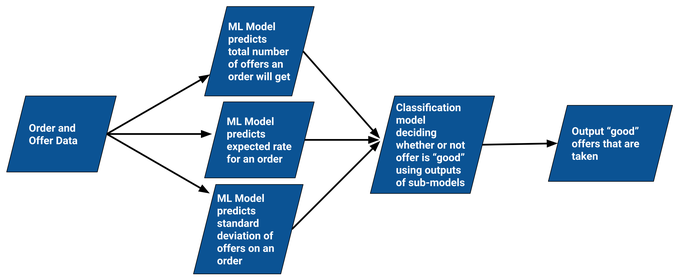
- The "**Lead Time Percentage**" violin-plots show that the longer into the lead time Flock waits to accept, the more offers they will receive. We can use this to weight our feature vector when predicting the number of offers on an order. Orders that were accepted late into the lead time more accurately represent the number of offers received.



Models

Our model comprises of 3 sub-models...

- **Number of Offers Prediction:** For a given order, predict how many offers N it will receive beforehand. A high N allows gambling with "over-rejecting" until a cheap offer comes along.
- **Expected Rate Prediction:** For a given order, predict the avg rate $rate_\mu$ of its expected offers; With each new offer $Offer^i$, $i \in [1, N]$, classify them as below avg (or "cheap") if $rate^i <$ predicted $rate_\mu$.
- **Expected Standard Deviation of Rate Prediction:** The Rate Average prediction model can be expanded with predicting the associated standard deviation $rate_\sigma$, so an offer can be classified as "very cheap" if $rate^i < rate_\mu - rate_\sigma$. Tolerance can be adjusted post-training with scalar λ: $rate^i < rate_\mu - \lambda\, rate_\sigma$.
- **Offer Classification Model:** Using the features from the 3 sub-models and the offer rate, we classified whether the offer was a "good" one and should be accepted.



Summary of Findings

- The **Rate Average Model** used linear regression, R2 = 85%.
- The **Rate Standard Deviation Model** used ordinal classification with Random Forest (classes=2). The ROC AUC score is 60% and the accuracy is 61-67%. The baseline model scored at 56% and lacked any lead-time weighting, or geography-based features.
- The **Number of Offers Prediction Model** has mean absolute error of 2.68. This was 10.2% better than our baseline model and 6.5% better than our non-sample weighted model.
- The **Offer Classification Model** was able to select offers with an average price 4.44% lower than what Flock Freight had selected.

Website!

GitHub!