Nima Yazdani
Keagan Benson
Benson Duong
Radu Manea

Quarter 2 Written Proposal

Background

Flock Freight acts as an intermediary in the shipping industry. Shipping carriers and shippers trust Flock Freight as a “freight broker”, providing a way to easily ship freight across the country, specializing in connecting multiple shippers’ loads with a carrier in what is known as a “Shared-Truckload” (STL). Flock Freight receives orders from shippers, which carriers will bid on to take the load. Flock Freight then decides which offer it will take on a certain order and pay the carrier to deliver the shipment. Oftentimes, the decision to accept or reject a carrier’s offer is very difficult. When trying to maximize margin, we must be wary of which offer we accept because there may be a future offer that is better than the one we choose.

Broad Problem Statement

In this project, the core issue is being able to optimize for the acceptance and reneging of delivery offers by carriers for Flock Freight’s orders.

We'd need to be able to determine that a given offer will be the best one seen compared to all future offers (basically gambling against unseen opportunities).

On the other hand, if we know a current offer isn't good, we'd also need to gamble the risk on reneging it for a better future offer, should one even arrive.

Further complicating the issue is that being indecisive for too long will make carriers uninterested and look for a different freight broker, thus losing clients.

We propose a model (or group of models) that should ultimately assist (to some extent) in a general task where given an order and an offer for it, should be able to determine whether that offer is an optimal selection, doing so in a fast, scalable way within milliseconds.

Narrowed Problem Statement

Further complicating the problem at hand, STLs are also important to consider, as other orders may come in that can be pooled after an offer has already been accepted for the original, individual order. These opportunities to optimize the offer acceptance method are vital for increasing margin and reducing cost. By using information about orders, as well as historical (prior) data on orders and the offers they received, we can build a model that follows previous work in the optimal stopping problem. Our model takes into account the context of how offers are received, as well as using order information to make a prediction on how much an offer should cost.

It is thus essential to provide a model and process that is reliable and scalable to Flock Freight's needs. This means that we must build off of the introduced secretary method for optimal stopping in order to include models for predicting cost, quality of offers, and number of offers. By implementing models for estimating all of these we can improve on the estimated cost already evaluated and the historically chosen offers. With our final stopping method concluded we will be able to write a paper describing how we optimize our baseline model to include more effective prediction methods far beyond what our original model/method accomplishes. Finally, we hope to develop an optimal stopping method that performs better than the historically chosen offers given to us in the data from quarter one.

Our new model should contain the following set of aspects and considerations that were duly noted over the course of Quarter 1 by our teammates:

- The Secretary Method should still regard undershooting as an adequate solution, even if it is not as good as a perfect match; overshooting is a scenario that must be avoided.
- Better Use of the estimated cost to find good offers potentially utilizing a quantile regression model to evaluate a threshold to identify the quality of each offer
- Better training of both estimated cost and estimated number of offers model by utilizing new features and testing out models other than logistic regression to see what performs best
- Employment of the potentially useful priors found and mentioned in the report paper, for example: the over-representation and dense coverage of orders with paths linked to zip codes in the Upper South / Mid Atlantic Region. This is potentially useful for the offer amount prediction model
- Employment of Thursday as a model feature, due to its unusually statistically significant prominence in the time related exploratory data analysis, mainly for tasks related to rate comparison.

Data

The data for this project will be based on the project data already used in the first quarter. This data contains all the sufficient information necessary for our models to adequately deliver results, and is adequate in scope enough - if not unavoidably inextensible any further due to Flock Freight's legal capacity to share company data.

Output

The primary output of this project will be prioritized as a research paper detailing comparison of model metrics and other model explainability.