Supply Chain Logistics Problem

Mechanistic Data Science Project

Aadil Khan

Real-world problem:

A real-world data set has been provided by Intel which has 9216 orders that need to be shipped with their supply chain network. These orders are sent from Intel's 15 warehouses from which they can be sent to 11 different origin ports. From these 11 different warehouses, various couriers are using ground transport or air transport to deliver the shipment to a destination port. These carriers charge different rates related to different weight bands. The dataset has further constraints regarding warehouse capacities, storing costs, etc.

The outcome from this mechanistic data science project would find the optimal warehouses, transport lanes, and couriers for the most cost-effective supply chain. This can be done using the 6 steps for solving a mechanistic data science problem.

Module 1: Multimodal data generation and collection

As mentioned, the data is provided by intel for its chip manufacturing facilities. This data has been provided to the Brunel University of London and is available at https://brunel.figshare.com/articles/dataset/Supply_Chain_Logistics_Problem_Dataset/7558679.

Module 2: Extraction of mechanistic features

The main order list has details regarding the Order ID, Order Date, Origin Port, Carrier, Service Level, Ship ahead day count, Ship Late Day count, Customer, Product ID, Plant Code, Destination Port, Unit quantity, and Weight. Furthermore, the rest of the files contains data regarding the Freight Rates, Warehouse Costs, Warehouse Capacities, etc. This data would be used to formulate the data science project.

Module 3: Knowledge-driven dimensional reduction

Some features such as type of carrier used, cost of transport, and warehouse cost might have a larger effect on the final cost and time of shipping the order. We reduce the dimensions of the problem by only considering the most important constraints and factors.

Module 4: Reduced-order surrogate models

We assume that the problem is linear to reduce the order of the model based on some physical principles.

Module 5: Neural networks for regression and classification

We do the regression analysis to estimate the relationship between the dependent variable/ outcome variable (cost) and the independent variables (warehouse storing rates, carries charges, etc.)

Module 6

This will tie the other modules together. We will be able to find the best routes for the supply of these orders to their customers.