Operations | Retail | Transportation optimization | Supervised Learning

Transit Time Prediction for E-Commerce Fulfillment Optimization and Carbon Emissions Reduction

BUSINESS PROBLEM

Consumers are purchasing an increasing amount of goods through digital channels as compared to brick and mortar and expect fast, reliable delivery. At the same time, society is facing the urgent challenge of climate change mitigation. A global sportswear retailer is investing in improving digital consumer experience while meeting its aggressive 2030 carbon reduction goals. This work focuses on improving critical operations decisions such as which distribution center and shipping speed to use in digital order fulfillment as a method to enhance consumer experience and reduce carbon emissions.

APPROACH

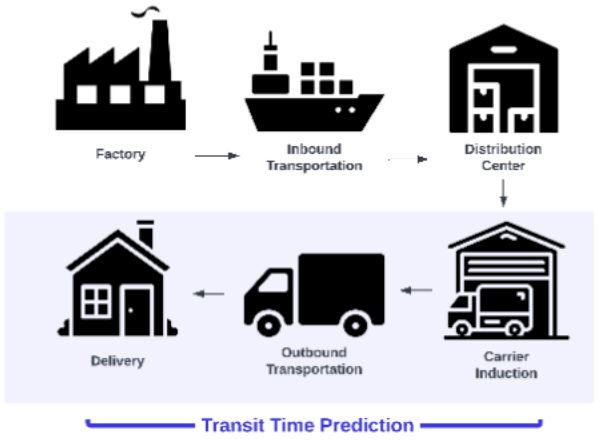
This work focuses on enhancing the accuracy of a highly influential input in the distribution center, shipping speed selection algorithm as a method to improve the quality of those important fulfillment decisions. Specifically, this work uses machine learning to predict transit time, or the number of days between third-party carrier network induction and parcel delivery.

DATA SOURCES

Two key datasets were identified and and procured: an internal North American digital order fulfillment dataset which provides attributes related to a parcel's journey through the supply chain (distribution center onward) and a carrier transit time estimate dataset that is provided by carriers to the organization as a method to communicate service levels.

Data Types and Format

The fulfillment dataset exists in a SQL-like data warehouse and includes categorical and numeric variables. The transit time dataset exists in a set of CSV files and includes numeric variables.



Author: Kathryn Angevine

IMPACT

The transit time prediction model developed in this work is associated with an accuracy of 67%. This is an improvement over current state estimates which are associated with an accuracy of 45%. A counterfactual analysis is conducted to assess the impact of improved transit time estimates on key digital fulfillment performance indicators. On a one month sample, the model enables improved fulfillment decisions. When compared to decisions made using the current state transit time estimates, the model enables fulfillment decisions that are associated with a 4.5% reduction in lead time, a 3% reduction in CO2 emissions and a 1.5% reduction in cost. This demonstrates how the transit time prediction model supports an enhanced digital consumer experience and the achievement of carbon reduction goals.

DRIVERS	Consumers are purchasing an increasing amount of goods online and as a result retailers are competing on and investing in their digital offerings. Additionally, society is becoming increasingly aware of the negative outcomes associated with climate change and businesses are now committing to carbon emissions reduction goals.
BARRIERS	Identifying source of truth data sets, processing data quality issues, and accessing adequate computational resources
ENABLERS	Access to expert-level mentors, strong team collaboration, high organization data fluency, and general data-driven company culture
ACTIONS	Interviewed 30+ stakeholders to understand problem space, formulated project scope, identified and procured necessary data, conducted exploratory data analysis, iterated and experimented with various model types and attributes, tested model on downstream systems, and conducted conversations with software engineering to discuss model productionization
INNOVATION	The approach consisted of many experimental iterations to determine favorable model characteristics. Experiments addressed model type (XGBoost, CatBoost, Random Forest, CART, etc) data set size and time period, winsorization, sampling, feature engineering, feature selection, etc.
IMPROVEMENT	Transit time estimate accuracy is now 67% instead of 45%. On a one month sample, the model enables fulfillment decisions that are associated with a 4.5% reduction in lead time, a 3% reduction in CO2 emissions and a 1.5% reduction in cost.
BEST PRACTICES	Start small and simple, iterate and add complexity over time as needed, identify a small but key group of stakeholders and keep them involved in the process, use visualizations and storytelling to gain buy in, focus on realizing incremental improvement as soon as you can, and don't get caught in an endless cycle of model improvement
OTHER APPLICATIONS	This model can be applied to other businesses that sell physical goods online and are trying to reduce carbon emissions and improve consumer experience through enhanced fulfillment processes.