

XPO Logistics – Data Science Machine Learning Challenge

Background:

Less-Than-Truckload (LTL) is a very complicated business. As part of operations, trips are created to deliver or pickup shipments from the customer locations. Effective trip planning is key for optimal operation of LTL business. For better trip planning, we need to estimate how long our drivers will spend at each customer location (called as Dwell Time).

Requirement:

- Use historical dwell time data at customer locations and create a predictive model to estimate dwell time (**DWELL_TIME**).
- Dataset is attached with email(~13K records).
- Each record in the dataset represents a customer stop.
- Train the model to predict the value of “**DWELL_TIME**” at a customer stop. Consider remaining columns as available input data for a customer stop.
- Make sure you include below steps but not limited to
 - 1) Data Preparation
 - 2) Choosing a model
 - 3) Training
 - 4) Evaluation
 - 5) Model performance improvement techniques
 - 6) Conclusion - approach, findings, how to further improve the model accuracy if more time is provided, etc.,
- Use Root Mean Square Error (RMSE) as a metric to validate your model. Do not spend too much time on improving RMSE. We are more interested to see your approach, methodology, understanding and skills in Machine Learning.

Language:

You can use python or R, which ever you are comfortable with.

Submission Format:

- Please share your code in one of the following formats with proper comments and description
 - .ipynb, .py, .r, .rmd, .html
- File name should be <Last_name>_<First_Name>_Solution
 - Ex: Yalam_Nishanth_Solution.py
- Note: Do not submit trained model.

Time Duration:

You can take up to 3.5 hrs to finish from the time you received the assignment.

Data Glossary:

- **DL,DP,HE,HK,HL,PU,SE,SL** - Activities performed at a customer location. Value represent count of shipments on each activity. 0 if the activity is not performed.
- **WEIGHT** - Total shipment weight
- **MOTORIZED_PIECES_COUNT** - No of pieces to move in the shipment
- **TOTAL_VOLUME_CUBIC_FOOT** - Total shipment volume
- **HAZMAT_FLAG** - Indicator for Hazardous material
- **FREEZABLE_FLAG** - Indicator for Freezable material
- **NODE_VISIT_CNT** - Number of times visited the customer location in the past
- **DOCK_AVAILABILITY_INDICATOR** - Indicator for Dock availability
- **FORKLIFT_AVAILABILITY_INDICATOR** - Indicator for forklift availability
- **CUSTOMER_LOCATION_TYPE** - Customer location type
- **DWELL_TIME** - Time spent at a customer location (Variable to predict)