TRANSRATER

PREDICT TRANSPORTATION COST EASIER

Peter Tian, Shuyang Lin, and Weixi Chen

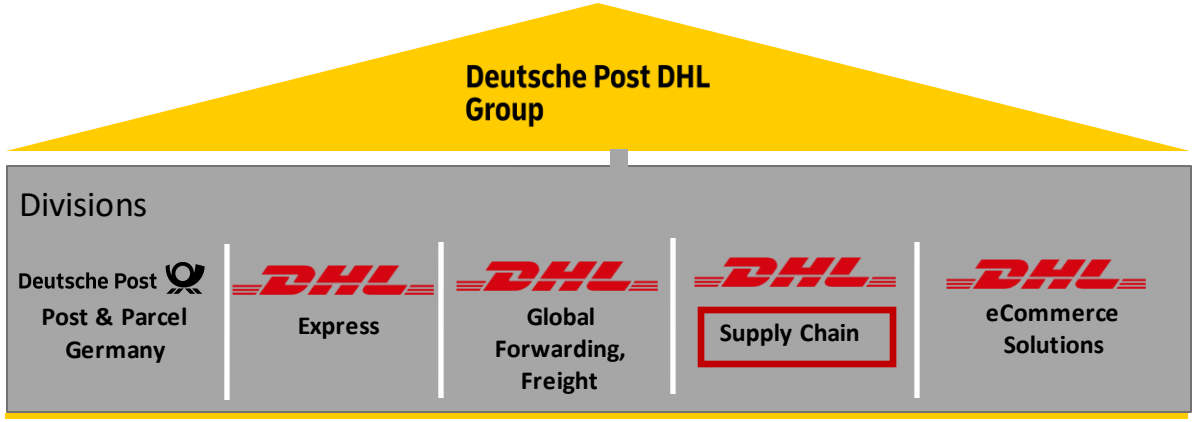
Supported By *Global Operations Science & Analytics Team*

DHL Supply Chain - Excellence. Simply Delivered.

October 2022

About DHL Supply Chain

DHL Supply Chain is a division of Deutsche Post DHL Group with a global network and an extensive logistics portfolio that deals with warehousing, transport and VAS for other companies



DHL Supply Chain is the

#1

Contract Logistics Provider

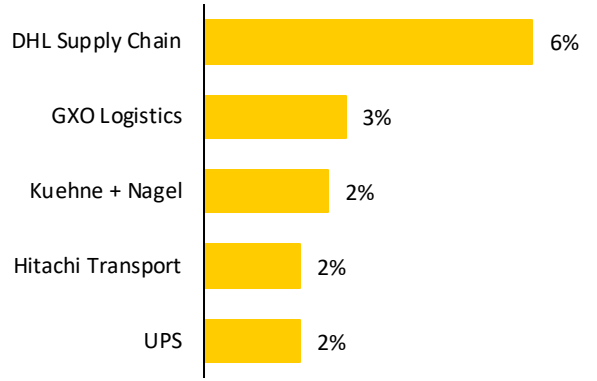
Managing Supply Chains to reduce complexities

50+
Countries Covered

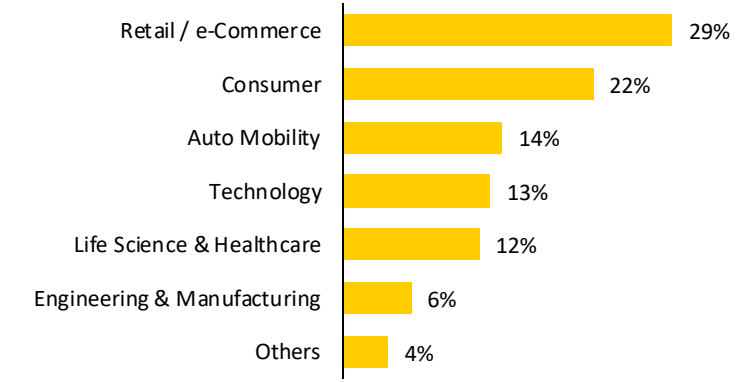
~160M
sq. footage of storage space

~ € 14bn
Revenue

Market Position



2021 Revenue Mix



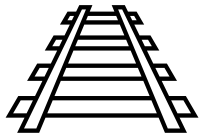
Contents

Project Overview

DHL Transportation Solution Team quotes third-party companies contract prices or spot prices to find feasible transportation solutions for client's shipments

TruckLoad Shipment (TL)

Less-Than-TruckLoad Shipment (LTL)



Intermodal Shipment (Railway)

DHL Transport Solution Team:

- The team considers quoted price in different data frames as a benchmark to find lowest price
- The team will save quoted price and historical shipments in the database

With available historical shipment data and quoted price data:

A Learning models is asked to **predict future transportation rates** by inputting available shipment information and quoted price.

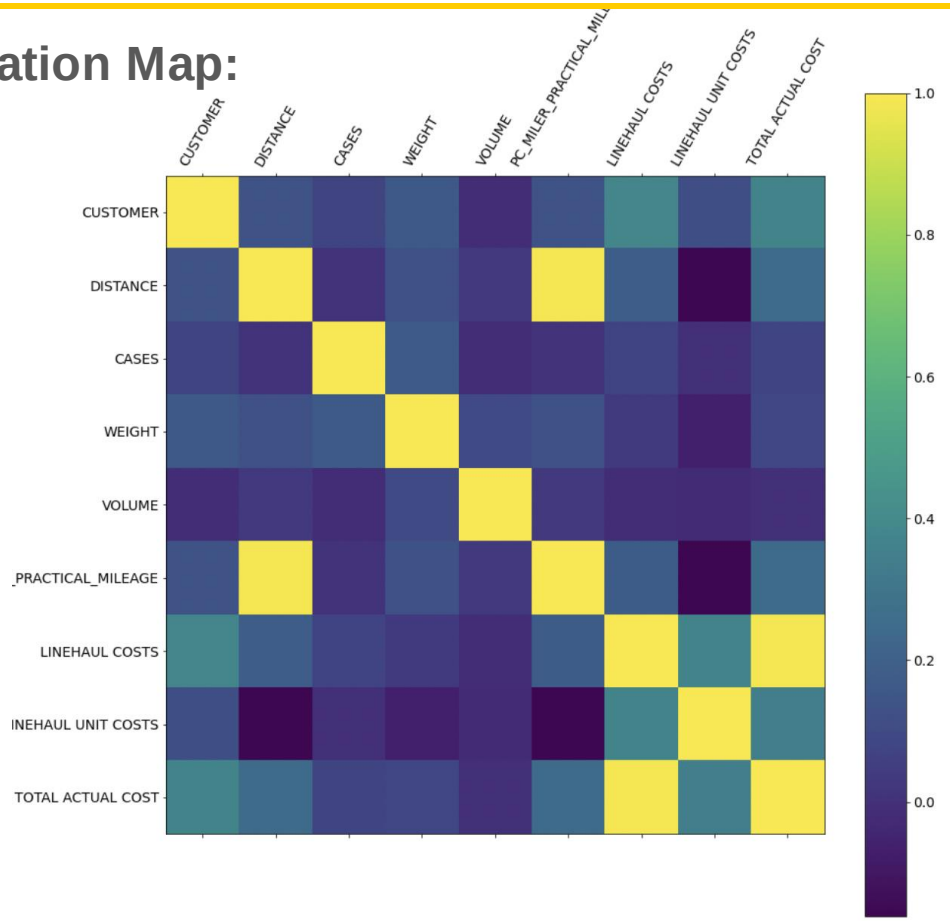
Data Description

Data is provided by DHL Operation and Analytics Team. All data is for study use and has be masked

Historical Shipment Data (MT_Data)
Description: The dataset includes 4 month historical shipment information
Column: 31
Rows: 4,963,508

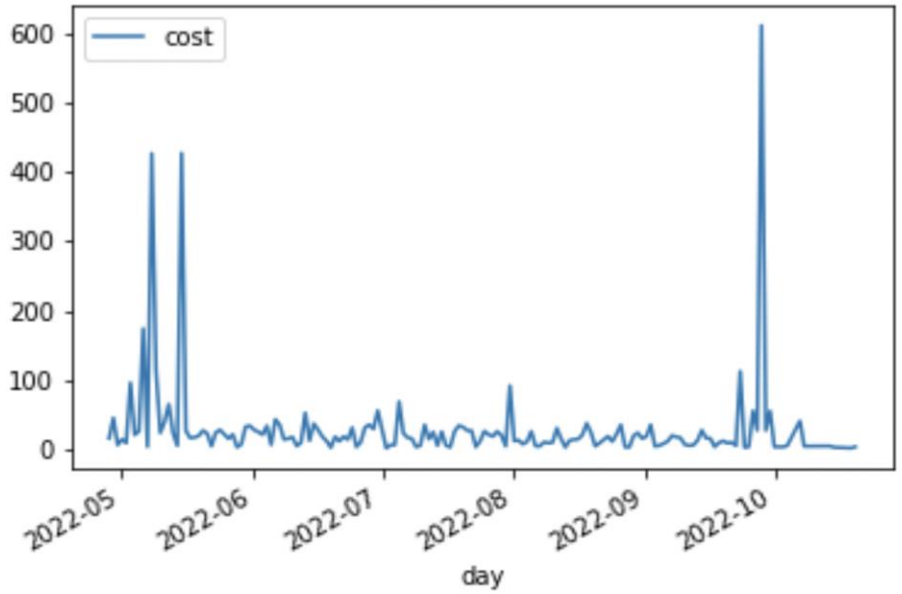
Historical Quoted Price Data (DAT_Data)
Description: The dataset includes 4 month historical quoted price data information
Column: 40
Rows: 63,486

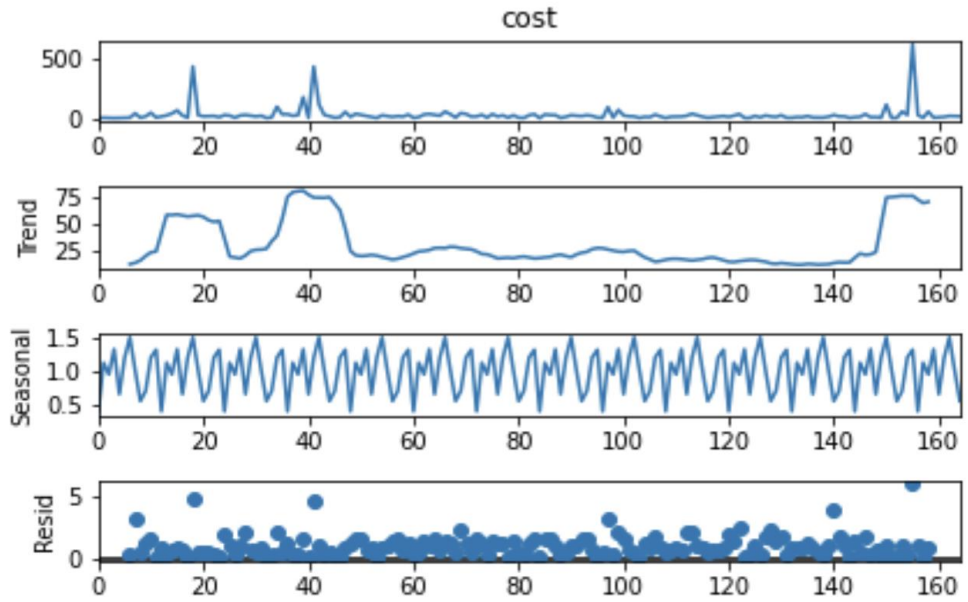
Correlation Map:



Exploratory Data Analysis

Data is provided by DHL Operation and Analytics Team. All data is for study use and has been masked



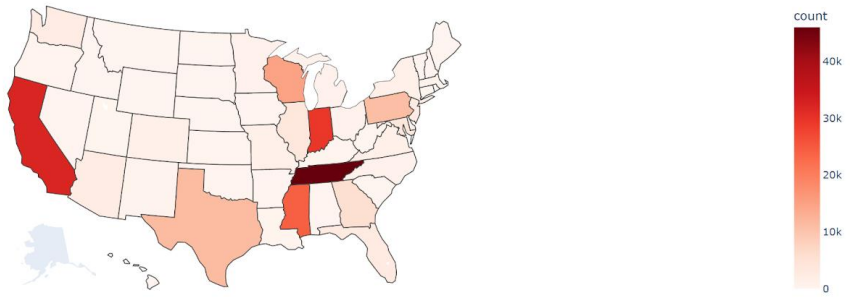


Exploratory Data Analysis

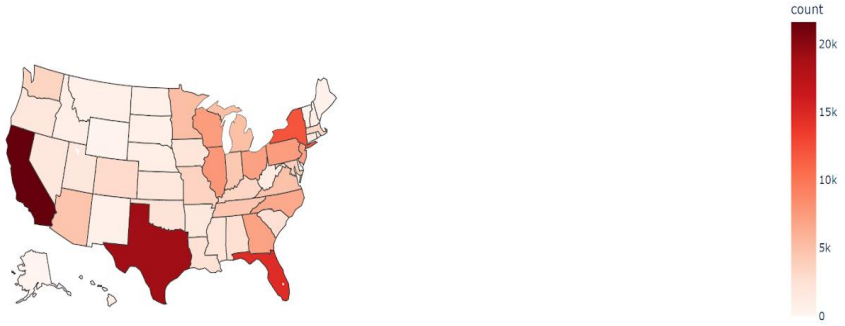
Data is provided by DHL Operation and Analytics Team. All data is for study use and has be masked

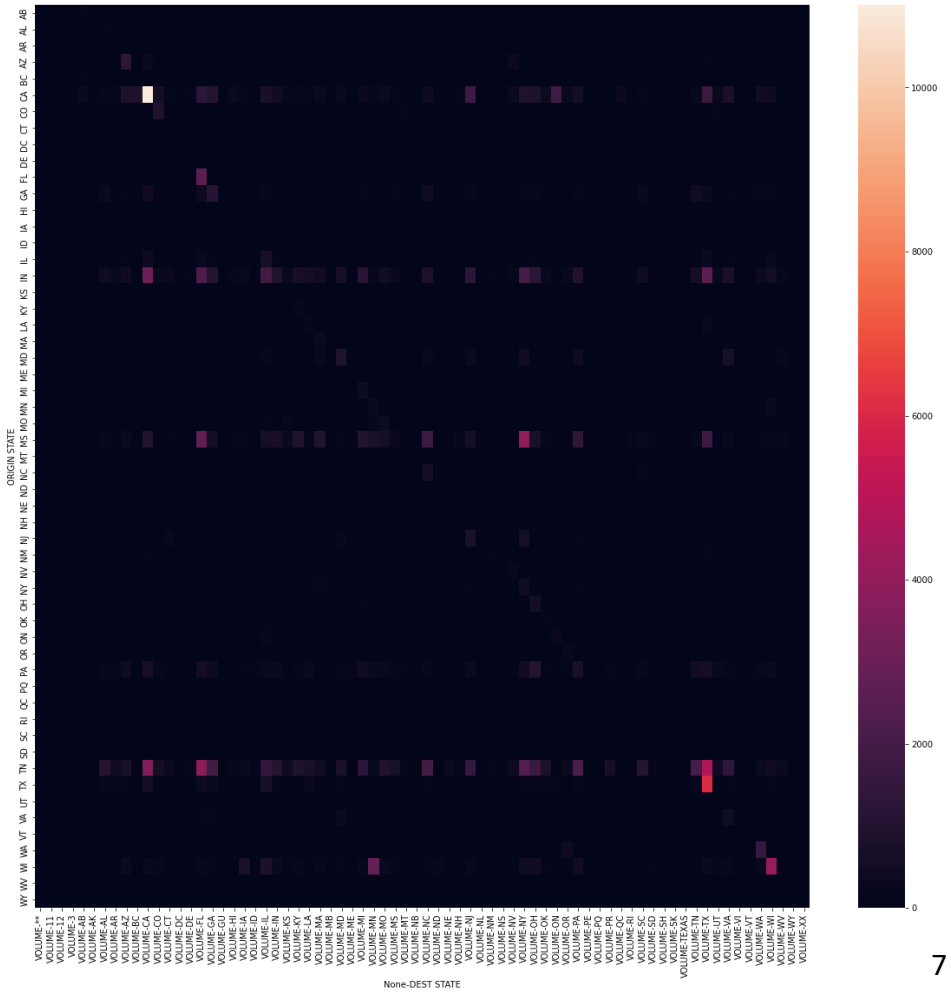
Shipment Map:

cases count of different origin state



cases count of different destination state





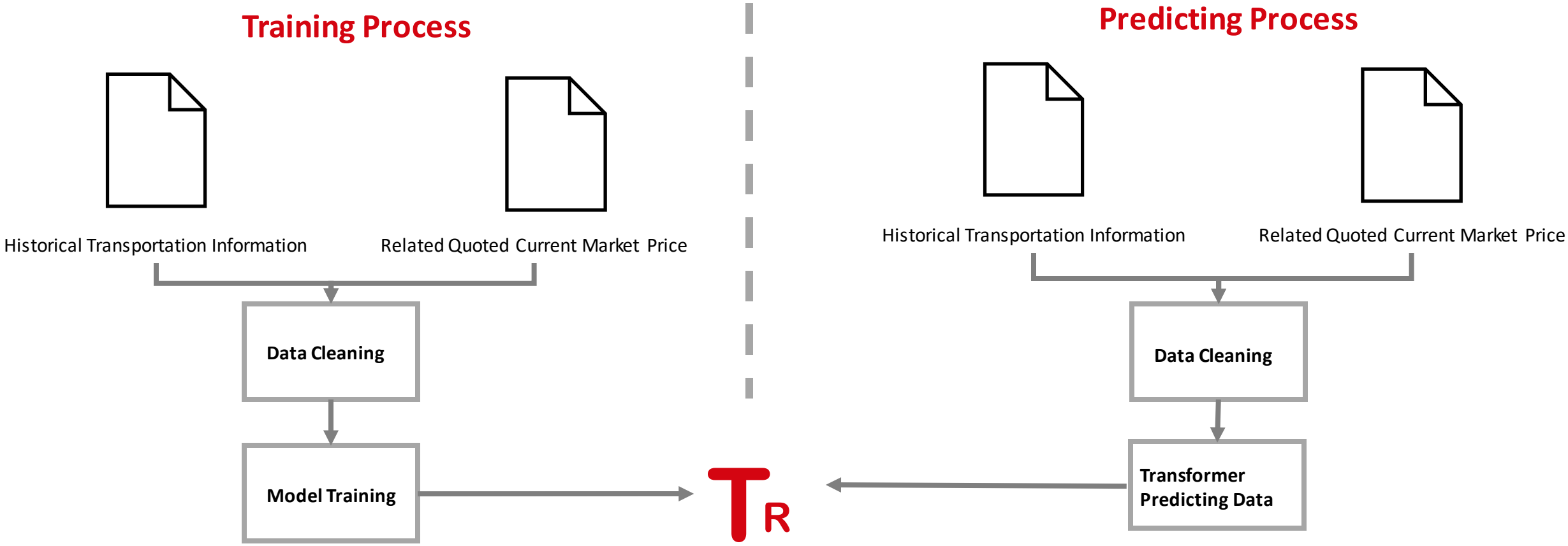
Data Cleaning

Data Cleaning is a critical working before we implementing model

STEPS	DESCRIPTION
Merge two datasets	Using Zip information and Appointed Shipping time as foreigner key to connect two tables
Narrow dataset by location	Collect records only for US and Canada Shipment
Splitting Raw Data	The original data is arranged into three for LTL, TL, and Railway datasets
Removing NA data	If a record has missing information in LINEHAUL COSTS , FUEL COSTS, ACC. COSTS, TOTAL ACTUAL COST, the record will be removed from model training.
Fix Zip related feature	Correctly zip into a correct US or Canada format
Correct Data	Compare Only keeps records whose shipment data is not later than quotes time
Splitting Raw Data	The original data is arranged into three for specific model training

Data Cleaning Working Flow

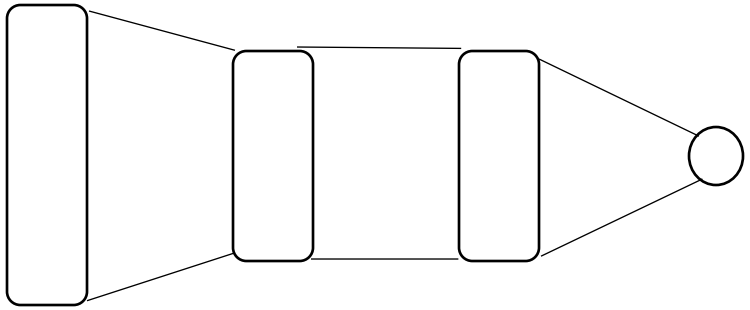
Data Cleaning is a critical working before we implementing model



NN Model Building

We use PyTorch to build a class to to train a sample model

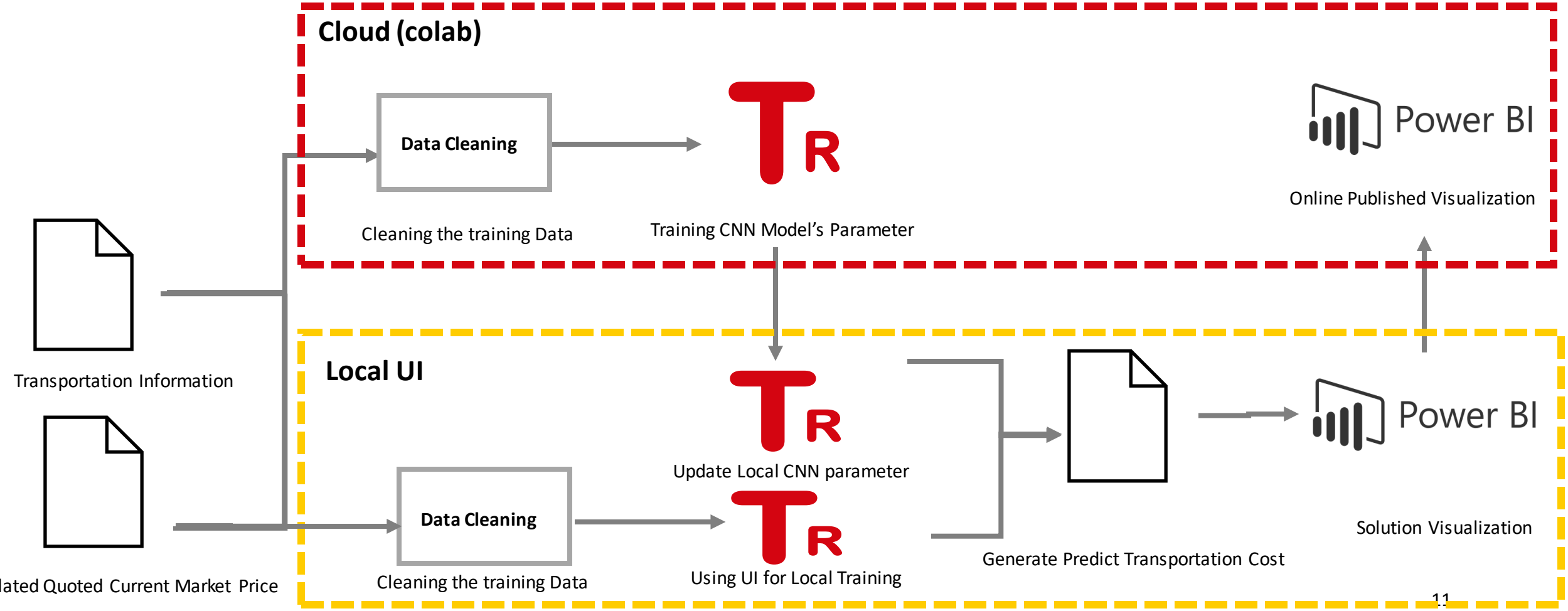
Step 1	Splitting Training data and Test data
Step 2	Building a Neural Network
Step 3	Setting Accuracy Measurement for Loss Calculation
Step 4	Setting Gradient Descent as Model Optimizer
Step 5	Building a Learning Epoch
Step 6	Record Test Error



Hyperparameter and Outcome	
Activation Function	ELU, ReLU
Hidden Layers	2
Hidden Layers Size	size of training row*2/3
Loss Measurement	SMAPE
Learning Epoch	10000

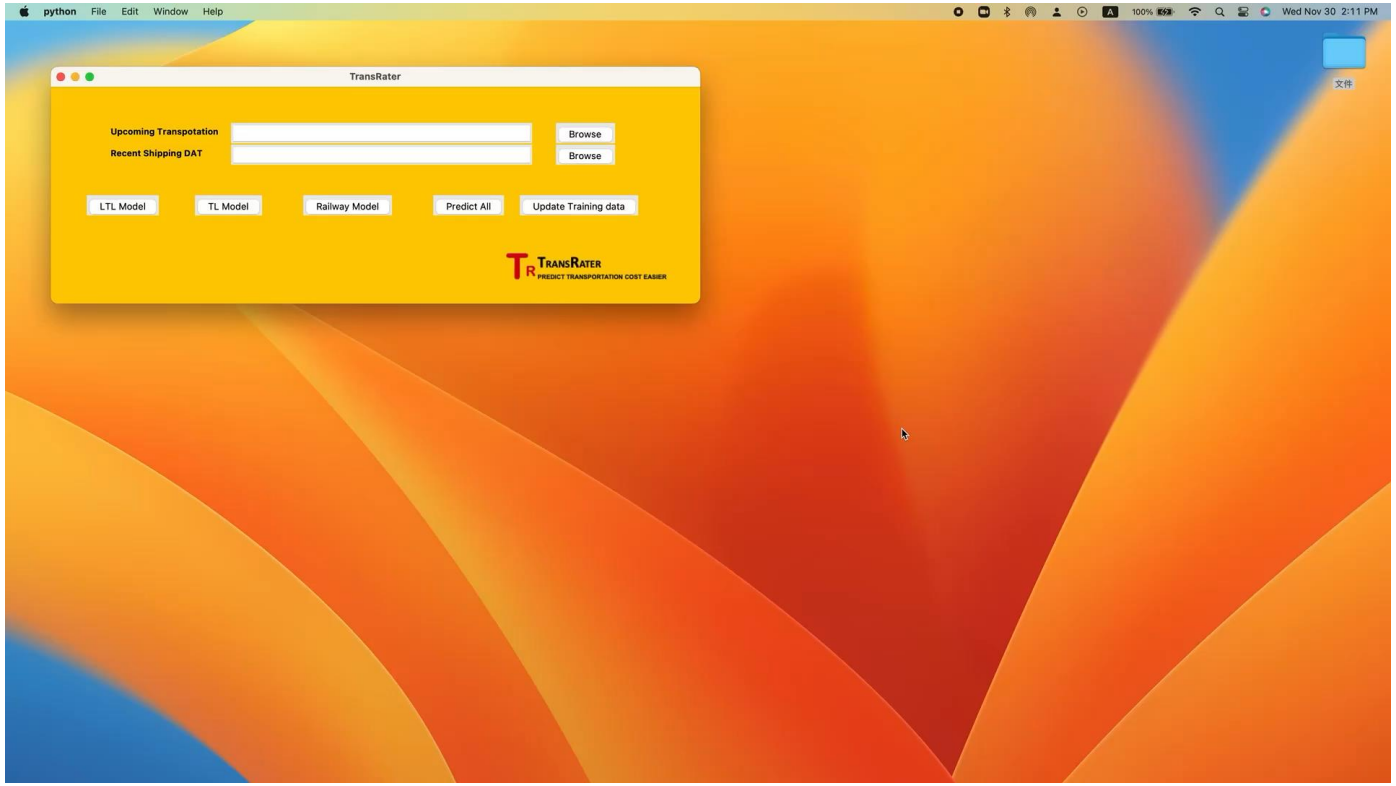
TransRater Completed Working Flow Overview

Rather than complete TransRater for the current class requirement, the complete working flow will include more steps



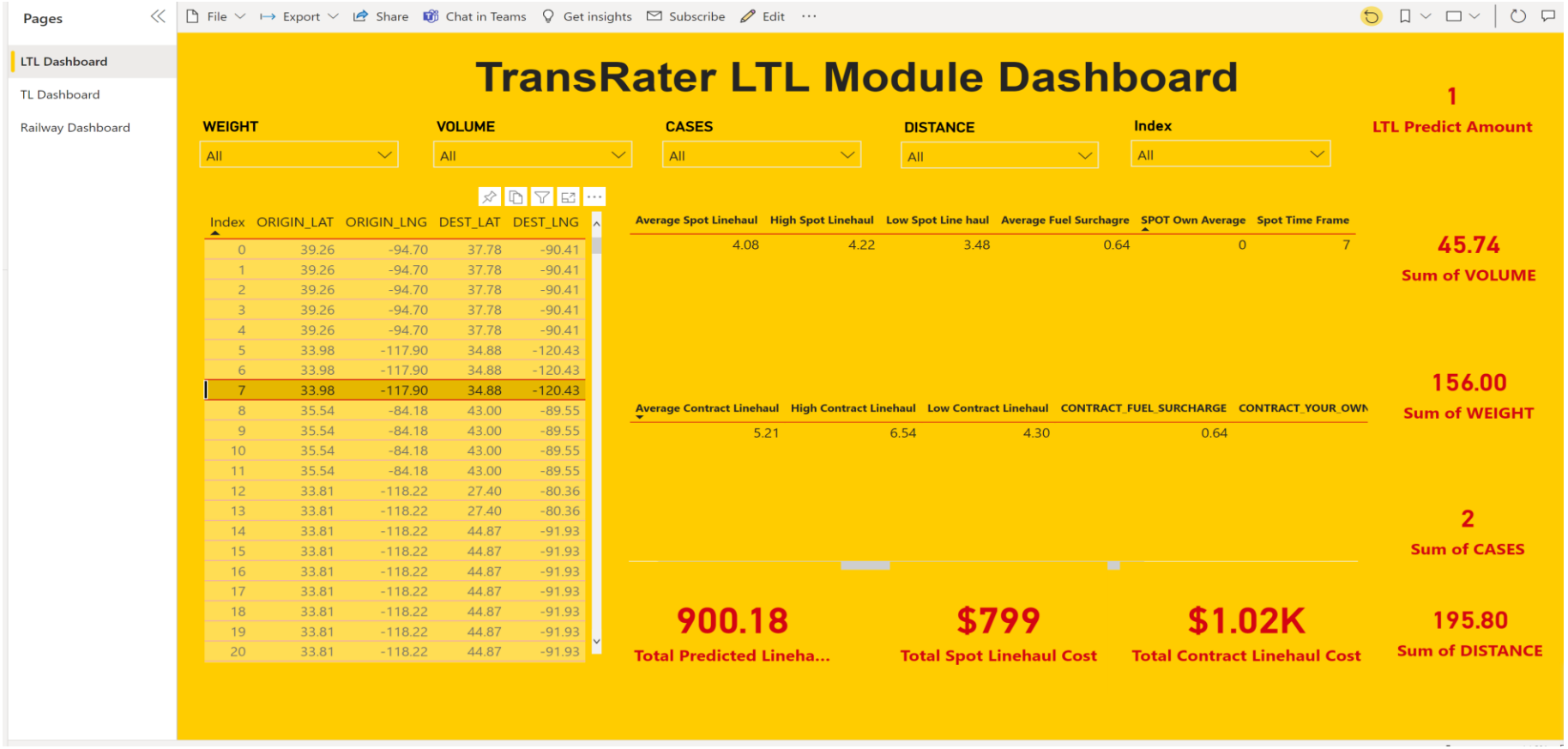
UI Demo

TransRater has a User Interface to help Transportation Solution Teams predict transportation quoted price



PowerBI Visulization

We use PowerBI to build a Dashboard for prediction visualization.



Demo

Challenge

Data Cleaning is a critical working before we implementing model

Challenge	Description
Data Cleaning	• The data size cause data cleaning very time consuming • The data need multiple step in data cleaning • We need multiple different standards to clean training data and predict data
Modeling	• Because of Colab limitation, our model can only learn 1/15 data • While we want the model available for local training, it is almost impossible to implement • Cannot find well brenchmark models for comparsion
UI	• The current version cannot display user each step's running time • The current version does not support customized predicting input • The current version does not support NN model's hyperparameter change in model training

Github

Link: https://github.com/PeterTian96/ML_Project_Transportation_Rate