Box Weight Volume Error Consignment Analysis RIVIGO

Background

Rivigo is a technology-driven logistics company that specializes in transporting consignments across different regions of the country for B2B clients. The company operates through its extensive network of 17 processing centers spread across India.

The process begins with pick-up agents collecting consignment boxes from clients at their warehouse or business locations. These agents update various consignment details, such as ***weight, box dimensions***, consignee name, address, and pin code, through a mobile application, which generates a **consignment note (8 digit unique consignment reference no)**.

Human interventions led to an increase in discrepancies in the weights and volume measurements by the agents while collecting the packages.

Agents mistakenly input incorrect weights or dimensions, such as recording 100 kg instead of 10 kg, or confusing measurement units like 30 x 30 x 20 inches as centimeters.

These inaccuracies led to improper **load planning**, inefficient **vehicle utilization**, and ultimately resulted in **revenue loss** and **extended cash collection cycles** (DSO was increased to 60 Days for some of the clients).

Problem Statement

The aim of this project is to develop a logical and construct a model capable of efficiently identifying consignments with inaccurately measured weights and volumes either during or before the load planning stage.

Based on the Analysis , the product team aimed to implement a QC (Quality Control) Validation System as a new product feature for real time check on consignment inaccuracy

Dataset

1. Approximately 8000 consignments were picked up daily across India, Approximately more than 2 Lakhs consignment in a month.
2. There are more than 1800 active clients spread across 15 major industries including - E Comm, Pharma, Automobile etc. and more than 16000 active delivery pincode. We have taken sample data from the clusters where no of cases of incorrect billIng were registered was very high
3. Mumbai and Delhi were major clusters, we started with 1 month’s data of delhi clusters. For delhi daily average pick up ranges from 1200-1500 consignments

Solution Design

1. **Features** :There was no labeled dataset, hence we started analyzing the sample data set with heuristic approach and developed some logic based on features like - density, length , breadth , heights, weight, types of clients, industry category, no of consignments
2. We have analyzed these features to understand the kind of boxes a client was providing in their consignment.
   Studied **patterns** for clients, used IQR Limits for density , Max Length , Length and weight frequency distribution, Control Charts Methods
3. We have developed logic to segregate the most possible outliers in the dataset , since it's a **classification** problem ,
   a. 0 - there is less chance of having error in weight or dimensions and
   b. 1 - there is a high possibility of having a weight or volume error.
4. **Flag Rate** : We run the heuristic on the previous days dataset and labelled them (*since last days consignments were the only one which were in the system and that can be verified across the logistics pipeline*) and sent the flagged samples to the ground for validations-Flag rate with which we started was ranging in between 15- 18 % (Approx 150-200 consignment)
5. **Hit Rate** : Out of Average 200 consignments flagged for validations- hit rate was observed at 80 %
6. **Refining Logic** : In case of False Positive Cases we again back to revisit the logic and tried to improve the hit rate % ; in this way we labeled the dataset for Delhi and Pune Clusters
7. **Site Visits** : We also visited the company sites , clients warehouses to understand the cause of error and how it can be controlled using tech
8. **FTRI** : After doing 3 months analysis and correction in Consignments we improved the FTRI % (*First Time Right Invoicing a KPI to check how much correct invoices are getting generated at first attempt*)

Product Development, Implementation & QC Team

1. **Validation system**: All the labeled data were passed to DS and Product Team for training the ML models and Automated Validation system feature was launched that helped to detect the wrong consignment at the time pickup and raise validation tickets to the agents for correcting them in real time by re measuring it.
2. **QC Team** : Post product feature launch - have coordinated over the QC performance across the cluster teams and managed the QC Executives, trained QC team on new product features.

Key Logic

1. Industry Level CFT -density - IQR
2. Client Level CFT density - IQR
3. Length Frequency Distribution
4. Weight per box Client Level Distribution
5. Client Type
6. Inch CM Measurement error flag
7. Default Entries (Max Permissible Limits , Highest Length, Lowest LBH entry etc)

Tools Used

Excel , MySQL, R programming

Role and Team Structure:

1. Data Analyst
2. Reporting Manager - Product Manager
3. 2- Executives for Validation follow ups and Consignment Corrections