

Case Study 7: Predictive modelling for Inbound Cartons

Production Support

Background

The client is largest Departmental store retailer in U.S., by revenue, and is one of most high profile company in terms of sales and operations.

Services

Data Analysis and Predictive Modeling, mainly driven from offshore.

Technologies / tools

R
Python
My SQL

Challenges

- Based on data present in application, very less visibility of Inbound cartons in upcoming days/next week. Resulting in extra labor cost or productivity loss.
- Next Day numbers being up to 50% off from expected cartons.
- Visibility of only up to 10 or 20% of next week expected cartons.

Solution

- After the POC for next day prediction model, enhanced model for prediction of next 1 to 10 days was developed using Python
- Extended the model to all DCs.
- Included the Carrier and Vendor performance features.

Benefits

- Main benefits of the solution implemented : Better labor planning
- It Improved the next day carton accuracy from 62% to 86%
- It Improved next week's carton accuracy from 13% to 53%

25%

More Accurate –
Next day

41%

More Accurate –
Next week

**Better
Labor
Planning**

Prediction modelling for Inbound Cartons

How it was achieved ?

Problem

The company hires labourers based on contract, labour planning is done depending on estimates of how many Cartons are going to arrive in DC in the coming days/week. This involved getting scheduled number of Cartons arriving for coming week from their application which acted as benchmark for our project. But the Scheduled was off by around 50% for 1 day prior, and went as low as 10-20% for 1 week prior which lead them to hire labours at last moment depending if higher number of Cartons arrived. This lead to higher loss for company as last moment labours are costly. Also sometimes labours were not utilized fully to lessor number of Cartons arriving than expected. Hence, the objective was to predict the number of Cartons arriving 1 to 10 days forward with as much accuracy as possible to reduce the effective labour cost for company.

Approach

Zensar did the POC first for 2 DCs, providing predictive model for next day. First exploratory data analysis was done to figure out whether data is good enough for predictive modelling and whether sufficient patterns are available. Post data analysis, predictive model was developed and initial results were presented to Macys which were very encouraging, resulting in converting it to full project.

Key features of Solution

- During POC, data issues were found and data was cleansed thru anomaly detection and removal.
- Features (Time based, Schedule based etc) were identified.
- Different algorithms were tried and GBM was applied providing best results.
- Cross Validation and Out of time testing was performed to test the model
- Feature Importance was calculated based on model tuning and testing results.
- Model was extended to other DCs (where DC specific features were also considered to provide better accuracy)
- From next day prediction, it was extend to 1 to 10 days Prediction.

Benefits

- During POC itself, with basic model, more than 10% better accuracy was achieved
- For next day, prediction was improved by 24%
- For next week, prediction was improved by 40%
- Suggestions were provide to client to keep data for longer period as it helps in model to provide better results and self learn from historical data.
- It overall helped client to reduce the labour cost and improve the planning and labour productivity.