Forecasting Demand of Perishable Products Using Machine Learning

Raymon van Dinter

ORL-33806
Data Driven Supply Chain Management
Wageningen University and Research

Abstract

Keywords: Data Driven Supply Chain Management, Machine Learning, Deep Learning, Regression

1. Introduction

Retailers that sell perishable products should manage their supply chain in order to have the least amount or waste or shortages as possible. A retailer in barbecue products was selected to be further investigated. This retailer recorded its demand on perishable products for 4 years. A model was created to simulate the operation of common retailers handling with perishable products. An attempt to improve usual demand prediction was sought after using machine learning models, fit with features of weather data from the Royal Dutch Meteorological Institute (KNMI), and with demand targets.

2. Retailer simulation

In order to investigate the operation of a common retailer of perishable products, a quantitative model was developed. This model kept in mind a daily routine:

- At the start of every day, the order up to level S will be set by predicting demand D. There are various methods to calculate S, which will be explained later in this section.
- Order quantity Q is determined by subtracting the daily starting inventory I from S.
- During the day, customers buy products. This demand is subtracted FIFO-wise from I. If demand is higher than stock, a shortage is counted.
- Next day starting inventory I is determined by adding Q to I FIFO-wise.
- Waste is computed by counting the expired products.
- At the end of the iteration, average shortage and waste are computed.

2.1. Determining D

In this simulation model, two methods of determining demand D were implemented. The first method was simulating D, the second was using a manual input (a `list`), which could be a predicted D using a machine learning model.

2.2. Computation of S

Order up to level S was computed by using the formula:

$$S = \mu_{R+L} + z(\sigma_{R+L}) \tag{1}$$

Where R means the daily ordering and L stands for the lead time. μ_{R+L} is the demand over $R + L$ days, where $z(\sigma_{R+L})$ is a safety stock that is added to the expected demand over review period plus lead time.

3. Prediction models

All demand prediction models are elaborated in this section.

3.1. *Analysis and Choice*

In order to forecast demand, several prediction models were created. These models included:

- Constant demand
- Mean average
- Linear regression
- Decision tree regression
- Support vector regression
- Random forest regression
- Gradient boosting regression
- Multi-layer perceptron regression

3.2. *Parameter settings*

In order to efficiently evaluate each machine learning model, the `GridSearchCV` object from the `sklearn` library has been used. Parameters for each of these models have been shown in

Model	Parameter values
Linear regression	{'normalize': False}
Decision tree regression	{'max_depth': 13, 'random_state': 0}
Support vector regression	{'C': 100, 'gamma': 0.01}
Random forest regression	{'max_depth': 8, 'max_features': 'auto', 'n_estimators': 500, 'random_state': 0}
Gradient boosting regression	{'learning_rate': 0.1, 'max_depth': 5, 'random_state': 0}
Multi-layer perceptron regression	{'alpha': 1e-06, 'hidden_layer_sizes': [10], 'max_iter': 1000000, 'solver': 'lbfgs, 'random_state': 0'}

Table 1: Model Parameters

3.3. *Quantitative comparison*

3.4. *Impact of including weather info*

4. Results

5. Conclusions and Recommendations