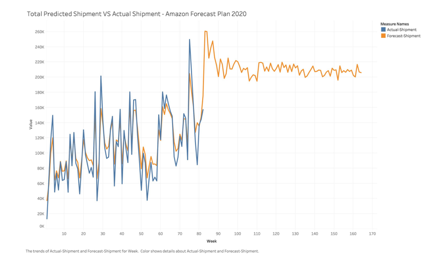
Demand Forecasting with a Segmented Approach

Unilever Team: Robert Morgan & RL Rosqueta
Program Mentor: Prof. Carine Simon



Shijian Chen
MBAn

Jun Jie Ong
MBAn

Stephen C. Graves
Faculty Advisor

About Unilever & Its Mission

Unilever is one of the world's leading suppliers of fast-moving consumer goods. Our products are sold in over 190 countries and used by 2 billion consumers every day. It aims to deliver improvement in activity spends growth impact, customer service and inventory levels.

Problem Statement

Retailers (Amazon & Walmart) order from Unilever cases of products (SKUs), Unilever needs to fulfil shipment orders within the average lead time of 4 weeks. Our project will be forecasting the shipments of cases of SKUs on a retailer level, with a model trained on rolling-based 1-week ahead forecast for Amazon (fast-lane priority) and a 4-week ahead forecast for Walmart. Forecast then is generated till end 2020.

Data Overview & Preprocessing

Internal Shipment

Retailer's Point Of Sales (POS)

Retailer's Inventory

Retailer's Promotions

Retailer	Number of SKUs	Size of Data	Time Period
Amazon	2949	80 weeks	2018W01 - 2019W23
Walmart	3369	105 weeks	2017W11 - 2019W11

We merged four datasets (Shipment, POS, Inventory and Promotions). We removed SKUs with zero shipments for the respective time period for each retailer, and filled in the missing weeks of shipments with 0 for each SKU.

Feature Engineering

Due to the small size of training data per SKU, clustering will help to increase the amount of training data and incorporate some inter-SKU similarities.

Additionally, as the sales team at Unilever highlighted the volatility of the shipments in the Amazon dataset, we have explored various concepts of volatility for clustering SKUs with similar volatility / time series patterns:

- Coefficient of Variation (**2.4** for Amazon and **1.3** for Walmart)
- Number of consecutive zero shipments
- Average rate of shipment

Optimal Lag Term Determination

We created historical features by taking 10 weeks of lag terms for each variable. We experimented with PACF and ACF for lag selection.

To choose the optimal lag terms, we performed LASSO to reduce the number of lag variables in the model and produced better accuracy.

Project Scope & Timeline

Spring 2019 — **Literature Review & Data Collection**
Trial with sample datasets on previously developed models on Target

June 2019 — **Data Preprocessing & Feature Engineering**
Liaising with customer facing teams to understand the features and variable importance

July 2019 — **Modelling & Optimization**
Perform clustering and implementing cluster-while-estimate approach

August 2019 — **Live Testing & Validation**
Testing of forecast results with demand planning system and field trip to Arkansas to validate results

Clustering & Modelling

K-medoid Clustering of Time Series:

As traditional attribute-based clustering method is not sufficient to separate contrasting time series, clustering using indicators of variability, which is inverse of dispersion factor (IOD), is selected to separate product shipment's time series data for model construction.

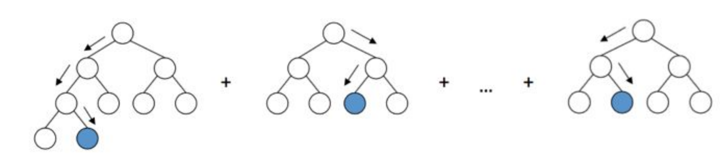


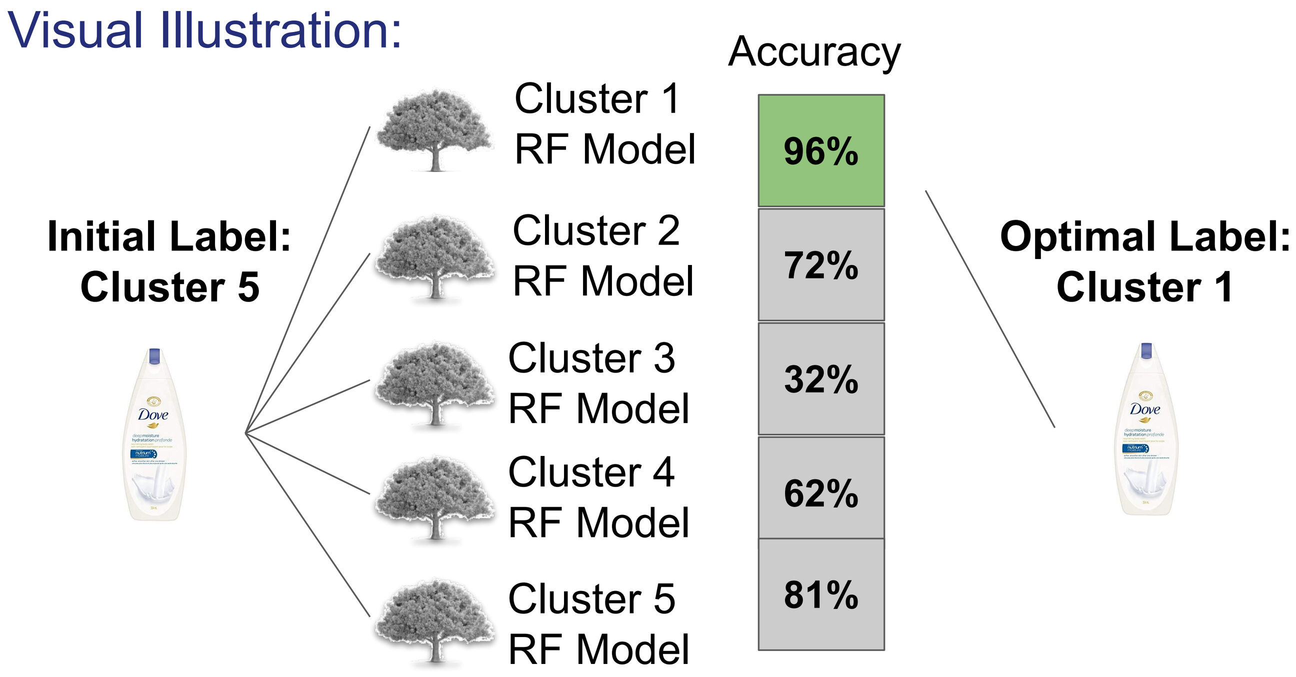
Figure a: Random Forest Model Concept

Model Selection: Random Forest

We tested out four models on Amazon and Walmart dataset, including Random Forest, gradient boosting, linear Lasso regression and Long-short term memory RNN. We focused on tree-based models because of interpretability and the need for the demand planners to intervene with respect to the correlation of variables to shipments. The results for Random Forest is 90% of the times better than the rest. Therefore Random Forest is selected as the main modelling framework for time series forecasting.

Cluster-While-Estimate Optimization & Heuristics

Visual Illustration:



Re-cluster SKUs for the best accuracy in validation set

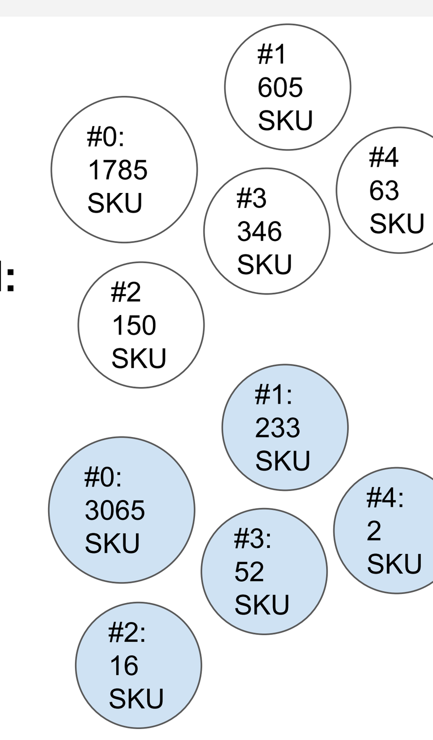


Figure b/c: Amazon Cluster Size (upper); Walmart Cluster Size (lower)

Heuristics Approximation:

Algorithm: Cluster-while-estimate Model (P) Optimal Solution Approximation via iterative re-estimation approach

Input: number of clusters k, product index i from 1 to I, total time stamps index t from 1 to T

a. Fix Zik = Zik_hat(t-1), solve (P) to obtain Fk_hat(t);
b. Fix Fk(t) = Fk_hat(t), solve (P) to obtain Zik(t);
c. Terminate if t = T or Zik_hat(t) = Zik_hat(t-1).

$$\min \sum_{i=1}^{n} L(y_i, \sum_{k=1}^{n} z_{ik} * f_k(x_i)) + \lambda * R(f_1, ..., f_n)$$

$$s.t. \sum_{k=1}^{l} z_{ik} = 1, i = 1, ..., n$$

$$z_{ik} \in 0, 1, i = 1, ..., n, k = 1, ..., l$$

Figure d: Cluster-while-estimate Optimization Formulation (Baardman, and Perakis, Leveraging Comparables for New Product Sales Forecasting, Dec 11, 2017)

Results for Amazon & Walmart Dataset

Amazon has a more volatile ordering pattern and would require both an immediate term (1-week ahead) and short term (4-weeks ahead) forecast. The rationale for using a cluster-while-estimate model is the resemblance of highly volatile time series with addition of new products.

Forecast Accuracy at SKU Level

Forecast Period	Unilever Current Accuracy	Random Forest Accuracy	Gradient Boosting Accuracy	Lasso Accuracy
1 week ahead	35%	55%	40%	36%
4 weeks ahead	N.A.	32%	26%	25%

Feature Importance

Rank of Importance	1	2	3	4
Variable Name	Retailer's Point of Sales	Internal Shipment	Retailer's Inventory	Retailer's Coupons

Cluster-while-estimate Improvement

Most Volatile Cluster (1785 SKUs)	Cluster-then-estimate	Cluster-while-estimate
Out-of-sample Forecast Accuracy	43%	72%

When breaking down the most volatile cluster with 1785 SKUs into 5 sub-clusters based on accuracy, cluster-while-estimate model helped to further improve accuracy by 29%. One explanation is that high volatile item behaves like new products, which is justified in the paper of Perakis et al. 2017.

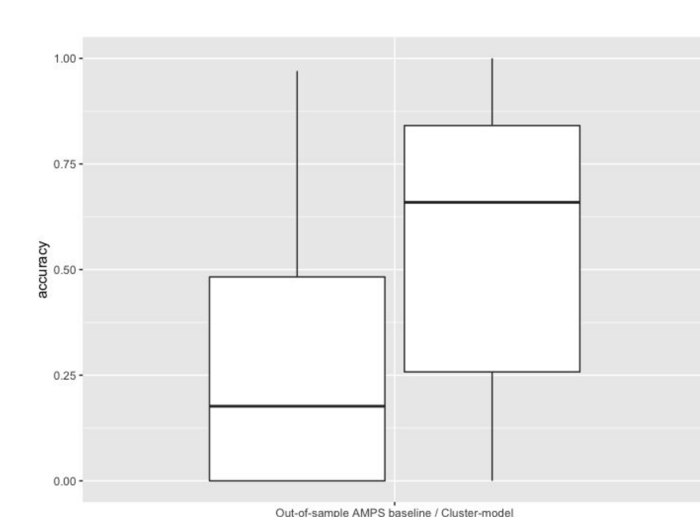


Figure e: Out-of-sample accuracy boxplot comparison between current Unilever baseline (left) and our model (right)

Walmart has a more stable ordering pattern and an average lead time of 4 weeks. Delivering a 4 week ahead forecast is crucial for the downstream demand planners.

Forecast Accuracy at SKU Level

Forecast Period	Unilever Current Accuracy	Random Forest Accuracy	Gradient Boosting Accuracy	Lasso Accuracy
4 weeks ahead	65%	79%	51%	49%

Feature Importance

Rank of Importance	1	2	3	4
Variable Name	Internal Shipment	Retailer's Point of Sales	Retailer's Inventory	Retailer's Sale Price

Business Impact & Contributions

14% - 20% Increase
in forecast accuracy across all SKUs for Walmart and Amazon Dataset

60% Reduction
In Demand Planning time on forecasting process that are rerouted to growth driving process

Expected improvement (to be measured) in service levels with customers

"Previously it took us more than 3 days to review and plan ahead with internal forecast, with MIT's ML baseline forecast, time has been reduced to 6 hours with 60% of all items."

- Unilever Customer Planning Analyst, Supply Chain