A Framework to Define, Predict and Evaluate Slow-Moving Grocery Items

Vivek Avlani, Rahul Gaadhe, Anant Gupta, John Kalathil, Matthew A. Lanham
Purdue University Krannert School of Management
vavlani@purdue.edu; rgaadhe@purdue.edu; gupta499@purdue.edu; jkalathi@purdue.edu; lanhamm@purdue.edu

PURDUE UNIVERSITY
Krannert School of Management

ABSTRACT

The project aims to evaluate the demand of slow-moving grocery items of a national grocery chain. Most of the companies rely on sophisticated machine learning algorithms to predict future sales. In this paper, a set of rules were defined to categorize product as slow moving and appropriate forecasting techniques such as SARIMA, LSTM, LightGBM and time series methods in the prophet package are implemented. Models were evaluated based on their SMAPE values. It was observed that the models' performance varied across the type of past demand of the products. This framework helped the retail company understand slow-moving products and improve response for demand.

INTRODUCTION

Most retailers are selling substantial amount of products, but not every product has the same turnover in terms of sales per week. To minimize capital tied in stocks and minimize cost of space, accurate forecasts of the demand is very essential. However, when a product experiences several periods of low or highly variable demand, assortment decisions and inventory control can be difficult resulting in excessive inventory costs or lost profits. Therefore, small improvement in forecasting demand can translate into significant savings. The prediction is challenging when the demand is intermittent due to a significant proportion of zero values.

The primary research questions that we are trying to address are:

- What is an appropriate way to define sparse demand product groups?
- Which modeling methods perform the best for each defined group?
- Which statistical performance measure is appropriate for each defined group and why?

LITERATURE REVIEW

Study	Algorithm	Results
Forecasting Intermittent Inventory Demands: Simple Parametric Methods vs Bootstrapping	Exponential Smoothing, Croston's Method, Syntetos-Boylan Approximation, WSS Method of Bootstrapping	Parametric methods require less computing power when demands for very large number of SKUs. WSS methods have advantages, however it might not be worth considering the complexity.
Intermittent Demand Forecast with Neural Networks	Croston Method, Time Series, Neural Network Dual and Rate Accuracy Metrix: MAE, MAPE, GMRAE, MASE	Based on realized service levels and trade-off curves between holding and backlog volumes the proposed neural network models out-perform conventional Croston's method and its modifications.
New Forecasting Techniques for Intermittent Demand, Based on Stochastic Simulation	Croston and SBA methods, MA & MD Method. Evaluation Matrix: Mean Error, MAD, MASE	MD method was the least biased and most efficient proposed forecasting technique might increase the quality of the decision making process in enterprises dealing with the problem of intermittent demand.

METHODOLOGY

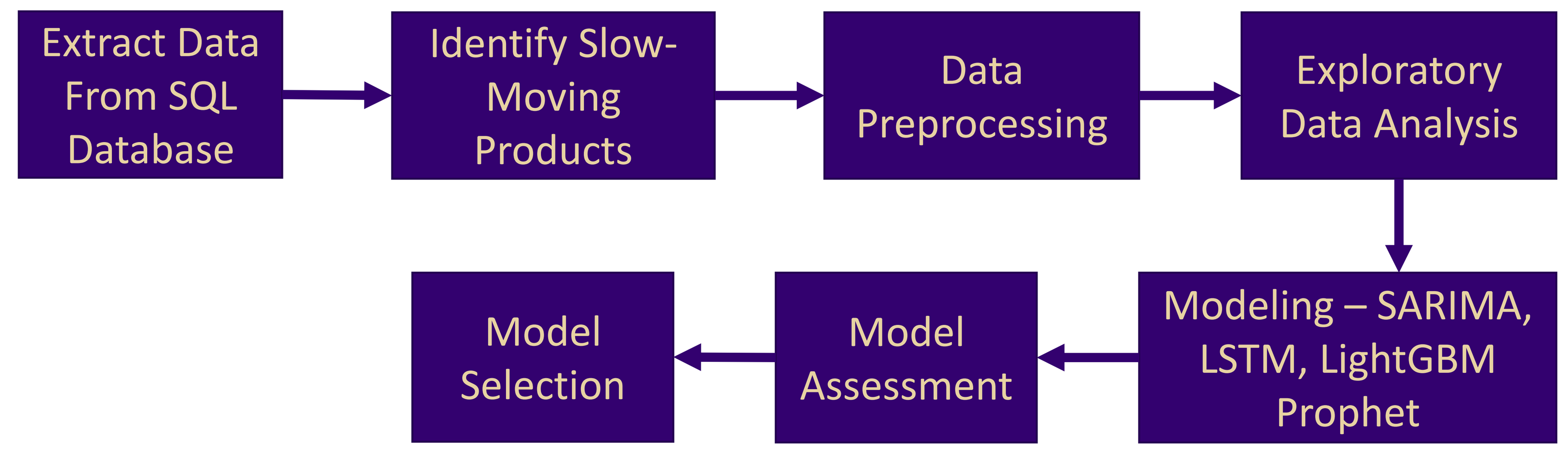

Figure 1: Analytical Workflow

Data - Data is provided and owned by a major retailer in the United States.

Identifying Sparse Selling Products - Products are classified as sparse selling if they don't sell a single quantity for at least 40% of the days in a year. The shaded buckets are identified as slow-moving.

Bucket	% of 'zero quantity sold'	Number of products
Bucket 1	0-20 %	34
Bucket 2	20-40 %	122
Bucket 3	40-60 %	363
Bucket 4	60-80 %	542
Bucket 5	80-100 %	846

Table 1: Products classified into buckets based on quantity sold

Feature Engineering - Based on variable 'quantity sold', multiple features are created such as lagged quantity, moving average quantity and sin-cos circular predictors.

Modeling - Each product is fit with various models including SARIMA, LSTM, LightGBM and Facebook's Prophet package.

Performance Measures - Given that about 50% of the quantities are zero, we have used SMAPE (Symmetric Mean Absolute Percentage Error) to avoid infinite value in MAPE.

RESULTS

Based on the categorization of the products as slow-moving, we found that about 1,751 products were slow-moving which is 91.8% of total products. The figure illustrates the performance of various models across the buckets.

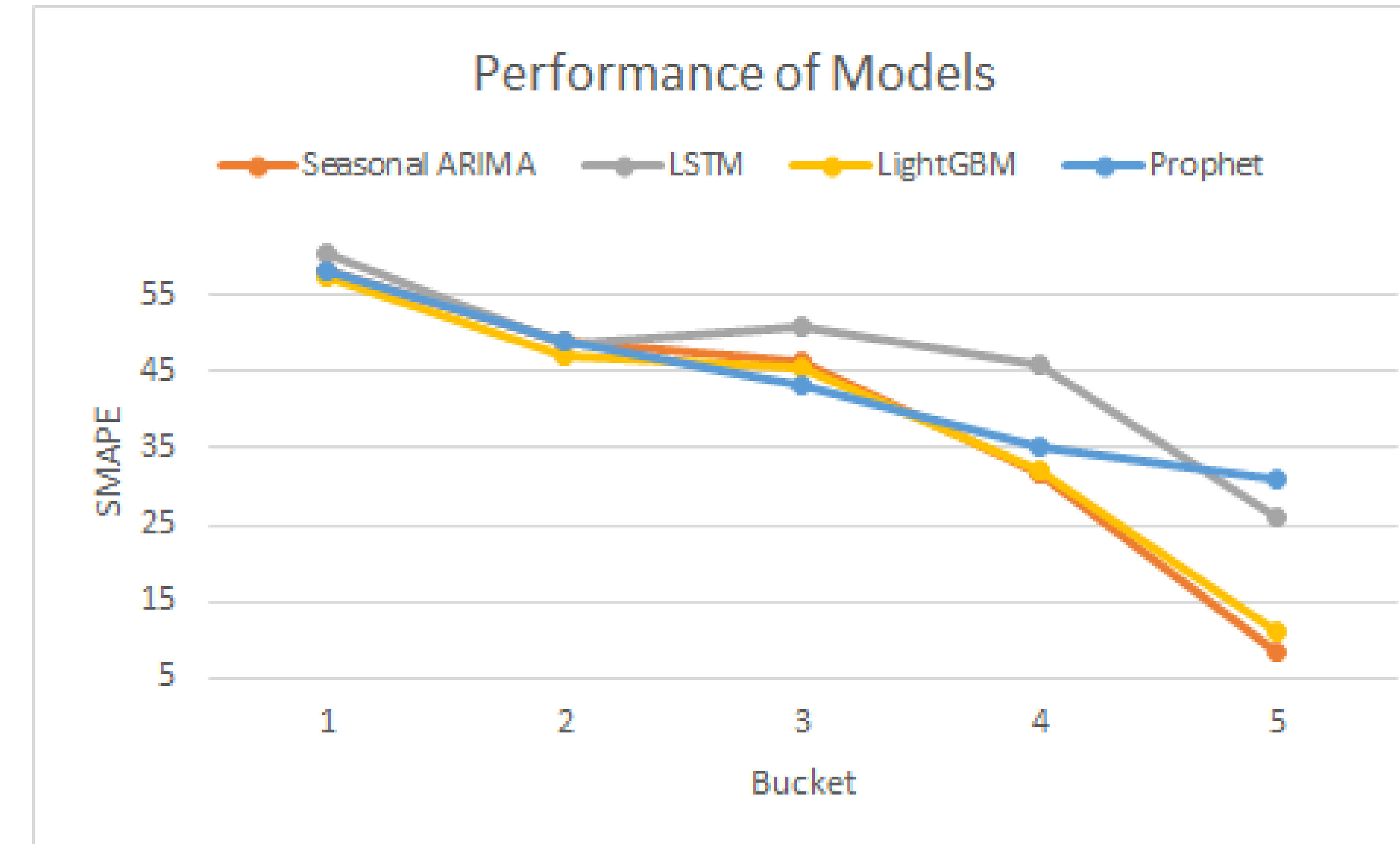

Figure 2: Model Performance

For buckets 1 and 2, all models performed equally well with SMAPE of around 58% and 47% respectively. For buckets 4 and 5, SARIMA and LightGBM performed best with SMAPE of 31% and 10% respectively.

Business Application: Good forecasts will enable the client to reduce inventory holding costs and help in optimal restocking strategies. The client struggles to forecast for products which have low and variable demand. The table below shows the results for best model with maximum accuracy for each bucket.

Bucket	Best Model
Bucket 1	Light GBM
Bucket 2	Light GBM
Bucket 3	Prophet
Bucket 4	Seasonal ARIMA
Bucket 5	Seasonal ARIMA

Table 2: Best model for each bucket

CONCLUSION

We have explored 4 methods to forecast slow moving demand – of these Prophet has given promising results for bucket 3 and Seasonal ARIMA and LightGBM for buckets 4 and 5.

The paper describes a way in which we could create a framework for evaluating and assessing the performance for various degrees of slow-moving goods. One solution never fits all the problems and it is the same case here where we recommend a different model for each bucket to optimize the results.