

TAP TO GO
BACK TO
KIOSK MENU

SAS[®] GLOBAL FORUM 2020

MARCH 29 - APRIL 1
WASHINGTON, DC



USERS PROGRAM

Abstract

Introduction

Methods

Results 1

Results 2

Conclusion

- Sales forecasting is an essential activity for the day to day operations of a retailer. The product forecasts are not only used for the purchase and allocation decisions by the merchants but also used as an input in the assortment optimization process for the stores.
- The goal of this presentation is to represent a non-traditional forecasting approach using machine learning techniques to generate the annual store and product-level sales forecast for a household merchandise and home furnishing retail chain. SAS Viya platform is leveraged for developing the machine learning models and the serverless technology of Google Big Query is leveraged for the data exploration and input preparation from the large volume of raw data.
- In this presentation, we lay out a modeling approach using store and product attributes that are not constrained by the availability of historical sales or time-series data. Machine learning algorithms such as Decision Trees, Random Forest, and Gradient Boosting which could be very easily auto-tuned in the Model Studio of SAS Viya were evaluated and applied for the sales prediction. SAS Visual Data Mining and Machine Learning along with SAS Data Preparation, SAS Visual Analytics, and SAS Model Manager were instrumental in orchestration of the forecast generation process.



Mallika Dey

A MACHINE LEARNING APPROACH FOR RETAIL SALES FORECASTING

Abstract

Introduction

Methods

Results 1

Results 2

Conclusion

Introduction

- Retailers are usually dependent on the availability of historical sales data and a time series modeling approach for generating future forecasts.
- This standalone approach is inadequate in predicting the sales for the new products and requires additional and often manual work of finding the right substitutable products.
- We demonstrate a machine learning-based approach to predict product sales forecast using product attributes, store attributes & demographics. The approach is applicable for new as well as existing products.
- Using SAS Viya Visual Data Mining and Machine Learning tool, we explored several machine learning techniques and applied the best fitting model for our use case.

Objective

The **Objective** of this e-poster is to show the applicability of the non-traditional machine learning approach in predicting new and existing product sales for retailers using SAS Viya.

Machine Learning Approach for Retail Sales Forecasting



A MACHINE LEARNING APPROACH FOR RETAIL SALES FORECASTING

Abstract
Introduction

Methods

Results 1

Results 2

Conclusion

Methodology: Data Preparation

Summarize Data
– store-sku level

Add attributes
and previous
year's sales

Define target
variable – current
year's sales

- Data Preparation: POS data for the last 3 years is joined with the store and product attributes for each store-product combination. Each year's sales are aggregated in a different column.
- Demographic data for stores is used to create store clusters based on customer profile and used as an attribute.
- A store-product level table is prepared by joining the columns for sales and the attributes and imported in SAS Viya VDMML.

Methodology: Model Training

- Different models are trained for different departments.
- Latest year's sales are used as the target variable and past years' sales and attributes are used as input variables.
- For new products, separate models are trained using only the attributes as input and latest year's sales as the target.
- Following types of models are trained and the test error is compared to select the best model:
 - Decision Trees
 - Random Forest
 - Gradient Boosting
- The best model for each department is used to generate the prediction for all store-product combinations.

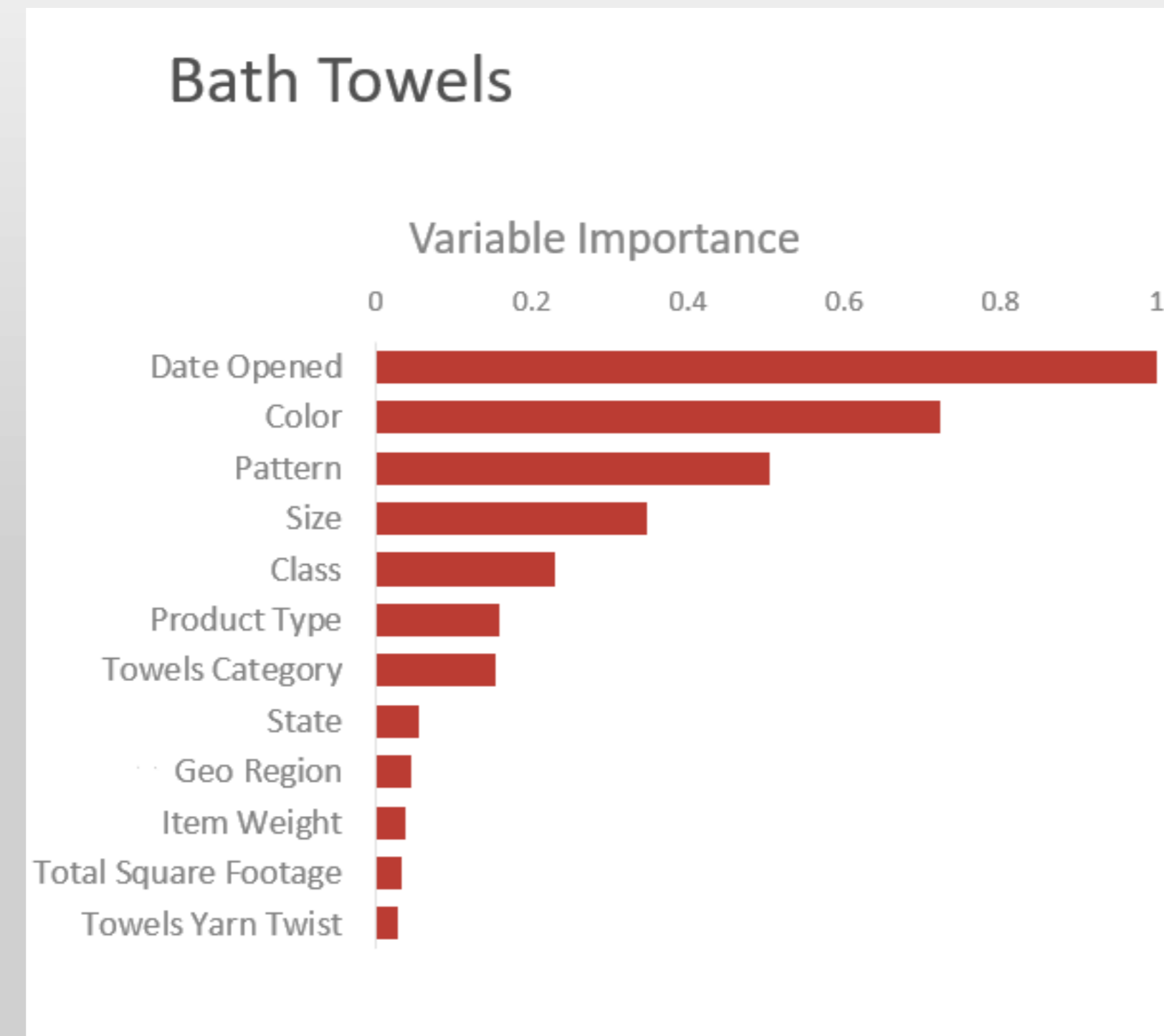
Train Random Forest,
Gradient Boosting &
Decision Trees
models

Select Best
model having
the lowest
RMSE for test

Predict sales for
all products
using the Best
model

A MACHINE LEARNING APPROACH FOR RETAIL SALES FORECASTING

- Abstract
- Introduction
- Methods
- Results 1**
- Results 2
- Conclusion



Results: Model Accuracy

- SKU level machine learning model accuracy ranged from 70% to 85%.
- For the existing products, model accuracy was compared with time series model accuracy.
- Accuracy of the Machine Learning model was either comparable or better by a few percentiles.
- For the new products, the accuracy rate was lower than that of the existing products.

Variable Importance

- Variable Importance chart from the model results helped us understand the important attributes of driving sales.
- Insights could be drawn from the variable importance results. E.g. for the bath towels category, the color of the towel is one of the most important factors for driving sales followed by size.



A MACHINE LEARNING APPROACH FOR RETAIL SALES FORECASTING

- Abstract
- Introduction
- Methods
- Results 1
- Results 2**
- Conclusion

Downstream Consumption of Predictions

- In addition to better accuracy for existing products' forecast, the machine learning approach made it possible to generate new product forecasts.
- Predictions for all store-SKU combinations are used in the store assortment optimization process.
- The models generate an overall yearly forecast for each store-sku combinations.



Weekly Forecast Using this Approach

- This approach can be used to generate weekly forecasts as well.
- Time series model can be trained in SAS Viya to generate a weekly seasonality index.
- The index can be applied to disaggregate the yearly forecast generated by the machine learning models.

A MACHINE LEARNING APPROACH FOR RETAIL SALES FORECASTING

Machine Learning Approach Helps Us Overcome Some of the Challenges in the Traditional Time Series Forecasting Approach

- Machine Learning approach can be used to generate forecasts for new products without waiting for a long time period for sales history to be generated.
- Cases when historical data is available, stand-alone machine learning approach or a combination of machine learning and time series approach leads to better accuracy of the forecasting solution.
- For use cases where the weekly or monthly forecast is not needed (e.g. allocation process for retailers), machine learning is a viable forecasting solution that offers decent accuracy.
- Success of the machine learning approach depends on the availability and quality of the product and store attributes data.

References

Leo Breiman, Jerome H. Friedman, Richard A. Olshen, and Charles J. Stone (1984). Classification and Regression Trees. Wadsworth International Group, Belmont, California

Leo Breiman. Random forests. Machine Learning, 45:5–32, 2001.

Real Carbonneau, Kevin Laframboise, Rustam Vahidov. “Application of machine learning techniques for supply chain demand forecasting”. European Journal of Operational Research 184 (2008) 1140–1154

Kharfan, Majd; Vicky Wing Kei Chan; “Forecasting Seasonal Footwear Demand Using Machine Learning.” (2018). Applied Science in Supply Chain Management at the Massachusetts Institute of Technology, DSpace@MIT - digital repository for MIT's research

The background of the banner features a scenic view of the Washington Monument at dusk, with a vibrant sunset sky in shades of pink, orange, and blue. In the foreground, there is a stone-lined canal with cherry blossom trees in full bloom on the left side. The text is overlaid on a dark teal rectangular area in the center.

SAS[®] GLOBAL FORUM 2020

USERS PROGRAM

MARCH 29 - APRIL 1 | WASHINGTON, DC | #SASGF

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Other brand and product names are trademarks of their respective companies.