

Anjali Hegde, Gajender Saharan, Jeffrey Chittaranjan, Ruthvik Arepalle, Shruti Gupta, Matthew A. Lanham

Purdue University Krannert School of Management

hegde11@purdue.edu; gsaharan@purdue.edu; jchittar@purdue.edu; rarepall@purdue.edu; gupta592@purdue.edu; lanhamm@purdue.edu

ABSTRACT

Most fashion retailers place inventory orders and investments well in advance. However, poor product performance can result in excess inventory which need to be liquidated through heavy promotions. It can hit their bottom line and tarnish their reputation.

We have developed a machine learning model that can be used to predict whether a newly launched product will be successful by fusing traditional sales data with cannibalization features. Our model can also be used to identify the right product characteristics, seasonal and geographic trends of successful products for future product launches.

INTRODUCTION

Currently, most fashion retailers liquidate non-performing products by giving high discounts which affect long term health of the brand and reduce profit margins. According to a 2018 report by Celect and Coresight Research, markdowns cost US non-grocery retailers \$300 Billion in revenue annually

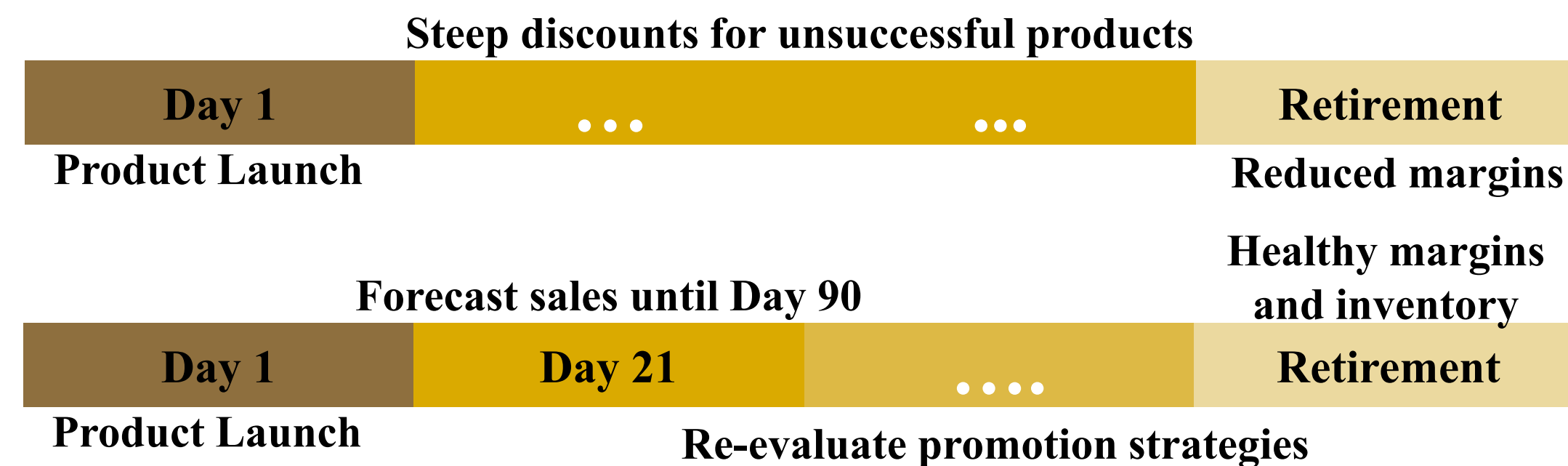


Fig 1. Current and proposed product lifecycle timeline

Research questions:

- Are the first 3 weeks of a product's launch predictive of determining 3-month sales ?
- Do high performing products have certain characteristics (colors, product categories, geographical trends, price, margins) that the less selling products lack?
- How much does the cannibalization impact the sales of newly launched products?

DATA

- SKU** Features such as size, color, product characteristics, launch date, retirement data, category, etc.
- POS** Transactional level data from POS in full-line stores – includes order date, SKU, store ID, price, margins, discount
- Web** Transactional level data retailer's website – includes order date, SKU, price, margins, discount
- Store** Geographical features of the full-line stores

METHODOLOGY

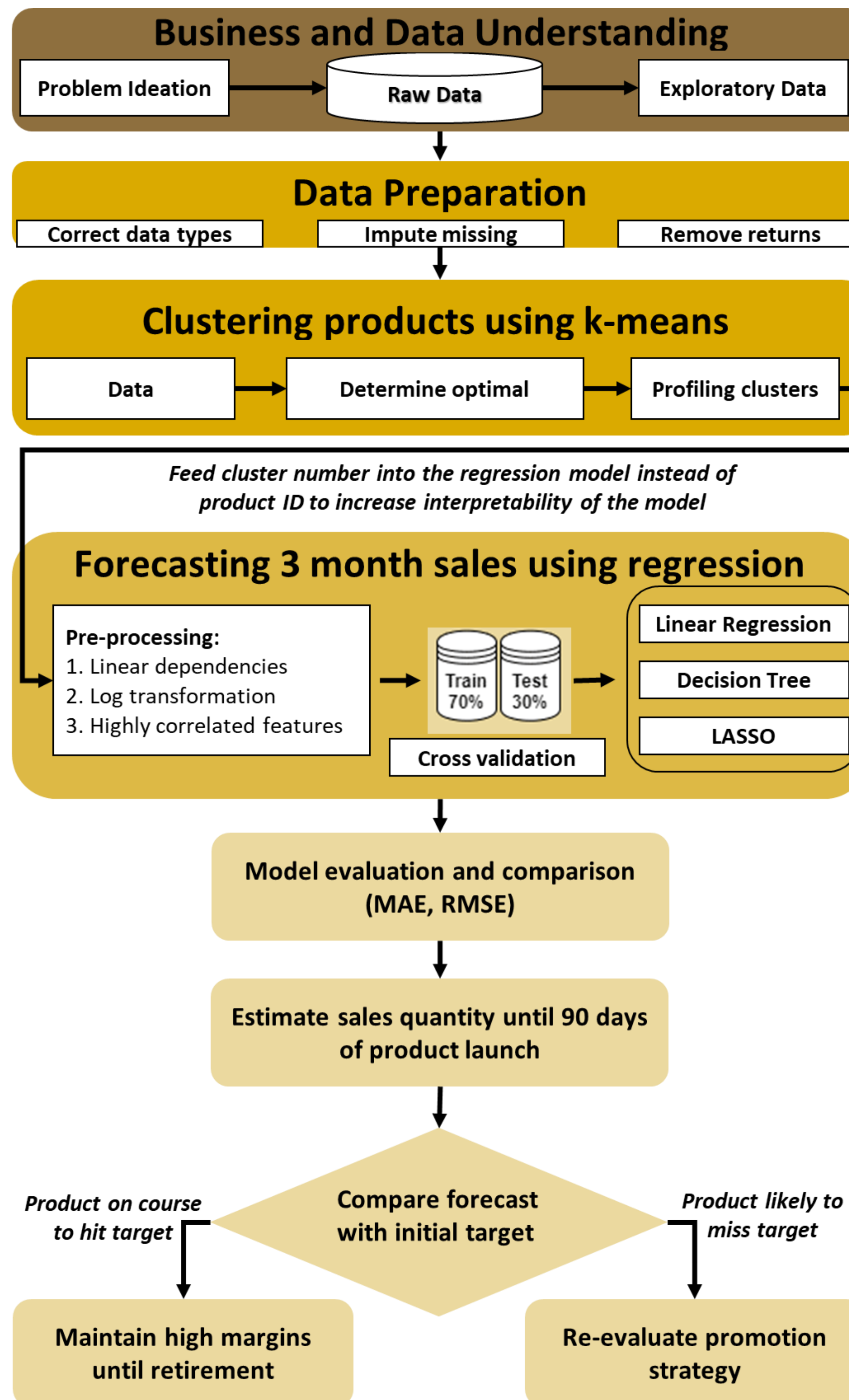


Fig 2. Process Flow

RESULTS

- Multiple Linear Regression was chosen since it generalized the best compared to other models and our business partners placed higher importance on interpretability rather than accuracy.
- The third week sales after launch is found to be very significant factor in estimating the sales quantity by the third month. A 10% increase in third week sales, led to 7.5% increase in 3-month sales

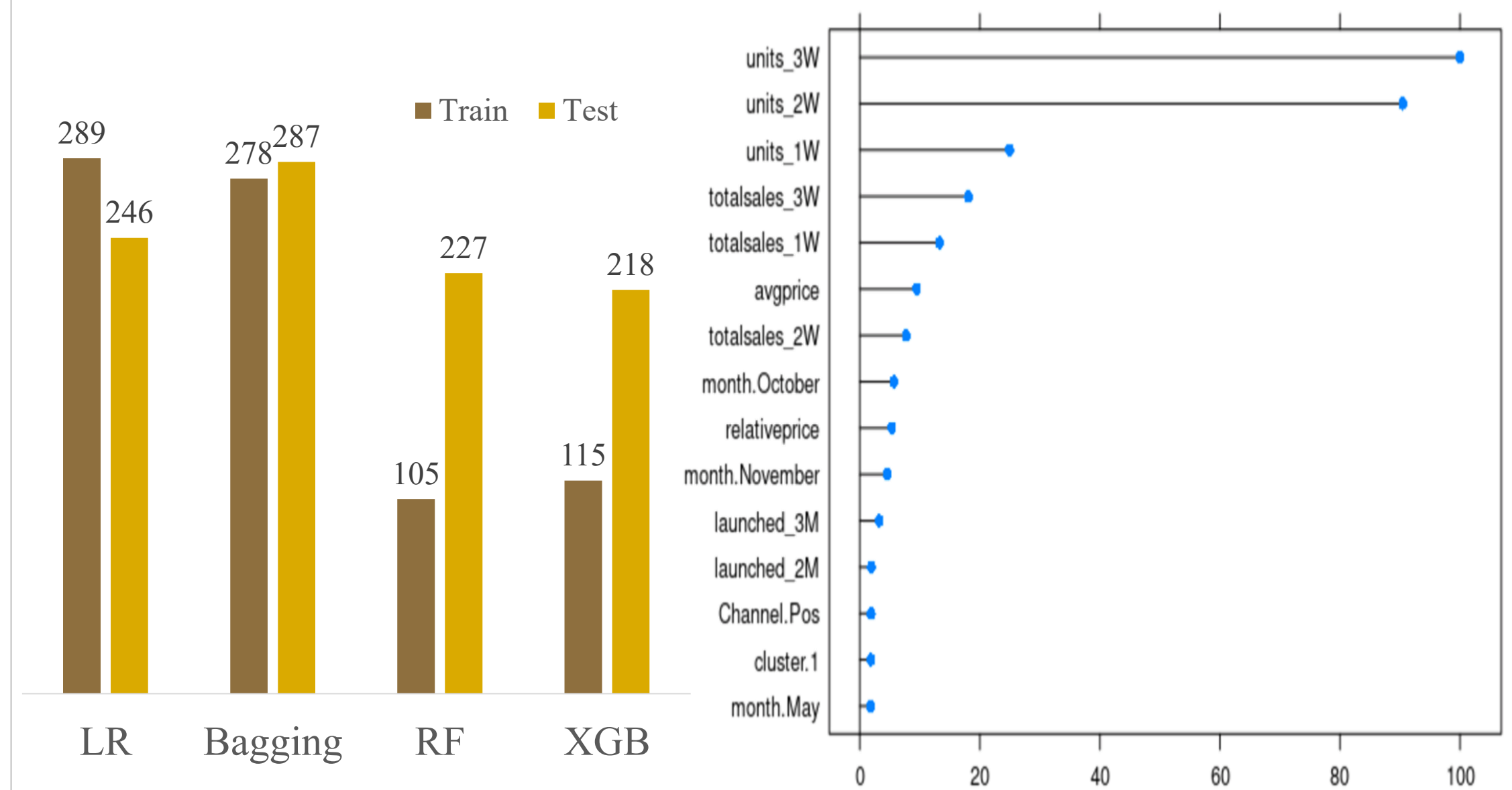


Fig 3. MAE Comparison of Models

Fig 4. Feature Importance Chart

CONCLUSIONS

The first three weeks of sales data, along with product characteristics, location details, launch season and cannibalization features are a good predictor of the product's success.

Products launched 30 days and 60 days prior to a new launch are found to have a negative impact on the new product performance.

For every product launched 60 days prior to a new product's launch, the sales of the newly launched product reduces by 7%.

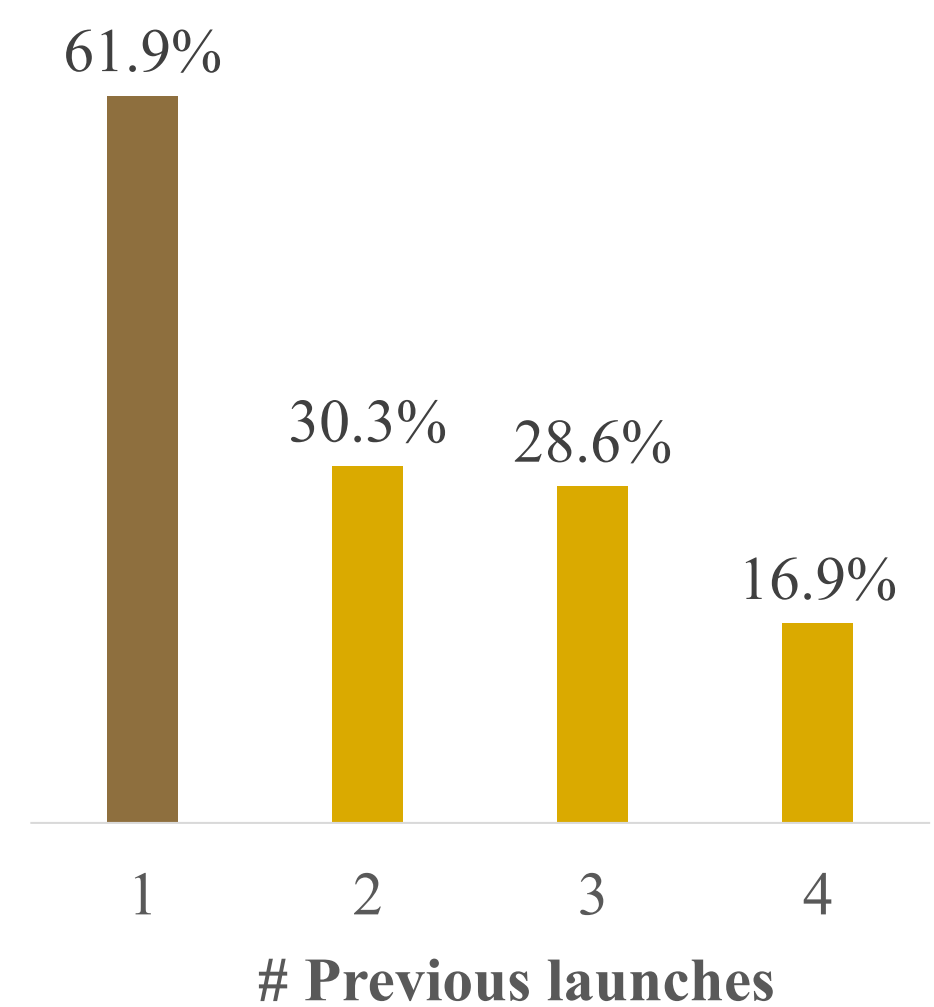


Fig 5. Effect of Cannibalization for category X

ACKNOWLEDGMENTS

We thank Professor Matthew Lanham and our industry partner for constant guidance on this project.