Demand Shaping: Promo Planning using Machine Learning

Shivam, Matthew A. Lanham
Purdue University Krannert School of Management
shivam@purdue.edu; lanhamm@purdue.edu

DAWN OR DOOM 2018 EMERGING TECHNOLOGY – RISKS AND REWARDS

DAWN OR DOOM THE NEW TECHNOLOGY EXPLOSION 150 YEARS OF GIANT LEAPS PURDUE UNIVERSITY

Abstract

What? We provide a promotional planning model which predicts future sales for products known to be on promotion, while knowing only total overall category promotional investment ($), and not individual promotional details.

How? We first use machine learning to forecast sales as arcuately as possible. We then measure the "uplift" in sales different than the traditional uplift modeling approach by making sales as a function of promotional spends.

Why? This company needed to understand the impact their promotional category investments were having on their business. Unfortunately, vital measurements were not recorded. Thus, we develop a design to provide insight into how to achieve this, while a proper design and future measurement collection procedure can be developed and implemented in the future.

Introduction

Stakeholders within a Demand Planning team from a multinational consumer goods company had been realizing lost sales due to aggressive forecasting in some markets. Supply of inventory increased but sales did not materialize, leading to inventory going bad, which is referred to as slow and obsolete inventory (SLOBs).

It was observed that the total estimated impact on net trade sales (NTS) due to SLOBs was 24%. Thus, the goal of this project is to optimize the inventory through advanced analytics and subsequently reduce the SLOBs for pilot markets in the APAC region.

Uplift Calculation (Model Output)

The below given table is a glimpse of the actual results obtained after the test data was provided to the model (trained earlier) in order to calculate the uplift:

SKUCategory	Month	Year	Actual Sales	Predicted Sales	BaseForecast	ActualSpends	Increased Spend (10%)	Uplift (%)
Baby Cream	4	2016	197,089	377,623	201,600	6,967,978	7,664,776	48%
Baby Cream	4	2017	1,924,922	1,769,003	1,420,791	7,979,290	8,777,219	-9%
Baby Lotion	4	2016	375,552	381,030	462,864	435,741	479,315	1%
Baby Lotion	4	2017	384,615	379,349	466,857	795,588	875,147	-1%
Baby Natural Oil	4	2016	2,184	2,904	18,000	-	-	25%
Baby Natural Oil	4	2017	979	2,704	18,000	228	251	64%
...by Oil	4	2016	3,533,679	3,977,827	4,588,620	14,617,026	16,078,728	11%

Figure 1. Representation of the working model

Literature Review

A popular technique to gauge the uplift in sales used across different industries is **Uplift Modelling**. It is a predictive modelling technique that directly models the incremental impact of a treatment (such as a promotion or direct marketing action) on the empirical data.

In our problem, to calculate the percentage change in sales due to promotion was failing because:

- The sales history data itself carries a promo factor (since when the data was recorded, there was an active promotion).
- There was no baseline (or without promo sales reference) to calculate the uplift percentage upon.

The model suggested herewith offsets these two limitations of Uplift Modelling.

Methodology

The design of the model that can be used in the demand planning process for the problem motivated is shown below

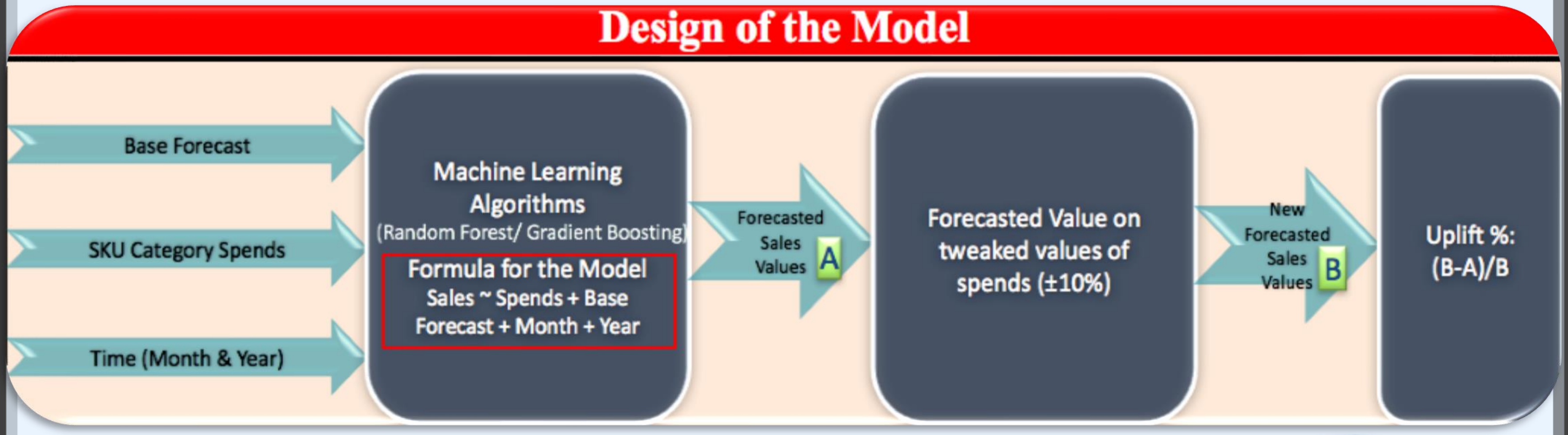


Figure 2. Model Design

Data

2 years SKU level data from BW
2 years SKU Category level data from Marketing team
SKU to SKU Category Mapping sheet

Data Cleaning & Pre-Processing

The data received from disparate sources were reshaped, mapped and collated. The figure alongside shows a glimpse of the data used.

SKU Category	Month	Sales (dzs)	Base Forecast (dzs)	Spends (₹)
Wash-proof	1	117,893	161,010	248,366
Top to Toe Bath	1	156,170	132,114	336,068
SF Secure Dry Wings	1	778,872	19,899,490	2,668,244
SF Secure Dry UT	1	553,461	70,733	1,550,383
SF Secure Cottony Regular/Wings	1	9,954,638	10,489,521	8,512,264
ORSL	1	6,616,350	4,957,419	3,012,499

Figure 3. Glimpse of Data

Feature Selection

1) The spend recorded each month is at a SKU category level only so aggregation of sales & base forecast after proper mapping (from SKU to its corresponding category) was done for each month.

Justification: *Running the model at SKU level would mean, disaggregating the spend data using assumptions (fair share allocation), which will lead to allocation of high investment to SKUs with high sales behavior within the same SKU category, giving erroneous predictions.*

2) The Month & Year were taken separately.

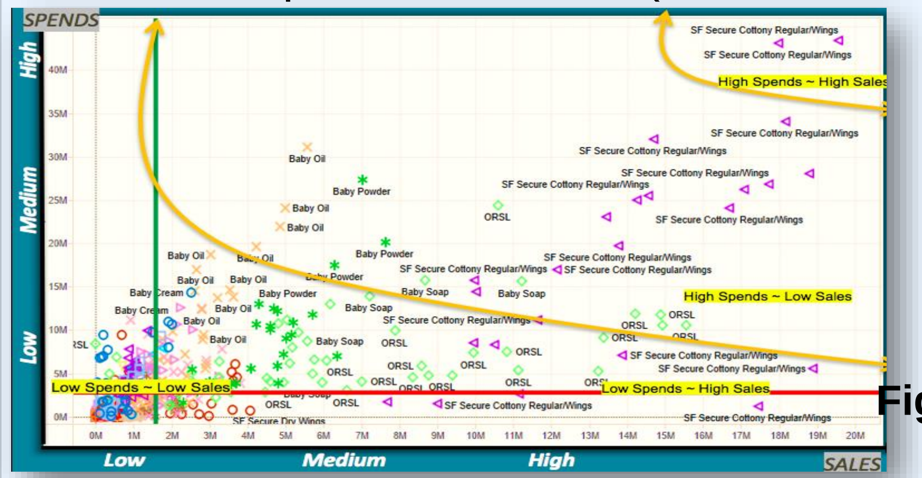
Justification: *To study trends (months) and seasonality (year) individually*

3) Only the Spends (Promo Investment), Sales & Base Forecast (Stat forecast from the Advanced Planning Optimizer) was taken into consideration for study.

Justification: *Total Forecast contains the Promo Balancing factors provided by the marketing team*

Methodology (Approach) Selection

As part of the EDA, the scatter plot of Sales vs Spends was studied along with Box Plot for possible outliers (as shown in the figures)



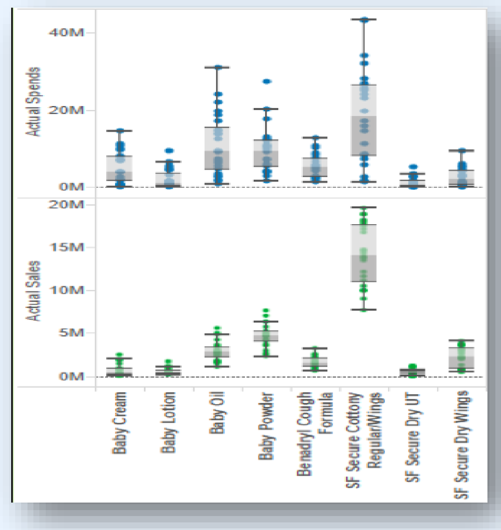


Figure 4. Scatter & Box Plots

Model Evaluation / Statistical & Business Performance Measures

- For checking the forecasting accuracy, MAPEs were used.

Results

Model accuracy across SKU categories w.r.t. the following algorithms: Random Forest, Gradient Boosting & Linear Regression (Multivariate) is shown in Figure 5.

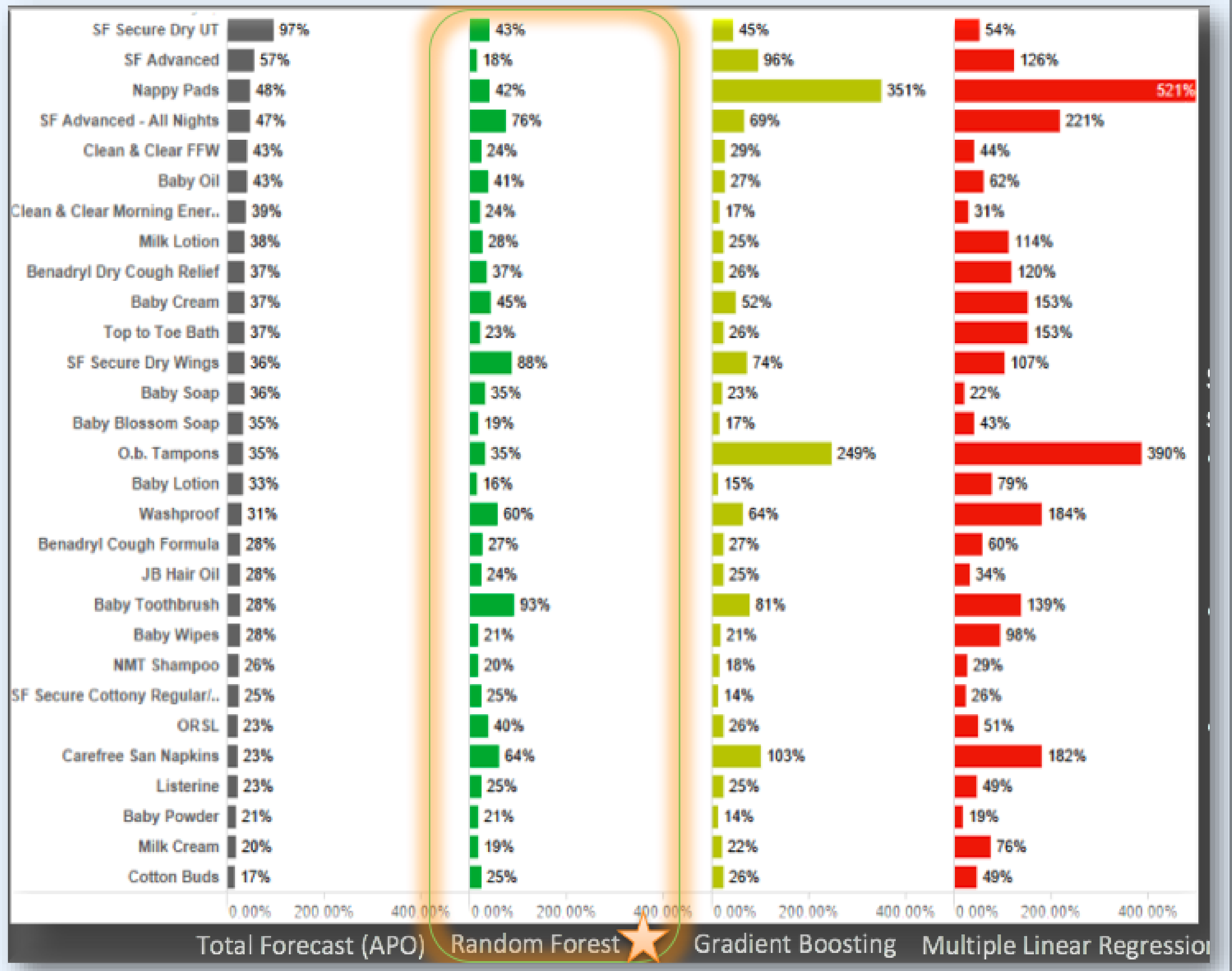


Figure 5. Model performance by category

The best results across all SKU Categories (on average) is given by the Random Forest model. Among the 30 SKU categories, the model suggested by us has better MAPEs across 21 (i.e. 70% of total) and outperforms their forecast.

Note: we were not able to show how the forecasts related to category spend due to proprietary reasons.

Conclusions

The scatter plot suggests the 4 demarcations (High Investment ~ Low Sales, Low Investment ~ High Sales, High Investment ~ High Sales & Low Investment ~ Low Sales). There is a need to focus on the products where investments are high but sales are not picking up and this model or application will help the demand planners in focusing on such products by comparing different forecasting accuracy metrics involving investment and selecting the optimum level of Investment for a smooth sell through performance.

Acknowledgements

We would like to thank the industry partner for allowing us to work on this project.