**Media Spend Cost Optimization**

**Team**

Our team members are:

1. Aditya Sanjay Pawar: https://github.com/aspawarsyredu
2. Himanshu Hegde: https://github.com/himanshu-hegde-syr
3. Keerthi Krishna Aiyappan: <https://github.com/KeerthiKrishna-Aiyappan>
4. Sri Venkata Subhramanya Abhishek Namana: https://github.com/snamana

Point of Contact: Himanshu Hegde

**Introduction**

**Context**

Marketing Mix Model, or Media Mix Model (MMM) is used by companies to measure how their media spending(TV, social media, email, newspapers etc.) contributes to sales, so as to optimize future budget allocation. Companies nowadays are investing millions of dollars into offline and online marketing channels to drive customer acquisition and sales. As the number of media channels increases, it gets harder and harder to track, analyze and quantify the impact of their marketing spend.

The sale of a product is influenced by various factors:

* Brand image
* Product distribution
* Product pricing
* Offline and online media spend
* Competitor pricing and media spend
* Discounts and promotions
* Seasonality
* External economic factors

When you have multiple factors affecting the demand for your product, it is crucial to understand and be able to quantify how each media channel you spend on (TV, social media, email, newspapers etc.) impacts your demand. This understanding is all the more essential while optimizing the budget for company media spend, by allocating spend across different media channels efficiently.

**What we are doing to solve this**

The solution we are proposing involves using Bayesian and Shapley based approaches to understand the marginal(individual) contribution of each media channel to sales, the diminishing effect media spend has, ad-stock lag effect of each media channel and the cost saved by optimizing media spend.

Our goal is to:

1. Understand the effectiveness of different media channels in driving product sales.
2. Quantify Ad-Stock effect of various media channels. Ad-stock is the lagged effect between advertisement launch and consumer purchase.
3. Assess the Diminishing Returns of each media channel. After a certain saturation point, increasing media spend will not result in increased product sales (i.e. it will result in diminishing sales returns).
4. Optimize future marketing decisions and quantify its impact. Once we have a clear understanding on the marginal impacts of each of the media channels, we can optimize our media spends given a certain budget.

**What is new in our approach?**

Traditional MMM models over the years have been built using Generalized Linear models and Bayesian Statistics. Over the past few years, Shapley(Shapley Additive exPlanations) based approaches have been highly used across different domains for calculating marginal contributions of different features of a non-linear model to the target variable. However, they have not been implemented across many MMM models. We intend to explore using Shapley to understand the relationship and impact of each of the media channels towards product demand and compare these insights to ones obtained by traditional methods.

**If you are successful, what difference will it make?**

A successful model will result in an optimized media allocation budget and strategy, which in turn will increase demand for products, yield more revenue and help reduce unnecessary spending.

Some examples of the insights are:

* TV ads contribute to 20% of sales revenue, while social media contributes to 10% of the sales revenue.
* Consumers start purchasing a product 2 weeks after TV ads, and 1 week after newspaper ads.
* Increasing the social media spend for product A above $10k yields no increase in sales returns.
* If we were to reduce the media spend budget to $30k, then investing $6k of this into newspaper ads, will see a 4% rise in product sales.

**Literature Review**

In today’s world, marketing mix modelling is predominantly done using additive techniques or multiple linear regression algorithms. The Adstock effect captures the prolonged and cumulative impact of advertising, enabling marketers to fine-tune campaign timing and effectiveness. Saturation Effects helps in identifying the point of diminishing returns on marketing investments, which is crucial for efficient budget allocation. Not all companies consider the adstock effect or saturation effect in their model building and that is something we will be doing. We will also incorporate seasonal and cyclical trends that might affect the sales of our product.

Edwin Ng, Zhishi Wang and Athena Dei[2021] used Bayesian and varying coefficient models to enhance Marketing Mix Modeling at Uber. Bayesian models were preferred over others as they were efficient at training on small amounts of data. As marketing data was often at the geographic/national level, the amount of data which is deemed relevant for model training is limited. In such scenarios, Bayesian models proved to be the most accurate.

Jessika Karlsson[2022] built marketing mix solutions using Bayesian Structural Time Series (BSTS), by incorporating prior beliefs into the distributions of the model parameters. The use of priors in Bayesian modelling also allowed them to incorporate prior domain knowledge regarding competitor product pricing into the model, hence strengthening the accuracy of the model.

Slava Kisilevich[2022] used tree-based ensembles and explained media channel performance using SHAP (Shapley Additive Explanations). SHAP is a game theoretic approach to explain the output of any machine learning model. It is a method for explaining individual predictions and answers the question of how much each feature contributes to this prediction. His methods showed a high degree of consistency with previously implemented Bayesian marketing mix models.

There are several open-source libraries like mamimo, Py-McMarketing and LightweightMMM that could be of use to us while we take this approach to modelling.

**Stakeholders**

1. Chief Marketing Officers (CMOs) and Marketing Teams: CMOs are under increasing pressure to demonstrate the value of marketing efforts, especially in challenging economic conditions. They need MMM to justify marketing spend by showing its impact on sales and to make informed decisions on budget allocation across various channels to optimize media performance.

2. Finance Departments: Finance teams are interested in understanding the return on investment (ROI) of marketing activities to ensure that budgets are used efficiently. They require clear, reliable metrics that link marketing spend to sales outcomes, helping in financial planning and forecasting.

3. Brand Managers: Individuals responsible for specific products or brands within a company need to understand how their marketing efforts are contributing to brand awareness, customer acquisition, and sales. They require granular insights into the effectiveness of different marketing tactics for their specific brand context.

4. Sales Teams: Sales departments benefit from understanding how marketing activities influence customer behavior and sales cycles. They need data that helps them align sales strategies with marketing efforts to maximize overall business performance.

5. Data Analysts and Data Engineers: These stakeholders are involved in the technical aspects of MMM, including data collection, model building, and analysis. They have access to the data, provide us with required data and business context, validate our insights and provide feedback.

6. Product Development Teams: Insights from MMM can inform product development by highlighting features or benefits that resonate with customers. These teams need information on how marketing for existing products impacts customer preferences and perceptions, which can guide the development of new products or improvements to existing ones.

**Data and Methods:**

**Data**

The marketing spend data belongs to a shampoo advertiser, provided by Neustar MarketShare. The data was consolidated from sources such as Kantar Media, IRI, ITG, JD Power, and Rentrak. It contains 4 years of weekly data, including weekly volume sales in ounces and media spend on major channels such as TV, magazines, display, YouTube, and search. Also available are retailer variables such as the average price per ounce, ACV weighted product distribution, discounts, and promotions.

The same dataset has been used by Yuxue Jin and Yueqing Wang, when they built a marketing solution for Google in 2017.

The columns in the data can be clubbed into 3 main categories:

**1. Media Variables**

Media Impression (prefix='mdip\_'): impressions of 13 media channels: direct mail, insert, newspaper, digital audio, radio, TV, digital video, social media, online display, email, SMS, affiliates, SEM.

Media Spending (prefix='mdsp\_'): spending of media channels.

**2. Control Variables**

Macro Economy (prefix='me\_'): CPI, gas price.

Markdown (prefix='mrkdn\_'): markdown/discount.

Store Count ('st\_ct')

Retail Holidays (prefix='hldy\_'): one-hot encoded.

Seasonality (prefix='seas\_'): month, with Nov and Dec further broken into to weeks. One-hot encoded.

**3. Target Variable**

Sales(‘sales’): Weekly sales in Dollars.

**Methods**

The solution will be built out in the following steps:

1. Data pre-processing
2. Exploratory Data Analysis
3. Feature Engineering
4. Modelling
5. Insight validation and feedback
6. Cost Optimization simulations

**Data pre-processing**

This step involves:

* Null/Missing value identification and treatment
* Assigning correct datatypes to features
* Ensuring Data Integrity- No duplicates, the level of data is standard across the data, etc.

**Exploratory Data Analysis**

Exploratory Data Analysis or EDA includes:

* Creating visual representations (e.g., histograms, box plots, scatter plots, etc.) to explore the distributions and patterns within the data
* Outlier analysis
* Correlation analysis
* Data summary across various hierarchies

**Feature Engineering**

We need to ensure that the features we have significantly explain product sales.

To do this, we need to:

* Select relevant features and create new features from pre-existing features if required.
* Encode categorical features
* Standardize continuous features (if required by the model)
* Handle missing values in the data, if any.

**Modelling**

We will be using a mixture of Bayesian models and Tree based Boosting methods with Shapley analysis to understand the marginal(individual) contribution of each media channel to sales, the diminishing effect media spend has, ad-stock lag effect of each media channel on sales.

These models' accuracy shall be tracked with various metrics, such as RMSE and MAPE.

**Insight validation and Feedback**

Reviewing model performance after each iteration, re-assessing the relevance of features and model parameters and re-iterating steps 1-4 to refine and improve the model.

This includes:

* Hyperparameter/parameter adjustment
* Modifying feature engineering
* Exploring alternative modelling approaches to address shortcomings of current model.

**Cost optimization simulations**

Once the champion model has been selected, run simulations for various budgeting decisions and calculate ROAS(Return on AD Spend) for various scenarios.

**Project Plan**



**Risks**

Some of the factors that affect the quality of an MMM model are:

1. Access to competitor data. Competitor product pricing and media spending are often unavailable. These variables will have to be accounted for in the model as unexplained variance in the target variable.
2. Mismatch in data hierarchy. Media spend and product level sales are often at different levels, media spend being at a geographic or national level, and product sales at a product level. This leads to smaller sized datasets, and granular level insights cannot be generated accurately at lower levels (individual campaigns).
3. As Marketing Mix Modelling do not always generate granular level insights, they are usually paired up with other processes like Multi Touch Attribution as part of a larger pipeline of analytics solutions that solve the problem of media spend optimization.
4. Complete visibility into offline media channels. Companies often do not have total visibility into how much visibility their offline media channels have. For example, a company might invest $1million on newspaper ads, but will not have access to what kind of readers are consuming that newspaper and in what regions.
5. The world of advertising is always changing, which means what worked yesterday might not work tomorrow. Our project needs to keep up with these changes by regularly updating our methods and information.

**References**

1. Edwin Ng, Zhishi Wang, Athena Dai [2019] Bayesian Time Varying Coefficient Model with Applications to Marketing Mix Modeling.
2. Jessika Karlsson[2022] Bayesian Structural Time Series in Marketing Mix Modelling.

**BA**

**When to use Power regression: when you have a learning curve**

**When / what is seasonality: Moving average smooths it out. It has a certain peripod**

**What does regression measure: Magnitude and direction of change.**

**What does correlation measure: Strength and direction of change.**

**Similarity between logit and probit: 0-1, predict prob, both are curves**

**Difference: Prob is normal distribution, Log is logistic distribution, sensitive**

**Whats a similarity between logit and NN: Both predict probs, 0-1**

**Difference: NN has hidden layers, Logit can have only 1 output while NN can have 0-1 or linear or multiple**

**~TYPE2~**

**MCQ: Excel, Access, Tableau, (Pictures and UI questions)**

**~TYPE3~**

**MCQ:**

**Identitfy the what the picture represents(correlation, linear regression, exponential regression, power regression, moving average, logit, probit, neural network, k-means clustering, Benford)**

**SQL: What is a filter in SQL,**

**What takes the table**

**What sorts things**

**Googale analytics: A lot of terminologies, organic search, direct, referral, emails,**

**~TYPE4 TYPE5~**

**MCQ: Regression Assumptions – Multiple Choice**

**Identify assumption violations, corrections (linearity, multi-collinearity, heteroscedasticity, serial correlation, outliers)**

**Strongest correlation:**

**Linear regression : p-values (0.05<) not significant, cannot interpret.So don’t interpret, R-square=0.789 what does it mean?, a unit change in house price wil affect x,y,z**

**Each additional squatfeet adds 61$ to the total (explain it in business terms not stats)**

**Don’t forget to check the significance first totalIf its given that the correlation is significant the continue with other interpretation.**

**Logit and probit interpretation and direction.**