Data Science Challenge

Task:

Task The challenge is set in the context of performance marketing and we want you to apply a bayesian mixed-media model (MMM) on our test dataset an interpret the insights from the model. Build the MMM with the latest PyMC package (https://www.pymc.io/).

Context We have a company X which runs an online shop. X advertises on seven different paid channels and has weekly costs in them. Marketing actions have usually not an immediate effect, ads and campaigns in one week influence usually sales in the coming weeks. Hence, the company is of course super interesting to understand how effective different channels are. In terms of channels think of TV, radio, billboards, but also online advertisement such as Google Ads, Facebook Ads, etc. So different channel can be expected to target different audiences at different times, and hence will have very different effects on future sales.

This is of course the perfect setting for an ambitious Data Scientist. ;-) Modelling the uncertainty and the delayed effects is of course key. We are working heavily with Bayesian models and would like here to test your understanding and approaches in this setting.

You will need to model the spend carry over effect (adstock). No need (for now) to overcomplicate the adstock shape effects with saturation or diminishing returns. Seasonality & trend might be interesting to be included in your model. Dataset MMM_test_data.csv start_of_week: first day of the week revenue: revenue generated in this week from sales spend_channel_1..7: marketing cost spend in this week in channel 1..7 Questions How do you model spend carry over? Explain your choice of prior inputs to the model? How are your model results based on prior sampling vs. posterior sampling? How good is your model performing? How you do measure it? What are your main insights in terms of channel performance/ effects? (Bonus) Can you derive ROI (return on investment) estimates per channel? What is the best channel in terms of ROI?

Introduction

This report presents the application of a Bayesian Mixed-Media Model (MMM) on a test dataset provided by Company X. The company runs an online shop and advertises on seven different paid channels. The goal is to understand the effectiveness of these advertising channels on sales. The dataset contains weekly advertising spend across these channels and the corresponding weekly revenue. The task involves modeling the carry-over effect of advertising spend (adstock), incorporating seasonality and trend, and evaluating the model's performance using PyMC.

As a data science student and beginner in this area, this task was challenging for me. I have almost four years of hands-on experience in Python and SQL, which gave me an edge when I started studying data science. Despite the challenge, I found the task very exciting and motivating, pushing myself beyond my limits. I love new projects, gaining new experiences, and exploring new fields.

Modeling spend carry over (Adstock)

What is spend carry over effect (Adstock):

The spend carry over effect, also known as adstock, is a concept used in marketing to describe how the effect of advertising spend is distributed over time. In other words, the impact of an advertising campaign does not only occur in the period when the spend happens but also carries over into subsequent periods.

Advantages of Adstock:

- We can measure the impact of the delayed effect of advertising.
- Budget allocation considering the long-term effects.
- Improving the advertising, recognizing the past effect.

Decay Rate :

The rate at which the impact of advertising decays over time is called the decay rate, often denoted by α. A higher α means the effect lasts longer, while a lower α α means the effect diminishes quickly.

Finding out the Decay rate:

There are two ways to find out the decay rate of adstock:

1). Using domain knowledge

From the insights of marketing experts.

2). Empirical estimation.

Using data to find the decay rate that best fits the observed effects of advertising on sales.

In this challenge, I followed the domain knowledge to find out the decay rate, as empirical estimation requires relevant data and is also time-consuming.

https://getrecast.com/adstock-rates/#:~:text=Adstocks%20by%20Channel,effects%20than%20others.

Formula for Decay Rate:

Given the half-life , the decay rate (α) can be calculated using the formula:

$$\alpha = 0.5^{(1/\text{ Halflife })}$$

Half life - The half-life is the time it takes for the impact of advertising to decay by half.

Channels	Half Life (weeks)	Half Life	Decay Rate
TV	2-6 weeks	(2+6)/2 =4	0.840896415
Radio	1-5 weeks	(1+5)/2=3	0.870550563
Newspaper	2-3 weeks	(2+3)/2=2.5	0.793700526
Print Magazines	4-8 weeks	(4+8)/2=6	0.870550563
Online (Upper Funnel)	2-4 weeks	(2+4)/2=3	0.840896415
Online (Lower Funnel)	1-2 weeks	(1+2)/2=1.5	0.707106781
Average	**0.820616877**		

Decay rate = 0.820616877

Choice of Prior Inputs

Choosing appropriate priors is an essential step in Bayesian modeling as it can influence the results of the analysis. So, we can go through the selection of priors in this analysis and provide a basic understanding of the Bayesian model initially.

Bayes' Theorem:is a way to update our beliefs about something based on new evidence. It combines our initial belief with the new evidence to give us a new and improved belief.

In this analysis, we chose beta to represent the impact of each advertising channel, intercept to capture the baseline revenue(It represents the revenue level when there is no advertising spend) without advertising, and sigma to model the variability in revenue that the model doesn't explain. These choices help create a comprehensive and accurate Bayesian model.

- **Beta Coefficients:** Normal distributions centered at 0 with a standard deviation of 1 reflect initial uncertainty about the impact of each channel.
- **Intercept:** Normal distribution centered at 0 with a standard deviation of 1 reflects initial uncertainty about the baseline revenue.
- **Sigma:** Half-normal distribution to ensure positivity reflects initial uncertainty about the variability in revenue.

Model Results: Prior vs. Posterior Sampling

Prior Sampling: Reflects initial beliefs about the parameters before seeing the data.

Posterior Sampling: Updates these beliefs based on the observed data, providing a more accurate estimate of the parameters.

Parameter	Prior Mean	Prior Std Dev	Posterior Mean	Posterior Std Dev
Beta[0]	0	1	0.8313	0.03
Beta[1]	0	1	-4.1163	0.04
Beta[2]	0	1	-0.0182	0.01
Beta[3]	0	1	0.4355	0.04
Beta[4]	0	1	0.2717	0.04
Beta[5]	0	1	1.1553	0.04
Beta[6]	0	1	0.5082	0.04
Intercept	**0**	**1**	**0.6924**	**1.0091**
Sigma	**1**	**0.5**	**701.5348**	**0.4917**

How good is your model performing? How do you measure it?

There are various tools to evaluate a Bayesian model.

Trace Plots: Used to check for convergence and good mixing of the MCMC chains.

Posterior Predictive Checks (PPC): Used to compare the model's predictions to the actual observed data.

Evaluation Metrics: Metrics like R-squared and RMSE can be used to measure the model's performance.

What are your main insights in terms of channel performance/ effects?

Parameter	Beta
Spend Channel 1	0.8313
Spend Channel 2	-4.1163
Spend Channel 3	-0.0182
Spend Channel 4	0.4355
Spend Channel 5	0.2717
Spend Channel 6	1.1553
Spend Channel 7	0.5082
Intercept	**0.6924**
Sigma	**701.5348**

- **Baseline Revenue (Intercept):** The baseline revenue is approximately $692.40 (if in thousands of dollars).
- **Unexplained Variability (Sigma):** The high sigma value suggests significant variability in revenue that is not explained by the model.

ROI

Channel	ROI Estimate
Channel 1	0.0075866
Channel 2	-0.01036351
Channel 3	-0.00260323
Channel 4	0.02206493
Channel 5	0.01707191
Channel 6	0.04286037
Channel 7	0.10313417

ROI Calculation: Calculated as the product of the beta coefficients and mean spends for each channel, normalized by the mean revenue.

- Channel 2 and Channel 3 have negative ROI, indicating that these channels are not generating sufficient revenue relative to their costs.
- Channels 1, 4, 5, and 6 show positive ROI, suggesting they are effective in driving revenue.
- Channel 7 stands out with the highest ROI, indicating it as the most effective channel in terms of return on investment.