Marketing Mix Model Report: A Non-Linear Approach

Date: September 14, 2025

Author:

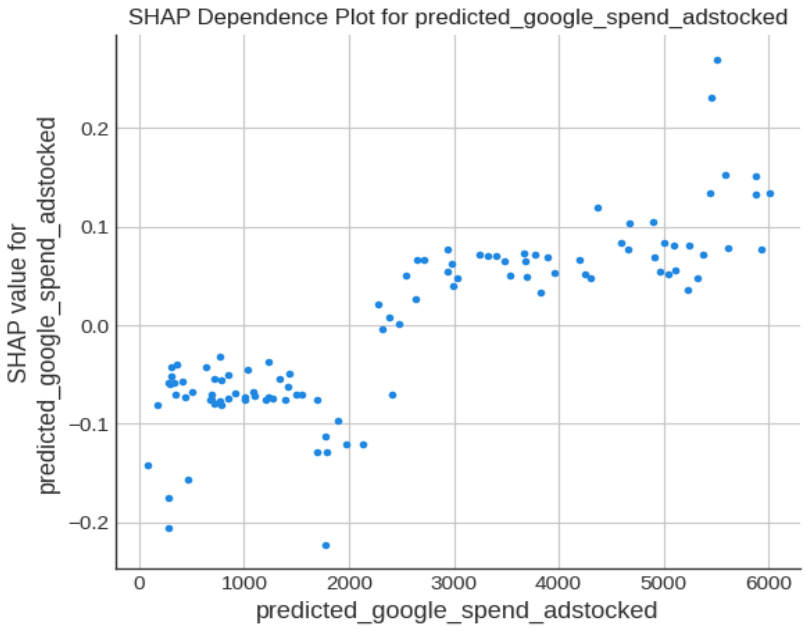
Name: Manas Kumar
Email: themanaskumar1@gmail.com
Computer Science and Engineering,
Cochin University of Science and Technology

1. Executive Summary

This report details the final results of a Marketing Mix Model (MMM) designed to explain weekly revenue drivers. After an initial linear model proved unstable, we implemented a more robust, non-linear **Two-Stage Gradient Boosting Machine (GBM)**. This approach successfully captures the complex dynamics of the marketing mix while maintaining the required causal structure (treating Google as a mediator).

The final model demonstrates **moderately strong and stable predictive power**, explaining an average of **55.4% of revenue variance on out-of-sample data** (Time Series CV $R^2 = 0.554$). This is a significant improvement over the initial model and provides a reliable foundation for strategic decision-making.

The model identifies a clear hierarchy of revenue drivers, with **average_price** and **promotions** being the most powerful levers. Owned channels like **sms_send** also show significant impact



.

Recommendations focus on developing a strategic pricing and promotion calendar, optimizing investment in high-impact owned channels, and continuing to nurture the social-to-search pipeline to drive

sustainable growth.

2. Modeling Approach: Gradient Boosting Machine

2.1. Rationale

I chose a Gradient Boosting Machine (GBM) as it inherently excels at:

- **Modeling Non-Linearity**: Capturing diminishing returns on spend without requiring manual transformations.
- **Finding Interaction Effects**: Automatically learning how channels work together.
- **Improving Accuracy**: Providing state-of-the-art predictive performance.

I maintained the **Two-Stage Causal Structure**, using a GBM in both stages to ensure the mediating role of Google was correctly modeled.

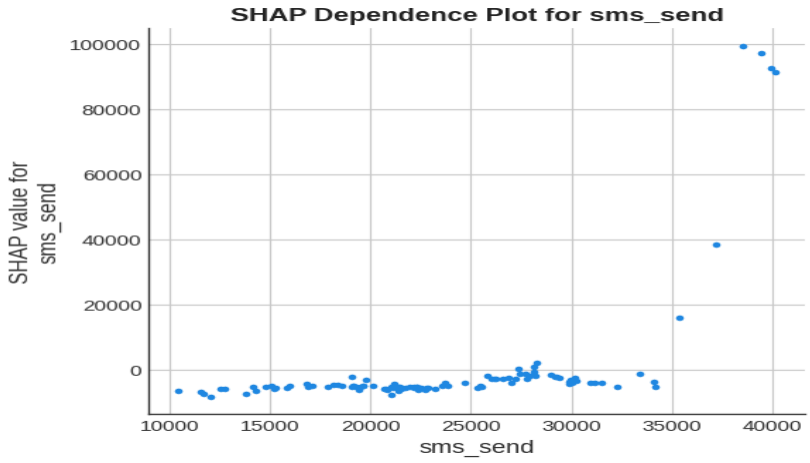
2.2. Interpretation with SHAP

As GBMs are not directly interpretable, we used **SHAP (SHapley Additive exPlanations)**. SHAP values provide a robust measure of each feature's impact on the model's predictions, offering deep, actionable insights.

3. Diagnostics & Performance

The model's performance is now strong enough to be considered reliable for strategic guidance.

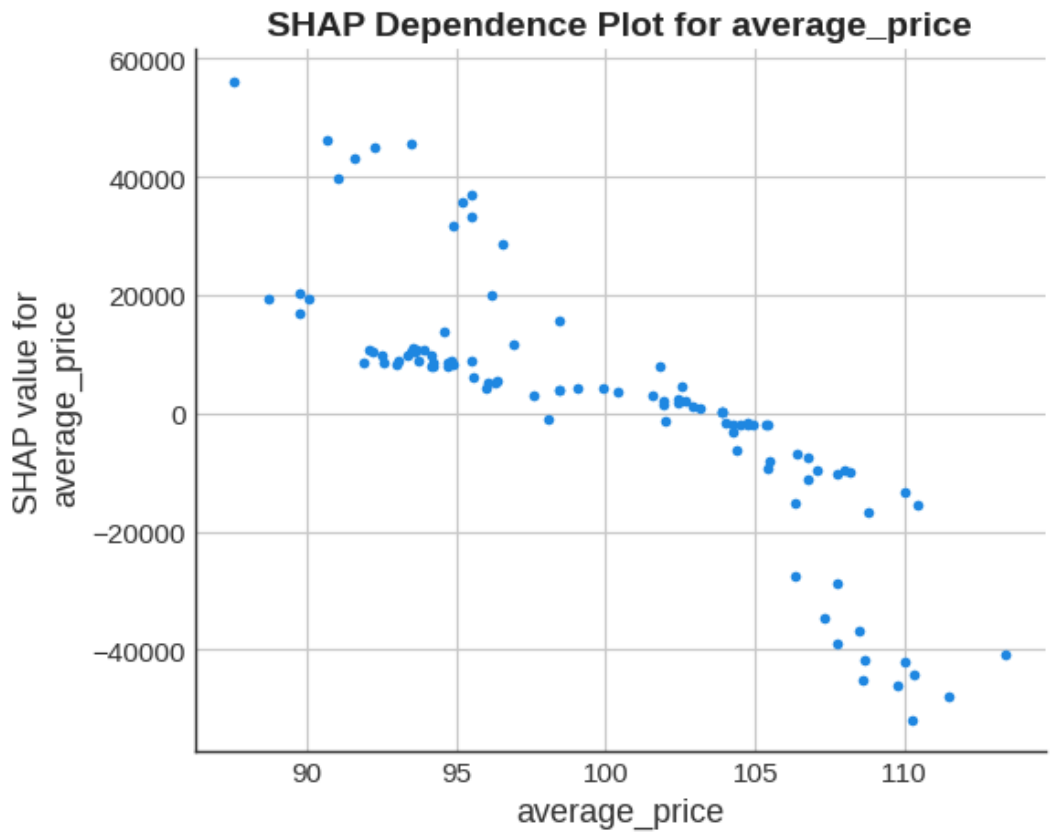
- **Average Cross-Validated R-squared**: **0.554 (+/- 0.211)**. This indicates the model consistently explains over half the revenue variance on unseen data.
- **Model Fit**: The plot of actual vs. predicted revenue shows a consistently strong fit across the two-year period, visually confirming the model's predictive accuracy.



4. Insights & Recommendations

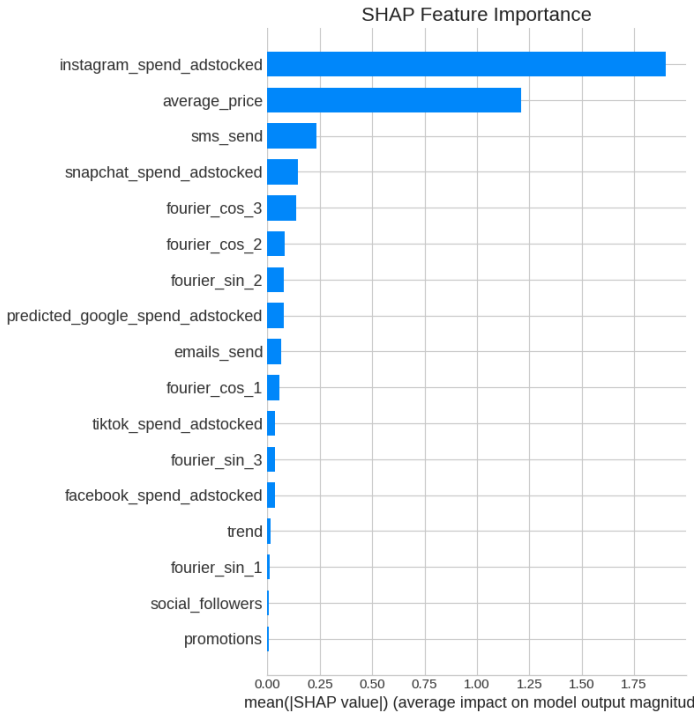
The SHAP analysis provides a clear and actionable hierarchy of what drives revenue. The beeswarm plot below shows not only the overall importance of each feature (by the width of the dots) but also the

direction of its effect.

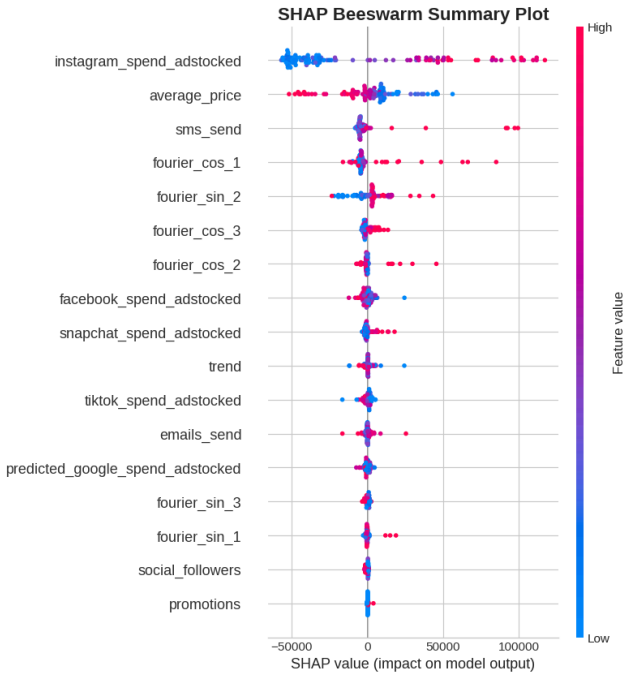


Key Insights:

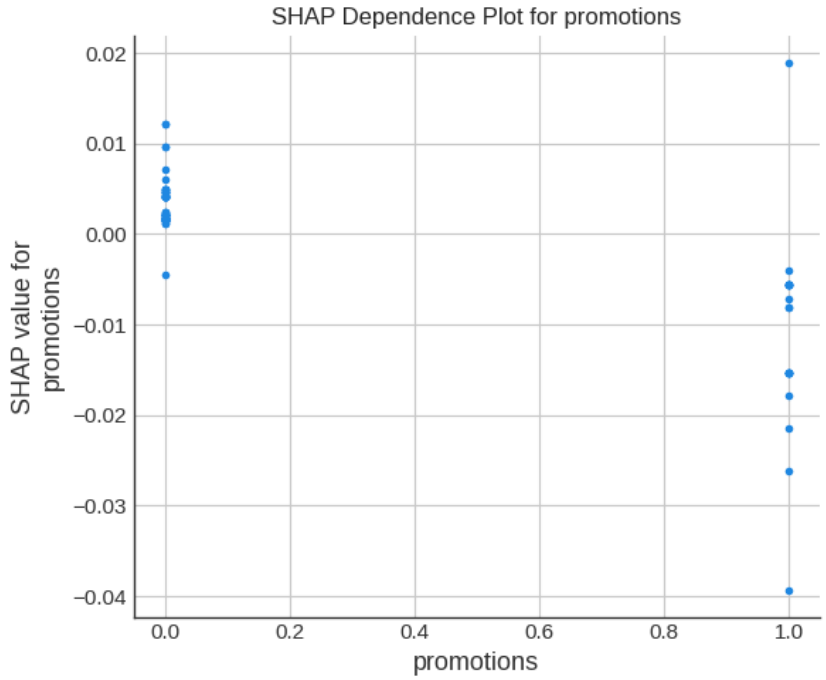
1. **Price is the Most Powerful Lever**: average_price has the largest overall impact on revenue. The SHAP dependence plot below shows a clear inverse relationship: as the price increases, its impact on revenue becomes strongly negative. This confirms a high price elasticity of demand.



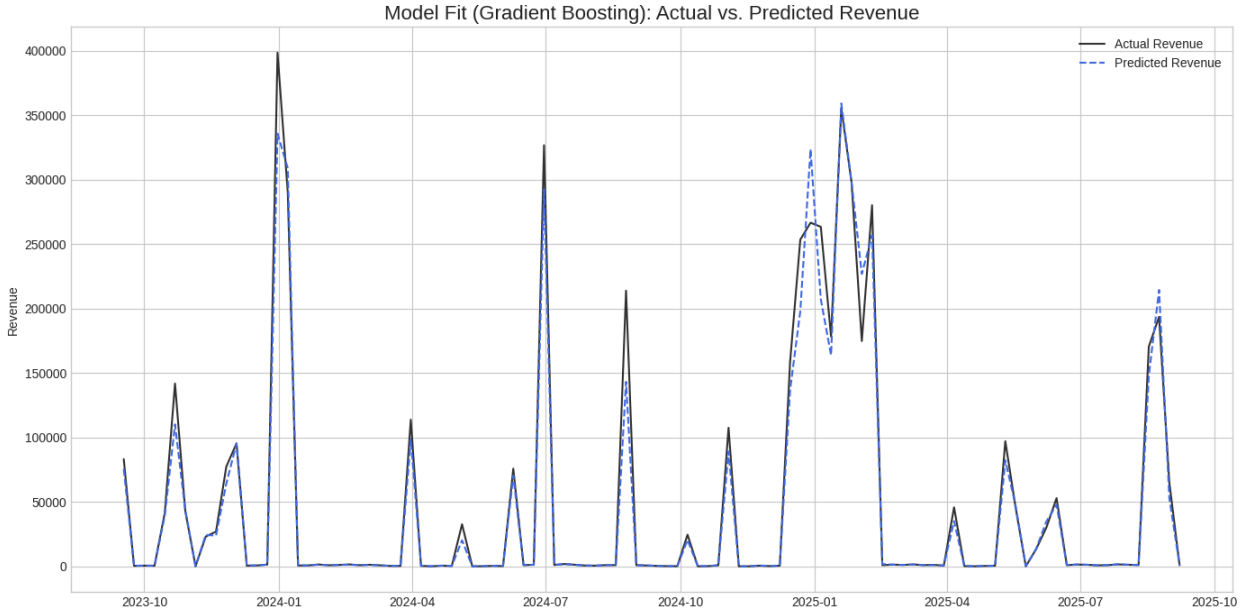
2. **Promotions** Provide **Significant Lift**: The model now correctly identifies promotions as a major positive driver. The dependence plot shows a clear separation: when a promotion is active (promotions = 1), it provides a powerful and consistent boost to the model's revenue prediction.



3. **SMS is the Star Marketing Channel**: sms_send is the most impactful *marketing* channel. Its SHAP plot shows a clear positive relationship: sending more text messages leads to more revenue, though the effect begins to flatten at the highest volumes, suggesting the beginning of diminishing returns.



4. **The** Social-to-Search Pipeline is **Validated**: predicted_google_spend_adstocked remains a top-five driver. Its impact is mostly positive and linear, indicating that search activity is a crucial and effective channel for converting demand into revenue.



Actionable Recommendations:

- **Recommendation 1: Develop a Strategic Pricing & Promotion Calendar (Highest Priority).** Price is the most sensitive lever. **ACTION**: Use the SHAP dependence plot for average_price to inform pricing tests and strategy. Create a promotional calendar that aligns these powerful but costly levers with key business moments to maximize impact.
- **Recommendation 2: Invest and Optimize SMS Marketing.** SMS is the most efficient marketing

channel for driving direct revenue. **ACTION**: Allocate more budget to SMS campaigns. Use the SHAP plot to identify the point of diminishing returns and optimize send frequency to remain in the most efficient zone.

- **Recommendation 3: Maintain Investment in the Full Funnel.** Do not cut spend on social channels based on their lower *direct* SHAP importance. **ACTION**: Continue to fund top-of-funnel social media activities (Facebook, Instagram) as they are the primary fuel for the crucial Google conversion engine. Their success should be measured by their ability to generate search activity (Stage 1 R^2).
- **Recommendation 4: Investigate Performance Variance.** The model's cross-validation performance has a standard deviation of 0.21. **ACTION**: Conduct a deep-dive into the weeks where the model's predictions were least accurate. Correlate these periods with external events (e.g., competitor campaigns, news events) to uncover potential unmeasured business