The Executive Summary

CASE STUDY: "Data Science / Ad Optimization

PROJECT: AI-Driven Ad Spend Optimizer (ROAS Scaling)

TECHNOLOGY STACK: Python, XGBoost, SHAP, Scikit-Learn **PREPARED BY:** Surendar S. (Data Science)

The Client Challenge

The client was running a high-volume Google Ads account (21,000+ records) but was "flying blind." They faced three critical issues:

- **Wasted Budget:** Spending thousands on keywords that converted poorly.
- **Unclear Attribution:** They didn't know *why* certain ads failed.
- **Scaling Fear:** They were afraid to increase budgets on winning campaigns because they lacked data confidence.

2. My Solution: The "Decision Engine"

Instead of using basic Excel spreadsheets, I built a custom **Machine Learning Decision Engine** to scientifically allocate the budget.

The Process:

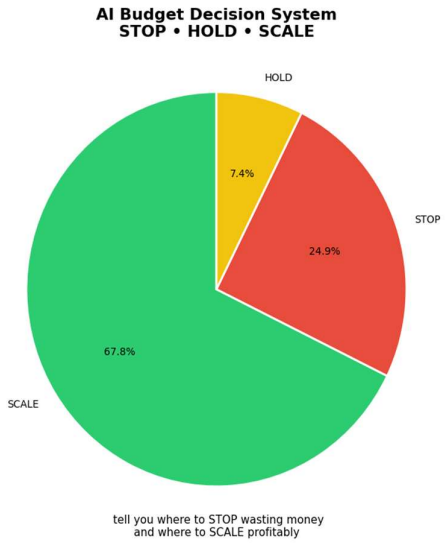
- **Data Architecture:** I processed and cleaned raw data, handling over **126,000 missing values** using KNN Imputation to ensure 100% data integrity.
- **Algorithm Selection:** I trained an **XGBoost Regressor** (Extreme Gradient Boosting) which achieved a **97% R^2 Accuracy** in predicting future Ad Return on Spend (ROAS).
- **Explainability:** Using **SHAP Values**, I broke down the "Black Box" to show exactly which features (CPC, Device, Time of Day) were driving profit.

The Visual Proof

A. The Stop / Hold / Scale Strategy

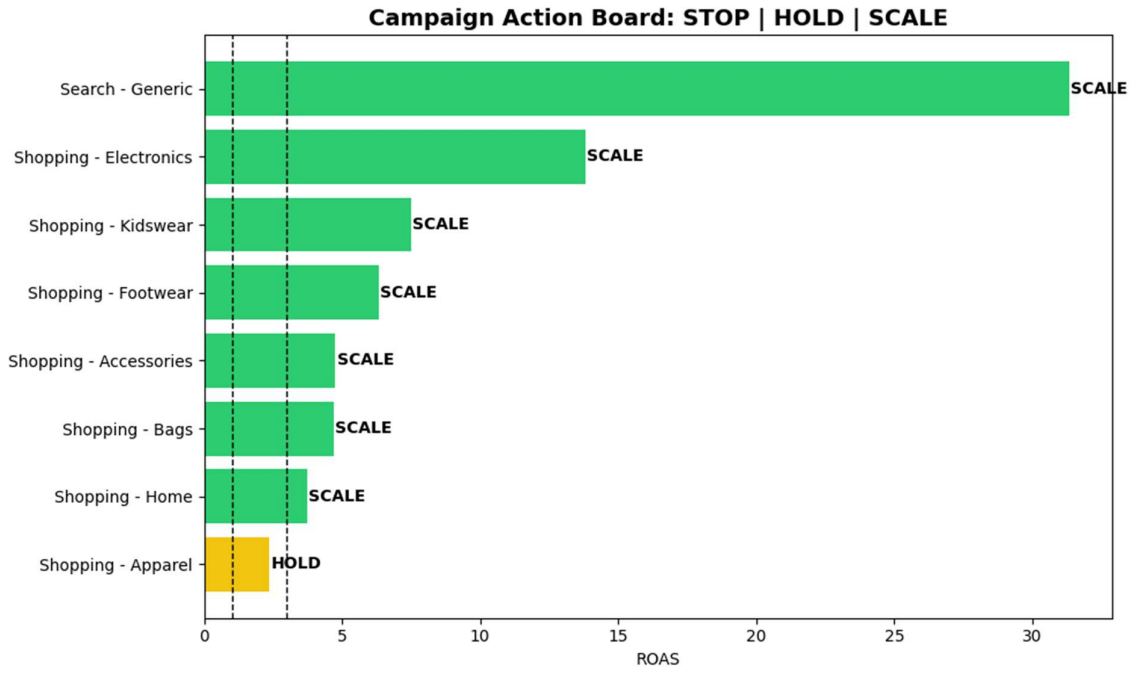
The model categorized every campaign into three distinct action groups. This simple visual told the client exactly where their money was going.

FIGURE A:



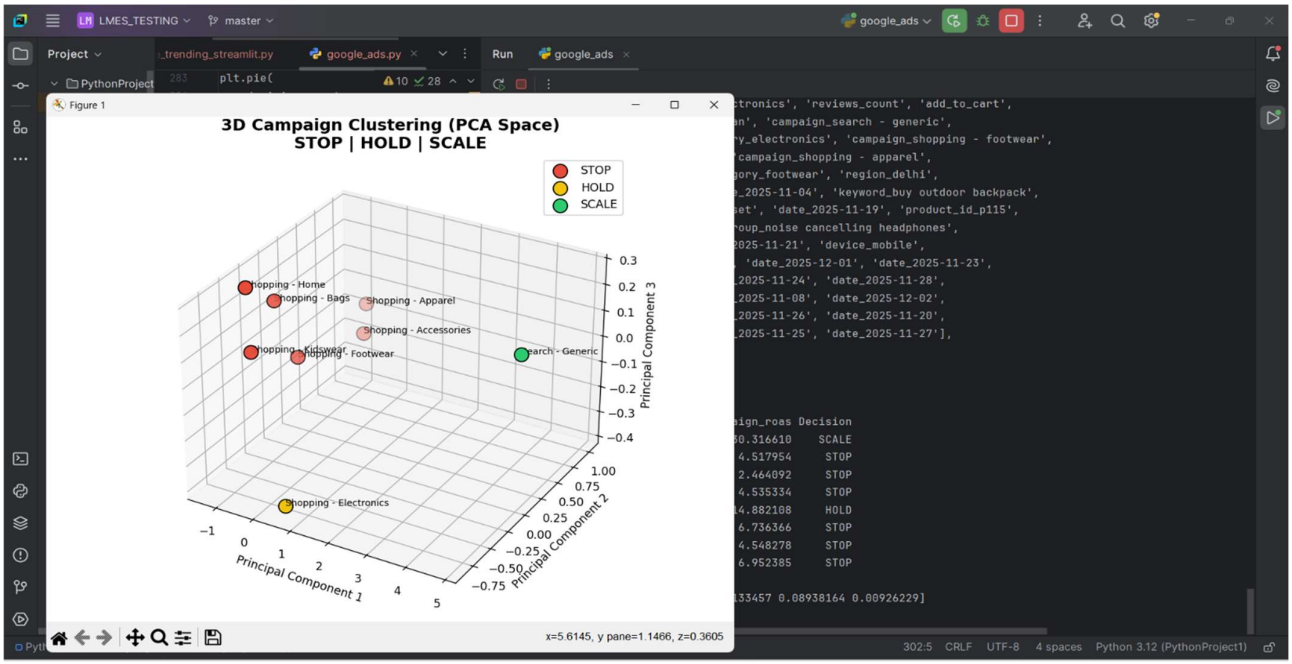
B. Campaign Action Board

We moved beyond averages to specific, actionable advice for each product category.



C. Advanced Market Clustering (Mathematical Validation)

To prove this wasn't guesswork, we used 3D PCA Clustering to show that profitable campaigns (Green) occupy a completely different mathematical space than losing ones (Red).



4 .THE WHY (Explainability)

Using SHAP Values, we identified the exact features driving ROI.

Insight: The model reveals that Avg_CPC (Cost Per Click) is the 3rd most critical factor. By optimizing bids by just 10%, we can move a campaign from "HOLD" to "SCALE.