LiME
Documentation

Contents

Introduction:

The tool is a comprehensive suite for media mix modeling offering nine intuitive tabs for various functionalities. Users can construct MMM models, analyze, optimize, and visualize spends, contributions, and scenarios across multiple media channels.

Login Credentials

Username: admin
Password: password

1. Data Preprocessing

- Upload a CSV or Excel file with appropriately labeled column names and cleaned data. Ensure there are no redundant columns, and the data has been treated for null or missing values.

1.1 Renaming and Classification

Use this section to rename variables if needed and reassign variable buckets. The bucket is initialized intuitively based on the variable name and might require changes. Click on the "Update Changes" button once complete and view the changes in the table generated after.

Note: Replace spaces with underscores and TBD

1.2 Select Time Variable

If your dataset contains multiple datetime variables, please select the appropriate date variable that is formatted correctly.

1.3 Base Price Selection

If the dataset includes a base price, select it from the dropdown menu. If it's unavailable, you can generate it by choosing 'Generate Base Price.' The base price will be calculated as a 4-week moving average of the non-promotional price.

1.4 Update Changes

After making all adjustments, click 'View Changes' to review the modifications in the table. Confirm changes and click 'Update Changes' to save all modifications.

Note:

2. Data Validation

This tab allows the user to Validate both media variables and non-media variables by analyzing them using relevant plots.

- Select Dependent Variable (will be used for all following plots for reference)

2.1 Media Channels

Choose a media channel for analysis and select impressions, clicks, or spends to generate a line plot. Once data is in line and you consider it valid, click 'Validate' to validate the variable. You can validate each variable individually or use 'Validate All' to validate all variables simultaneously.

Note: Unvalidated variables will not be used for model building or any subsequent analysis

2.2 Non-Media Variables

Similarly, select and analyze any non-media variables present in the dataset, if applicable

2.3 Exploratory Data Analysis

In this section, the user can generate a correlation heatmap by selecting variables from the provided checkboxes. You can choose individual or all variables to view the heatmap.

Additionally, two buttons are available for in-depth exploratory data analysis (EDA):

'Generate Profile Report': Produces a pandas profiling report containing univariate analysis for all variables in the dataset.
'Generate s Report': Creates a comprehensive bivariate analysis report using the open-source Sweetviz package, downloadable for local use.

3. Transformations

Use this section to set ranges for decay and lag parameters for media variables and lags for non-media variables, if necessary.

- Select the global range for adstock and lag transformations via a slider.
- Modify individual variable ranges by selecting options from the dropdown slider if needed.
- Confirm the selections through the summary table generated below.
- Apply the chosen transformations by clicking the 'Apply Transformations' button.
- Adstock will be applied incrementally (0.05 intervals) within the specified range.
- Lag will be applied incrementally (increments of 1) within the selected range.
- After applying transformations, users can select transformed variables for each variable based on their corresponding correlation values from the table.
- Upon completion of selections, click on "Submit Selection" to save the chosen variables.
- If no explicit selections are made, the default setting chooses the top three transformations per variable, considering their correlation with the target. Click on "Submit" to finalize the selections.

4. Model Build and Performance

- This tab allows user to implement transformations and construct personalized models by customizing selected parameters.
- Select a range for adstock and lag transformations using a slider.
- Apply transformations by clicking the 'Apply Transformations' button.
    - Adstock is applied incrementally (0.1 intervals) within the chosen range.
    - Lag is applied incrementally (increments of 1) within the selected range.
- After applying transformations, use the 'Create Combinations of Variables' button.
    - Variable combinations ensure each model iteration contains impressions or transformed clicks from every channel.
- Choose the number of iterations and click 'Build Models.'
- View a table displaying R2, AdjR2, and MAPE for the top 10 valid models.
- Click on any model to view its detailed results and analysis
- Model result is displayed for test data
- Click on “Save this model to tune” and type a name and enter to save the model

Valid Model: After each iteration of model building, the model will be saved if the following criteria are met: all coefficients of the channels are positive, and at least 80% of the variables exhibit statistical significance

5. Model Tuning

This tab allows user to tune the selected model from the previous page.

Event Flags: Helps in quantifying the impact of specific occurrences of events.

- Specify the event's time duration, assign it a name, and save the flag by pressing the Enter key.

Select Parameters to Apply:

Users can choose a saved flag alongside various parameters to apply to the selected model. Upon selecting, clicking "Build model with selected parameters and Flags" applies the chosen settings, showcasing the results of the tuned model.

6. Model Result Overview

This tab offers a comparative analysis of spends and contributions across channels both overall and on a weekly basis. Users can delve into individual channel analysis by selecting a specific channel to explore its breakdown.

7. Build Response Curves

This tab visualizes response curves that depict the relationship between contributed revenue and spending for each channel. Users can interactively adjust curve parameters to achieve a better fit based on their analysis of

Sigmoid curve used:

$$f(x) = \frac{K}{1+b \cdot e^{-a \cdot (x-x_0)}}$$

- x: The input variable represents the independent variable.
- K: The saturation level or the maximum value that the curve approaches as x approaches infinity.
- b: Governs the steepness of the curve. Larger values of 'b' result in a steeper curve.
- a: Determines the rate of change of the curve. Higher values of 'a' mean a faster initial rise or fall in the curve.
- x0: Represents the midpoint of the curve along the x-axis. It's the value of 'x' at which the curve's slope is the steepest.

If a user wishes to adjust any parameter of the curve, they can use the respective widgets to make modifications. After tweaking, by clicking "Update Parameters," these adjusted values will be applied for optimization specifically tailored to that channel's analysis.

8. Scenario Planner

This tab facilitates spend optimization and simulation across all channels/tactics. Users have the option to either optimize all channels collectively or selectively choose specific channels for optimization based on a new budget.

> **Note:** Optimization runs on test data ranging from 6th October 2023 to 5th November 2023.

- **Enter Updated Spend Changes:** Input the revised percentage change in total spends/budget. To optimize for the same budget, enter 0. Choose the channels you wish to optimize.
- **Channel Selection**: Optimize specific channels by selection or optimize all channels with the "Optimize All Channels" widget.
- **Optimization Process**: Click the "Optimize" button to derive optimized results based on the selected spend adjustments.
- **Save Optimized Scenarios:** Save the results in the "Save Scenario" section by assigning a name to the scenario for future reference.

9. Saved Scenario

In this tab, users can both view and download saved scenarios, accessing comprehensive details such as NRPU, ROI, MROI, and spends per NRPU. Multiple scenarios, tailored to different budgets, can be saved on the scenario planner and reviewed individually on this page for thorough analysis.

10.Optimized Result Analysis

This tab offers an extensive visual representation of the optimized results derived from the model. Users can generate these visualizations for each saved scenario from the "Saved Scenarios" page. The visualizations include:

- Percentage spends split for the selected budget across each channel.
- Overall Effectiveness and Efficiency metrics.
- Additionally, users can drill down into individual channels to view various metrics from the dropdown menu for detailed analysis.