

Multi- Touch Attribution model and Marketing spend Optimization

Business Objective

Marketing is a technique of getting potential clients or customers interested in one's products and services. Producers can reach out to their customers via various marketing channels. It looks simple by definition but is indeed a much more challenging task.

Advertisers use various online marketing channels to reach consumers, and they typically want to know how much each channel contributes to their marketing success. This is what is known as multi-channel attribution.

In many cases, advertisers approach this problem using simple models that help them understand the importance of each marketing channel. A few of attribution models are as follows:

- i. Single Touch attribution models like First Touch Conversion, Last Touch Conversion
- ii. Multi-Touch attribution models like Linear Touch Conversion
- iii. Probabilistic models like Markov chains, etc

We will try building all these models in our project to understand how these models will help us find which channels will lead us to more conversions.

In this project, we aim to improve the advertising ROI by quantifying the actual value of the multi-faceted advertising campaigns. This will enable business stakeholders to make decisions based on the millions of converting click paths by isolating the impact of every touchpoint.

Data Description

The data is in the form of a CSV file with a data size of 586737 rows and 6 columns. The columns are as follows:

- Cookie - Anonymous customer-id
- Time - Date and time when the visit took place
- Interaction - Categorical variable indicating the type of interaction that took place
- Conversion - indicating whether a conversion took place, 0: not converted, 1: converted
- Conversion value - Value of the potential conversion event
- Channel (target variable) - The marketing channel that brought the customer to our site

Aim

The project aims at building multiple attribution models on the given dataset to discover channels leading to greater customer conversions.

Tech stack

- Language - Python
- Libraries – numpy, matplotlib, seaborn, itertools, gekko, pandas-profiling

Approach

1. Importing the required dependencies and libraries.
2. Import the dataset.
3. Exploratory Data Analysis (EDA) –
 - Create EDA report using pandas profiling python module
4. Building Single Touch Attribution Models
 - Last Touch Attribution model
 - First Touch Attribution model
 - Last non direct Touch Attribution model
5. Building Multi Touch Attribution Models
 - Linear Attribution model
 - Position based (U-shaped) Attribution model
 - Position Decay Attribution model
6. Probabilistic Attribution model
 - Markov Attribution model
 - Shapley Value model
7. Results
 - Tables – an average of all the models
 - Graphs – plot the models
8. Build a Budget Optimization Engine

Modular code overview

```
input
|_attribution_data.csv

src
|_engine.py
|_ML_Pipeline
    |_data_prep.py
    |_first_touch.py
    |_last_non_direct.py
    |_last_touch.py
    |_linear.py
    |_markov.py
    |_optimization.py
    |_plot_data.py
    |_position_decay.py
    |_shapley.py
    |_u_shaped.py

lib
|_Attribution_modelling_and_Budget_Optimization.ipynb

output
|_graphs for all the models
```

Once you unzip the modular_code.zip file you can find the following folders within it.

1. input
2. src
3. output
4. lib
 1. Input folder - It contains all the data that we have for analysis. Our input data is in the form of a CSV file names, "attribution_data.csv."
 2. Src folder - This is the most important folder of the project. This folder contains all the modularized code for all the above steps in a modularized manner. This folder consists of:
 - engine.py
 - ML_PipelineThe ML_Pipeline is a folder that contains all the functions put into different python files, which are appropriately named. These python functions are then called inside the engine.py file.

3. Output folder – The output folder contains graphs for all the eight models we created.
4. Lib folder - It contains the original ipython notebook that we saw in the videos. The ppt used for the explanation during the videos is provided inside the reference folder.

Project Takeaways

1. Understanding the business problem.
2. Importing the dataset and required libraries.
3. Performing basic Exploratory Data Analysis (EDA).
4. Understanding what Single Touch attribution models are
5. Build a Last-Touch attribution model
6. Build a First-Touch attribution model
7. Build a Last-non-direct Touch attribution model
8. Understanding what Multi-Touch attribution models are
9. Build a Linear attribution model
10. Build a Position-based attribution model
11. Build a Position Decay attribution model
12. Understanding what Probabilistic attribution models are
13. Build a Markov attribution model
14. Build a Shapley value model
15. Understand how to evaluate the results for all the models
16. Build a budget optimization engine for every channel