Maximizing ROI in Digital Advertising with Optimization and Machine Learning

Changhee Han, Minhee Kim, Yuna Seo

cxh210037@utdallas.edu, mxk210141@utdallas.edu, yxs220012@utdallas.edu

Abstract

Bid planning in display advertising has become a significant marketing strategy in the ad auction market, determining ad publishing and creating revenue from ad delivery. Previous works have experimented with bidding strategy optimization using machine learning and linear programming. In our work, we introduce a new utility function that considers ad prioritization and network effects. Additionally, we use XGBoost and CatBoost classifiers to represent the response rate of ad delivery. Our experiment identified the specific ad delivery conditions that have the highest utility. Based on these results, we planned the entire budget.

Introduction

In today's digital age, digital advising has become an essential element for companies reach their target audience such as paid search, social, and display. However, with a limited budget, it's critical to maximize our campaign return to ensure a high ROI. In this paper, we explore the use of optimization techniques, specifically, the Gurobi, to allocate the advertising budget across different digital channels and audience types. In addition, we use the XGBoost and CatBoost for machine learning algorithm to predict the click-through rate(CTR) of different advertising strategies. Our goal is to maximize the campaign return while minimizing costs to guide marketing teams in making data-driven decisions for optimal campaign performance.

Solutions

Budget optimization model

According to Liu, M. et al (2019), a bid strategy in RTB can be formulated as equation 1, the optimization problem with constraints. In this equation, I denotes the set of impressions in the entire advertisement records. The winning function, $w()$, represents the winning probability from the advertisement auction, and $bid()$ means the bidding strategy. $p(x_i)$ indicates the response rate given ad feature x_i, where consists of high-dimensional feature of campaign ad itself. $cost(i)$ means the cost of acquiring impressions from users, and the cost cannot exceed the budget limitation.

$$argmax \sum_{i \in I} w(bid(i)) \times p(x_i)$$

$$s.t. \sum_{i \in I} w(bid(i)) \times cost(i) \leq Budget$$

(1)

The optimization objective is to maximize the valuation from advertisement delivery, which is the product of winning probability with user response in the given bidding strategy. Thus, the optimization aims to find the optimal bid strategy solution to achieve this objective. In our work, we used this formula as a base model and added the utility function.

Ad campaigns have different purposes and approaches based on the target characteristics (Ren, K., et al, 2016 October). We formulated this concept with the utility fuction, $R(\cdot)$, which generates the expected return with the different campaigns, ad types, and audience types (equation 2) within the same budget constraint. Each t, a denotes the component of ad type, T, and audience type, AUD.

$$argmax \sum_{t \in T} \sum_{a \in AUD} \sum_{i \in I} R(i_{t,a}) \times w\left(bid(i_{t,a})\right) \times p\left(x_{i_{t,a}}\right)$$

$$s.t. \sum_{t \in T} \sum_{a \in AUD} \sum_{i \in I} w\left(bid(i_{t,a})\right) \times cost(i_{t,a}) \leq Budget \quad (2)$$

Utility function

The utility function can be defined as the effect from the ad delivery. In our work, we defined the effect as the outcome of each campaign without considering the ad type and audience type. This is because the advertiser pursues the visible output, such as product purchases. Furthermore, from the ad company's perspective, we set the prioritization of the outcome from ad delivery (equation 3).

$$R_{visible} = \alpha \cdot \#webvisit + \beta \cdot \#collateralview + \gamma \cdot \#productview + \delta \cdot \#formcomplete$$

$$\alpha < \beta < \gamma < \delta \quad (3)$$

Furthermore, we defined another ad effect - the network effect - by considering the word-of-mouth (WoM) marketing that occurs with ad consumption. In our work, we experimented with the WoM effect using video completes, social likes, and social shares, which are variables that can be shared with the online community (Shy, O., 2001). We set up a multiple linear regression with these three variables as indicator variables and the click-through rate (CTR) of the next day as the target variable.

	Coeffcients			
	$Intercept_{t-1}$	$videocompletes_{t-1}$	$sociallikes_{t-1}$	$socialshare_{t-1}$
CTR_t	3.3e-0.1***	-6.9e-0.7***	2.93e-04	-2.85e-03

Based on this result, we formulated the network effect term and added it to the utility function (equation 4).

$$R_{network} = \lambda_1 videocompletes + \lambda_2 sociallikes + \lambda_3 socialshare + intercept$$

$$R = R_{visible} + R_{network} \quad (4)$$

Winning Function

We formulated the winning function based on the dollar spent on ad delivery, with the function representing a binary value based on the cost. Based on the bid auction assumption, cost occurs when the bid strategy wins the auction. Therefore, we can interpret cost occurrence as a win in the bid auction (equation 5) (Zhang, W., et al, 2014 August).

$$w\left(bid(i)\right) = \begin{cases} 1, if\ cost\ occur \\ 0, otherwise \end{cases} \quad (5)$$

Response Model

We built a probabilistic model which determines the response rate from variables. Previous literatures used various machine learning frameworks to develop the model. In our work, we utilized XGboost and Catboost as our classifer model (equation 6).

$$p(x_i) = CTR_{prediction} \in [0:1] \quad (6)$$

Results

Machine learning model

Our modeling process followed preprocessing, training, and model evaluation. In our dataset, we removed null data in the channel type column and duplicated data, and transformed categorical ad data into dummy variables. To train the model, we split the data into a train set and a test set with an 80:20 ratio. Finally, we trained the data with our classifier models. In the XGBoost model, we set the max_depth to 500,

n_estimators to 100, and learning rate to 0.01. In the CatBoost model, we used a max depth of 16, epoch of 100, and learning rate of 0.01. The train result is shown in Table 1. With the trained model, we generated predicted CTR.

Table 1. Model evaluation

Classifier	Model Performance				
	Accuracy	Precision	Recall	F1 Score	AUC
XGBoost	0.80	0.57	0.23	0.33	0.59
Catboost	0.89	0.76	0.71	0.73	0.78

Optimal budget plan

We solved equation 2 using the Mixed Integer Linear Programming (MIP) method with the GurobiPy package. This optimization method maximized the utility based on the ad delivery plan while adhering to limited budget constraints. The ad delivery plan included the channel type and audience type, with a total of 8 pairs identified in our dataset. We allocated a budget range of $20 to 50% of the total budget ($1 million) for each channel and audience type. The results are presented in Table 2.

Table 2. optimal budget plan

Ad delivery		Expected utility		Cost ($1,000)	
Channel Type	Audience Type	XGboost	CatBoost	XGboost	CatBoost
Search	1	1329.71	1587.71	410.0	500.0
Display	3	848.71	870.71	285.0	289.0
Display	5	290.70	372.00	140.0	183.0
Search	5	138.71	85.70	85.0	46.0
Social	2	6.71	6.71	20.0	20.0
Display	4	6.71	6.71	20.0	20.0
Social	2	6.71	6.71	20.0	20.0
Social	3	6.71	6.71	20.0	20.0

Based on these results, we identified the top three pairs with high expected utility as (search, audience 1), (display, audience 3), and (display, audience 5). Both machine learning models confirmed the effectiveness of these pairs. Additionally, we found that four pairs, (social, audience 2), (display, audience type 4), (social, audience 2), and (social, audience 3), have the minimum cost, indicating that spending the minimum amount on these pairs can yield optimal results.

Conclusions

We introduced an optimization program with a utility function and machine learning model. In our experiment, we identified the top three ad delivery plans and the optimal budget plan. However, our method predicted the response rate with ad features and did not include customer information. Ad consumption is affected by customer preference. Additionally, our framework has two separated processes: the machine learning model and optimization. In the RTB environment, we cannot deny the optimization effect on the training process. Therefore, for future work, we can integrate the machine learning process and optimization work together, which can improve budget planning. (Bergman, D., et al, 2022)

References

Bergman, D., Huang, T., Brooks, P., Lodi, A., & Raghunathan, A. U. (2022). Janos: an integrated predictive and prescriptive modeling framework. INFORMS Journal on Computing, 34(2), 807-816.

Liu, M., Li, J., Yue, W., Qiu, L., Liu, J., & Qin, Z. (2019, September). An Intelligent Bidding Strategy Based on Model-Free Reinforcement Learning for Real-Time Bidding in Display Advertising. In *2019 Seventh International Conference on Advanced Cloud and Big Data (CBD)* (pp. 240-245). IEEE.

Ren, K., Zhang, W., Rong, Y., Zhang, H., Yu, Y., & Wang, J. (2016, October). User response learning for directly optimizing campaign performance in display advertising. In Proceedings of the 25th acm international on conference on information and knowledge management (pp. 679-688).

Shy, O. (2001). The economics of network industries. Cambridge university press.

Zhang, W., Yuan, S., & Wang, J. (2014, August). Optimal real-time bidding for display advertising. In Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 1077-1086).