Youngju Yoo
Internship Presentation

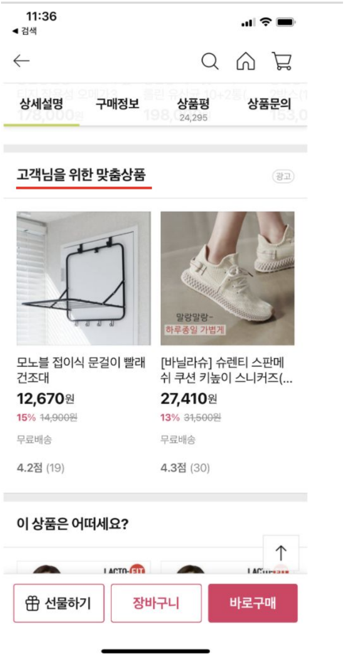
@Youngju Yoo

2022/02/11

Performance Analysis based on requested categories and respond categories

MOLOCO

Background



In the case of PDP(product Detail Page), the GS SHOP requests **selected items** for each Sponsored AD for each **context item**.

Context item

The item that made a request.

`focal-elf-631.commerce_prod.selection_request`

Selected item

Respond item

`focal-elf-631.commerce_prod.selection_response`

Purpose

Context item(request item) category

↑↓

Selected item(response item) category

Does the difference between the requested item categories and the response categories have an impact on advertising performance?

Method

- Set two groups and compare the performance(CTR, CR, ROAS) between them
  - A: (1) Respond to the request, (2) impression occurred, and (3) the requested item’s category and the response items' category were the same.
  - B: (1) Respond to the request, (2) impression occurred, and (3) the requested item’s category and the response items' category were different.
- Also consider the categories 1, 2, 3(large, medium, small) to analysis

```
t_cmp_result AS(
  SELECT
    t_cmp_click.mtid_per_item,
    t_cmp_click.request_category,
    t_cmp_click.response_category,
    t_cmp_click.is_click,
    t_cmp_click.Ad_spending,
    t_cmp_click.click_timestamp,
    IF(purchase.mtid IS NOT NULL, TRUE, FALSE) AS is_purchase,
    SUM(IF(purchase.mtid IS NOT NULL, purchase.revenue.amount, 0)) AS revenue
  FROM
    t_cmp_click
  LEFT OUTER JOIN
    `focal-elf-631.dcr_prod.attributed_purchases` AS purchase
  ON
    t_cmp_click.mtid_per_item = purchase.mtid
    AND TIMESTAMP_SUB(purchase.event_at, INTERVAL 7 DAY) <= t_cmp_click.click_
    AND purchase.event_at > t_cmp_click.click_timestamp
  GROUP BY 1,2,3,4,5,6,7
)
```

colab

Result

Category difference type	impression	Imp ratio(%)	CTR(%)	CR(%)	Ad_spending	revenue	D7_ROAS(%)
Different category	4,424,126	66.24	3.81	1.23	10,155,390	55,437,741	545.89
Same first category	826,415	12.37	5.49	1.24	2,772,340	19,569,365	705.87
Same second category	266,224	3.98	6.08	1.49	964,210	6,924,164	718.11
Same third category	728,819	10.91	7.78	1.86	3,635,670	40,992,989	1127.52
Different category with empty	408,179	6.11	3.32	1.71	827,700	13,262,607	1602.34
Same category with empty	24,417	0.36	2.89	2.41	58,830	1,357,235	2307.04

- If the requested item's category and the response item's category were the same, the performance was better in all of CTR, CR, D7 ROAS than the different category case.
- Also, the more sub-categories are the equal, the better performance in all of CTR, CR, D7 ROAS is.

Discovery of Evergreen Seller Cluster

MOLOCO

Background & purpose

- If we know the typical characteristics of **sellers with good performance**, it will be helpful to track the performance of the RMP product and onboard new clients.

- We know high churn-rate sellers show low performance and low churn-rate sellers show a growth trend from Minho’s pre-analysis

- Define the evergreen sellers and divide evergreen sellers into some clusters and analyze the differences among clusters

Method

1. Define the evergreen sellers

2. Cluster the evergreen sellers

- Data Exploration
  - Check which data we can use
- Research Methodologies
  - Research which clustering algorithm is the best suitable for this task
- Feature engineering
  - define features to be used in clustering
- Find Clusters
  - cluster sellers into several seller groups and compare between groups
- Implications
  - Find the implications (significant differences and its reason) between groups

colab

Result

Definition of Evergreen Seller

Seller who operates campaigns until now and whose valid date(daily campaign spending over ₩1,000) ratio is over 0.7

Row	seller_title	seller_id	seller_end	seller_duration	cnt	ratio
1	(주)컨트롤케이앤제이	1031583	2022-01-11	252	238	0.9444444444444444
2	(주)인씨엘	1028688	2022-01-11	255	246	0.9647058823529412
3	엑스앤오	1040941	2022-01-11	255	253	0.9921568627450981
4	주식회사 다솜인터내셔널	1003350	2022-01-11	179	177	0.9888268156424581
5	금성덴탈	1044013	2022-01-11	255	238	0.9333333333333333
6	(주)커커	1034555	2022-01-11	250	232	0.928
7	주식회사네이쳐리빙	1011216	2022-01-11	253	253	1.0
8	주식회사 한스갤러리	1031744	2022-01-11	93	87	0.9354838709677419
9	빼다 BBEDA	1045023	2022-01-11	252	245	0.9722222222222222
10	피피컴퍼니	1045385	2022-01-11	199	197	0.9899497487437185
11	(주)서울코리아	1038168	2022-01-11	245	240	0.9795918367346939
12	피카소	1042769	2022-01-11	246	228	0.926829268292683
13	(주)디바인바이오	1040333	2022-01-11	253	235	0.9288537549407114
14	D102 메리움점	1043443	2022-01-11	255	252	0.9882352941176471
15	가온누리	1030686	2022-01-11	230	225	0.9782608695652174
16	(주)가담에스앤에이치	1028599	2022-01-11	227	209	0.920704845814978

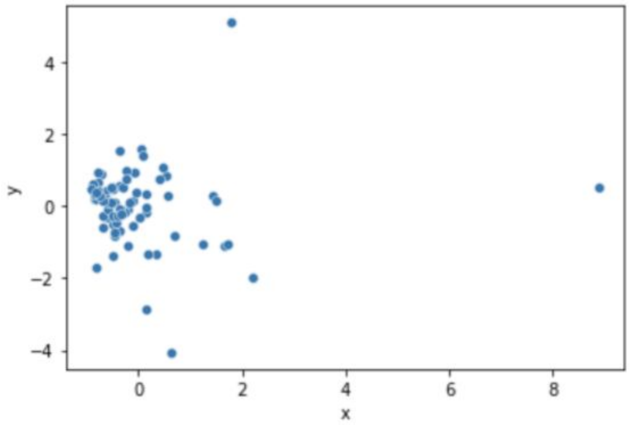
Total 72 evergreen sellers

Result

Cluster evergreen sellers based on seller feature

- Seller features
  - Valid date ratio
  - Avg. # of campaign per month
  - Avg . # of items per campaign
  - Avg. daily budget
  - Avg. Item unit price
  - # of unique categories

PCA algorithm



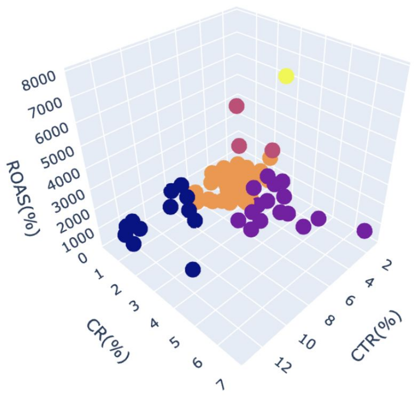
- The distribution of sellers is concentrated except for a few
- Even after seller clustering, there is no evergreen seller group with distinct characteristic
- New approach : group the sellers by their performance(CTR, CR, ROAS)

Result

Cluster evergreen sellers based on seller performance

- Seller Performance
  - CTR(%)
  - CR(%)
  - ROAS(%)

Normalize &
K-means algorithm

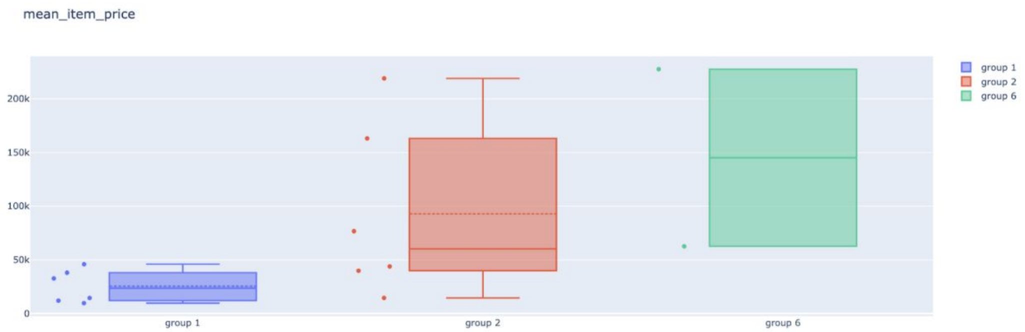


Result

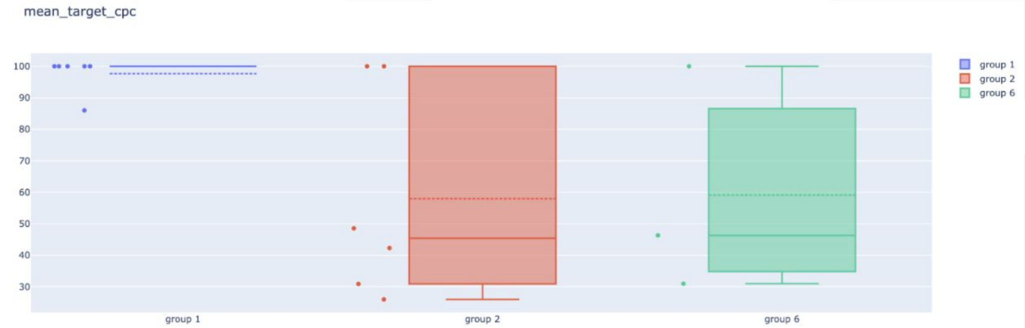
- 3 effective seller group
  - Group1 : High CR, Low CTR
  - Group2 : Low CR, High CTR
  - Group6 : highest ROAS

Result

- 3 effective seller group
  - Group1 : High CR, Low CTR
  - Group2 : Low CR, High CTR
  - Group6 : highest ROAS



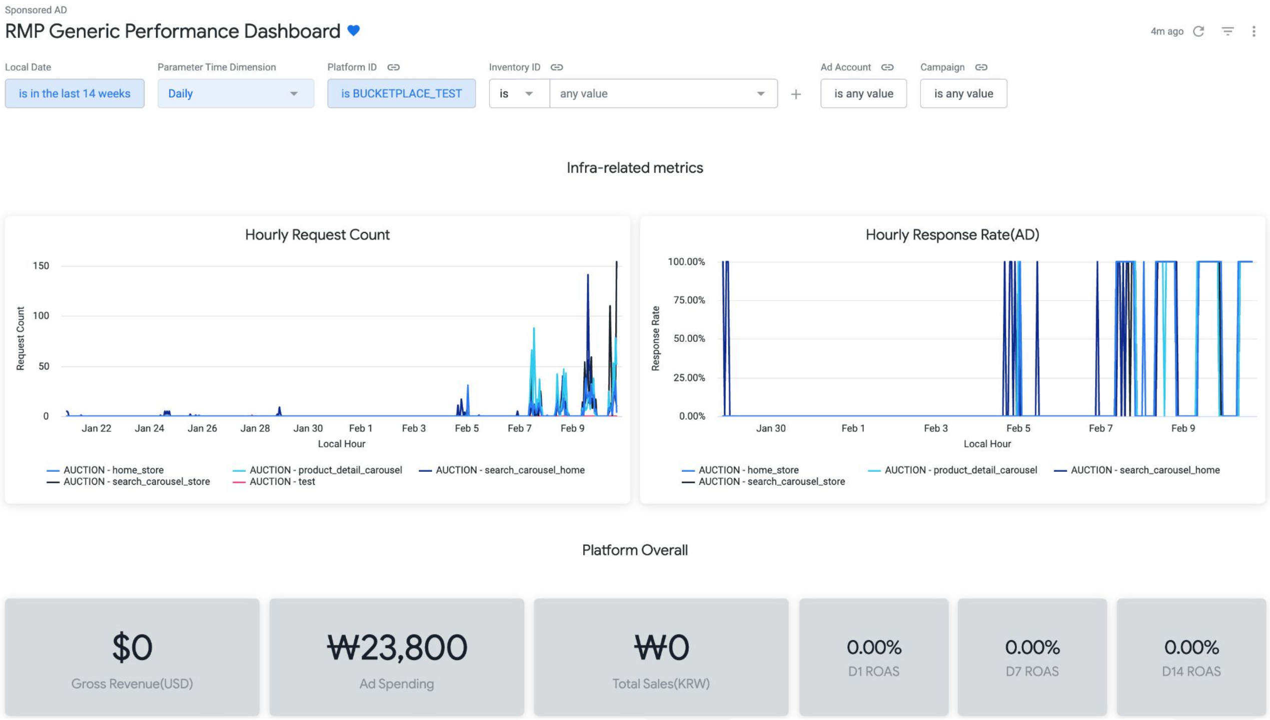
Seller who operates campaign with high price of items show low CR



Seller who operates campaign with low target cpc show high CR

[Bucketplace] Sponsored AD Looker Dashboard

Result



RMP Generic Performance Dashboard

RMP Value Pricing Investigation

Background & Purpose

platform_id	mtid	mtid_per_item	rank	cpc	score
GSSHOP	ChA--MqJGhCjb5bSRcmwDTuEMbvkZAGGhQIARoQvtCk47t5SOiS1vwKGNdIlCACKgAyAA	ChA6NQaIN3RDh5wXCf6B0FYNEMbvkZAGGhQIARoQvtCk47t5SOiS1vwKGNdIlCACKgAyAA	0	100.0	0.0703380182385446
GSSHOP	ChA--MqJGhCjb5bSRcmwDTuEMbvkZAGGhQIARoQvtCk47t5SOiS1vwKGNdIlCACKgAyAA	ChCN2FmFKTFNf49d8ISxwjF0EMbvkZAGGhQIARoQvtCk47t5SOiS1vwKGNdIlCACKgAyAhgB	1	100.0	0.066044382750988
GSSHOP	ChA--MqJGhCjb5bSRcmwDTuEMbvkZAGGhQIARoQvtCk47t5SOiS1vwKGNdIlCACKgAyAA	ChAvGXuEkEBEIlbhJJDdnLXhEMbvkZAGGhQIARoQvtCk47t5SOiS1vwKGNdIlCACKgAyAhgC	2	100.0	0.05926716700196266
GSSHOP	ChA--MqJGhCjb5bSRcmwDTuEMbvkZAGGhQIARoQvtCk47t5SOiS1vwKGNdIlCACKgAyAA	ChDSCCYDAGlDHYwinPEnuKgcEMbvkZAGGhQIARoQvtCk47t5SOiS1vwKGNdIlCACKgAyAhgD	3	100.0	0.05584614723920822
GSSHOP	ChA--MqJGhCjb5bSRcmwDTuEMbvkZAGGhQIARoQvtCk47t5SOiS1vwKGNdIlCACKgAyAA	ChBt3tQ65YJOy70s99AjaLCaEMbvkZAGGhQIARoQvtCk47t5SOiS1vwKGNdIlCACKgAyAhgE	4	100.0	0.04614824801683426
GSSHOP	ChA--MqJGhCjb5bSRcmwDTuEMbvkZAGGhQIARoQvtCk47t5SOiS1vwKGNdIlCACKgAyAA	ChAmy-3E2ONFXYZxlqXIo4F4EMbvkZAGGhQIARoQvtCk47t5SOiS1vwKGNdIlCACKgAyAhgF	5	100.0	0.037515074014663696
GSSHOP	ChA--MqJGhCjb5bSRcmwDTuEMbvkZAGGhQIARoQvtCk47t5SOiS1vwKGNdIlCACKgAyAA	ChCj3AYRgoxIR6H77DHWHdXJEMbvkZAGGhQIARoQvtCk47t5SOiS1vwKGNdIlCACKgAyAhgG	6	70.0	0.040931664407253265
GSSHOP	ChA--MqJGhCjb5bSRcmwDTuEMbvkZAGGhQIARoQvtCk47t5SOiS1vwKGNdIlCACKgAyAA	ChARvp4EEP9NplOyly459jpfEMbvkZAGGhQIARoQvtCk47t5SOiS1vwKGNdIlCACKgAyAhgH	7	70.0	0.034544359892606735
GSSHOP	ChA--MqJGhCjb5bSRcmwDTuEMbvkZAGGhQIARoQvtCk47t5SOiS1vwKGNdIlCACKgAyAA	ChBFVZXAtVlJl7GsRGmU2WDPEMbvkZAGGhQIARoQvtCk47t5SOiS1vwKGNdIlCACKgAyAhgI	8	100.0	0.019597411155700684
GSSHOP	ChA--MqJGhCjb5bSRcmwDTuEMbvkZAGGhQIARoQvtCk47t5SOiS1vwKGNdIlCACKgAyAA	ChD9YbSN9mxEELWnaYhibUoZEMbvkZAGGhQIARoQvtCk47t5SOiS1vwKGNdIlCACKgAyAhgJ	9	100.0	0.016598261892795563
GSSHOP	ChA--MqJGhCjb5bSRcmwDTuEMbvkZAGGhQIARoQvtCk47t5SOiS1vwKGNdIlCACKgAyAA	ChCZ5JuJJYJOB7L-AuYjitfUEMbvkZAGGhQIARoQvtCk47t5SOiS1vwKGNdIlCACKgAyAhgK	10	100.0	0.014208987355232239
GSSHOP	ChA--MqJGhCjb5bSRcmwDTuEMbvkZAGGhQIARoQvtCk47t5SOiS1vwKGNdIlCACKgAyAA	ChBI1ogjr5JEsb0odFPSkY0hEMbvkZAGGhQIARoQvtCk47t5SOiS1vwKGNdIlCACKgAyAhgL	11	100.0	0.013976757414638996
GSSHOP	ChA--MqJGhCjb5bSRcmwDTuEMbvkZAGGhQIARoQvtCk47t5SOiS1vwKGNdIlCACKgAyAA	*null*	12	20.0	0.02286849357187748

Value price : CPC x relevency_score → Top 6 value price are selected

Value price formula maximizes sellers' Ad spending compared to impressions → Check whether the high-ranked items showed higher **(Ad spending) / (impression count).**

If there's no relationship, propose a new value-pricing formula

Method

1. decide conditions to gather data: time period, platform (GS SHOP), inventory id, etc.

2. gather data (by joining the internal auction table with history_imp/click tables)

3. check the relationship between rank and (Ad spending)/(impression count)

4. If there's no relationship, propose a new value-pricing formula instead of CPC x relevancy_score.

[colab]()

```
result_by_rank AS(
SELECT
  internal_auction.rank,
  COUNT(internal_auction.mtid_per_item) AS imp_cnt,
  COUNT(t_click.click_info.Ad_spending) AS click_cnt,
  COUNT(purchase.mtid) AS purchase_cnt,
  SUM(t_click.click_info.Ad_spending) AS Ad_spending,
  SUM(purchase.revenue.amount) AS revenue
FROM
  internal_auction
LEFT OUTER JOIN
  t_click
ON
  internal_auction.mtid_per_item = t_click.click_info.mtid
LEFT OUTER JOIN
  `focal-elf-631.dcr_prod.attributed_purchases` AS purchase
ON
  internal_auction.mtid_per_item = purchase.mtid
  AND purchase.platform_id = "GSSHOP"
WHERE
  internal_auction.mtid_count = 13
GROUP BY
  1 ),
```

```
SELECT
  *,
  SAFE_DIVIDE(Ad_spending,
    imp_cnt) * 100 AS spending_per_imp,
  SAFE_DIVIDE(click_cnt,
    imp_cnt) *100 AS CTR,
  SAFE_DIVIDE(purchase_cnt,
    click_cnt) *100 AS CR,
  SAFE_DIVIDE(purchase_cnt,
    imp_cnt) *100 AS purchase_per_imp,
  SAFE_DIVIDE(revenue,
    Ad_spending) * 100 AS ROAS
FROM
  result_by_rank
ORDER BY
  rank
```

Result

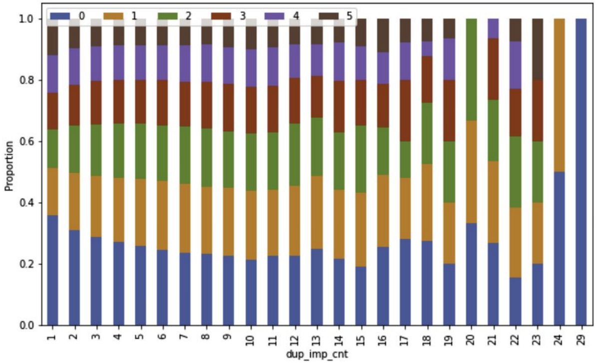
rank	imp_cnt	click_cnt	purchase_cnt	Ad_spending	revenue	spending_per_imp	CTR	CR	ROAS	purchase_per_imp
0	1079498	43850	1000	3320350.0	3.5606913E7	307.583	4.062	2.281	1072.384	0.093
1	251946	26997	466	2001740.0	1.842108E7	794.512	10.715	1.726	920.253	0.185
2	180663	14608	239	1028170.0	8851709.0	569.109	8.086	1.636	860.919	0.132
3	166469	13049	181	897160.0	5726649.0	538.935	7.839	1.387	638.309	0.109
4	155658	11817	157	806630.0	5324500.0	518.207	7.592	1.329	660.092	0.101
5	152099	11445	148	761700.0	5427467.0	500.792	7.525	1.293	712.547	0.097

- The higher the rank, the higher CTR and ad spending per impression except rank0
- In rank0, the lowest CTR and ad spending per impression but the highest ROAS. We need to find the cause for extremely low CTR and ad spending per impression in rank0
- Also, since it shows great correlation among rank, CTR, and, ad spending per impression with CPC x relevancy score formula, we don’t have to propose a new value-pricing formula

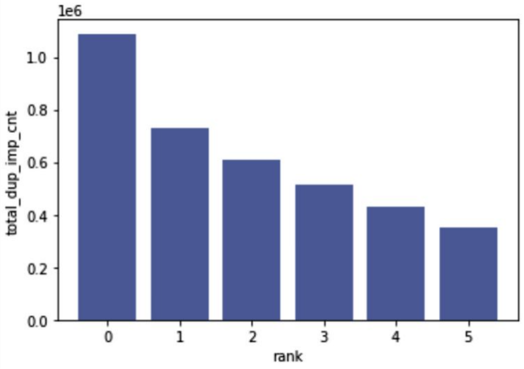
Additional analysis

Find the cause of extremely low performance of rank0

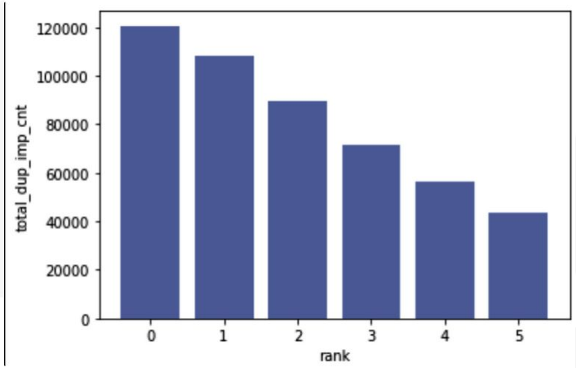
- We supposed that the lowest CTR of rank0 is due to high impression caused by high repeated impression count.
- Repeated Impression : Hourly repeated impressions of the same item to the same user
- But repeated impression occurs evenly at all rank → repeated impression is not the cause of the low performance of rank0



proportion of rank for each duplicate impression count



total duplicate impression count for each rank



total duplicate impression count over 5 for each rank

It will be continued with jeonghyun,,

Investigating the different characteristics of new models

Background & Purpose

- The ML team has run two experiments for GSSHOP HOME sponsored ads
    - Baseline model = recommends items based on user signals data
    - CTR optimization model = recommends items based on maximized CTR
    - Sales amount optimization model = recommends items based on maximized ROAS
- Have to check the characteristics of those different ranking scores
- Investigate the metrics'(CTR, CR, ROAS, etc.) trends by the rank in internal auctions

Result

Baseline model

rank	imp_cnt	click_cnt	purchase_cnt	Ad_spending	revenue	spending_per_imp	CTR	CR	ROAS	purchase_per_imp
0	69541	2096	38	169930.0	3066500.0	244.359	3.014	1.813	1804.567	0.055
1	29745	1265	13	100620.0	339647.0	338.275	4.253	1.028	337.554	0.044
2	25110	664	5	48160.0	59560.0	191.796	2.644	0.753	123.671	0.02
3	24594	577	4	41570.0	65340.0	169.025	2.346	0.693	157.181	0.016
4	24431	535	7	39070.0	170028.0	159.92	2.19	1.308	435.188	0.029
5	24206	524	7	34550.0	132573.0	142.733	2.165	1.336	383.713	0.029

- The higher rank, the higher CTR except rank0

Result

CTR optimized model

rank	imp_cnt	click_cnt	purchase_cnt	Ad_spending	revenue	spending_per_imp	CTR	CR	ROAS	purchase_per_imp
0	69339	2242	16	184460.0	451019.0	266.026	3.233	0.714	244.508	0.023
1	30096	1405	12	112600.0	480823.0	374.136	4.668	0.854	427.019	0.04
2	25472	662	8	49230.0	148447.0	193.271	2.599	1.208	301.538	0.031
3	24560	541	5	39090.0	263250.0	159.161	2.203	0.924	673.446	0.02
4	24169	540	7	38580.0	189420.0	159.626	2.234	1.296	490.98	0.029
5	24067	533	4	35210.0	85420.0	146.3	2.215	0.75	242.602	0.017

- The higher rank, the higher CTR as model trained except rank0

Sales amount optimized model

rank	imp_cnt	click_cnt	purchase_cnt	Ad_spending	revenue	spending_per_imp	CTR	CR	ROAS	purchase_per_imp
0	55677	1407	16	112830.0	779231.0	202.651	2.527	1.137	690.624	0.029
1	23892	1025	9	80170.0	335392.0	335.552	4.29	0.878	418.351	0.038
2	19969	545	5	40270.0	89449.0	201.663	2.729	0.917	222.123	0.025
3	19536	459	2	34340.0	62720.0	175.778	2.35	0.436	182.644	0.01
4	19286	446	4	32590.0	134403.0	168.983	2.313	0.897	412.406	0.021
5	19408	456	2	31090.0	22710.0	160.192	2.35	0.439	73.046	0.01

- Expected result was the higher the rank, the higher ROAS, but result came out different. We should test again after gathering more data.

Internship review

Great time at MOLOCO !!

THANK YOU!

www.moloco.com