

Федеральное государственное автономное образовательное учреждение высшего
образования
"Национальный исследовательский университет
"Высшая школа экономики"

Образовательная программа "Прикладная математика"
бакалавр

ОТЧЕТ
по проектной работе

Оптимальное распределение расходов между рекламными источниками в
интернет-маркетинге.

Выполнили студенты группы БПМ-193

Сабитов Эльфат Рамильевич
Солдатов Алексей Валерьевич

Руководитель проекта

Иванов Фёдор Ильич

(должность, ФИО руководителя проекта)

(оценка)

(подпись)

(дата)

Москва - 2020
25 мая 2020

Содержание

1	Введение	2
2	Используемые обозначения	2
3	<i>М&М</i> - Метрики и Методы	2
4	Алгоритм	3
5	Итог	5
6	Ссылка на код	5
	Список литературы	5

1 Введение

Двадцать первый век - век интернета, в котором не найти места, не затронутого рекламой. Как всем прекрасно известно, её основная цель - привлечение покупателей, но с этим приходит необходимость анализировать, связанные с этим данные. В нашем проекте был произведен прогноз таких данных на месяц при помощи имеющейся статистики по кампаниям.

2 Используемые обозначения

- BID_i - количество оптимизируемых фраз
- CR_i - процентное содержание конверсий в некотором количестве кликов за определённое время.
- $CR_{ур}$ - коэффициент конверсии более высокого уровня.
- $CPC_i = k_i * BID_i$ - средняя цена клика по i фразе, а k - некоторый коэффициент
- CV_i - количество конверсий по i фразе
- $CL_i = w_i * \sqrt{BID_i}$ - количество кликов по i фразе, а w - так же некоторый коэффициент
- CPA_i - стоимость конверсии по i фразе

3 М&М- Метрики и Методы

Если вы начнете глубже изучать интернет маркетинг и его особенности, то неизбежно столкнётесь с определёнными метриками, наиболее важные из которых мы рассмотрим далее.

Конверсия - одна из самых часто встречающихся метрик, показывающая эффективность работы построенной модели оптимизации контекстной рекламы: чем больше конверсия, тем лучше для рекламодателя. Конверсия по своей сути показывает какая часть пользователей, посетивших сайт, совершила на этом сайте целевые действия. Целевым действие в свою очередь называется любая активность на сайте компании.

Пулинг - если вам не хватает статистики, то вам стоит обращаться к нему



$$CR = \frac{CV + 1}{CL + \frac{1}{CR_{up}}}$$

Эта простая версия пулинга спасет вас и ваш прогноз, если вы обделены статистикой по слову здесь и сейчас. Как оно работает? - Всё просто. Мы должны взять имеющееся количество кликов CL и конверсий CV, а потом взять еще коэффициент конверсий для следующего уровня пулинга. Т.е. при подсчете коэффициента конверсий по фразе мы можем обратиться к коэффициенту конверсий группы объявлений, но в нашем случае это был сразу CR кампании.

Ставка -это цена, которую компания готова заплатить, если по её объявлению пользователем был совершен клик. Рекламные системы не берут плату за размещение объявления на сайте, плата взимается лишь в случае перехода на сайт компании. Данный переход не гарантирует совершения пользователем конверсии, поэтому необходимо предсказывать их возможное количество для оптимального распределения бюджета между рекламными системами.

Для стабильной работы всех метрик сайту следует использовать специальные инструменты. Как пример можно взять инструменты Яндекс. Метрики, которые позволяют отслеживать кол-во посетителей, конверсии и вообще вести хорошую статистику вашего сайта. Мы в данной работе получаем метрики в виде таблицы Excel с статистикой для фраз и кампаний за различные временные промежутки. Обращение к ним осуществляется посредством библиотеки `orenpuxl`.

4 Алгоритм

Для начала мы высчитываем коэффициент конверсий в масштабах кампании

	A	B	C	D	E	F	G	H	I	J	K	L
1	campaign id	clicks fact	clicks predict	sum fact	sum predict	CV fact	CV predicted	CPC fact	CPC predicted	CPA fact	CPA predicted	
2	16846504	77	55	8 495,00 P	4 353,00 P	4	3	110,32 P	79,15 P	2 123,75 P	1 451,00 P	
3	16913681	804	448	121 524,00 P	29 743,00 P	22	10	151,15 P	66,39 P	5 523,82 P	2 974,30 P	
4	17614076	37	37	4 517,00 P	4 746,00 P	3	3	122,08 P	128,27 P	1 505,67 P	1 582,00 P	
5	17614778	100	150	11 511,00 P	5 058,00 P	7	11	115,11 P	33,72 P	1 644,43 P	459,82 P	
6	17615323	28	5	3 652,00 P	292,00 P	1	1	130,43 P	58,40 P	3 652,00 P	292,00 P	
7	17615391	33	27	2 438,00 P	1 870,00 P	1	1	73,88 P	69,26 P	2 438,00 P	1 870,00 P	
8	17615460	15	6	2 036,00 P	810,00 P	0	0	135,73 P	135,00 P	0,00 P	0,00 P	
9	17615564	43	42	3 256,00 P	1 899,00 P	2	2	75,72 P	45,21 P	1 628,00 P	949,50 P	
10	20887923	5	5	310,00 P	248,00 P	0	0	62,00 P	49,60 P	0,00 P	0,00 P	
11	20887956	14	13	404,00 P	529,00 P	0	0	28,86 P	40,69 P	0,00 P	0,00 P	
12	21788563	0	0	0,00 P	0,00 P	1	1	0,00 P	0,00 P	0,00 P	0,00 P	
13	21788573	2	2	106,00 P	106,00 P	0	0	53,00 P	53,00 P	0,00 P	0,00 P	
14	22148117	99	100	3 096,00 P	2 288,00 P	3	3	31,27 P	22,88 P	1 032,00 P	762,67 P	
15	22149059	106	86	14 384,00 P	12 067,00 P	4	4	135,70 P	140,31 P	3 596,00 P	3 016,75 P	
16	22149592	23	20	1 405,00 P	774,00 P	0	0	61,09 P	38,70 P	0,00 P	0,00 P	
17	22150151	15	15	939,00 P	750,00 P	0	0	62,60 P	50,00 P	0,00 P	0,00 P	
18	22150456	0	0	0,00 P	0,00 P	0	0	0,00 P	0,00 P	0,00 P	0,00 P	
19	22150620	0	0	0,00 P	0,00 P	0	0	0,00 P	0,00 P	0,00 P	0,00 P	
20	22152389	16	16	1 471,00 P	1 077,00 P	0	0	91,94 P	67,31 P	0,00 P	0,00 P	
21	22394083	56	56	2 735,00 P	1 793,00 P	2	2	48,84 P	32,02 P	1 367,50 P	896,50 P	
22	31425540	761	731	14 614,00 P	10 591,00 P	2	1	19,20 P	14,49 P	7 307,00 P	10 591,00 P	
23	33788979	14	14	320,00 P	260,00 P	0	0	22,86 P	18,57 P	0,00 P	0,00 P	
24	41002217	26	26	596,00 P	721,00 P	1	1	22,92 P	27,73 P	596,00 P	721,00 P	
25	summary	2274	1854	197809	79975	53	43	86,99 P	43,14 P	3 732,25 P	1 859,88 P	
26												
27												

$$CR_{campaign}^j = \frac{CV_{campaign}^j}{CLICKS_{campaign}^j}$$

Ищем коэффициент конверсий для j-ой кампании, где $CV_{campaign}^j$ - конверсия j кампании, а $CLICKS_{campaign}^j$ - клики j кампании

Далее ищем коэффициент конверсий фразы, используя пулинг.

$$CR_{word}^i = \frac{CV_{word}^i + 1}{CL_{word}^i + \frac{1}{CR_{campaign}^j}}$$

При $Word_i \in Campaign_j$

Т.е. используем статистику кампании, в которую входит слово. При этом мы используем неполную статистику по фразам за некоторый промежуток времени (CV_{word}^i и CL_{word}^i).

Необходимо вычислить коэффициенты k и w для каждой кампании по формулам, которые будут использоваться в дальнейших вычислениях средней цены клика и количества конверсий по i фразе.

$$k = \frac{CPC}{CPA \cdot CR_{campaign}}$$

$$w = \frac{CLICKS_{campaign}}{\sqrt{CPA_{campaign} \cdot CR_{campaign}}}$$

Снова смотрим неполную статистику для слов за некоторый промежуток времени

$$CL_{word}^i = w * \sqrt{BID_{word}^i}$$

$$CPC_{word}^i = k * BID$$

Считаем клики и их среднюю цену i фразам

$$CV_{word}^i = CL_{word}^i * CR_{word}^i$$

Получаем конверсии по i фразе.

Наша цель - получить прогноз по следующим метрикам CPC, CL, CV для кампаний. Чтобы получить результаты, нужно просуммировать данные метрики для каждого i слова, входящего в j кампанию.

$$CPC_{pred}^j = \sum_i CPC_{word}^i | word_i \in campaign_j$$

$$CL_{pred}^j = \sum_i CL_{word}^i | word_i \in campaign_j$$

$$CV_{pred}^j = \sum_i CV_{word}^i | word_i \in campaign_j$$

campaign_id_new	clicks_result	cpc_result	cv_result
16913681	4215594,552	209,0624	117013,43
16846504	24534,20704	377,05215	1272,7833
31425540	2727784,744	150,7319	7266,4352
17614076	6487,943913	136,21023	423,60478
17614778	29384,19556	45,891603	2013,1531
17615323	2818,897902	476,55564	113,41429
17615391	15462,08505	769,53907	1076,2763
17615564	17384,12566	382,41887	805,50147
22149059	118775,6018	217,06082	5168,7863
22394083	32837,33114	120,7976	1165,1752
22148117	44417,5744	63,694816	1319,3353
22148117	134,5794234	57,79	4,0781643

5 Итог

Наша модель действительно позволяет сделать прогноз таких метрик, как средняя цена клика *cpc_result*, количество кликов *clicks_result*, количество конверсий рекламной кампании *cv_result*, что делает её применимой в анализе данных в интернет-маркетинге и контекстной рекламе. Данные метрики позволяют компании более эффективно распределять расходы на рекламу, получая наилучший эффект от неё.

6 Ссылка на код

Солдатов Алексей Валерьевич: <https://github.com/Alex193/HSE>

Сабитов Эльфат Рамильевич: https://github.com/MarioAuditore/Project_HSE_freshman

Список литературы

- [1] <https://habr.com/ru/company/calltouch/blog/341390/>
Calltouch - "Оптимизация ставок: зависимость между ценой клика и установленной ставкой"
- [2] <https://habr.com/p/337888/>
Calltouch - "Использование различных метрик для кластеризации ключевых запросов"
- [3] <https://www.mccme.ru/free-books/llang/newllang.pdf>
"Набор и вёрстка и системе LATEX"