

AiCon

全球人工智能与机器学习技术大会

从点击率提升到用户长期行为优化

浅谈人工智能技术在电商搜索的落地应用

洪亮劼

Etsy数据科学主管

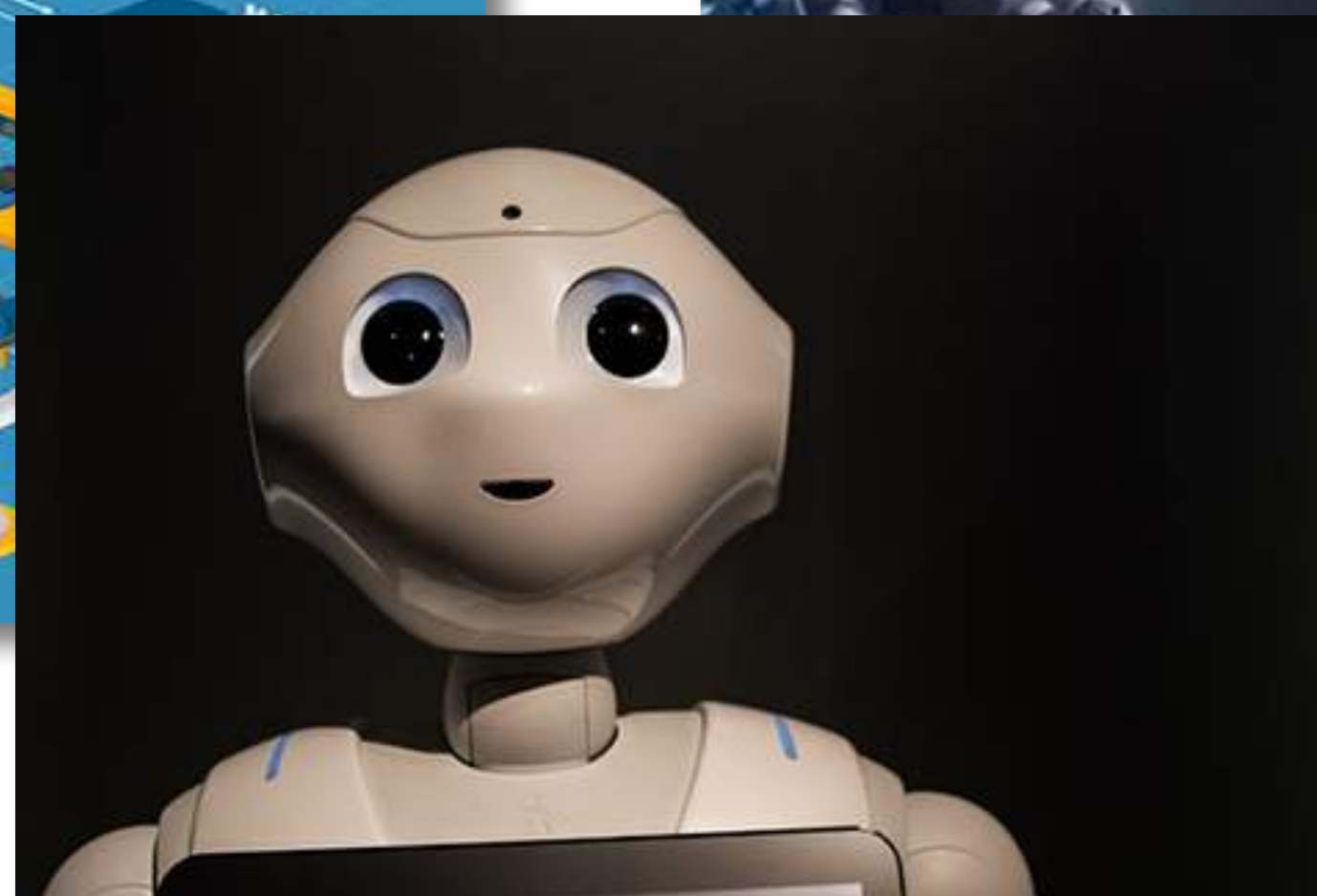
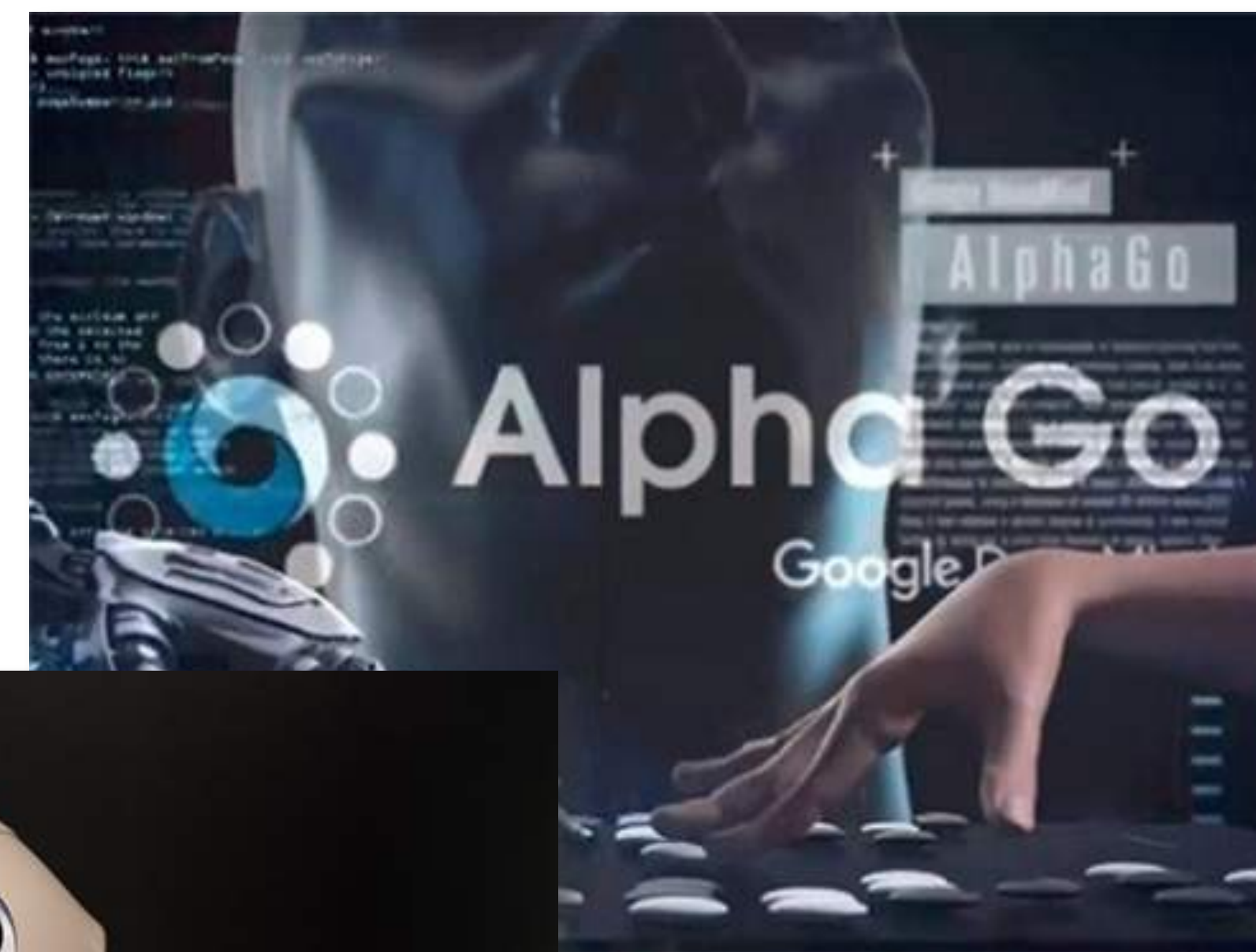
TABLE OF CONTENTES

人工智能产品的系统性提升

从点击率到用户长期行为优化

人工智能在电商搜索的应用

人工智能对各行业的影响



人工智能产品系统性提升

人工智能落地难不难？



人工智能产品系统性提升

人工智能落地难不难？

- 碰运气，照搬



人工智能产品系统性提升

人工智能落地难不难？

- 碰运气，照搬
- 没有理解产品场景



人工智能产品系统性提升

人工智能落地难不难？

- 碰运气，照搬
- 没有理解产品场景
- 难在不能系统性提升



人工智能产品系统性提升

什么是系统性提升？



人工智能产品系统性提升

怎么利用数据驱动的人工智能技术来不断系统性地提升产品的品质？



人工智能产品系统性提升

怎么利用数据驱动的人工智能技术来不断系统性地提升产品的品质？



人工智能产品系统性提升

利用数据驱动的人工智能技术来不断系统性地提升产品的品质。

- 是方法论



人工智能产品系统性提升

利用数据驱动的人工智能技术来不断系统性地提升产品的品质。

- 是方法论
- 是系统工程



人工智能产品系统性提升

利用**数据驱动**的人工智能技术来不断**系统性**地提升产品的品质。

- 是方法论
- 是系统工程
- 不是算法和模型的简单叠加



人工智能产品系统性提升

产品的长期目标



人工智能产品系统性提升

产品的长期目标

- 和整个公司目标密切相关



人工智能产品系统性提升

产品的长期目标

- 和整个公司目标密切相关
- 需要对用户进行长时间观测



人工智能产品系统性提升

产品的长期目标

- 和整个公司目标密切相关
- 需要对用户进行长时间观测
- 非常不敏感



人工智能产品系统性提升

产品的长期目标

- 和整个公司目标密切相关
- 需要对用户进行长时间观测
- 非常不敏感
- 一般无法进行在线实验



人工智能产品系统性提升

产品的中期目标



人工智能产品系统性提升

产品的中期目标

- 不一定是公司长期目标



人工智能产品系统性提升

产品的中期目标

- 不一定是公司长期目标
- 需要对用户进行一段时间观测



人工智能产品系统性提升

产品的中期目标

- 不一定是公司长期目标
- 需要对用户进行一段时间观测
- 相对不是很敏感



人工智能产品系统性提升

产品的中期目标

- 不一定是公司长期目标
- 需要对用户进行一段时间观测
- 相对不是很敏感
- 一般可以进行在线实验



人工智能产品系统性提升

产品的短期目标



人工智能产品系统性提升

产品的短期目标

- 一定不是公司长期目标



人工智能产品系统性提升

产品的短期目标

- 一定不是公司长期目标
- 可以很快对用户观测而得到



人工智能产品系统性提升

产品的短期目标

- 一定不是公司长期目标
- 可以很快对用户观测而得到
- 比较很敏感



人工智能产品系统性提升

产品的短期目标

- 一定不是公司长期目标
- 可以很快对用户观测而得到
- 比较很敏感
- 一般可以进行在线实验



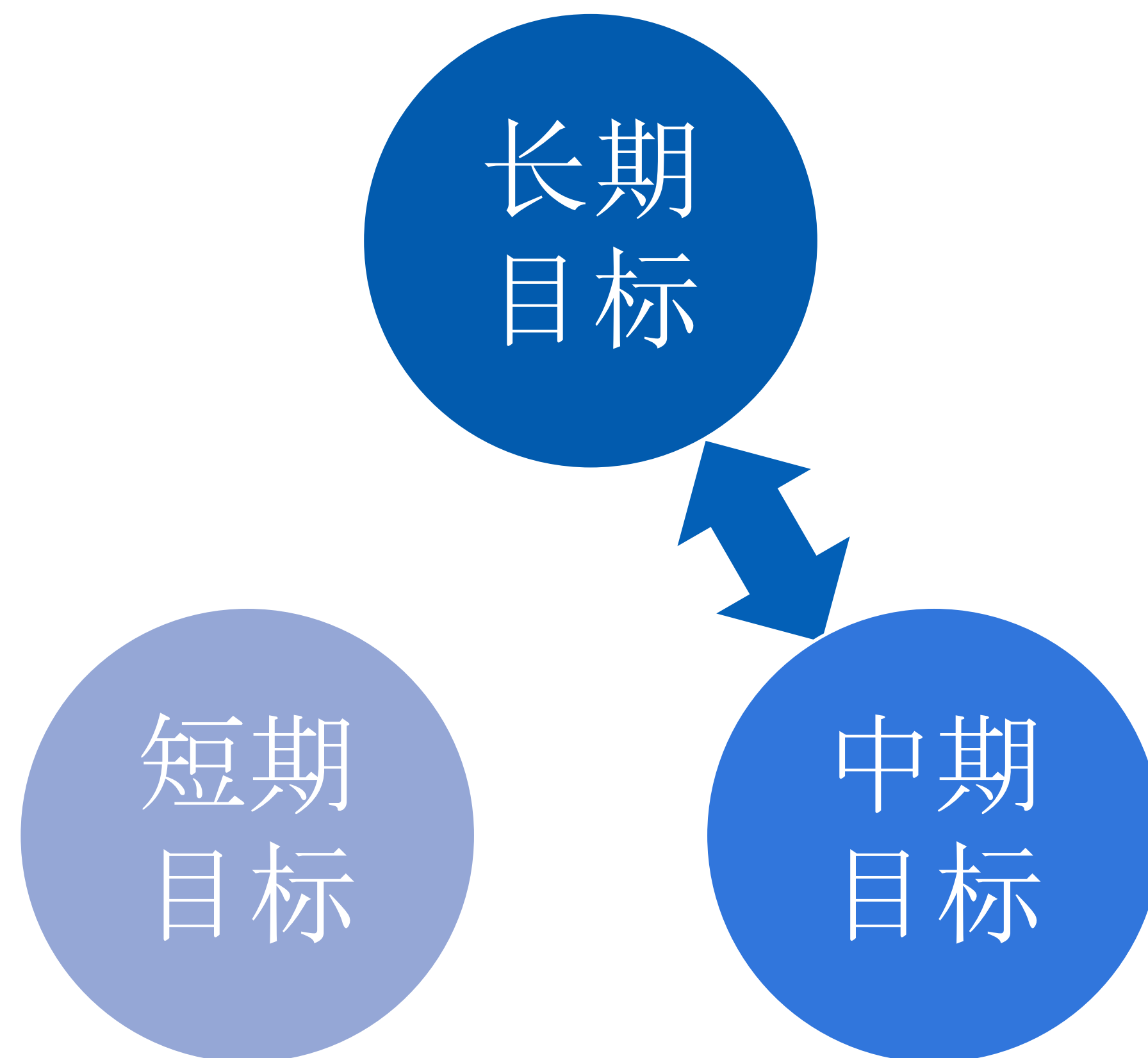
人工智能产品系统性提升

长期
目标

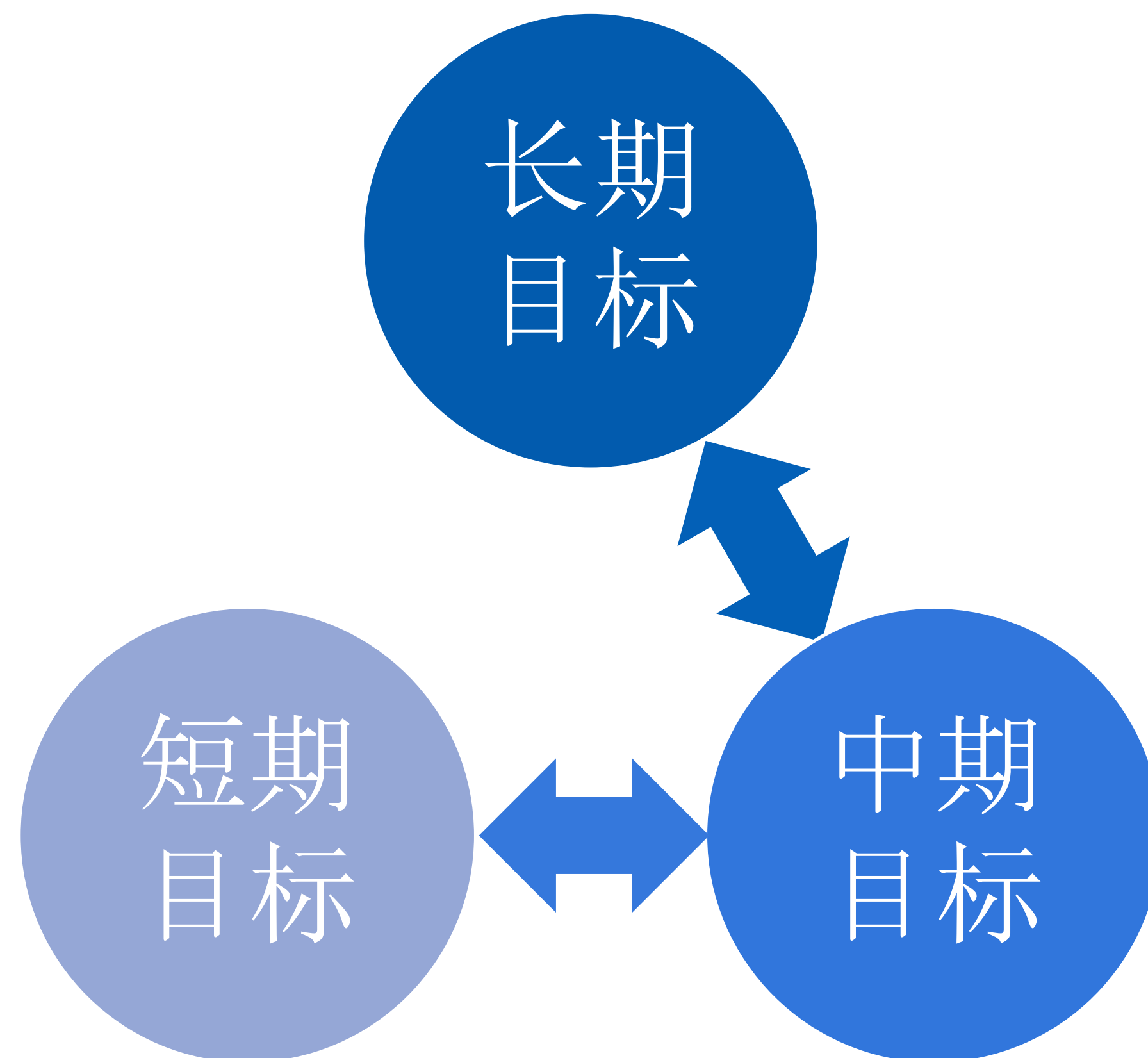
短期
目标

中期
目标

人工智能产品系统性提升



人工智能产品系统性提升



人工智能产品系统性提升

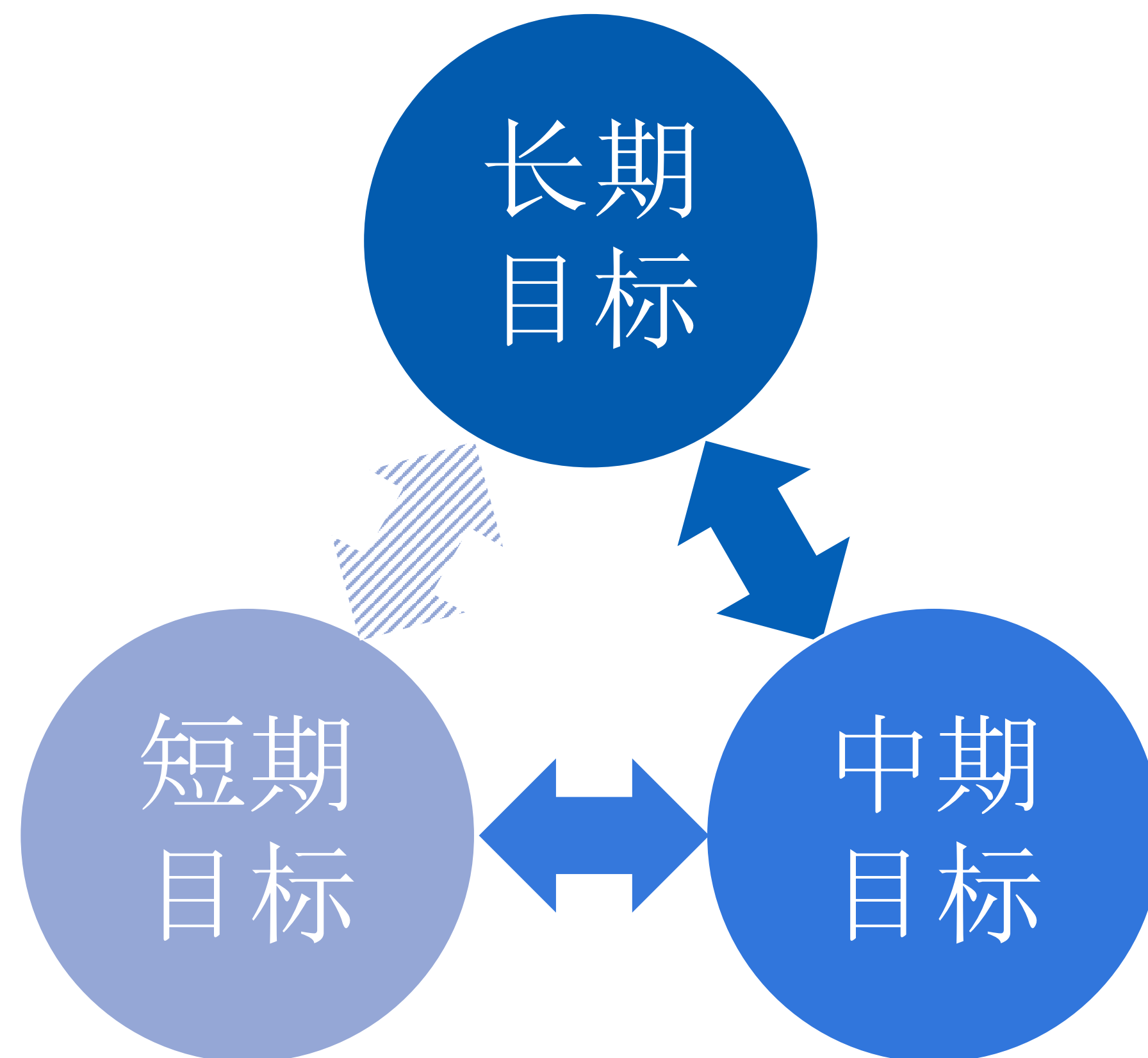


TABLE OF CONTENTES

人工智能产品的系统性提升

从点击率到用户长期行为优化

人工智能在电商搜索的应用

电商搜索产品

Amazon search results for "canon camera". The page shows a search bar with "canon camera" entered. Below the search bar, there are navigation links and a list of product categories. The main content area displays several camera bundles, including a Canon EOS Rebel T6 Bundle and a Canon EOS 5D Mark IV bundle. Each bundle includes a camera, lens, and accessories. The prices are \$446.99 and \$5,348.00 respectively. There are also filters for international shipping, Amazon Prime, and price ranges.

JD.com search results for "李斯特" (Liszt). The page shows a search bar with "李斯特" entered. Below the search bar, there are navigation links and a list of product categories. The main content area displays a list of books and music-related products, including "李斯特 (LISZT)" and "李斯特 (LUSHENGSHI)". There are filters for brand, author, and genre. The page also shows a "精品推荐" (Special Recommendation) section.

eBay search results for "xiaomi". The page shows a search bar with "xiaomi" entered. Below the search bar, there are navigation links and a list of product categories. The main content area displays several Xiaomi smartphones, including a 5.0" Xiaomi Redmi 4 and a Xiaomi Mi A1. Each listing includes a product image, title, price, and shipping information. There are filters for categories, price, and shipping options.



电商搜索产品



电商搜索产品

The screenshot shows the Etsy search results for 'harry potter'. The search bar at the top contains 'harry potter' and a 'Search' button. Below the search bar are navigation links for 'Sell on Etsy', 'Register', and 'Sign'. The main content area displays a grid of product listings. On the left, there is a sidebar with filters for 'All categories', 'Shipping', 'Special offers', 'Shop location', 'Item type', and 'Price (\$)'. The product listings include:

- Harry Potter Poster Hogwarts Expre... (USD 15.00)
- Hogwarts Canvas Art, Harry Potter ... (USD 20.00)
- Wizard Glasses with Lightning Bolt ... (USD 19.99)
- Harry Potter Music Box - En ... (USD 32.21)
- SUPER SOFT Novelty Word Socks, ... (USD 12.49)
- Harry Potter Quote Coasters. 22 To ... (USD 11.21)
- Harry Potter YETI Cup Harry Potter ... (USD 19.00)
- House Sorting Bath Bomb | Harry P... (USD 4.99)

A grid of various Harry Potter merchandise items, including:

- When In Doubt Go To The Library! Women's Harry Potter Shirt/ Harry Potter Shirt/ Vacation Shirt/ Hogwarts/ Women's Shirt/ Hermione Shirt (USD 5.99)
- Albus Dumbledore Harry Potter Qu... (USD 12.99)
- Harry Potter Party Theme Gender R... (USD 14.95)
- Harry Potter Machine Embroidery D... (USD 1.95)
- Harry Potter Inspired Custom Laser ... (USD 29.99)
- Book Lover Gift - After all this time ... (USD 24.00)
- Harry Potter Grim Tea Cup Set (Incl... (USD 14.99)
- When In Doubt Go To The Library! ... (USD 18.90)

电商搜索产品

购买欲望强

- 需要快速帮助用户定位所需的商品（因素：价格、品牌、品质、商家等）



电商搜索产品

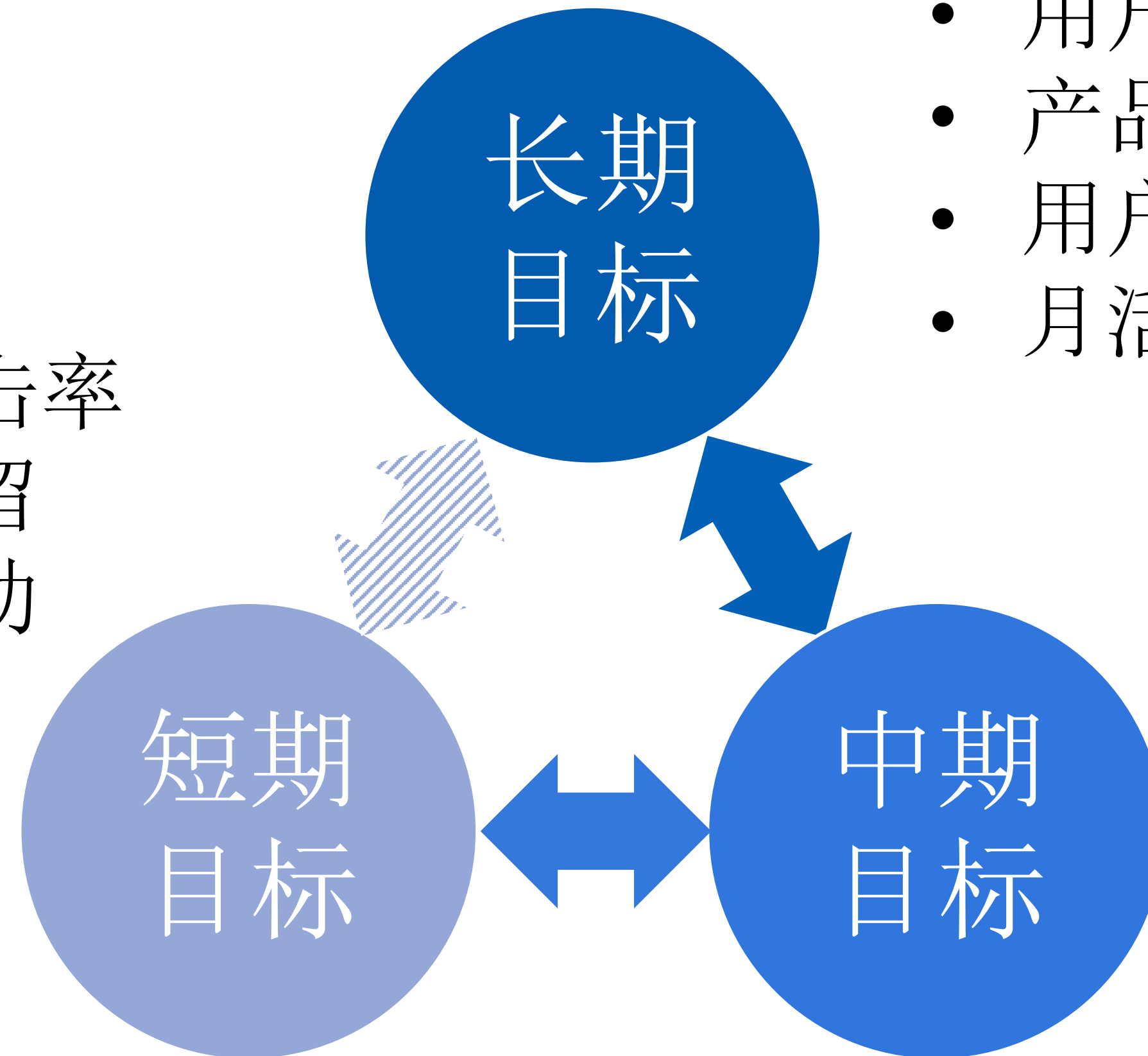
购买欲望弱

- 需要帮助用户发掘新的兴趣、新的话题，但同时不能让用户感觉无聊



从点击率到用户长期行为优化

- 用户模块、页面点击率
- 用户模块、页面驻留
- 用户模块、页面互动



- 用户留存率
- 产品年增长率
- 用户长期满意度
- 月活跃用户数

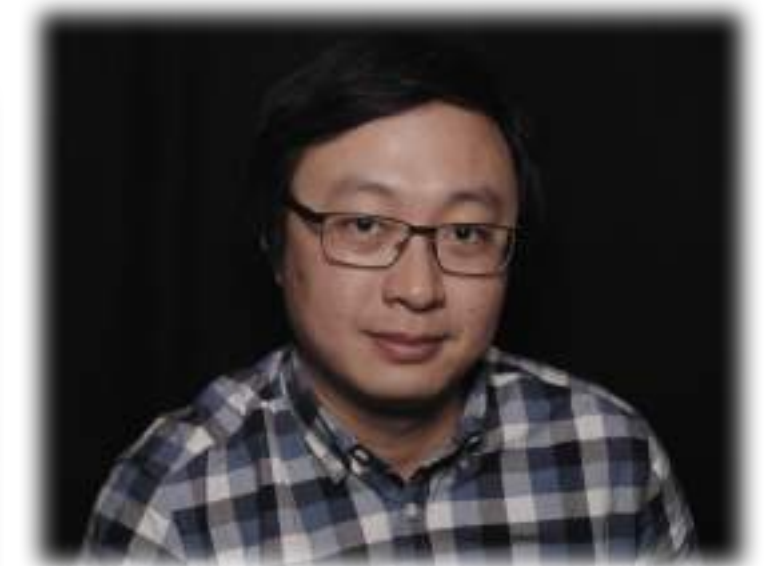
- 用户会话期销售额
- 用户会话期满意度
- 用户会话期驻留时间

TABLE OF CONTENTES

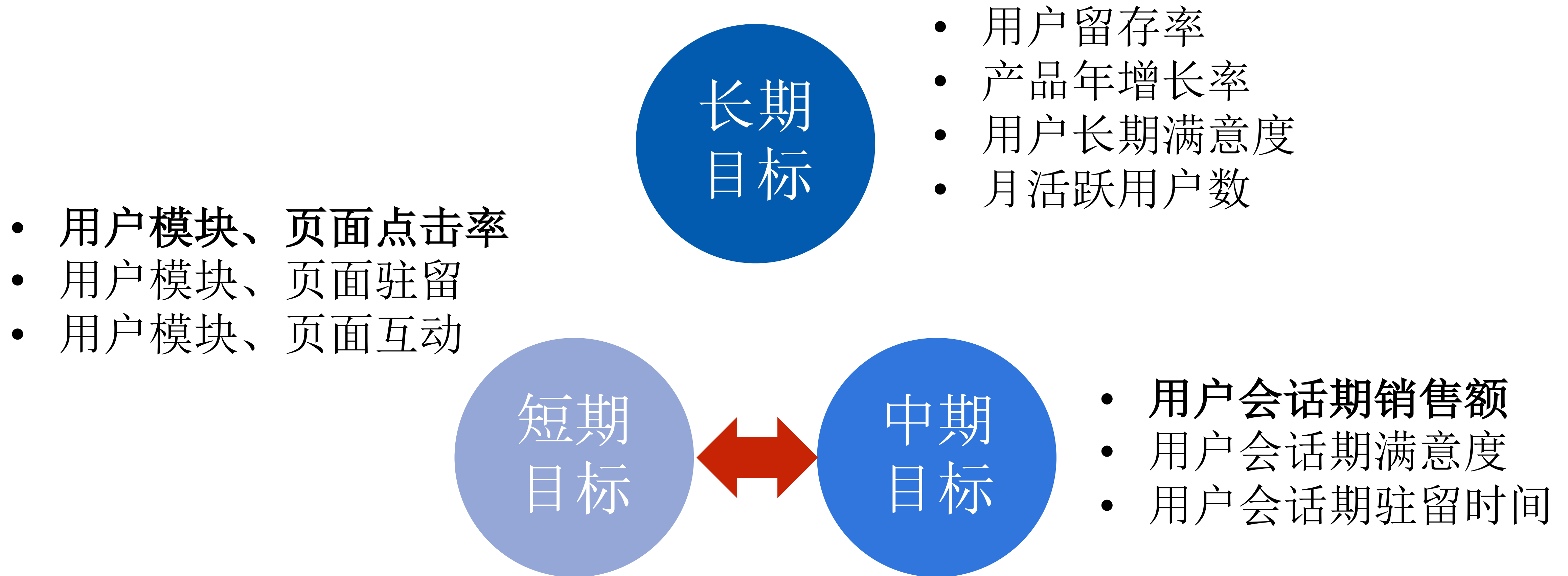
人工智能产品的系统性提升
从点击率到用户长期行为优化
人工智能在电商搜索的应用

人工智能在电商搜索的应用

利用数据驱动的人工智能技术来
不断系统性地提升搜索的品质



人工智能在电商搜索的应用



人工智能在电商搜索的应用

主要思想：

希望通过用户在搜索页面以及后续商品页面的行为来最大化用户会话期的销量（GMV）

人工智能在电商搜索的应用

用户会话期销售额（GMV）建模

人工智能在电商搜索的应用

用户会话期销售额 (GMV) 建模

$$GMV = \sum_{\underbrace{\forall s \in S}_{\text{A search session}}} \sum_{\underbrace{\forall i^s}_{\text{An item in } s}} \underbrace{Price(i^s)}_{\text{Price of } i^s} \underbrace{Pr(\Phi = 1 | i^s, q^s)}_{\text{Prob of purchase}}$$

人工智能在电商搜索的应用

用户会话期销售额 (GMV) 建模

$$GMV = \sum_{\underbrace{\forall s \in S}_{\text{A search session}}} \sum_{\underbrace{\forall i^s}_{\text{An item in } s}} \underbrace{Price(i^s)}_{\text{Price of } i^s} \underbrace{Pr(\Phi = 1|i^s, q^s)}_{\text{Prob of purchase}}$$

$$Pr(\Phi = 1|i, q) = \underbrace{Pr(\Psi = 1|i, q)}_{\text{click model}} \underbrace{Pr(\Phi = 1|i, q, \Psi = 1)}_{\text{purchase model}},$$

人工智能在电商搜索的应用

用户会话期销售额（GMV）建模

- 用排序学习中（Learning to Rank）的配对法（Pairwise Learning）来直接优化基于点击的列表指标（NDCG）
- 用二元分类（Binary Classification）来优化购买概率

人工智能在电商搜索的应用

用排序学习中（Learning to Rank）的配对法（Pairwise Learning）来直接优化基于点击的列表指标（NDCG）

$$\max_e N_{max}^{-1} \sum_{i=0}^m \frac{2^{l(i)}}{\log(1 + \tilde{r}(i))}$$

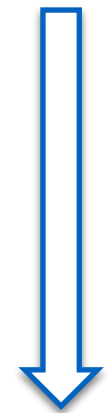


$$\max_e N_{max}^{-1} \sum_{i=0}^m \frac{2^{l(i)}}{\log(1 + \sum_{j=1, j \neq i}^m \sigma(\tilde{P}r(r(i) < r(j))))}$$

人工智能在电商搜索的应用

用二元分类 (Binary Classification) 来优化购买概率

$$\mathcal{L}_p = \sum_{i=1}^N \log\{1 + \exp[-l'_i(w_p x_i)]\} + \|w_p\|^2$$



$$\mathcal{L}_p = \sum_{i=1}^N Price(i) \log\{1 + \exp[-l'_i(w_p x_i)]\} + \|w_p\|^2$$

人工智能在电商搜索的应用

用户会话期销售额（GMV）建模

$$Pr(\Phi = 1|i, q) = \underbrace{Pr(\Psi = 1|i, q)}_{\text{click model}} \underbrace{Pr(\Phi = 1|i, q, \Psi = 1)}_{\text{purchase model}}$$

从搜索结果列表中点击 从商品页面决定是否购买

- 通过训练集和验证集来学习两个模型参数
- 通过两个模型的概率预测的乘积来进行排序预测

人工智能在电商搜索的应用

线下实验数据集

Sessions	Queries	Items	Avg. Items per Session
334,931	239,928	6,347,251	19.0
Keywords	Buyers	Sellers	Avg. Items per Query
631,778	270,239	550,025	26.5

人工智能在电商搜索的应用

模型特征

Relevance	Low Level	Sum of TF
		Sum of $\text{Log } TF$
		Sum of Normalized TF
		Sum of $\text{Log Normalized } TF$
		Sum of IDF
		Sum of $\text{Log } IDF$
		Sum of ICF
		Sum of $TF-IDF$
		Sum of $\text{Log } TF-IDF$
		$TF-\text{Log } IDF$
		$Length$
	$\text{Log } Length$	
	High Level	$BM25$
		$\text{Log } BM25$
LM_{DIR}		
LM_{JM}		
LM_{ABS}		
Revenue	$Price$	
	$Price - \text{Cat. Mean}$	
	$(Price - \text{Cat. Mean}) / \text{Cat. Mean}$	

人工智能在电商搜索的应用

比较模型和算法

Click	RankNet [2]	RNet
	RankBoost [11]	RBoost
	AdaRank [39]	ARank
	LambdaRank [3]	LRank
	ListNet [4]	LNNet
	MART [13]	MART
	LambdaMART [38]	LMART
Purchase	SVM [5]	SVM
	Logistic Regression [28]	LR
	Random Forest [22]	RM
Both	Weighted Purchase [44]	WT
	LMART+RM	LMRM
	LETORIF	LETORIF

人工智能在电商搜索的应用

算法效果

Category	Method	Click NDCG@5			Purchase NDCG@5			Revenue NDCG@5		
		Train	Vali	Test	Train	Vali	Test	Train	Vali	Test
Click	RNet	0.1743	0.1731	0.1378	0.1672	0.1721	0.1676	0.1692	0.1700	0.1356
	RBoost	0.2150	0.1768	0.1323	0.2150	0.1768	0.1715	0.2150	0.1768	0.1311
	ARank	0.1718	0.1711	0.1351	0.1718	0.1711	0.1706	0.1718	0.1711	0.1358
	LRank	0.1694	0.1688	0.1360	0.1678	0.1711	0.1672	0.1713	0.1719	0.1366
	LNNet	0.1665	0.1703	0.1355	0.1601	0.1682	0.1620	0.1646	0.1696	0.1348
	MART	0.2700	0.1758	0.1380	0.2155	0.1803	0.1796	0.2696	0.1688	0.1408
	LMART	0.3056	0.1777	0.1412	0.3056	0.1777	0.1717	0.3056	0.1777	0.1370
Purchase	SVM	0.1785	0.1772	0.1336	0.1831	0.1754	0.1755	0.1816	0.1752	0.1320
	LR	0.1978	0.1739	0.1310	0.1978	0.1739	0.1782	0.1978	0.1739	0.1332
	RM	0.3359	0.1698	0.1363	0.3329	0.2305	0.1798	0.3327	0.1685	0.1376
Both	WT	0.1970	0.1682	0.1334	0.1815	0.1763	0.1761	0.1781	0.1648	0.1375
	LMRM	0.2943	0.2597	0.1354	0.3087	0.2530	0.1688	0.2943	0.2594	0.1332
	LETORIF	0.1765	0.1550	0.1351	0.2731	0.1841	0.1801	0.2039	0.1698	0.1494

人工智能在电商搜索的应用

算法效果

Category	Method	Rev@1	Rev@2	Rev@3	Rev@4	Rev@5	Rev@6	Rev@7	Rev@8	Rev@9	Rev@10
Click	RNet	4.47	4.69	4.89	4.91	5.06	5.23	5.21	5.33	5.46	5.55
	RBoost	4.57	4.69	4.69	4.76	4.97	5.17	5.23	5.36	5.49	5.57
	ARank	4.37	4.66	4.76	4.90	5.06	5.20	5.33	5.47	5.59	5.67
	LRank	4.38	4.61	4.74	4.86	5.07	5.25	5.42	5.42	5.67	5.78
	LNet	4.30	4.59	4.78	4.99	5.16	5.35	5.49	5.61	5.63	5.63
	MART	4.62	4.72	4.86	5.04	5.26	5.47	5.47	5.64	5.74	5.86
	LMART	4.46	4.54	4.73	5.10	5.31	5.56	5.75	5.90	6.01	6.14
Purchase	SVM	4.41	4.54	4.76	4.77	4.95	5.16	5.34	5.50	5.64	5.77
	LR	4.29	4.65	4.65	4.69	4.74	4.81	4.94	4.97	5.11	5.11
	RM	4.52	4.82	4.86	5.02	5.18	5.33	5.50	5.66	5.79	5.92
Both	WT	4.52	4.69	4.80	4.85	5.01	5.07	5.23	5.32	5.35	5.41
	LMRM	4.42	4.50	4.72	5.08	5.23	5.41	5.57	5.60	5.73	5.85
	LETORIF	4.58	4.90	5.08	5.47	5.64	5.85	6.02	6.19	6.40	6.54

人工智能在电商搜索的应用

小结

- 一个类似的模型在线上提高了GMV达10%
- 一个电商搜索模型的一小步探索
- 论文正在修改中

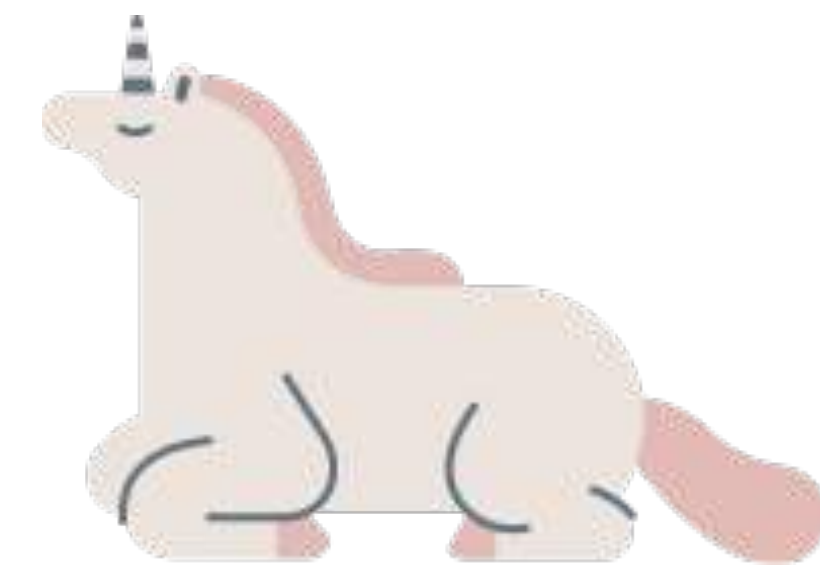


TABLE OF CONTENTES

人工智能产品的系统性提升
从点击率到用户长期行为优化
人工智能在电商搜索的应用

Thanks!