第二十七届全国信息检索学术会议
CCIR 2021
中国·大连

基于因果分析的搜索纠偏方法

徐 君　　张 骁

中国人民大学高瓴人工智能学院

提 纲

- 搜索中的偏差
- 因果推断简介
- 基于因果的搜索纠偏方法
- 实验设置方法
- 总结与未来发展方向

搜索与推荐在模型训练数据上的差异

- 搜索：专家标注数据

五级标签
Perfect、Excellent、Good、Fair、Bad

- 推荐：用户反馈

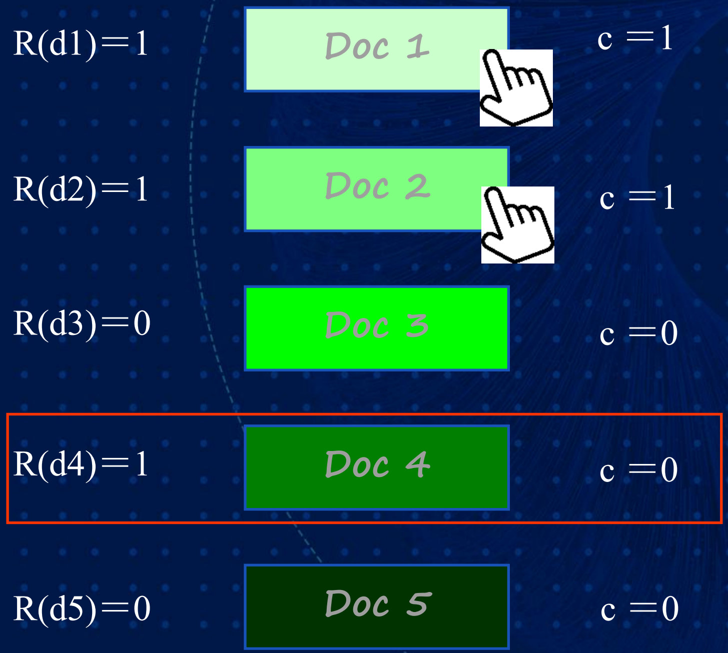
点击、收藏、喜欢、点赞、购买……

为何搜索引擎不直接采用用户反馈数据训练模型？

专家标注与用户反馈的核心差异：偏差

	优点	缺点
 专家标注	✓ 几乎无偏	× 成本昂贵 × 数量稀少 × 非个性化
 用户反馈	✓ 成本低 ✓ 数量多 ✓ 反映用户真实喜好	× 含有多种偏差 (位置偏差、样本选择偏差、信任偏差等等)

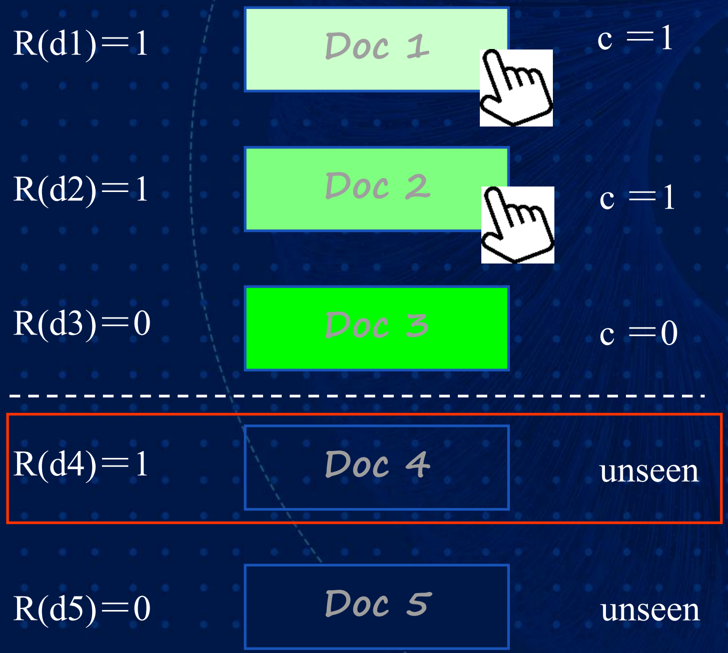
位置偏差



(a)

- 位置偏差是指用户很少会去浏览排在后面的文档，从而导致排名靠后的相关文档上的点击较少

样本选择偏差

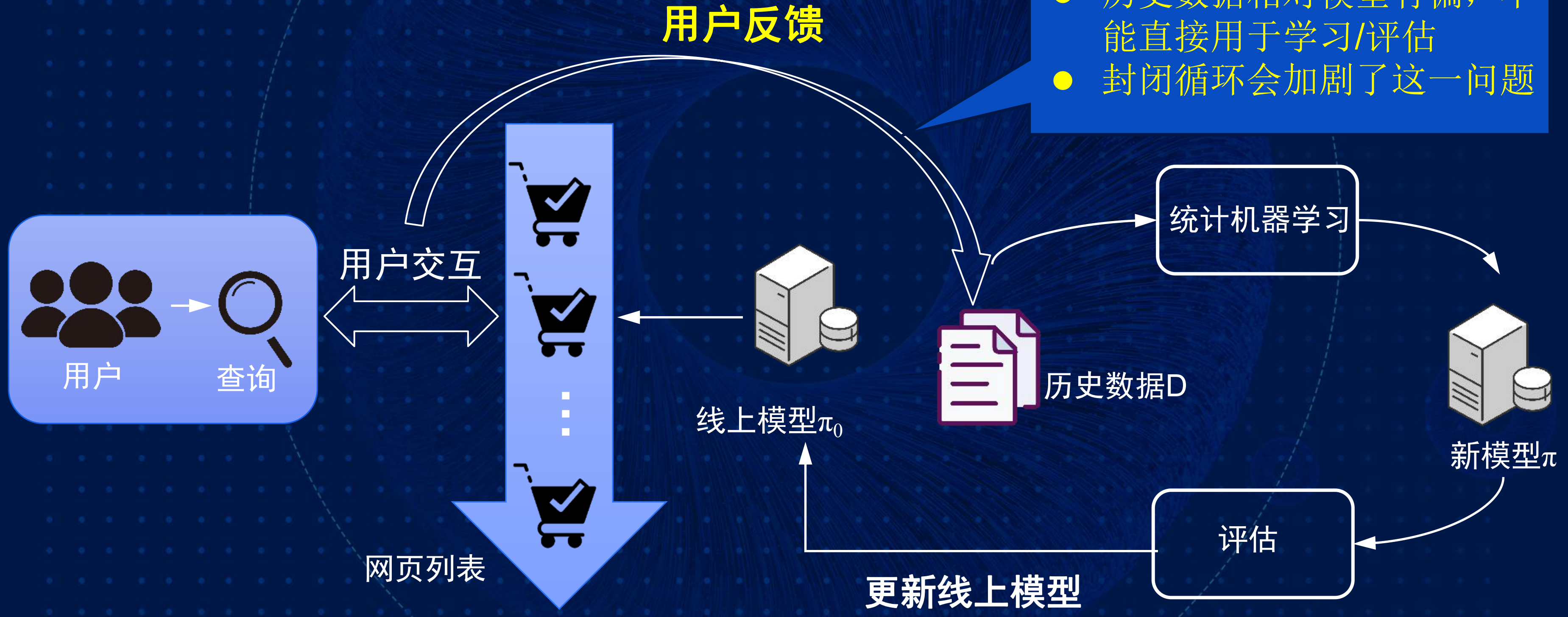


(a)

- 样本选择偏差是指由于用户设备大小的限制导致搜索引擎只能展示排序靠前的部分文档，因此被截断的相关文档没有任何点击。

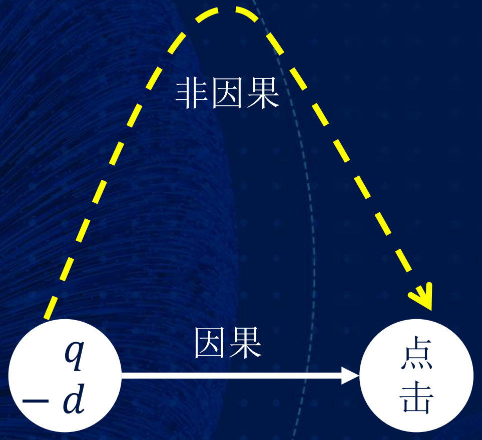
直接使用带偏差的反馈数据训练搜索模型

- 历史数据相对模型有偏，不能直接用于学习/评估
- 封闭循环会加剧了这一问题



如何使用有偏的反馈训练无偏的排序？

- 核心问题：消除用户反馈中的偏差
- 因果推断成为重要的纠偏工具
  - 观测数据：(q-d对，用户点击)
  - 纠偏：观测到的用户的点击中
    - ✓ 有多少因素是由q-d之间的相关性导致的？（因果关系，无偏）
    - ✓ 有多少因素是由其他混淆变量导致的？（非因果关系，偏差）



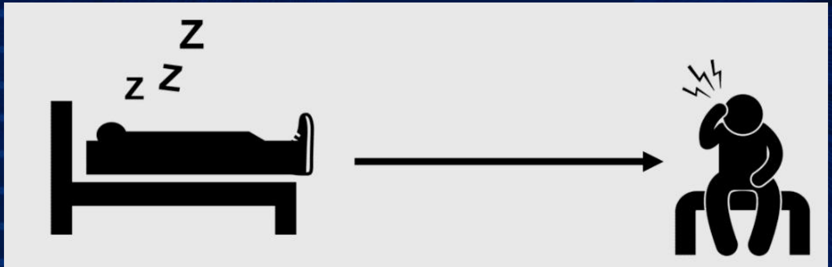
提 纲

- 搜索中的偏差
- 因果推断简介
- 基于因果的搜索纠偏方法
- 实验设置方法
- 总结与未来发展方向

一个因果推断的例子

- 因果：干预变量T（*treatment*）的变化如何导致另一个结果变量Y（*outcome*）改变（在保持其它变量不变的情况下）
- 例如：穿鞋睡觉是否会导致头痛？

$T=1$：穿鞋睡觉
$T=0$：脱鞋睡觉



$Y=1$：头痛
$Y=0$：不头痛

来源：“Brady Neal， Introduction to Causal Inferenc
https://www.bradyneal.com/causal-inference-cours

最直接的方法

- 对某一个实验对象i
  - 先让i穿鞋睡觉（$T=1$），观测他是否头痛$Y_i(1)=Y_i|_{do(T=1)}$
  - 再让i脱鞋睡觉（$T=0$），观测他是否头痛$Y_i(0)=Y_i|_{do(T=0)}$
  - 对实验对象i而言，变量T对Y的因果效应为$Y_i(1)-Y_i(0)$

    $Y_i(1)-Y_i(0)=0$：无因果效应

    $Y_i(1)-Y_i(0)\neq 0$：有因果效应
- 对N个实验对象
  - 逐一计算每一个实验对象的因果效应
  - 求平均

实验对象i	$Y_i(0)$	$Y_i(1)$	因果效应
1	0	1	$1-0=1$
2	1	1	$1-1=0$
…	…	…	…
N	1	0	$0-1=-1$
			Mean

应用：信息检索中的离线评价

- 实验对象i：第i个查询、对应的文档集合以及相关性标签
  - $T=0$：使用基线模型对文档进行排序，计算$NDCG$值，记为$Y_i(0)$
  - $T=1$：使用新模型对文档进行排序，计算$NDCG$值，记为$Y_i(1)$
  - 第i个查询：不同的模型对$NDCG$的因果效应为

    $Y_i(1)-Y_i(0)=0$： T无因果效应，新模型未改变查询i的效果

    $Y_i(1)-Y_i(0)\neq 0$： T有因果效应，新模型改变了查询i的效果
- 测试集中有N个查询
  - 逐一计算每一个查询的因果效应（$NDCG$值之差）
  - 求平均

上述方式的优缺点

- 优点：了解在每一个测试对象（查询）上的个体因果效应
- 缺点：要求每一个对象能够反复试验（同时测试$T = 0$和$T = 1$）
  - 对查询i，能同时计算基线模型的$NDCG$和新模型的$NDCG$，并保证两次计算互不干扰
  - 不适合在线评价：在线用户只能对$T = 0$或$T = 1$的结果进行反馈（点击）
    - ✓ 让用户再看一次查询i的结果，上一次的结果会影响其点击行为
    - ✓ 实验对象i不想多次被安排穿鞋/脱鞋睡觉，同时前一次脱鞋/穿鞋睡觉的行为可能影响他后一次是否头痛

因果分析中的基本假设

对象 i	处理变量T	$Y_i(0)$	$Y_i(1)$	因果效应
1	1	?	1	?
2	0	1	?	?
…	…	…	…	?
N	1	?	0	?

- 不能观测到单个样本的所有处理结果
  - 对实验对象i，只能观测到$Y_i(0)$或$Y_i(1)$
  - 观测到实验对象i穿鞋睡觉醒来后头痛（$Y_i(1) = 1$，事实），则观测不到他脱鞋睡觉醒来后是否头痛（$Y_i(0) = ?$，反事实）
  - 用户输入查询i，展示了基线模型的结果用户未点击（$Y_i(0) = 0$，事实），则不知道如果展示新模型的结果用户是否会点击（ $Y_i(1) = ?$，反事实）

困难：$Y(1)$和$Y(0)$不可以比较

实验对象i	$Y_i(0)$	$Y_i(1)$	因果效应
1	0	1	$1-0=1$
2	1	1	$1-1=0$
…	…	…	…
N	1	0	$0-1=-1$

同时观测到$Y(1)$和$Y(0)$，可以逐个比较

实验对象i	处理变量T	$Y_i(0)$	$Y_i(1)$	因果效应
1	1	?	1	?
2	0	1	?	?
…	…	…	…	?
N	1	?	0	?

$Y(1)$和$Y(0)$只能观测到一个，不可以逐个比较

解决方案1：随机

- 给定N个实验对象，对每一个实验对象i
  - 扔硬币随机决定$T=0$或者$T=1$
  - 如果$T=0$，则让i脱鞋睡觉，观察他醒来后是否头痛$Y_i(0)$
  - 如果$T=1$，则让i穿鞋睡觉，观察他醒来后是否头痛$Y_i(1)$

实验对象i	处理变量T	Y	$Y_i(0)$	$Y_i(1)$	因果效应
1	1	1	?	1	?
2	0	1	1	?	?
3	0	0	?	0	?
…	…		…	…	?
N	1	0	?	0	?

解决方案1：随机（续）

- 依据T的取值，将实验对象分成两组，计算平均处理效应(average treatment effect, ATE)

$$\mathbb{E}\left(Y_i(1) - Y_i(0)\right)$$

实验对象i	处理变量T	Y	$Y_i(0)$	$Y_i(1)$	因果效应
1	1	1	?	1	?
2	0	1	1	?	?
3	0	0	?	0	?
…	…		…	…	?
N	1	0	?	0	?

$T = 1$

$T = 0$

为什么随机可以得到因果？

实验对象i	$Y_i(0)$	$Y_i(1)$	因果效应
1	0	1	$1-0=1$
2	1	1	$1-1=0$
…	…	…	…
N	1	0	$0-1=-1$

$Y_i(1)$和$Y_i(0)$同时存在，可以逐个比较

实验对象i	处理变量T	Y	$Y_i(0)$	$Y_i(1)$	因果效应
1	1	1	?	1	?
2	0	1	1	?	?
3	0	0	?	0	?
…	…		…	…	?
N	1	0	?	0	?

$Y_i(1)$和$Y_i(0)$只存在一个，不可以逐个比较

但可以比较$Y_i(1)$和$Y_i(0)$的期望（平均）

应用：信息检索中的在线评价

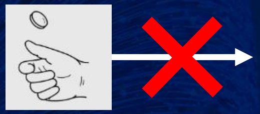
- 在线A/B测试
  - 试验用户流量：$i = 1, \cdots, N$
  - $T = 0$：基线模型；$T = 1$：新模型
  - $Y = 1$：用户点击；$Y = 0$：用户不点击
- 用户流量随机分割成 A/B 两组
  - 试验组：$T = 1$，使用新模型进行排序，观测用户是否点击$Y_i(1)$
  - 对照组：$T = 0$，使用基线模型进行排序，观测用户是否点击$Y_i(0)$
  - 比较试验组和对照组的点击率差（平均因果效应），决定是否发布

随机的优缺点

- 优点
  - 每一个对象只需要测试一次，符合在线测试的要求
- 不足
  - 实验研究：先设计实验，再搜集实验数据
  - 无法得知在单个对象上的细粒度因果效应
  - 需要对测试对象施加随机处理
    - ✓ 实际操作中代价较高（A/B测试往往只在极少部分流量上进行）
    - ✓ 在某些场景下存在伦理问题（如医疗中，无法随机给病人试药）

是否可以从“非实验研究”数据中分析因果？

实验对象i	处理变量T	Y	$Y_i(0)$	$Y_i(1)$
1	1	1	?	1
2	0	1	1	?
3	0	0	?	0
…	…		…	…
N	1	0	?	0

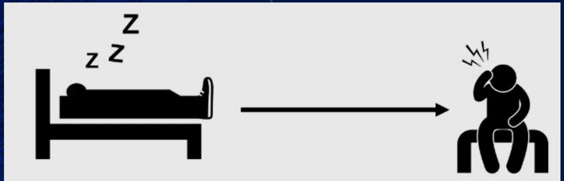


T

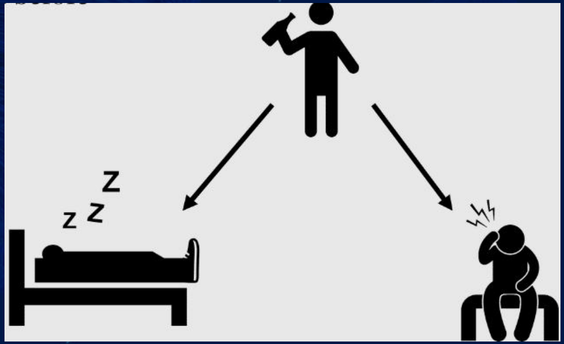
- 与实验数据的区别：干预变量T不能被控制，只能被观测
- 简单的方法：仍然按照“实验研究”的计算方法，基于表格中统计平均因果效应：$\mathbb{E}\big(Y_i(1)-Y_i(0)\big)$

简单方法存在问题

- 可能存在一个干扰变量C
  - $C=1$：酗酒睡觉；$C=0$：清醒睡觉
  - 喝酒$C=1$可同时导致两个后果
    - ✓ 穿鞋睡觉
    - ✓ 醒来后头痛
  - 穿鞋睡觉与头痛呈现相关，但并不存在因果关系
- 非实验数据可受到未知因素C的干扰
  - 可能存在无穷多的未知干扰变量



以为的因果关系



真正的因果关系

干扰变量C的影响：不一致的数据分布

$T=1$: 穿鞋睡觉

酗酒 酗酒 酗酒

酗酒 酗酒 酗酒 清醒

酗酒 酗酒 酗酒

$E[Y|T=1]$

$T=0$: 脱鞋睡觉

清醒 清醒 清醒

酗酒 清醒 清醒 清醒

清醒 清醒 清醒

$E[Y|T=0]$

- 数据分布不一致，导致两个分组所得的统计量不可直接比较
- $E[Y|T=1]-E[Y|T=0]$不能体现因果效应

与随机实验产生的数据进行比较

$T=1$: 穿鞋睡觉

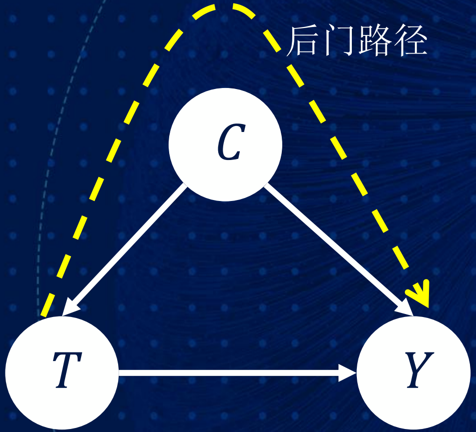
$T=0$: 脱鞋睡觉

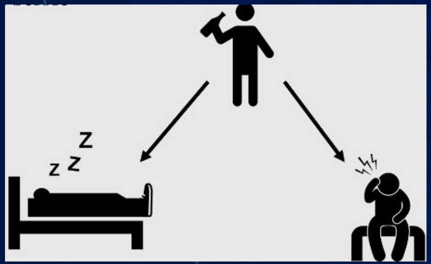
$E[Y|T=1]$

$E[Y|T=0]$

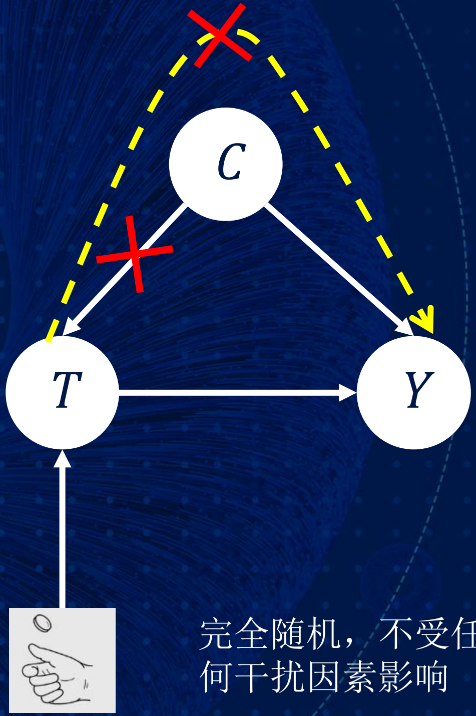
- 随机实验消除了C的影响，两个组所得到的统计量可比
- $E[Y|T=1]-E[Y|T=0]$表示平均因果效应

基于因果图的解释





“非实验研究”日志数据



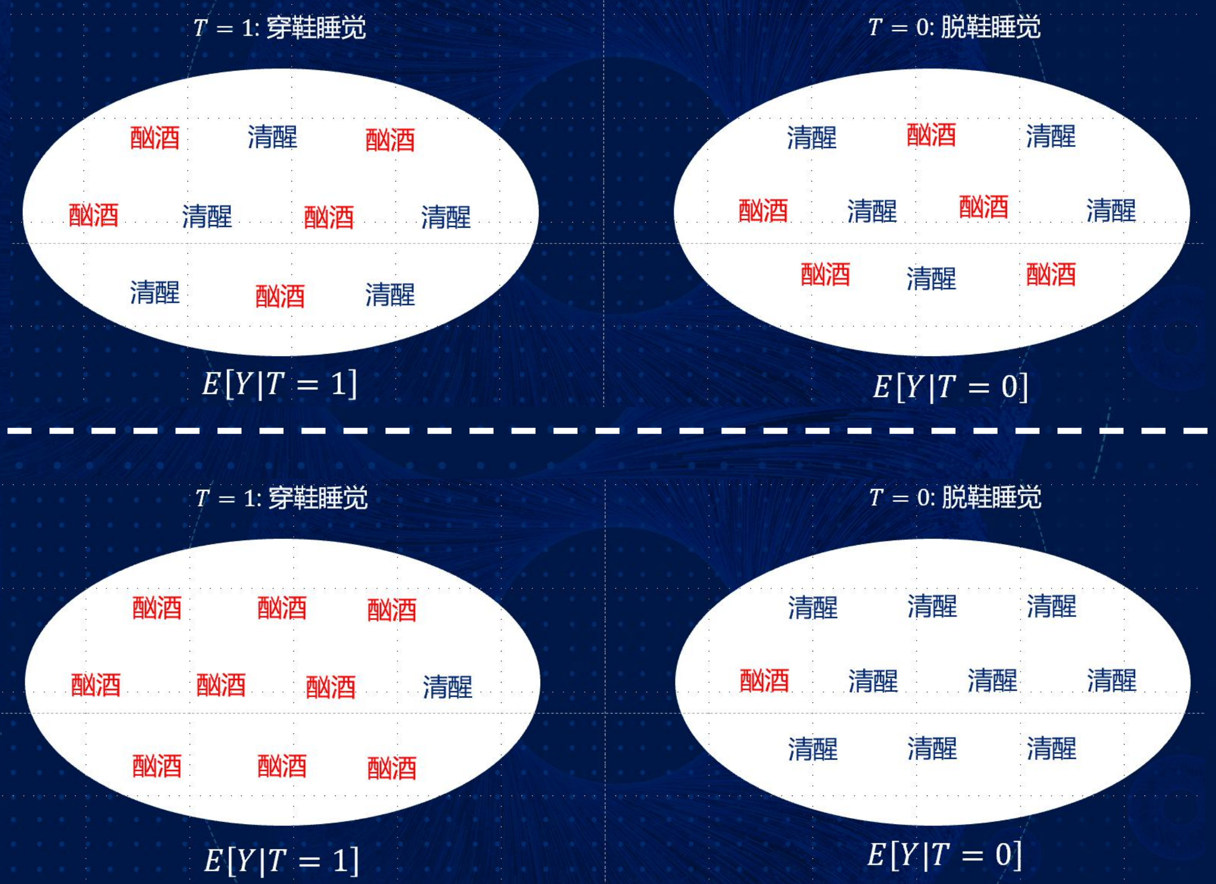
完全随机，不受任何干扰因素影响

随机实验

问题的关键：在一致的数据分布上进行比较

一致的数据分布，可以比较

不一致的数据分布，不能比较



从另外一个角度去分组数据：按C分组

$C = 1$: 酗酒

穿鞋 穿鞋 穿鞋

穿鞋 穿鞋 穿鞋 脱鞋

穿鞋 穿鞋 穿鞋

$C = 0$: 清醒

脱鞋 脱鞋 脱鞋

穿鞋 脱鞋 脱鞋 脱鞋

脱鞋 脱鞋 脱鞋

- 假设不存在C以外的影响数据分布的因素
  - 每一组内部$T = 1$（穿鞋）和$T = 0$（脱鞋）的平均因果效应可以比较

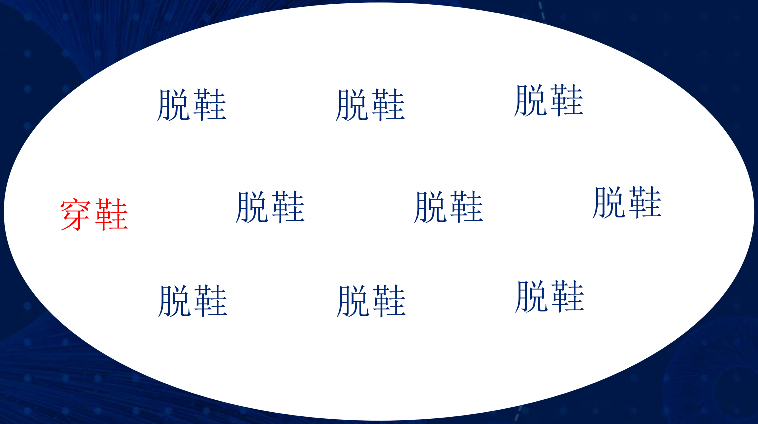
分别评估每一组的平均因果效应

$C = 1$: 酗酒

$$E[Y|T = 1, C = 1] - E[Y|T = 0, C = 1]$$

$C = 1$的条件下T的因果效应

$C = 0$: 清醒



$$E[Y|T = 1, C = 0] - E[Y|T = 0, C = 0]$$

$C = 0$的条件下T的因果效应

总体的因果效应

$C = 1$: 酗酒

穿鞋 穿鞋 穿鞋
穿鞋 穿鞋 穿鞋 脱鞋
穿鞋 穿鞋 穿鞋

$$E[Y|T=1,\ C=1] - E[Y|T=0,\ C=1]$$

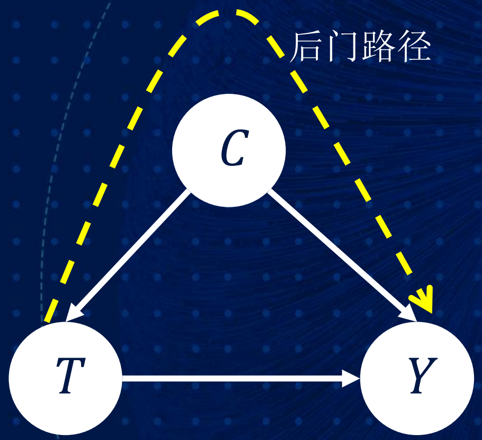
$C = 0$: 清醒

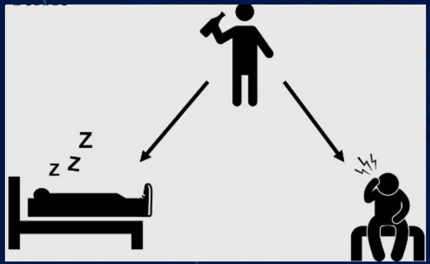
脱鞋 脱鞋 脱鞋
穿鞋 脱鞋 脱鞋 脱鞋
脱鞋 脱鞋 脱鞋

$$E[Y|T=1,\ C=0] - E[Y|T=0,\ C=0]$$

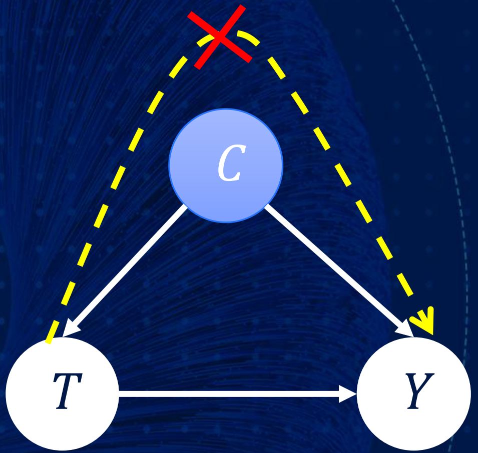
总体因果效应：$E_C\left[E[Y|T=1,\ C] - E[Y|T=0,\ C]\right]$

基于因果图的解释





“非实验研究”日志数据



C成为被观测到的变量，

以C为条件阻断了后门路径

另一个因果的例子：辛普森悖论

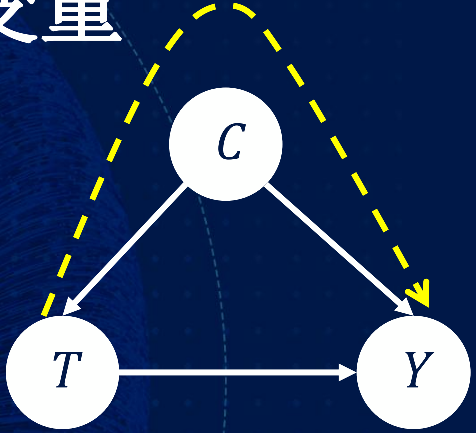
死亡率（Y）	轻症($C=0$)	重症($C=1$)	总体
治疗方案A ($T=0$)	15% (210/1400)	30% (30/100)	16% (240/1500)
治疗方案B ($T=1$)	10% (5/50)	20% (100/500)	19% (105/550)

- A/B两种治疗方案哪个更好？
  - 治疗方案A好：总体统计数据看，A ($T=0$)的死亡率较低（16% < 19%）
  - 治疗方案B好：
    - ✓ 对于轻症病人($C=0$)，治疗方案B ($T=1$)的死亡率比较低 (10% < 15%)
    - ✓ 对于重症病人($C=1$)，治疗方案B ($T=1$)的死亡率也比较低 (20% < 30%)

来源：“Brady Neal， Introduction to Causal Inferenc
https://www.bradyneal.com/causal-inference-cours

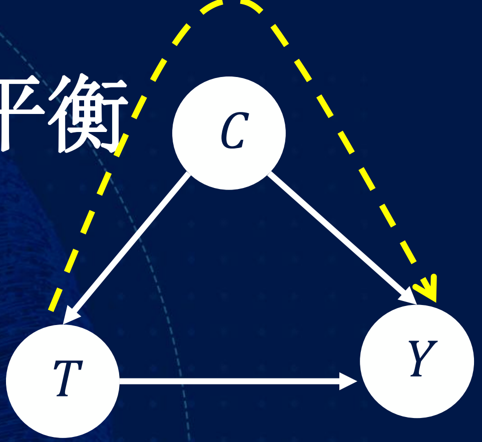
对辛普森悖论的因果解释：混淆变量

死亡率（Y）	轻症($C=0$)	重症($C=1$)	总体
治疗方案A ($T=0$)	15% (210/1400)	30% (30/100)	16% (240/1500)
治疗方案B ($T=1$)	10% (5/50)	20% (100/500)	19% (105/550)



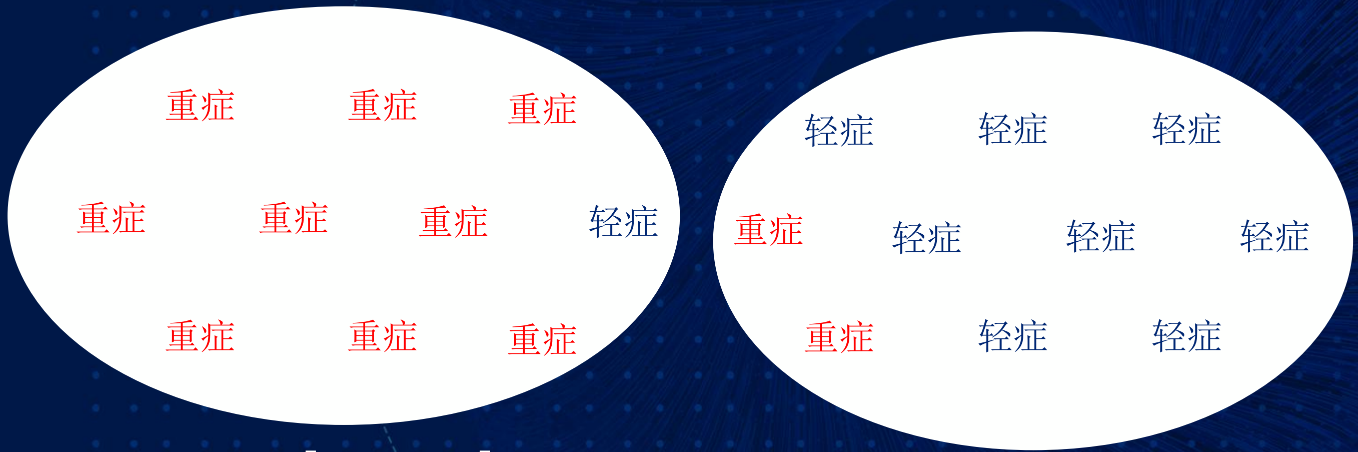
- 症状的轻重C会同时影响治疗方案的选择和死亡率
  - 如：A治疗方案简单便宜，轻症的病人倾向选择A，重症病人倾向B
  - 病症的轻重影响了数据的分布，进而导致两组间的比较不可信

混淆变量导致观测到的数据分布不平衡



$T=1$: 选择$\mathcal{B}$方案

$T=0$: 选择A方案



$E[Y|T=1]$

$E[Y|T=0]$

- 虽然$\mathcal{B}$方案治疗效果更好，但是治疗的都是重症病人，总体死亡率自然就会高，比较不公平！

公平的比较（方案一：多次治疗）

- 抛弃现有数据，重新进行实验。对同一个病人
  - 先用A方案治疗一次，观测治疗效果
  - 再用B方案治疗一次，观测治疗效果
- 现实中不可行！
  - 第一次治疗好了，没有必要进行第二次治疗
  - 第一次治疗失败了，无法进行第二次治疗

公平的比较（方案二：随机）

$T=1$: 选择B方案

轻症 重症 轻症

重症 轻症 重症 轻症

重症 轻症 重症

$E[Y|T=1]$

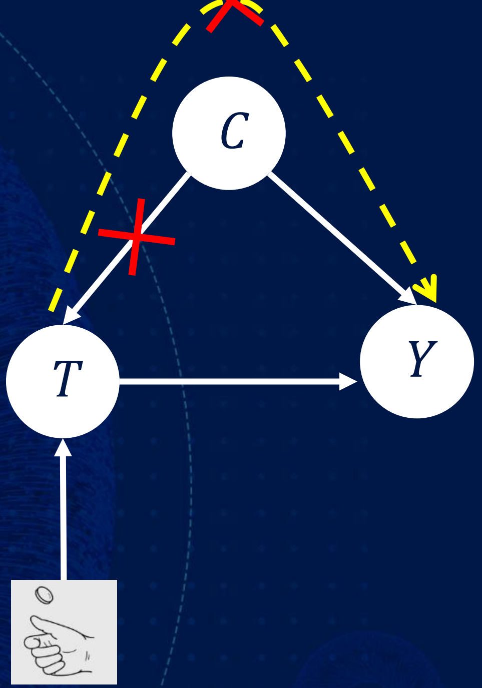
$T=0$: 选择A方案

重症 轻症 重症

重症 轻症 重症 轻症

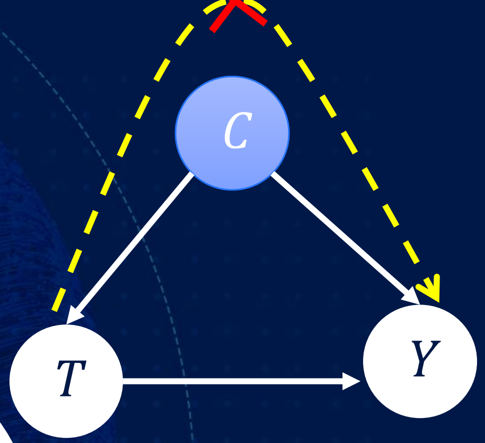
轻症 重症 轻症

$E[Y|T=0]$



- 重新进行实验，不管病症轻重随机选择治疗方案
  - 存在伦理问题

公平的比较（方案三：分组）



$C=1$: 重症

方案B 方案B 方案B

方案B 方案B 方案B 方案A

方案B 方案B 方案A

$C=0$: 轻症

方案A 方案A 方案A

方案B 方案A 方案A 方案A

方案A 方案A 方案A

$$E[Y|T=1,\ C=1]$$
$$E[Y|T=0,\ C=1]$$

$$E[Y|T=1,\ C=0]$$
$$E[Y|T=0,\ C=0]$$

- 按照重症和轻症进行分组，计算每一个组内不同方案的效果

公平的比较（方案二：分组）

死亡率（Y）	轻症$(C=0)$	重症$(C=1)$	总体	因果
治疗方案A $(T=0)$	15% (210/1400)	30% (30/100)	~~16% (240/1500)~~	$\frac{1450}{2050}\times 15\% + \frac{600}{2050}\times 30\% \approx 0.194$
治疗方案B $(T=1)$	10% (5/50)	20% (100/500)	~~19% (105/550)~~	$\frac{1450}{2050}\times 10\% + \frac{600}{2050}\times 20\% \approx 0.129$

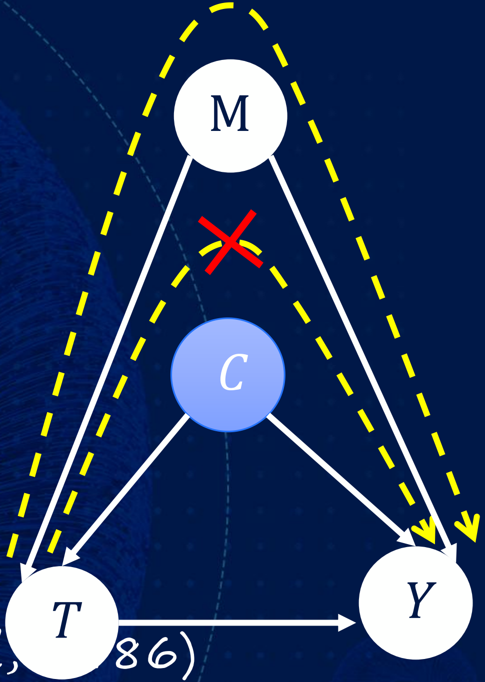
- 治疗效果
  - 总人数1400+100+50+500=2050，其中1450人轻症，600人重症

$$E(Y|do(T=t)) = P(C=0)E(Y|T=t,C=0) + P(C=1)E(Y|T=t,C=1)$$

- 方案比较
  - $E(Y|do(T=1)) - E(Y|do(T=0)) = 0.129 - 0.194 < 0$
  - 方案B治疗效果更好！

上述方法的优缺点

- 优点：从“非实验数据”中发现因果关系
  - 廉价，可以大量搜集：例如直接搜集搜索用户点击
  - 在某些场景下规避了干预实验带来的伦理问题
- 缺点：较强的假设
  - “No causation without manipulation” (Holland, 86)
  - 有可能存在未知的干扰变量M，使得后门路径依然存在！
  - 两个组需要有重叠（即：有部分轻症病人也选择了B方案，同时部分重症病人也选择了A方案）
  - 需要对数据生成的机制有全面的理解



因果分析的能力边界

表 1.1　一个简单的模型分类法。最详细的模型（顶部）是一个机械的或物理的模型，通常涉及一组微分方程。在表格的另一端（底部），有一个纯粹的统计模型。这个模型可以从数据中学习，但是除了建模之间的联系，它提供的洞察力往往很小。因果模型可以被看作介于两者之间的描述，从物理现实主义中抽象出来，同时保留回答某些干预或反事实问题的能力。参见 Mooij 等人（2013）讨论物理模型与结构因果模型之间的联系。6.3 节将讨论干预方式

模型	独立同分布情况下进行预测	在不断变化的分布或干预情况下进行预测	回答反事实问题	获得物理洞察力	从数据中学习
机械 / 物理，例如 2.3 节	是	是	是	是	?
结构因果模型，例如 6.2 节	是	是	是	?	?
因果图模型，例如 6.5.2 节	是	是	否	?	?
统计模型，例如 1.2 节	是	否	否	否	是

Jonas Peters, Dominik Janzing, and Bernhard Sch¨olkopf. Elements of Causal Inference Foundations and Learning Algorithms.

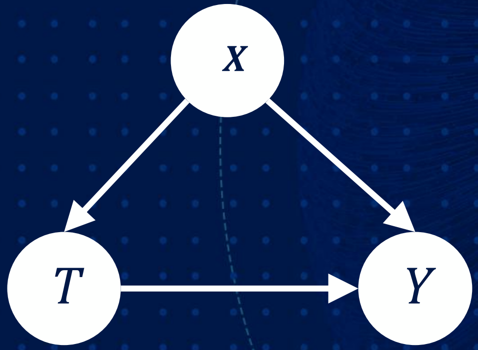
本部分小结

- 搜索中存在各种偏差，因果推断成为搜索纠偏的一种重要手段
- 基本的因果推断方法
  - 如果单个样本可以干预多次：获得对象级别的细粒度因果效应
  - 不能观测到单个样本的所有干预结果：随机设置干预值，获得平均因果效应
  - 不能进行干预实验：加入假设阻断后门路径，获得平均因果效应

第二部分

- 搜索中的偏差
- 因果推断简介
- 基于因果的搜索纠偏方法
- 实验设置方法
- 总结与未来发展方向

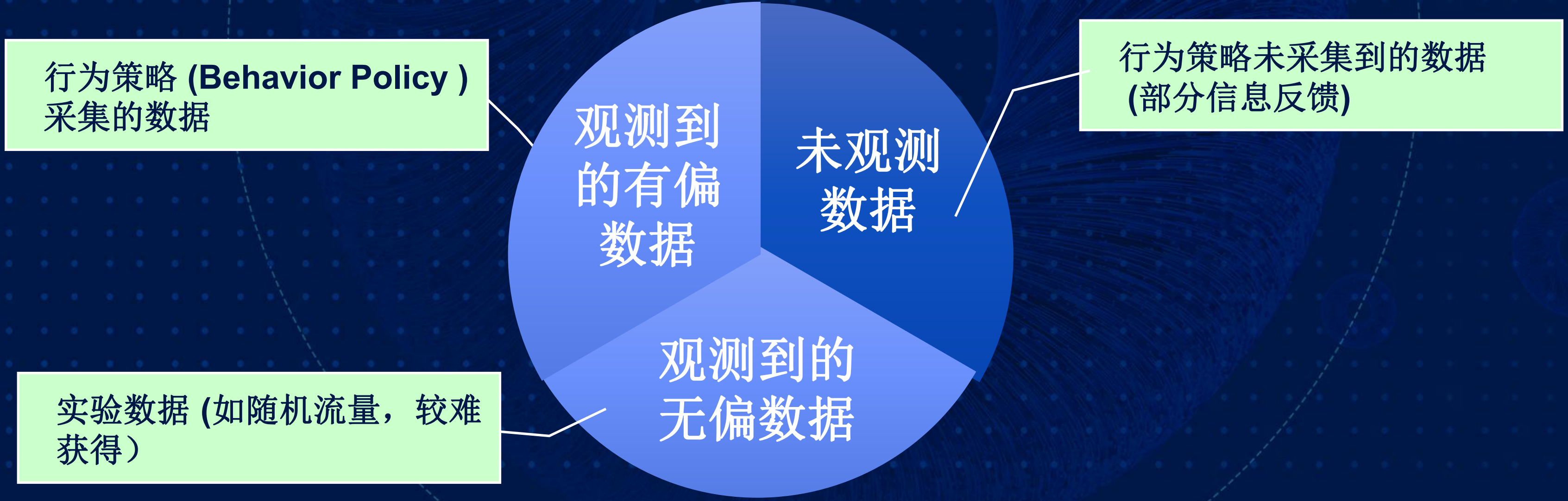
搜索任务中的因果图



- 可能存在一个混淆变量X
  - 干预变量 T：展示位置、是否曝光、流行度等
  - 结果变量 Y：用户反馈（是否点击）
  - 混淆变量 X：查询、文档的特征
- 结果变量 Y是一种隐式反馈，真实的relevance隐含在 Y中

搜索任务中的数据

- 用于训练搜索模型的数据：（查询，文档；用户反馈）



搜索纠偏的因果方法总览

- 利用因果推断方法消除偏差的影响，实现搜索纠偏

面向搜索的因果推断方法

逆倾向得分	插补方法
双稳健方法	其他方法

Zhenhua Dong, Jun Xu, Ji-Rong Wen, etc. “*Counterfactual learning for recommender system*”, RecSys 2020.

因果推断→搜索纠偏

存在混淆变量 C 影响数据分布

$C=1$: 酗酒

穿鞋 穿鞋 穿鞋

穿鞋 穿鞋 穿鞋 脱鞋

穿鞋 穿鞋 穿鞋

$C=0$: 清醒

脱鞋 脱鞋 脱鞋

穿鞋 脱鞋 脱鞋 脱鞋

脱鞋 脱鞋 脱鞋

因果推断→搜索纠偏

存在混淆变量 C 影响数据分布

$C=1$: 酗酒

穿鞋 穿鞋 穿鞋

穿鞋 穿鞋 穿鞋 脱鞋

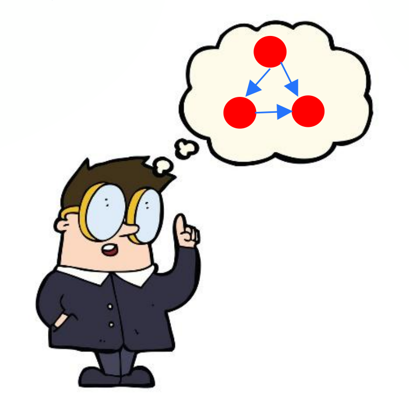
穿鞋 穿鞋 穿鞋

$C=0$: 清醒

脱鞋 脱鞋 脱鞋

穿鞋 脱鞋 脱鞋 脱鞋

脱鞋 脱鞋 脱鞋



本质：平衡数据分布

逆倾向得分 (Inverse Propensity Score, IPS)

- 基本原理：通过赋权平衡数据分布
- 数学表示：

观测数据的个数

$$\theta_{\mathrm{IPS}} = \arg\min_{\theta \in \Theta} \sum_{i=1}^{\boxed{N}} \frac{\ell\left(y_i, \hat{y}_i^{\theta}\right)}{\boxed{p_i}} + \lambda \mathcal{R}(\theta)$$

倾向得分 (Propensity Score): 该条数据接收处理（T=1）的概率

- 适用于观测到的有偏历史数据
- 该条数据接收处理的概率越小 → 倾向得分越小 → 逆倾向得分1/p越大

Thorsten Joachims, Adith Swaminathan et.al. Unbiased learning-to-rank with biased feedback. In: Proceedings of the 10h ACM International Conference on Web Search and Data Mining (WSDM 2017), page 781–789.

逆倾向得分 (Inverse Propensity Score, IPS)

- 倾向得分的表达：以排序中的位置偏差为例

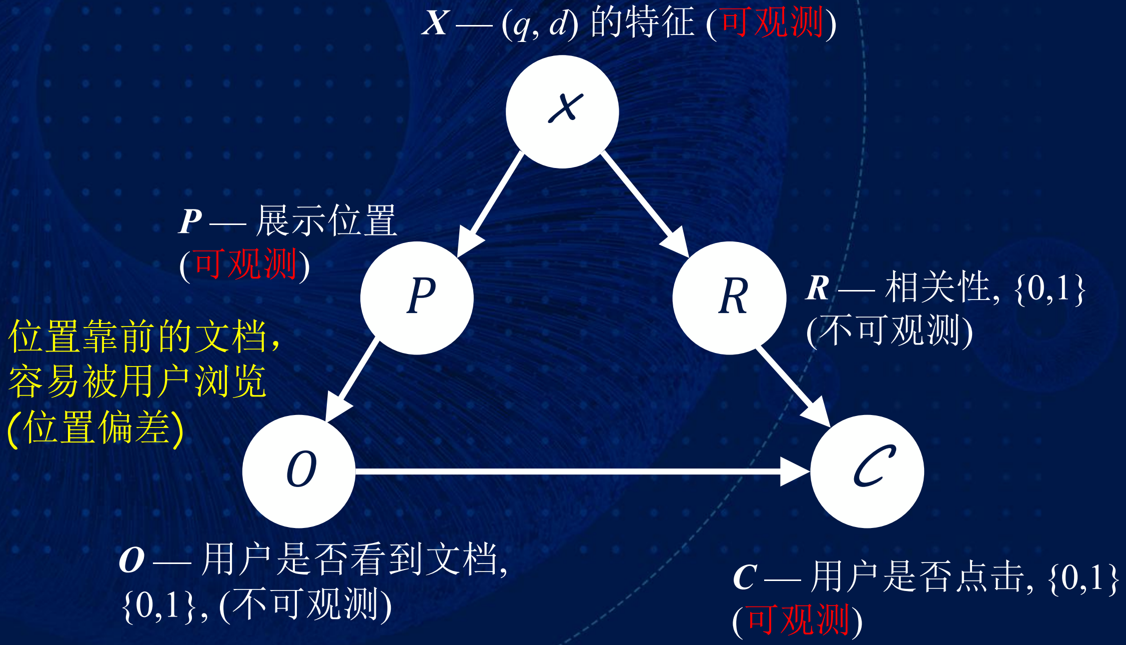
排序任务的目标：

给定 $\boldsymbol{q}$ — 查询，$\boldsymbol{d}$ —候选文档集合

找到合适的排序函数 f_θ：

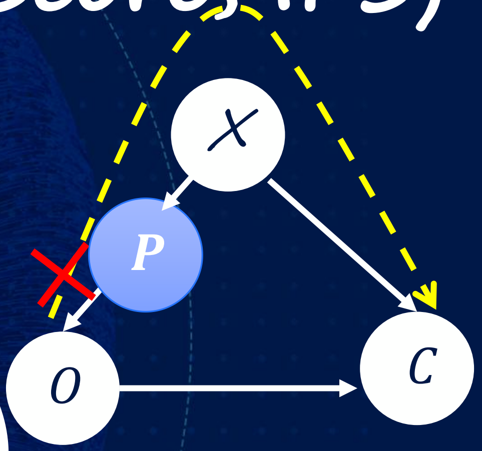
$$f_{\boldsymbol{\theta}}(d_i|\boldsymbol{q},\boldsymbol{d}) : d_i \rightarrow p$$

获得好的排序度量指标。



逆倾向得分 (Inverse Propensity Score, IPS)

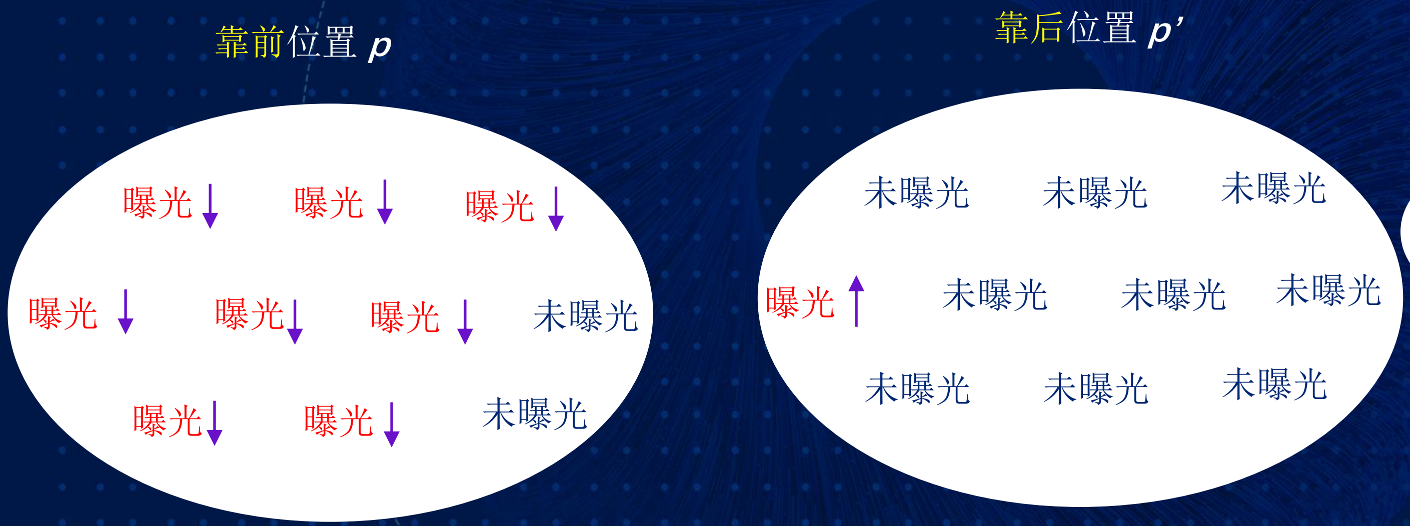
- 位置偏差导致的分布不平衡问题

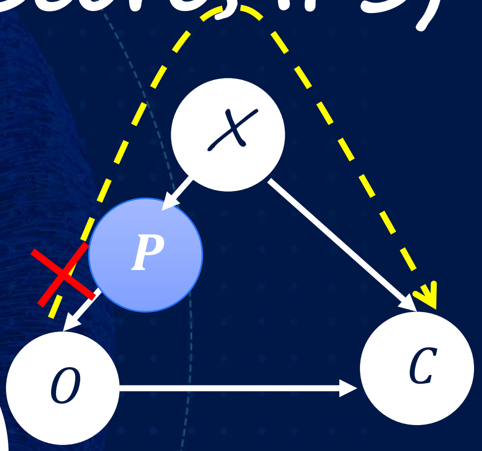


- 分组、IPS的共同本质：平衡实验组与对照组的差异。
- 将逆倾向得分设置为：1/p = 1/Pr(O=1)，即曝光概率的倒

逆倾向得分 (Inverse Propensity Score, IPS)

- 位置偏差导致的分布不平衡问题





- 分组、IPS的共同本质：平衡实验组与对照组的差异。
- 将逆倾向得分设置为：$1/p = 1/Pr(O=1)$，即曝光概率的倒

逆倾向得分 (Inverse Propensity Score, IPS)

- 倾向得分的表达：以排序中的位置偏差为例
  - 全候选文档上的原始目标函数：

$$\boldsymbol{\theta} = \arg\min_{\boldsymbol{\theta}\in\Theta} \mathcal{L}(\boldsymbol{\theta}) := \sum_{d_i\in\boldsymbol{d}} \Delta\left(r_i, f_{\boldsymbol{\theta}}(d_i|\boldsymbol{q},\boldsymbol{d})\right) + \lambda\mathcal{R}(\boldsymbol{\theta})$$

  - 曝光文档上的经 *IPS* 修正的目标函数：

$$\boldsymbol{\theta}_{\mathrm{IPS}} = \arg\min_{\boldsymbol{\theta}\in\Theta} \mathcal{L}_{\mathrm{IPS}}(\boldsymbol{\theta}) := \sum_{\boxed{i:o_i=1}} \frac{\Delta\left(r_i, f_{\boldsymbol{\theta}}(d_i|\boldsymbol{q},\boldsymbol{d})\right)}{\boxed{\Pr(o_i=1)}} + \lambda\mathcal{R}(\boldsymbol{\theta})$$

曝光数据

倾向得分

逆倾向得分 (Inverse Propensity Score, IPS)

- 倾向得分的表达：以排序中的位置偏差为例

  - 引入点击模型

$$\underbrace{\Pr(C=1)}_{\text{用户点击}} = \underbrace{\Pr(O=1)}_{\text{用户观察到文档}} \times \underbrace{\Pr(R=1)}_{\text{文档-查询相关性}}$$

  - 引入排序度量（*additive measure*）

$$\Delta\left(r_i, f_{\boldsymbol{\theta}}(d_i|\boldsymbol{q},\boldsymbol{d})\right) = \mathrm{rank}(d_i|\boldsymbol{q},\boldsymbol{d},f_{\boldsymbol{\theta}}) \times r_i$$

即：相关的文档排在靠后位置 → 大损失

$$\begin{aligned}
\mathcal{L}_{\mathrm{IPS}}(\boldsymbol{\theta}) &= \sum_{i:o_i=1} \frac{\Delta\left(r_i, f_{\boldsymbol{\theta}}(d_i|\boldsymbol{q},\boldsymbol{d})\right)}{\Pr(o_i=1)} + \lambda\mathcal{R}(\boldsymbol{\theta}) \\
&= \sum_{i:o_i=1} \frac{\mathrm{rank}(d_i|\boldsymbol{q},\boldsymbol{d},f_{\boldsymbol{\theta}}) \times r_i}{\Pr(o_i=1)} + \lambda\mathcal{R}(\boldsymbol{\theta}) \\
&= \sum_{i:o_i=1\wedge r_i=1} \frac{\mathrm{rank}(d_i|\boldsymbol{q},\boldsymbol{d},f_{\boldsymbol{\theta}})}{\Pr(o_i=1)} + \lambda\mathcal{R}(\boldsymbol{\theta}) \\
&= \sum_{\boxed{i:c_i=1}} \frac{\mathrm{rank}(d_i|\boldsymbol{q},\boldsymbol{d},f_{\boldsymbol{\theta}})}{\Pr(o_i=1)} + \lambda\mathcal{R}(\boldsymbol{\theta})
\end{aligned}$$

转化为点击数据上的损失函数

逆倾向得分 (Inverse Propensity Score, IPS)

- 理论保证：

Theorem 1. $\mathbb{E}[\mathcal{L}_{\text{IPS}}(\boldsymbol{\theta})] = \mathcal{L}(\boldsymbol{\theta})$.

Proof.

$$
\begin{aligned}
\mathbb{E}[\mathcal{L}_{\text{IPS}}(\boldsymbol{\theta})] &= \mathbb{E}_O\left[\sum_{i:o_i=1} \frac{\Delta\left(r_i, f_{\boldsymbol{\theta}}(d_i|\boldsymbol{q},\boldsymbol{d})\right)}{\Pr(o_i=1)}\right] + \lambda\mathcal{R}(\boldsymbol{\theta}) \\
&= \sum_{i:o_i=1} \mathbb{E}_O\left[\frac{\Delta\left(r_i, f_{\boldsymbol{\theta}}(d_i|\boldsymbol{q},\boldsymbol{d})\right)}{\Pr(o_i=1)}\right] + \lambda\mathcal{R}(\boldsymbol{\theta}) \\
&= \sum_{d_i\in\boldsymbol{d}} \mathbb{E}_O\left[o_i \times \frac{\Delta\left(r_i, f_{\boldsymbol{\theta}}(d_i|\boldsymbol{q},\boldsymbol{d})\right)}{\Pr(o_i=1)}\right] + \lambda\mathcal{R}(\boldsymbol{\theta}) \\
&= \sum_{d_i\in\boldsymbol{d}} \Pr(o_i=1) \times \frac{\Delta\left(r_i, f_{\boldsymbol{\theta}}(d_i|\boldsymbol{q},\boldsymbol{d})\right)}{\Pr(o_i=1)} + \lambda\mathcal{R}(\boldsymbol{\theta}) \\
&= \sum_{d_i\in\boldsymbol{d}} \Delta\left(r_i, f_{\boldsymbol{\theta}}(d_i|\boldsymbol{q},\boldsymbol{d})\right) + \lambda\mathcal{R}(\boldsymbol{\theta}) \\
&= \mathcal{L}(\boldsymbol{\theta}).
\end{aligned}
$$

期望线性性

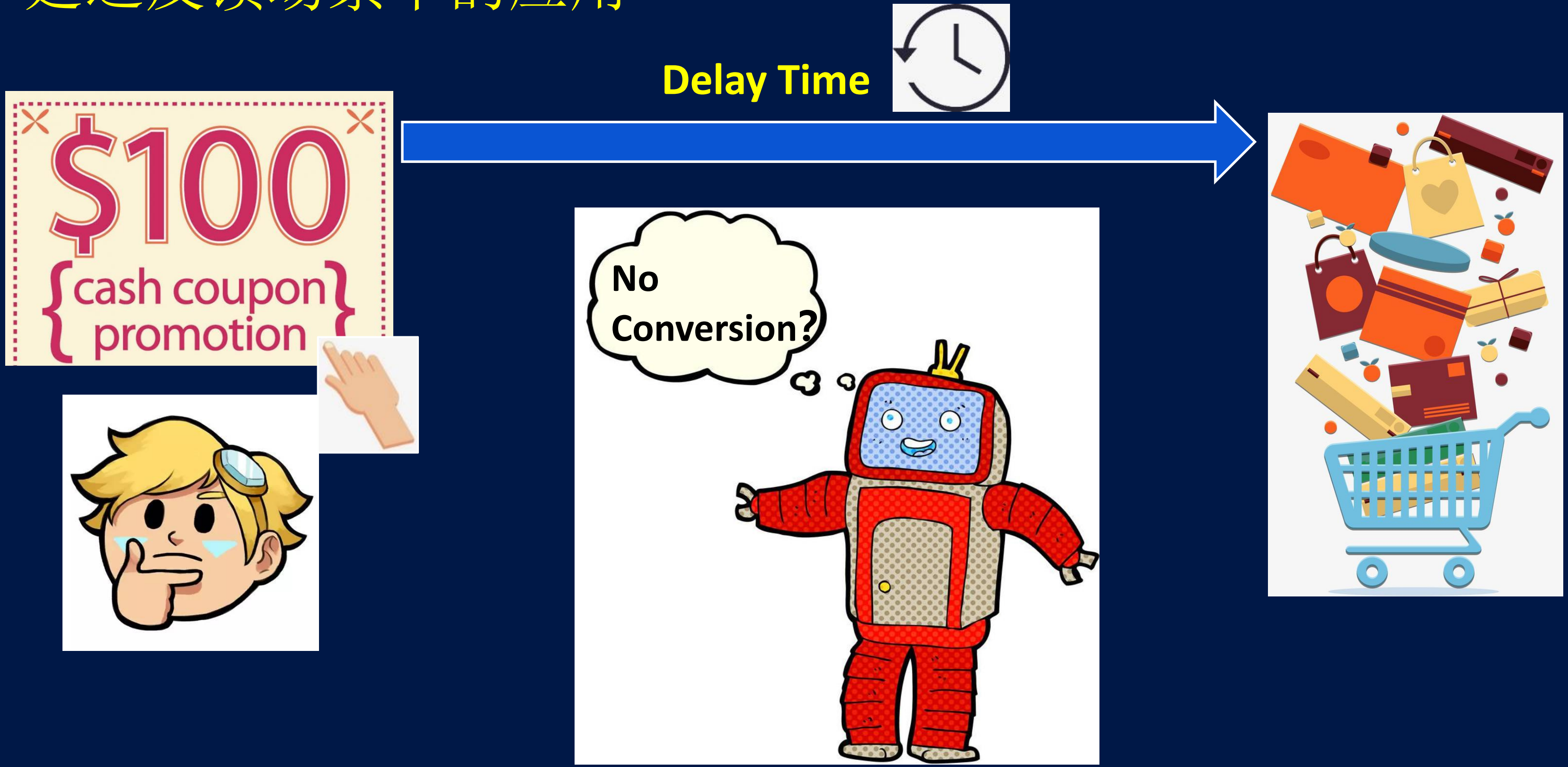
引入指示变量

期望定义

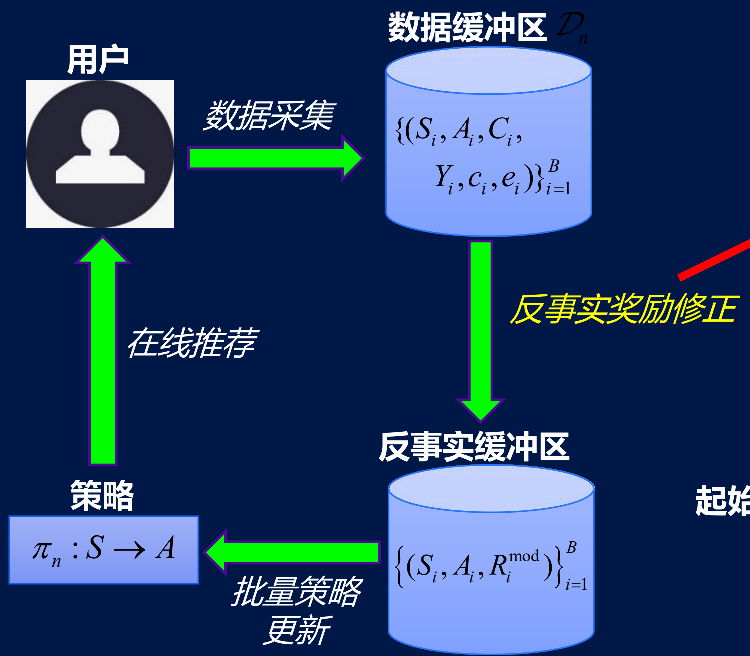
仅需点击数据就可以获得全数据上的无偏估计。

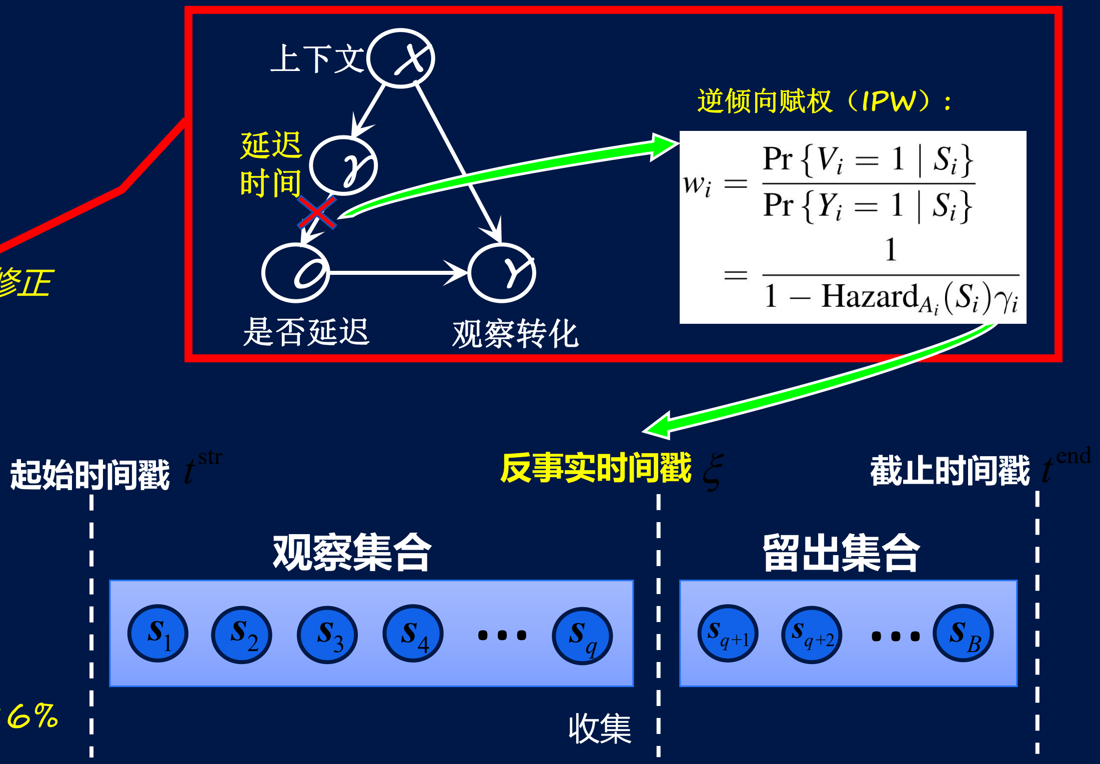
逆倾向得分 (Inverse Propensity Score, IPS)

延迟反馈场景中的应用



逆倾向得分 (Inverse Propensity Score, IPS)





- 在实际产品数据离线测试CVR提升了3.86%

Xiao Zhang, Jun Xu, Jirong Wen, etc. *"Counterfactual Reward Modification for Streaming Recommendation with Delayed Feedback"*, SIGIR 2021.

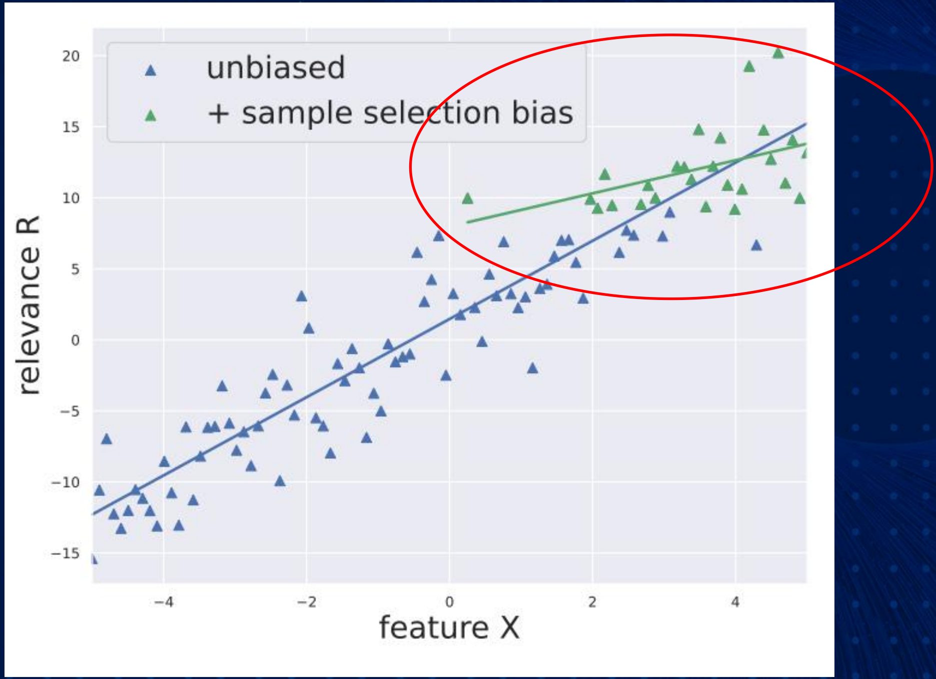
逆倾向得分 (Inverse Propensity Score, IPS)

- IPS方法的缺点：
  - 估计偏差：倾向得分在分母上，微小误差会有较大影响
  - 高方差：倾向得分估计的高方差 → 模型参数的逼近的低收敛率
  - 非零倾向得分：倾向得分不能为 0 → 随机的行为策略

$$\boldsymbol{\theta}_{\mathrm{IPS}} = \underset{\boldsymbol{\theta}\in\Theta}{\arg\min} \sum_{i=1}^{N} \frac{\ell\left(y_i, \hat{y}_i^{\boldsymbol{\theta}}\right)}{\boxed{p_i}} + \lambda\mathcal{R}(\boldsymbol{\theta})$$

倾向得分 (Propensity Score)

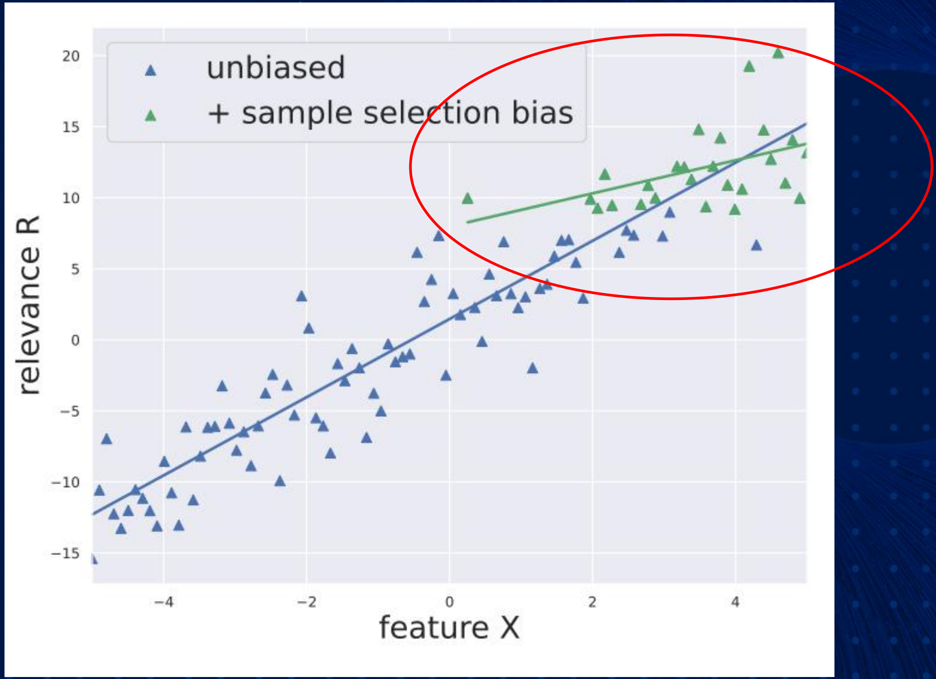
IPS其他扩展：Heckman两阶段法



样本选择偏差→回归偏差

Zohreh Ovaisi, Ragib Ahsan, Yifan Zhang, Kathryn Vasilaky, Elena Zheleva: Correcting for Selection Bias in Learning-to-rank Systems. WWW 2020: 1863-1873

IPS其他扩展：Heckman两阶段法



样本选择偏差→回归偏差

分两步建模：

1. 应用 *Probit* 模型回归选择该样本的概率

2. 修正的最小二乘回归对样本回归（Y = a*X + 修正项）

Zohreh Ovaisi, Ragib Ahsan, Yifan Zhang, Kathryn Vasilaky, Elena Zheleva: Correcting for Selection Bias in Learning-to-rank Systems. WWW 2020: 1863-1873

插补方法 (Imputation)

- 基本原理：估计未观测数据 → 平衡数据分布
- 数学表示：

全样本上

$$\boldsymbol{\theta}_{\mathrm{IMP}} = \arg\min_{\boldsymbol{\theta}\in\Theta} \sum_{i=1}^{\boxed{T}} \ell\left(\boxed{g_i}, \hat{y}_i^{\boldsymbol{\theta}}\right) + \lambda\mathcal{R}(\boldsymbol{\theta})$$

插补标签

  - 适用于未观测到的数据
  - 假设：观测数据对未观测数据具有泛化性（有交集、反馈机制服从某分布）

Stef Van Buuren. Flexible imputation of missing data (2nd edition). CRC Press, 2018.

插补方法 (Imputation)

为什么不通过删除数据达到平均数据分布的效果？

插补方法 (Imputation)

为什么不通过**删除数据**达到**平均数据分布**的效果？

不适用于非随机丢失（MNAR, Missing not at Random），如：存在混淆变量

插补方法 (Imputation)

常见的插补方法

- 均值替代：均值就被用于插补缺失数据
- 多重替代：每条缺失数据替代多次
- 近似替代：观测数据中最相似的数据替代
- 回归插补：从观测数据外推

插补方法 (Imputation)

常见的插补方法

- 均值替代：均值就被用于插补缺失数据
- 多重替代：每条缺失数据替代多次
- 近似替代：观测数据中最相似的数据替代

统计分布

- 回归插补：从观测数据外推

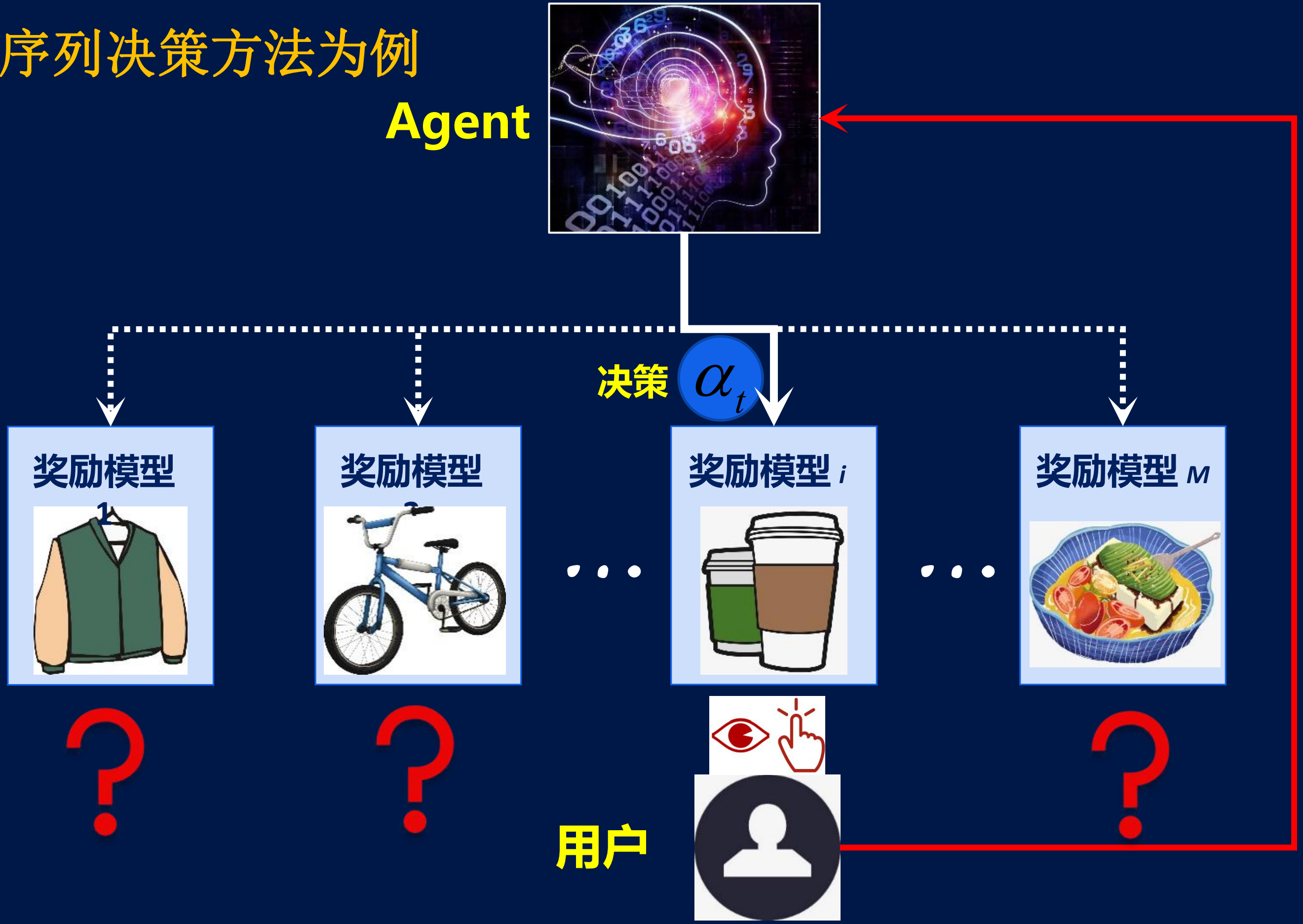
模型泛化

插补方法 (Imputation)

- 以排序中的位置偏差为例

插补方法 (Imputation)

以序列决策方法为例



不完全信息反馈:

$$R_t = R(s_t, a_t, D_t)$$

当前决策

部分信息反馈 (Bandit 反馈):

仅接受当前决策的反馈。

插补方法 (Imputation)

以序列决策方法为例

Agent

不完全信息反馈：

$$R_t = R(s_t, a_t, D_t)$$

当前决策

决策 α_t

部分信息反馈 (Bandit 反馈)：

仅接受当前决策的反馈。

奖励模型 1

?

奖励模型 2

?

…

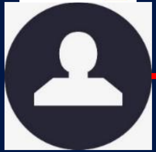
奖励模型 i

…

奖励模型 M

?

用户



训练数据

反馈未知

反馈已知

反馈未知

反馈未知

反馈未知

插补方法 (Imputation)

定理 (后悔界, Xiao Zhang, Jun Xu et al., submitted to ICLR 2022)

令 $\eta \in (0,1)$。
第 n 幕, 若奖励 $\{R_{n,b}\}_{b\in[B]}$ 独立且有上界 C_R,
则在每一回合 $b \in [B]$ 中对于动作 $A \in \mathcal{A}$ 以至少 $1-\delta$ 的概率有

$$\left|\langle \bar{\boldsymbol{\theta}}_A^n, \boldsymbol{s}_{n,b}\rangle - \langle \boldsymbol{\theta}_A^*, \boldsymbol{s}_{n,b}\rangle\right| \le \underbrace{\left[\lambda\|\boldsymbol{\theta}_A^*\|_2 + \nu + \gamma^{\frac{1}{2}}\eta^{-\frac{1}{2}}C_{\mathrm{Imp}}\right]}_{\text{偏差}} \underbrace{\left[\boldsymbol{s}_{n,b}^\top \left(\boldsymbol{\Psi}_A^n\right)^{-1} \boldsymbol{s}_{n,b}\right]^{\frac{1}{2}}}_{\text{方差}},$$

逐渐减小的
额外偏差

低方差

插补方法 (Imputation)

- 插补方法的缺点：
  - 如何估计：插补模型需要具有泛化能力
  - 如何评价：插补模型的真实标签未知（插补标签准确性的评价）

全样本上

$$\boldsymbol{\theta}_{\mathrm{IMP}}=\underset{\boldsymbol{\theta}\in\Theta}{\arg\min}\sum_{i=1}^{\boxed{T}}\ell\left(\boxed{g_i},\hat{y}_i^{\boldsymbol{\theta}}\right)+\lambda\mathcal{R}(\boldsymbol{\theta})$$

插补标签

一般需要和其他纠偏方法结合

双稳健方法 (Doubly Robust Methods)

- 基本原理：观测数据赋权 + 未观测数据估计→ 平衡数据分布

(IPS + Imputation)

- 数学表示：

$$\boldsymbol{\theta}_{\mathrm{DR}} = \underset{\boldsymbol{\theta} \in \Theta}{\arg\min} \sum_{i=1}^{T} \Bigg[\underbrace{\frac{\ell\left(y_i, \hat{y}_i^{\boldsymbol{\theta}}\right) - \ell\left(g_i, \hat{y}_i^{\boldsymbol{\theta}}\right)}{\Pr(o_i = 1)} \times \mathbb{I}(o_i = 1)}_{\text{逆倾向得分 (IPS)}} + \underbrace{\ell\left(g_i, \hat{y}_i^{\boldsymbol{\theta}}\right)}_{\text{插补项 (Imputation)}} \Bigg] + \lambda \mathcal{R}(\boldsymbol{\theta})$$

Cassel, C. M., S¨arndal, C. E., and Wretman, J. H. Some results on generalized difference estimation and generalized regression estimation for finite populations. Biometrika, 63:615–620, 1976.

双稳健方法 (Doubly Robust Methods)

- 基本原理：观测数据赋权 + 未观测数据估计→ 平衡数据分布

(IPS + Imputation)

- 数学表示：

全样本上　　Baseline　　只在观测数据上 (o=1)

$$\boldsymbol{\theta}_{\mathrm{DR}} = \arg\min_{\boldsymbol{\theta}\in\Theta} \sum_{i=1}^{T} \left[\underbrace{\frac{\ell\left(y_i, \hat{y}_i^{\boldsymbol{\theta}}\right) - \ell\left(g_i, \hat{y}_i^{\boldsymbol{\theta}}\right)}{\Pr(o_i = 1)} \times \mathbb{I}(o_i = 1)}_{\text{逆倾向得分 (IPS)}} + \underbrace{\ell\left(g_i, \hat{y}_i^{\boldsymbol{\theta}}\right)}_{\text{插补项 (Imputation)}} \right] + \lambda \mathcal{R}(\boldsymbol{\theta})$$

Cassel, C. M., S¨arndal, C. E., and Wretman, J. H. Some results on generalized difference estimation and generalized regression estimation for finite populations. Biometrika, 63:615–620, 1976.

双稳健方法 (Doubly Robust Methods)

- 以排序中的位置偏差为例

双稳健方法 (Doubly Robust Methods)

- **理论保证**

令IPS估计的误差为 $\varepsilon_{\mathrm{IPS}} = 1 - \frac{p}{\hat{p}}$ ，

插补估计的误差为 $\varepsilon_{\mathrm{IMP}} = g(\boldsymbol{x}) - g_{\mathrm{true}}(\boldsymbol{x})$ ，

则期望奖励估计的误差为: $\left|\mathbb{E}[V_{\mathrm{DR}}] - V_{\mathrm{true}}\right| = \left|\mathbb{E}_{\boldsymbol{x}}\left[\varepsilon_{\mathrm{IPS}} \times \varepsilon_{\mathrm{IMP}}\right]\right|$

Miroslav Dudík, John Langford, Lihong Li: Doubly Robust Policy Evaluation and Learning. ICML 2011: 1097-1104

双稳健方法 (Doubly Robust Methods)

- 理论保证

  令IPS估计的误差为 $\varepsilon_{\mathrm{IPS}} = 1 - \frac{p}{\hat{p}}$，

  插补估计的误差为 $\varepsilon_{\mathrm{IMP}} = g(\boldsymbol{x}) - g_{\mathrm{true}}(\boldsymbol{x})$，

  则期望奖励估计的误差为: $\left|\mathbb{E}[V_{\mathrm{DR}}] - V_{\mathrm{true}}\right| = \left|\mathbb{E}_{\boldsymbol{x}}\left[\varepsilon_{\mathrm{IPS}} \times \varepsilon_{\mathrm{IMP}}\right]\right|$

两误差乘积 → 两种估计之一精确即可

Miroslav Dudík, John Langford, Lihong Li: Doubly Robust Policy Evaluation and Learning. ICML 2011: 1097-1104

双稳健方法 (Doubly Robust Methods)

- 双稳健方法的缺点：
  - 如何折衷：*IPS*项和插补项的折衷
  - 如何优化：倾向得分估计和插补项估计需要同时优化

$$\boldsymbol{\theta}_{\mathrm{DR}} = \arg\min_{\boldsymbol{\theta}\in\Theta} \sum_{i=1}^{T} \left[\underbrace{\frac{\ell\left(y_i, \hat{y}_i^{\boldsymbol{\theta}}\right) - \ell\left(g_i, \hat{y}_i^{\boldsymbol{\theta}}\right)}{\Pr(o_i = 1)} \times \mathbb{I}(o_i = 1)}_{\text{逆倾向得分 (IPS)}} + \underbrace{\ell\left(g_i, \hat{y}_i^{\boldsymbol{\theta}}\right)}_{\text{插补项 (Imputation)}} \right] + \lambda \mathcal{R}(\boldsymbol{\theta})$$

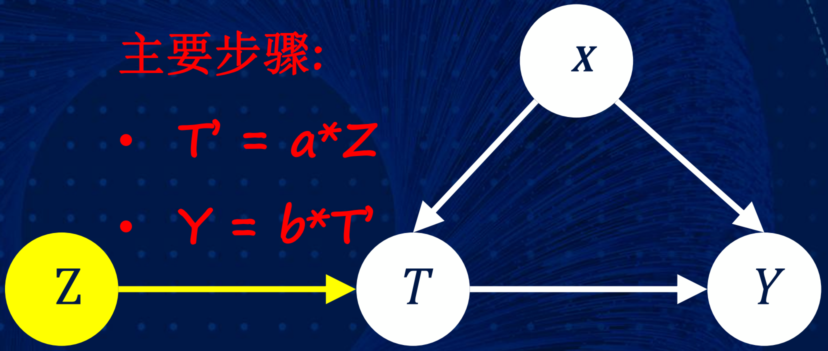
其他方法

1. 工具变量方法
2. 似然分解方法

…

主要步骤:

- $T' = a*Z$
- $Y = b*T'$



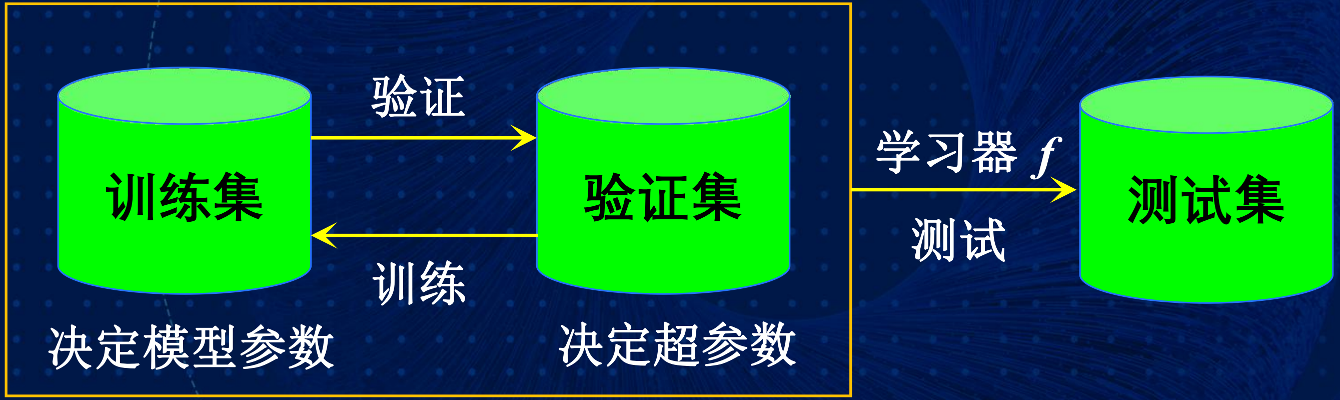
工具变量方法

Jason S. Hartford, Greg Lewis, Kevin Leyton-Brown, Matt Taddy: Deep IV: A Flexible Approach for Counterfactual Prediction. ICML 2017: 1414-1423

提 纲

- 搜索中的偏差
- 因果推断简介
- 基于因果的搜索纠偏方法
- 实验设置方法
- 总结与未来发展方向

有监督学习的实验设置



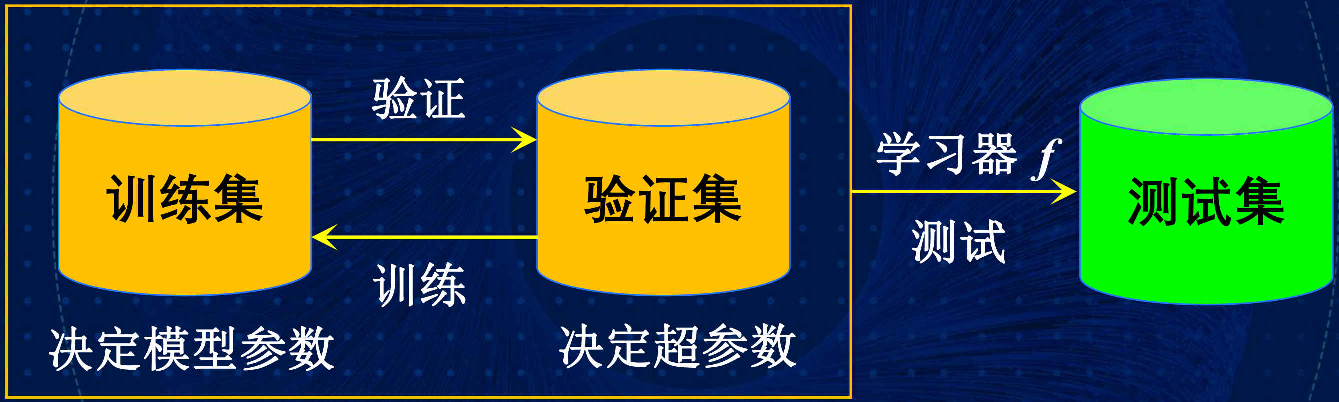
- 独立同分布 (*i.i.d.*)：训练集、验证集、测试集来源于相同数据分布。

搜索场景中的数据

	优点	缺点
无偏数据  人工标注数据	✓ 几乎无偏	× 成本昂贵 × 数量稀少 × 非个性化
有偏数据  用户交互数据	✓ 成本低 ✓ 数量多 ✓ 真实反映用户喜好	× 含有多种偏差 (位置偏差、样本选择偏差、信任偏差等等)

因果纠偏方法实验设置

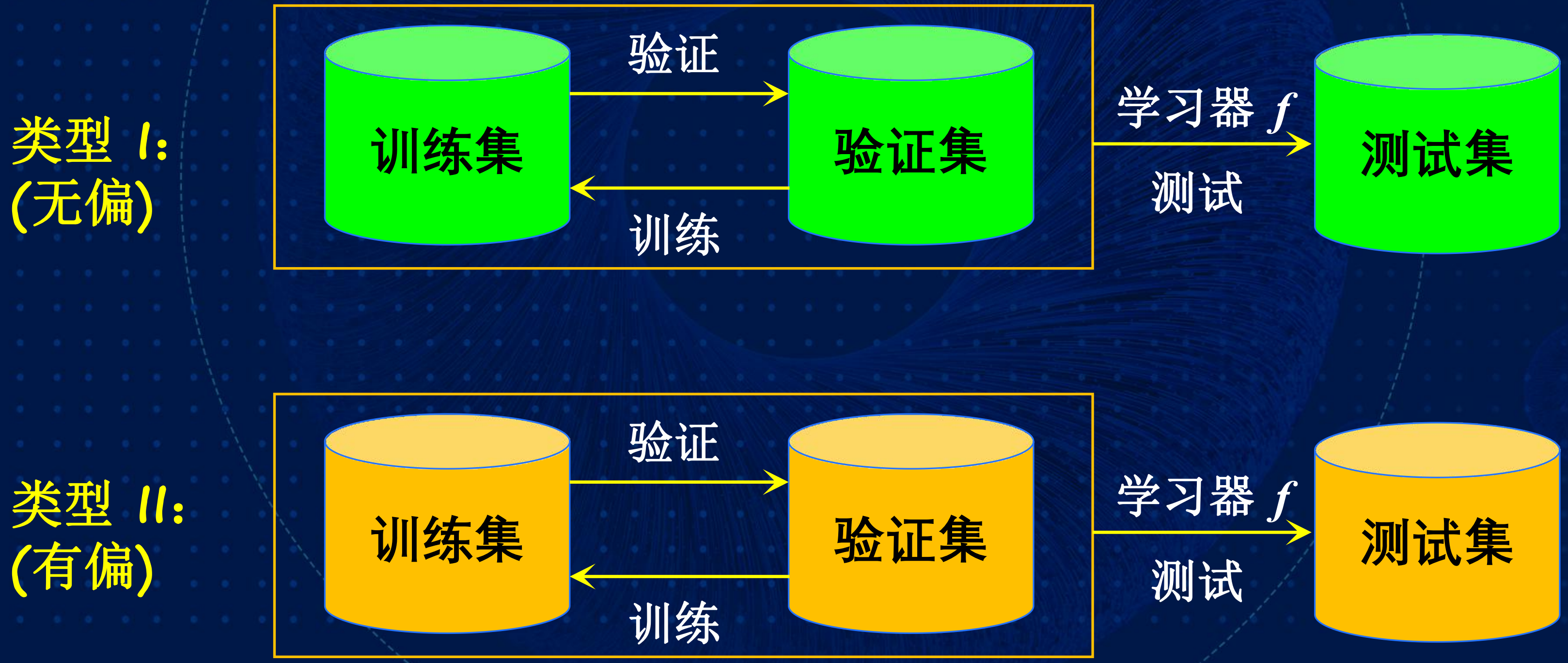
理想中的实验设置



- 训练集 → 有偏的用户交互数据
- 验证集 → 有偏的用户交互数据
- 测试集 → 无偏的人工标注数据

因果纠偏方法实验设置

现实中的公开搜索数据集



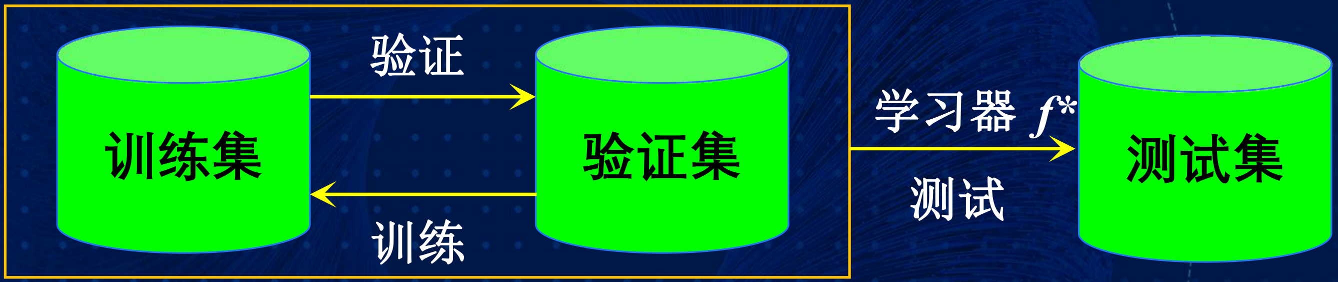
因果纠偏方法实验设置

要求数据集必须既有用户反馈，又有人工标注，而这样的数据很难获得。

因果纠偏方法实验设置

现实中的公开搜索数据集

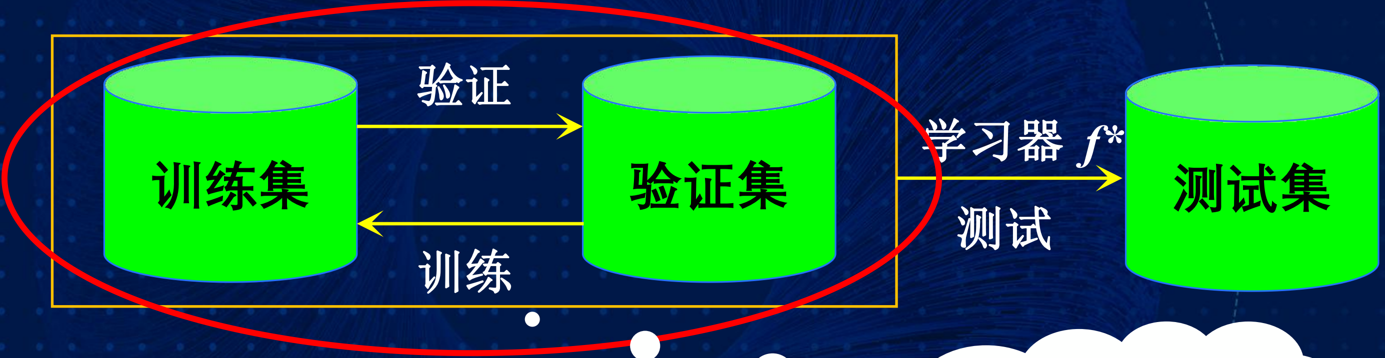
类型 l:
(无偏)



因果纠偏方法实验设置

现实中的公开搜索数据集

类型 I:
(无偏)

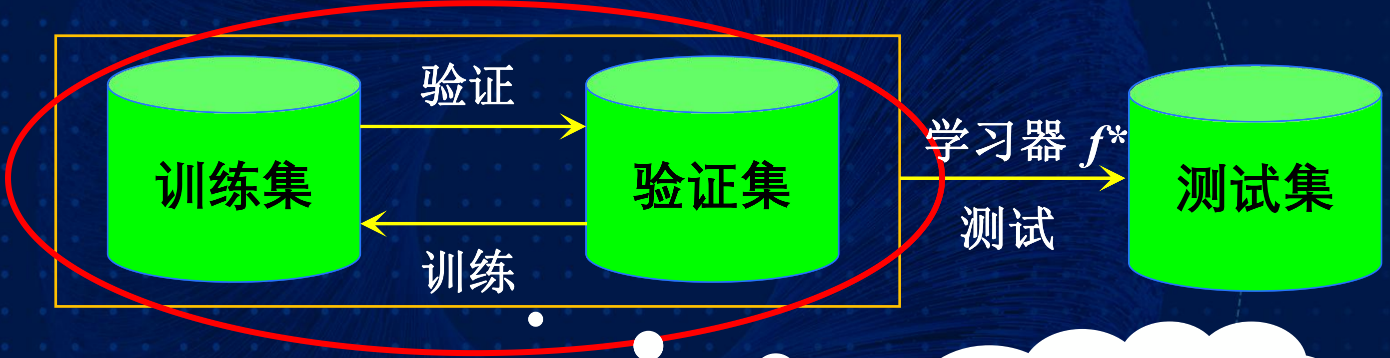


如何构造有偏的训练、验证集?

因果纠偏方法实验设置

现实中的公开搜索数据集

类型 I:
(无偏)

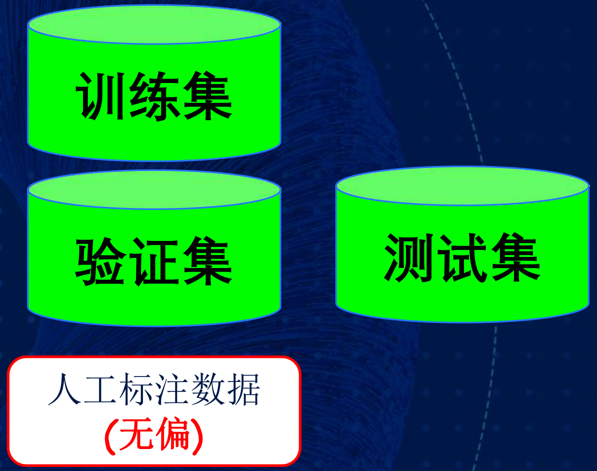


如何构造有偏的训练、验证集?

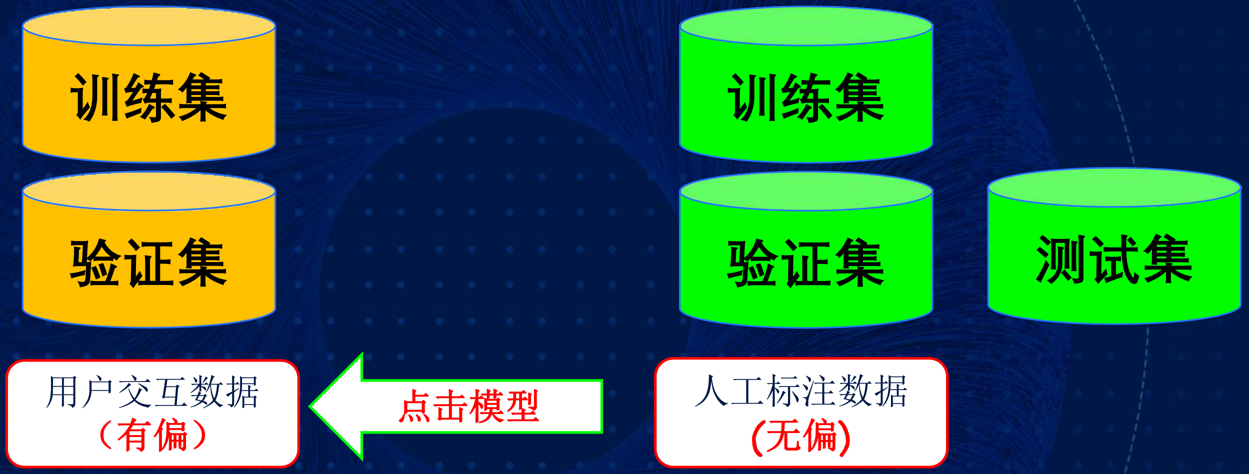
解决方法：半合成实验

（Semi-Synthetic Experiment）

半合成实验流程



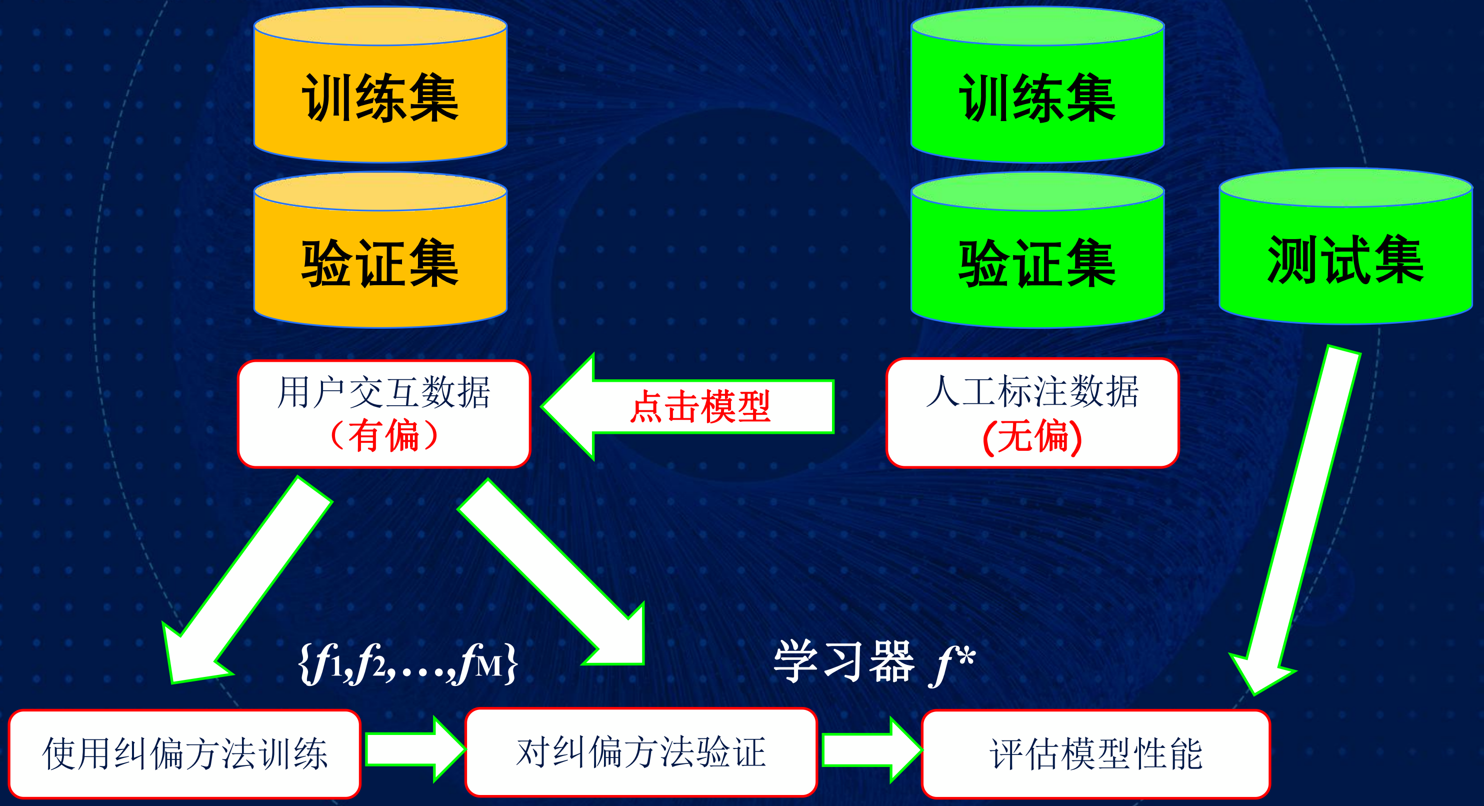
半合成实验流程



半合成实验流程


训练集
验证集
训练集
验证集
测试集
用户交互数据
（有偏）
点击模型
人工标注数据
(无偏)
使用纠偏方法训练
对纠偏方法验证

半合成实验流程


训练集
验证集
训练集
验证集
测试集
用户交互数据
（有偏）
点击模型
人工标注数据
(无偏)
$\{f_1,f_2,\ldots,f_M\}$
学习器 f^*
使用纠偏方法训练
对纠偏方法验证
评估模型性能

半合成实验流程

- 半合成实验（Semi-synthetic Experiment）：

模拟用户交互，使用公开的人工标注数据集合成用户交互数据来进行训练和验证。

➢ 关键问题：如何模拟用户交互行为？点击模型（PBM, UBM等等）

例如：使用PBM模拟带有位置偏差的用户点击

$$P(C = 1) = P(E = 1)P(R = 1)$$

假设$P(E = 1) = (\frac{1}{k})^{\eta}$，其中$k$是某文档所在的位置，从而模拟出每一个文档被点击的概率。以这个点击概率的伯努利分布生成点击数据。

因果纠偏方法实验设置

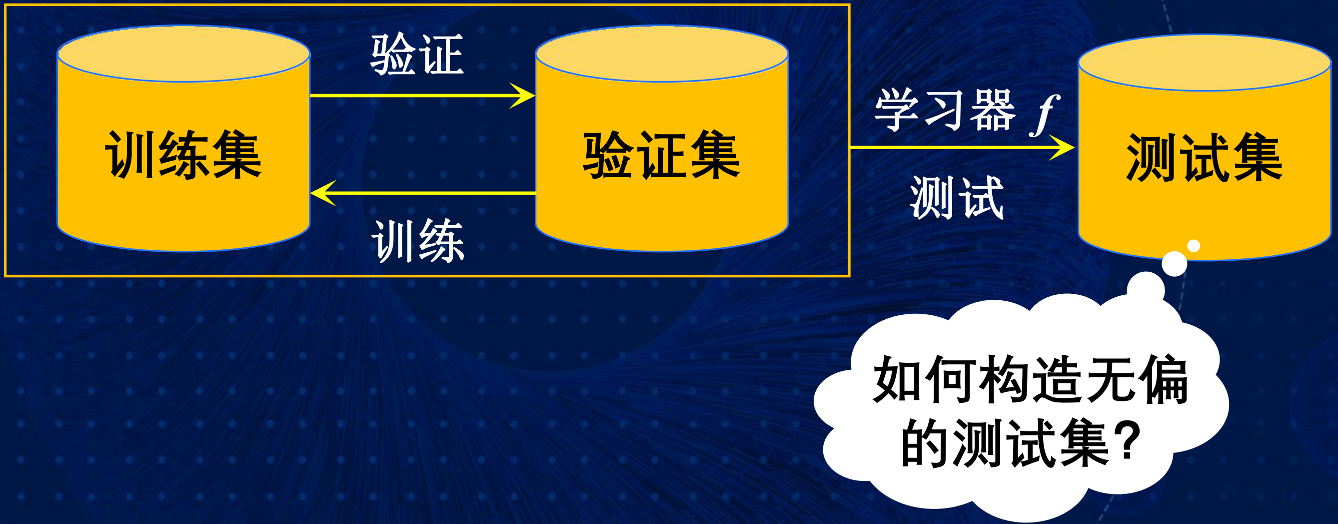
现实中的公开搜索数据集

类型 II:
(有偏)

因果纠偏方法实验设置

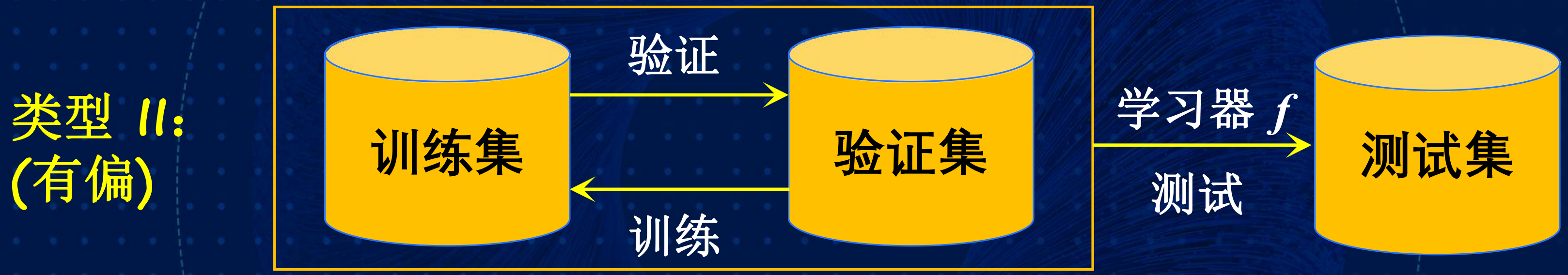
现实中的公开搜索数据集

类型 *II*:
(有偏)



因果纠偏方法实验设置

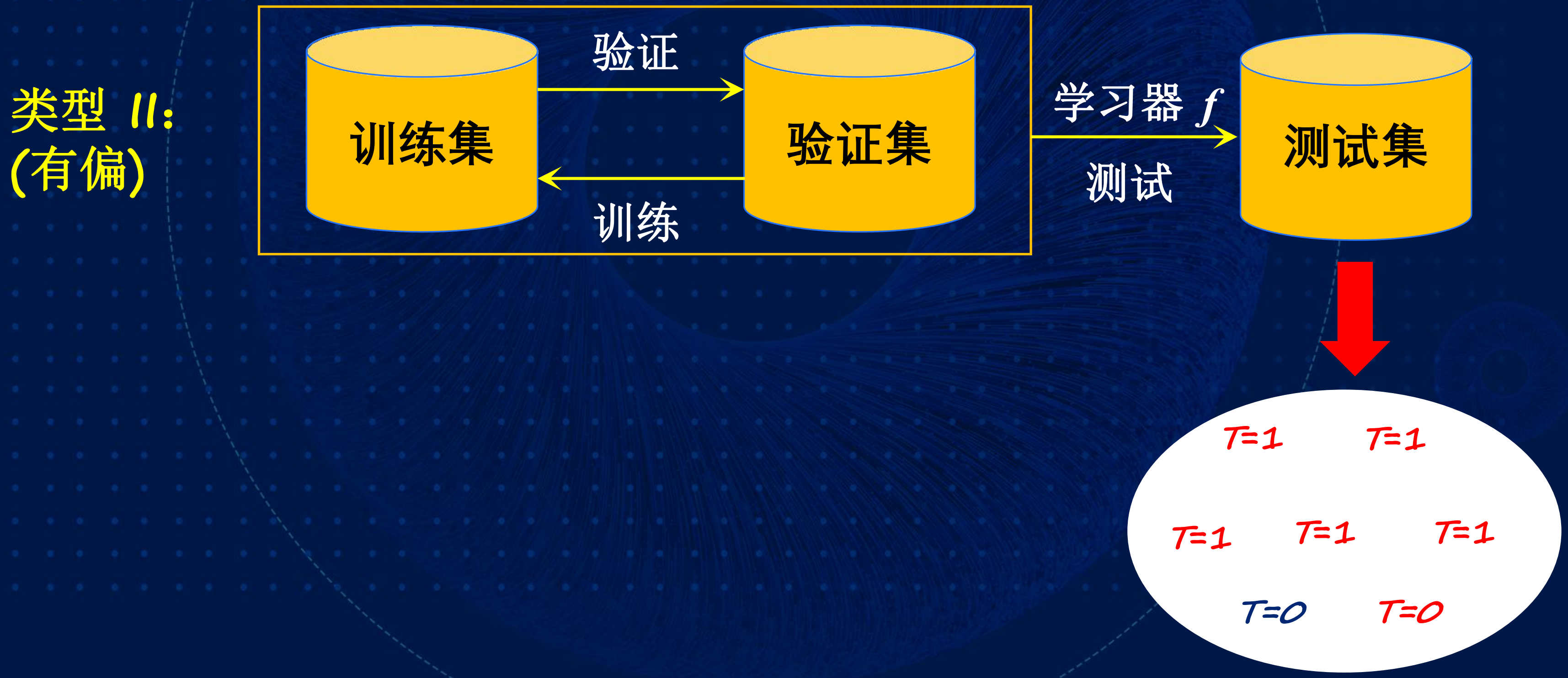
现实中的公开搜索数据集



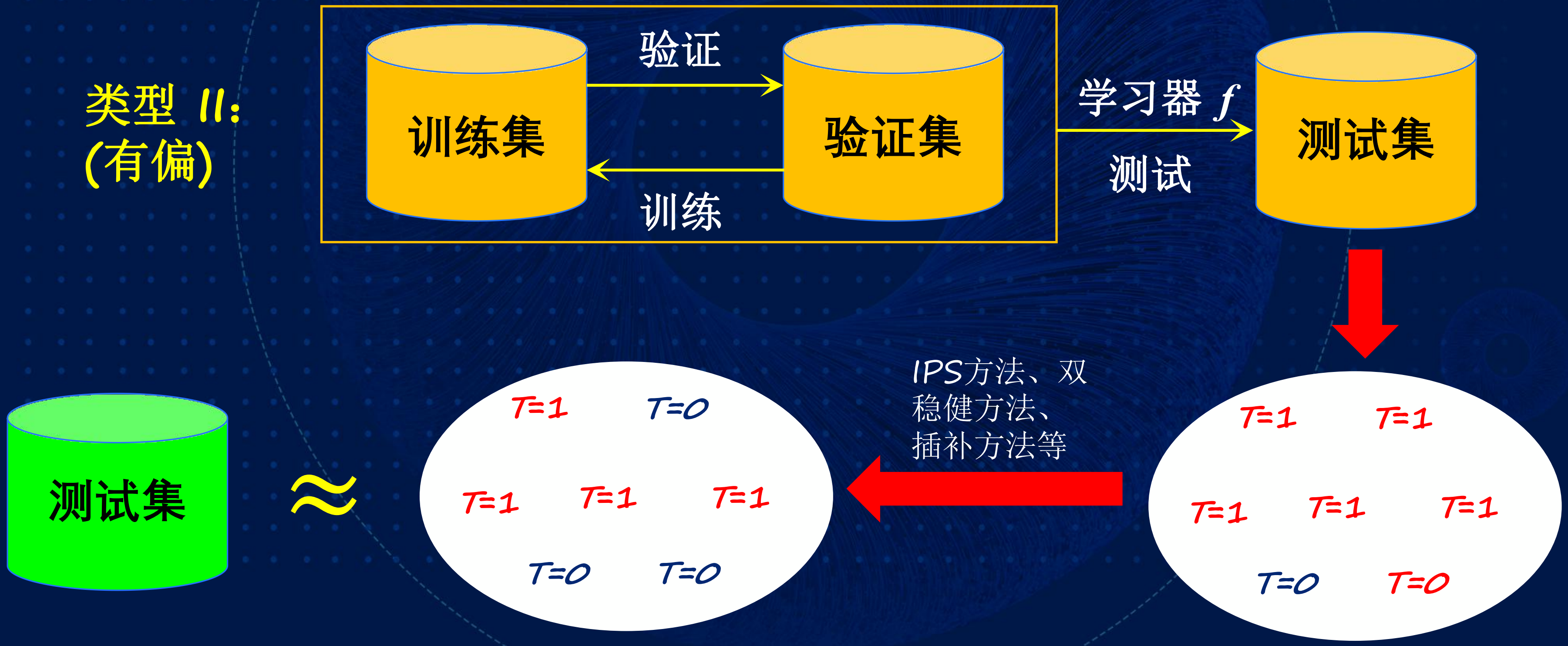
解决方法：反事实评价

（Counterfactual Evaluation）

反事实评价实验设置

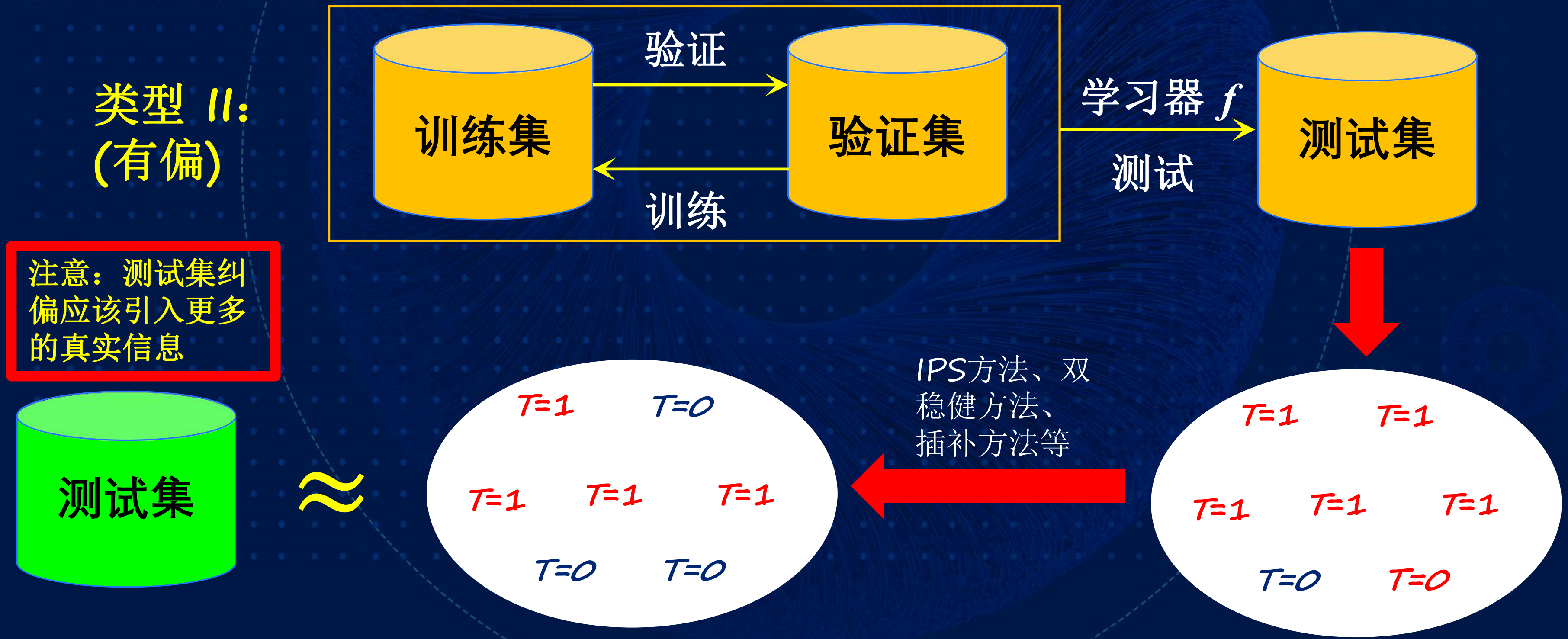


反事实评价实验设置


类型 II:
(有偏)
训练集
验证
验证集
训练
学习器 f
测试
测试集
IPS方法、双
稳健方法、
插补方法等
T=1
T=0
T=1
T=1
T=1
T=0
T=0
T=1
T=1
T=1
T=1
T=1
T=0
T=0
测试集
≈

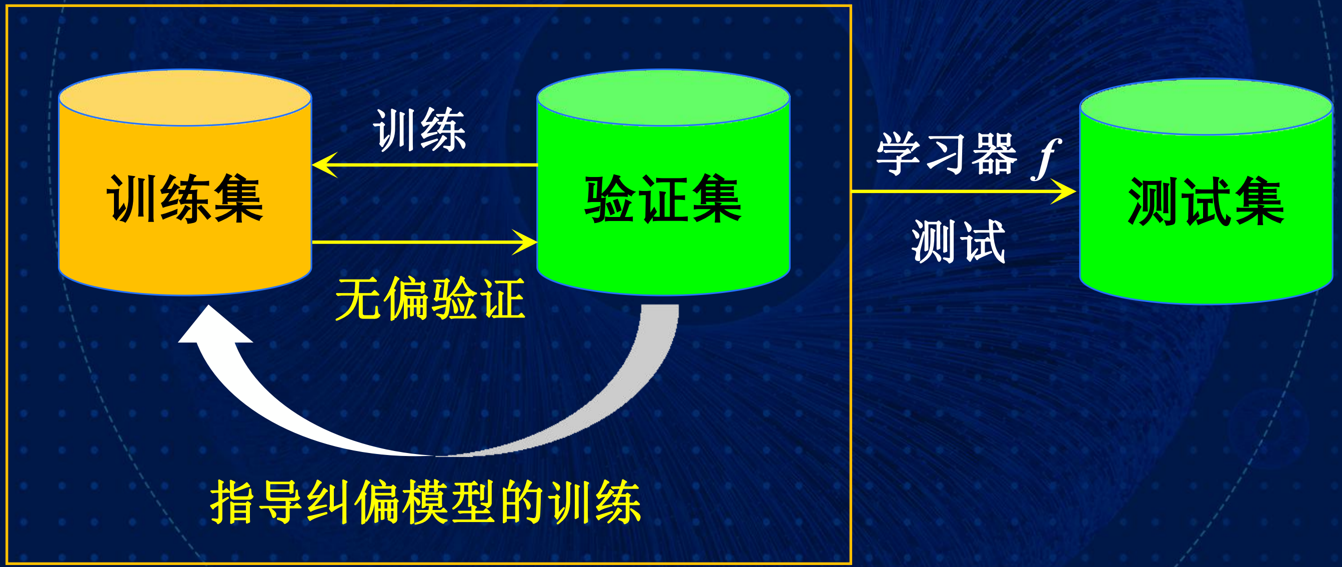
反事实评价实验设置

类型 II:
(有偏)



实验设置：其他进展

有少量用户标注数据可以作为验证集



Jiangxing Yu, Hong Zhu, Chih-Yao Chang, Xinhua Feng, Bo-Wen Yuan, Xiuqiang He, Zhenhua Dong: Influence Function for Unbiased Recommendation. SIGIR 2020: 1929-1932

数据集&评价指标

- 常用的搜索公开数据集

YaHooC14B，WEB30K，WEB10K，MQ2007

- 常用的评价指标

nDCG@K，MRR，MAP

https://webscope.sandbox.yahoo.com/
https://www.microsoft.com/en-us/research/project/mslr/
http://www.bigdatalab.ac.cn/benchmark/bm/dd?data=MQ2007

实验设置中的挑战性问题

- 理想数据难获得
- 离线评价与在线评价的差异性
- 评价指标的测评

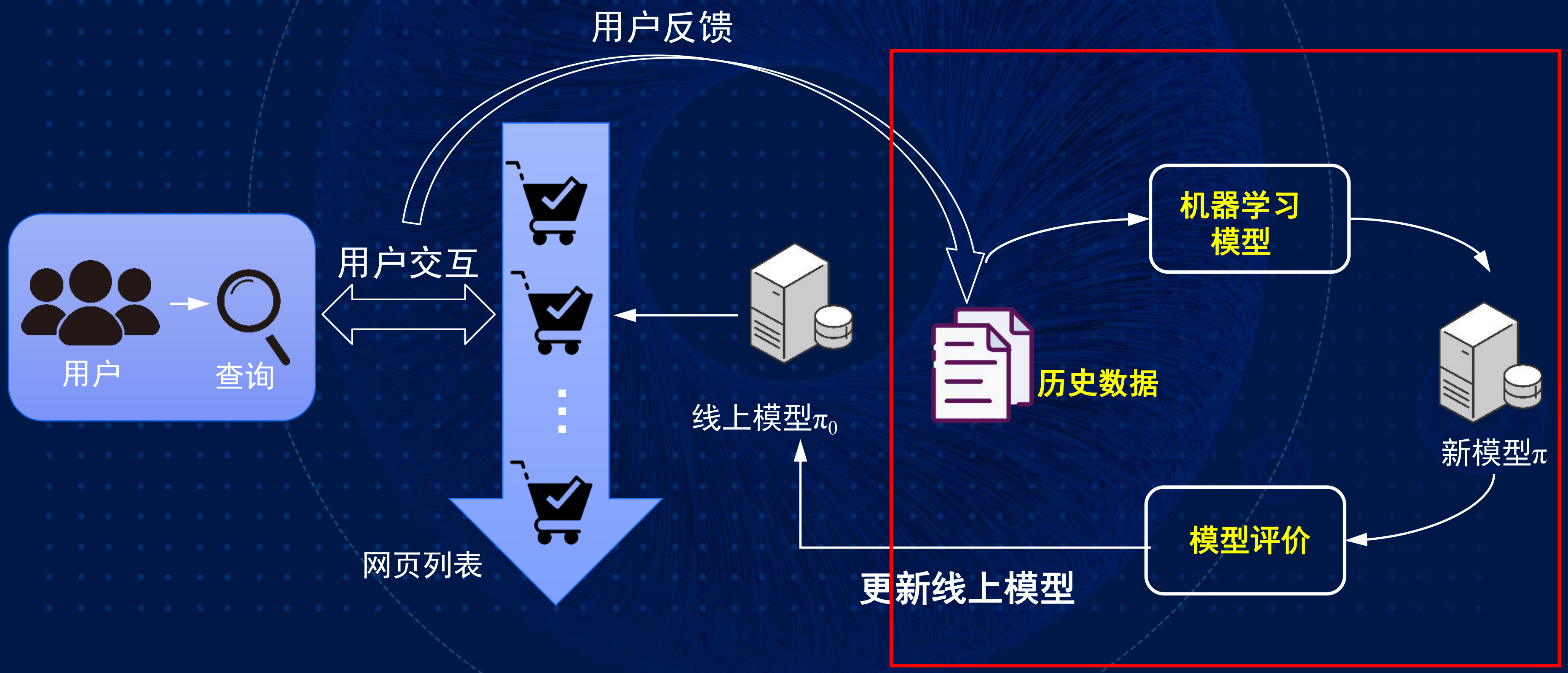
提 纲

- 搜索中的偏差
- 因果推断简介
- 基于因果的搜索纠偏方法
- 实验设置方法
- 总结与未来发展方向

本部分小结

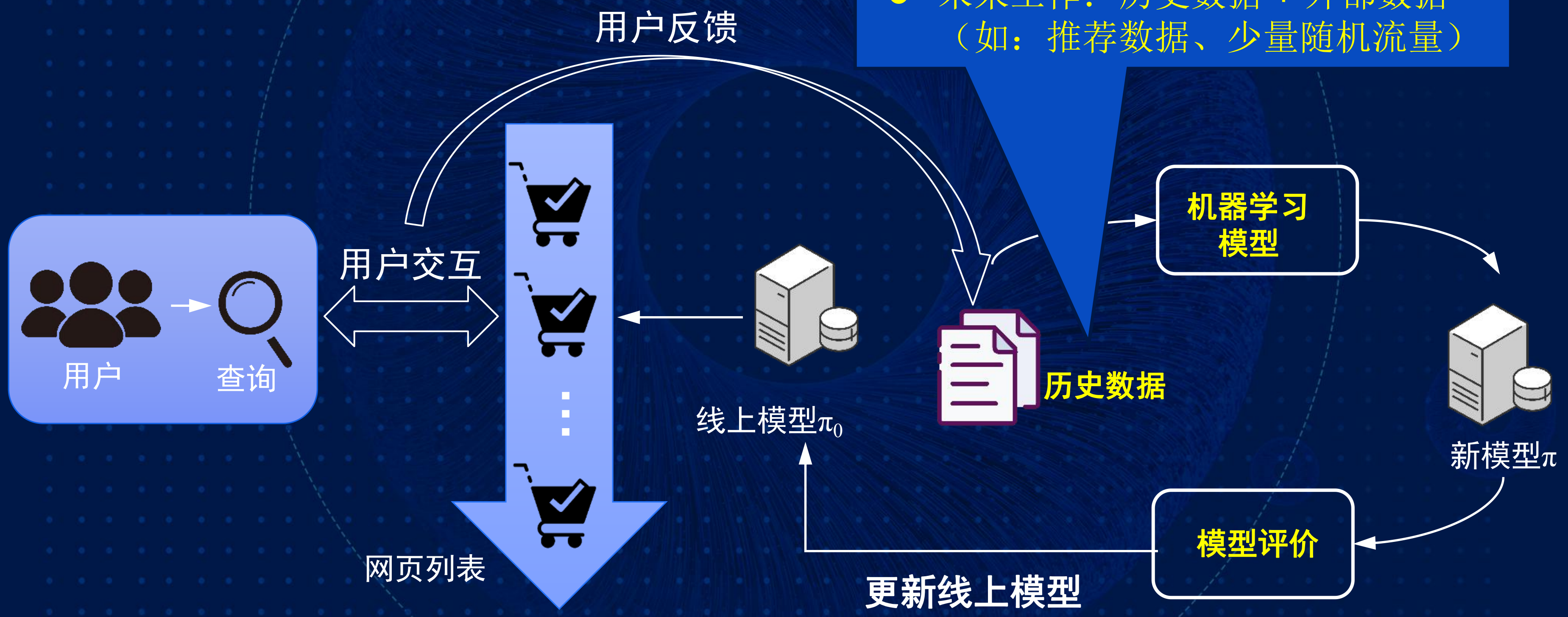
- 基本的因果的搜索纠偏方法的本质：平衡数据分布
- 几类基于因果的搜索纠偏方法（IPS, Imputation, DR 等）
- 因果纠偏的实验设置

未来发展方向



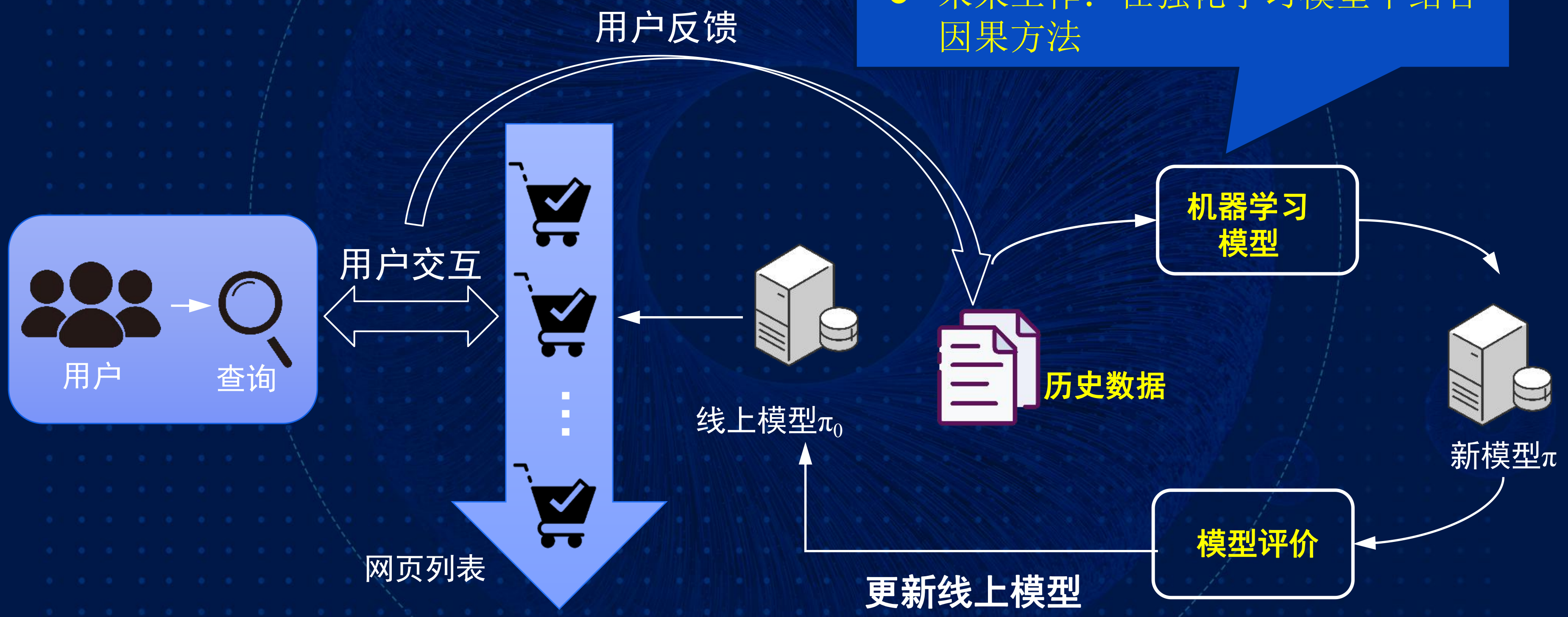
未来发展方向

- 已有工作：历史数据均衡化改造
- 未来工作：历史数据 + 外部数据（如：推荐数据、少量随机流量）



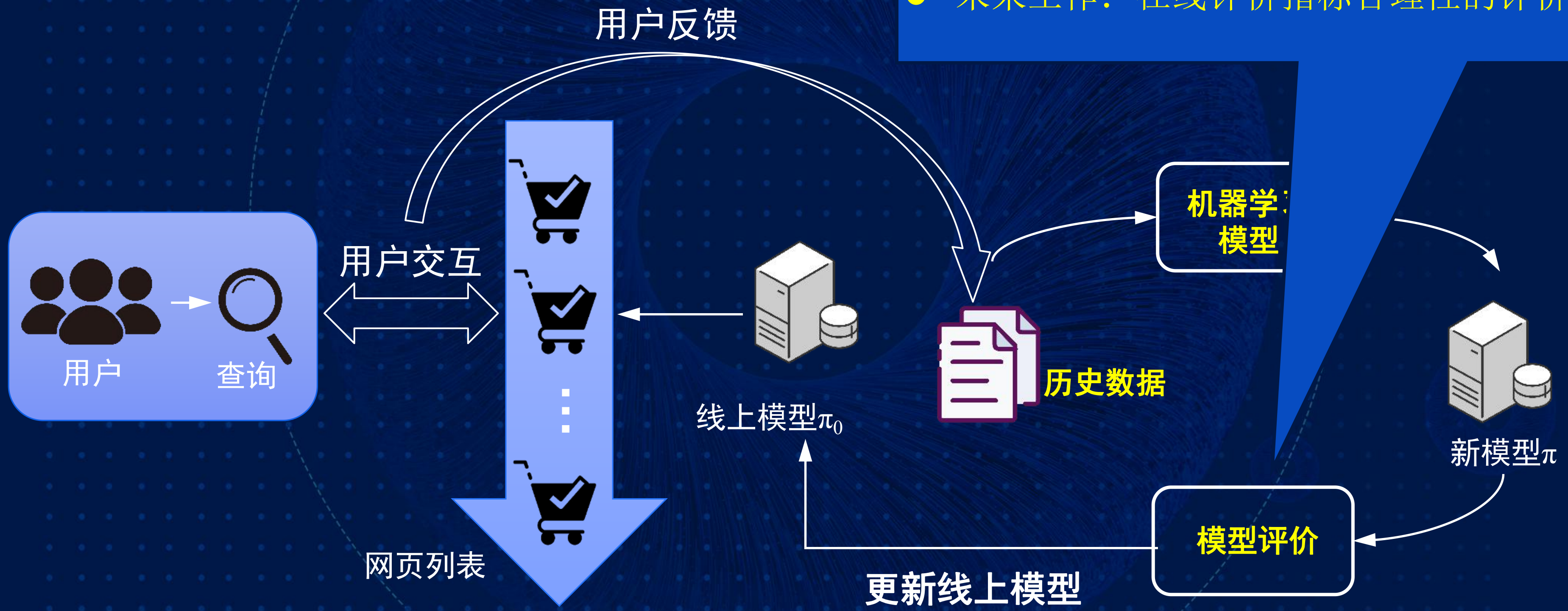
未来发展方向

- 已有工作：统计机器学习模型纠偏
- 未来工作：在强化学习模型中结合因果方法



未来发展方向

- 已有工作：基于因果纠偏的离线评价
- 未来工作：在线评价指标合理性的评价



谢 谢！