

大规模在线学习在OPPO商业算法的设计与实践

OPPO-数智工程系统-算法平台部 / ROY

InfoQ 企业会员

企业数字化传播一站式服务

InfoQ 企业会员是为满足企业在中国开发者群体中的品牌曝光需求而推出的一款矩阵化资源包。可为企业提供包括“企业号服务”、“企业动态宣发”、“品牌展示通道”在内的多项专属权益与服务，助力企业高效触达开发者群体，提升数字化时代影响力。



企业号服务

深度触达 300 万中高端开发者



企业动态宣发

新媒体矩阵覆盖百万粉丝



品牌展示通道

线上平台 10 万+ 流量曝光



1

OPPO商业化系统

2

OPPO实时系统设计

3

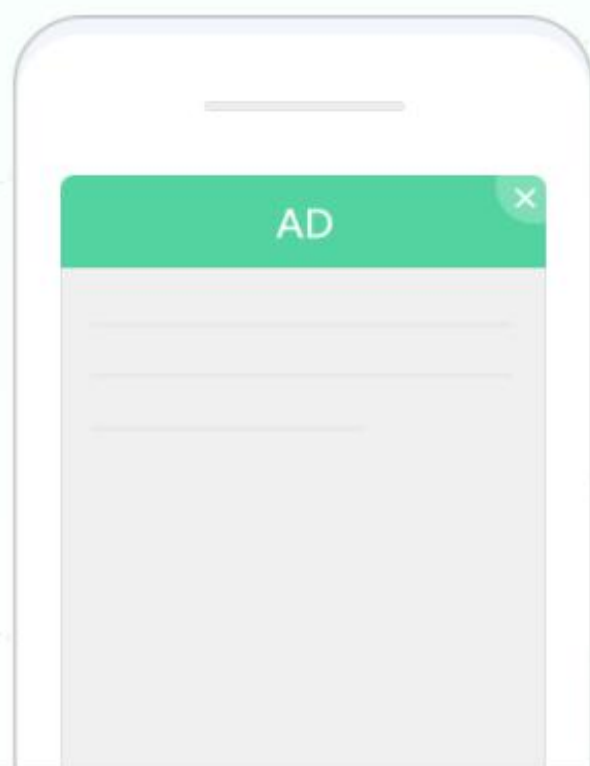
系统开发与实践

4

总结与经验

OPPO商业化广告业务

联盟广告



Banner广告

以通栏式或矩形式出现在应用中

使用场景：用户停留较久或者访问频繁的页面

优势：适用媒体范围广，接入成本低，自然展现于媒体页面中，用户认可度高

插屏广告

基于用户使用场景切换

使用场景：应用内视图切换、游戏过关或失败、图书翻页、应用退出等

优势：交互场景丰富，减少对应用界面的占用，深受广告主青睐，预算能力充足





原生广告

与内容高度融合的广告样式

使用场景：列表页、焦点图、动态更新页、内容页等

优势：广告内容融入媒体环境，深受大牌媒体青睐，提升变现能力成就优秀用户体验

开屏广告

应用开启后全屏展现的广告样式

使用场景：应用启动时

优势：视觉效果震撼，提升媒体品牌价值渗透，广告点击率高，变现能力出色



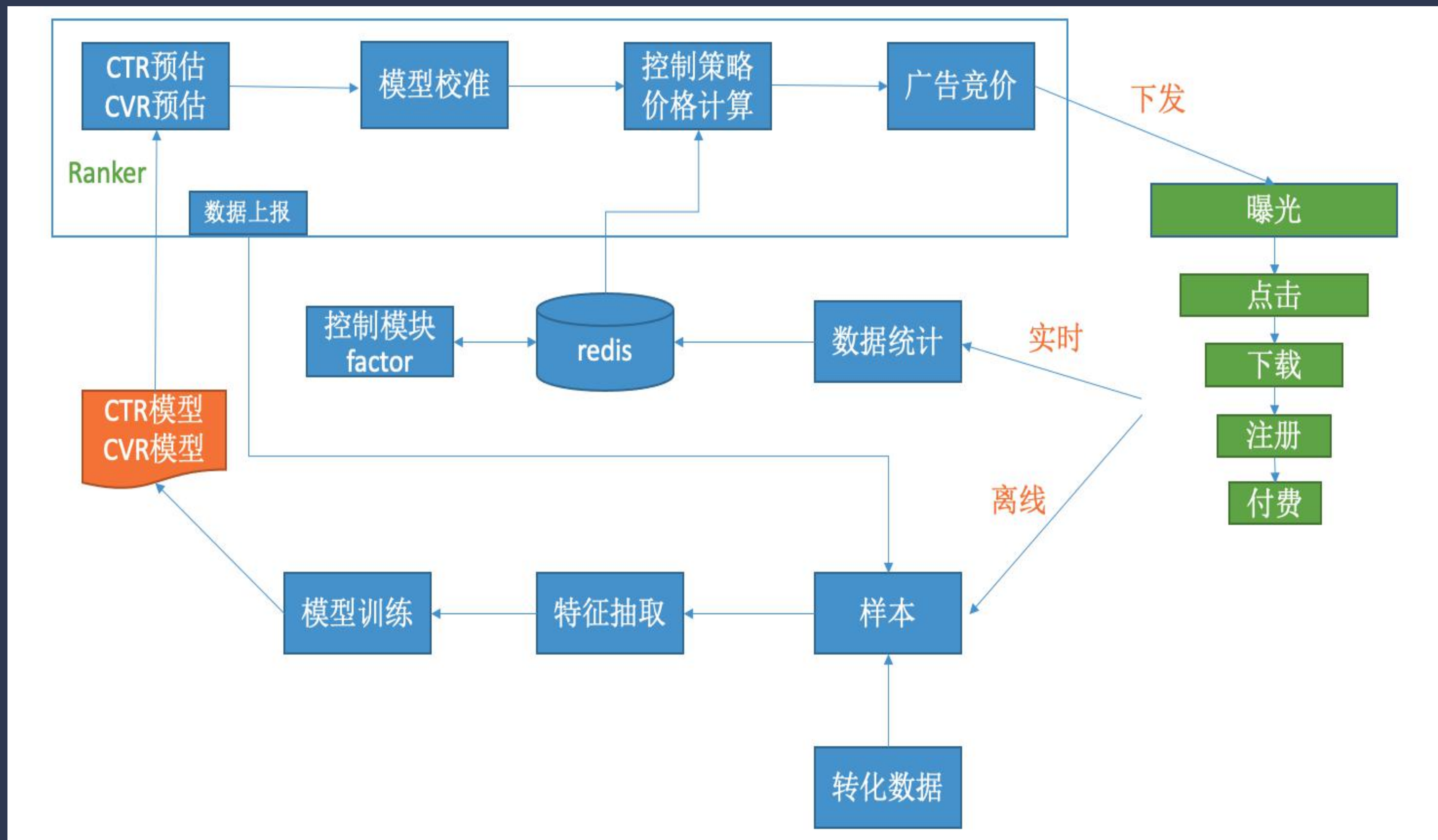
OPPO商业化主要消耗产品

依赖CTR/CVR模型进行预估

- 1、oCPX类广告，以ocpc为主
- 2、CPC广告
- 3、外部的DSP投放CPM/CPC广告

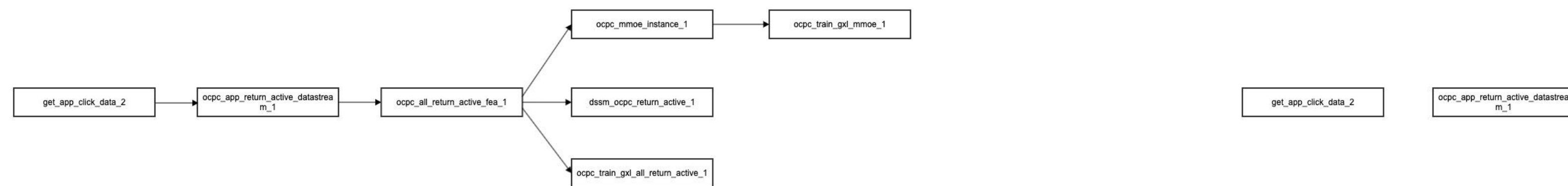
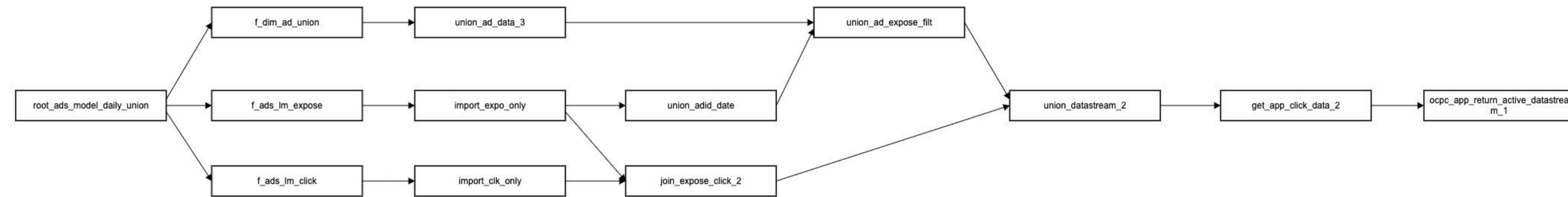
OPPO 商业广告系统架构

- Ranker作为核心服务，接入流量，调用CTR/CVR模型
- 样本任务离线执行，天级别和曝光，点击，转化数据拼接，生产样本，并训练模型



OPPO天级模型架构问题

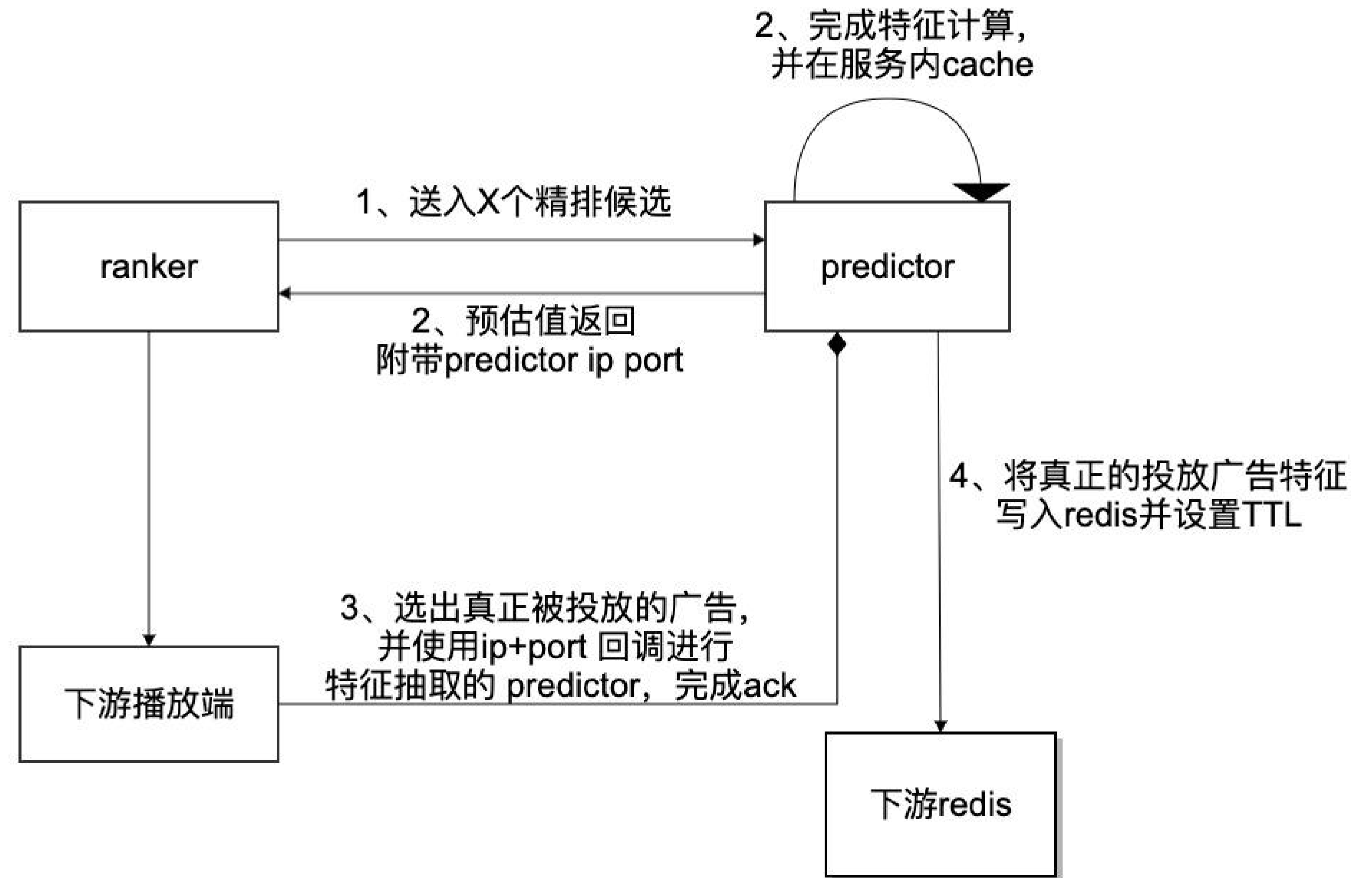
1. 由多个MR任务组成，任务维护难度大
2. 样本产出时间，受到资源，和流量大小的双重影响。
3. 对实时特征的响应慢，attention类的模型结构也难以拿到收益。
4. 离线和在线完成两次特征抽取，带来了线上特征不一致问题



OPPO流式样本设计方案

设计目标

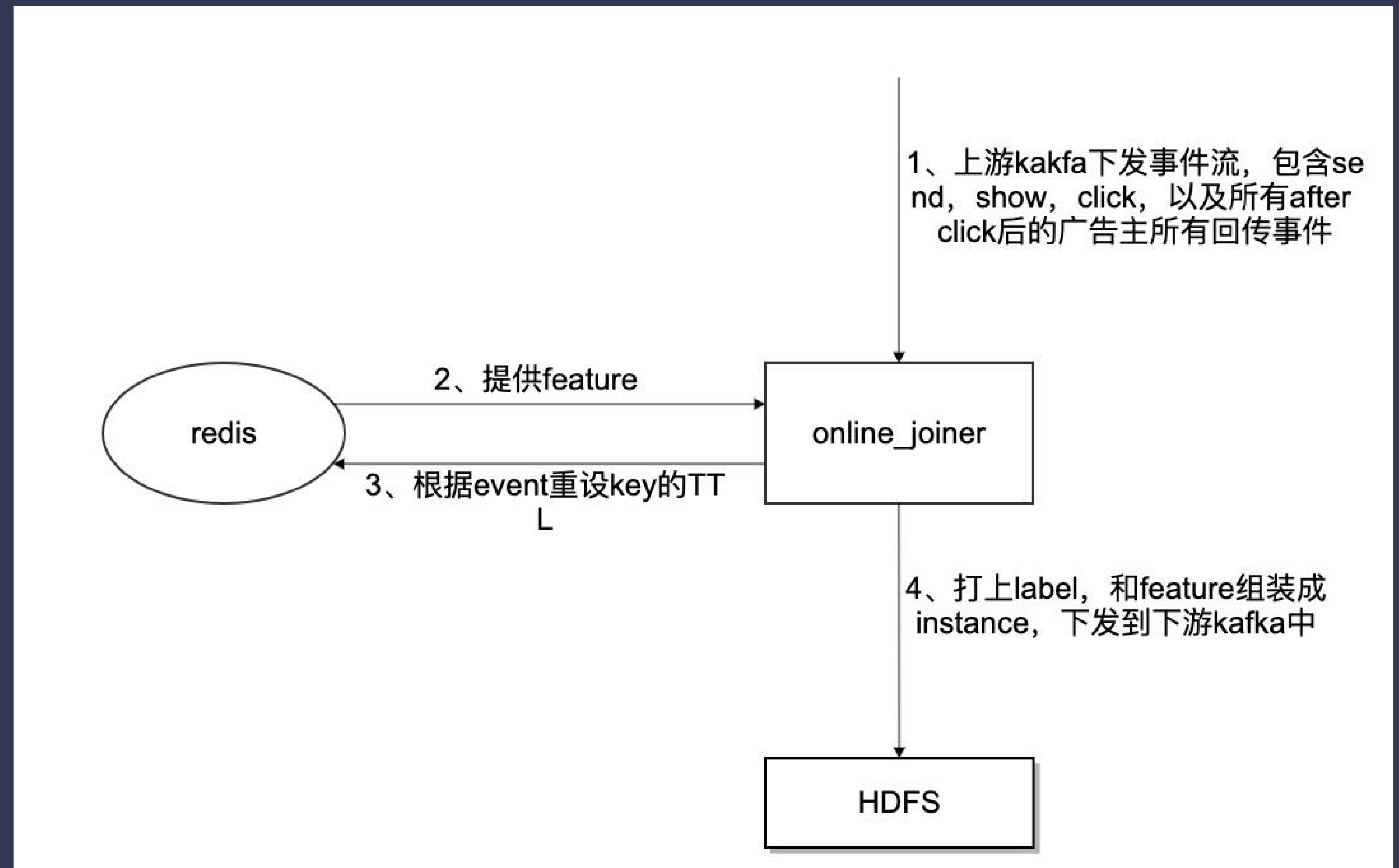
模型训练和serving两个阶段，严格使用同一份特征，即一份流量，只会计算一次特征



OPPO流式样本设计方案

设计目标

事件发生的时间点，即是样本产出的时间点



Delay Feedback问题

关键问题

负例事件到达时，它到底是不是一个正例，是需要一个等待窗口（归因窗口）才知道。

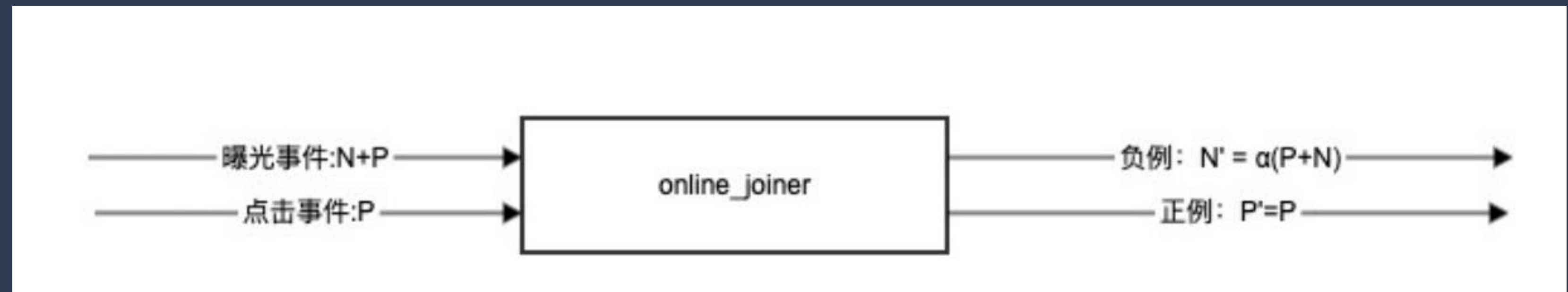
解决办法

按照流式的设计目标，事件发生的时刻，即是样本产生的时刻，当负例事件到达时，和feature拼接，当做一个负样本发送，当正例事件达到时，同样处理逻辑，拼接feature，发送样本

Delay Feedback问题

实际样本

假设以 α 的概率进行负采样，假设真实的正例数量为 P ，负例数量为 N ，实际得到的样本正例 $P' = P$ ， $N' = \alpha(P+N)$



Delay Feedback问题

纠偏方案

《Modeling Delayed Feedback in Display Advertising》：需要真实发生时间D，当前已使用时间E，对时间回流时间进行建模，不符合流式事件来即发样本的需求

- X : 特征
- $Y \in \{0, 1\}$: 当前时刻 conversion 是否已经发生了
- $C \in \{0, 1\}$: conversion 最终是否会发生
- D : 回传延迟的真正时间
- E : 当前已过去的时间

$$P(C = 1|X = x) = p(x) = \frac{1}{1 + \exp(-w_c x)}$$

$$P(D = d|X = x, C = 1) = \lambda(x) \exp(-\lambda(x)d)$$

- 对于当前观察到conversion的样本，其似然函数为公式(3)即，

$$p_1 = \lambda(x_i) \exp(-\lambda(x_i)d_i) * p(x_i)$$

- 对于当前未观察到conversion的样本，其似然函数为公式(4)即

$$p_0 = 1 - p(x_i) + p(x_i) * \exp(-\lambda(x_i)e_i)$$

Delay Feedback问题

纠偏方案

Importance Sampling: 需要假设 $p(x) = f(x)$ ，用模型代替真实观测结果，符合流式样本需求

$$E_p[\log f_\theta(y|x)] = E_b\left[\frac{p(x, y)}{b(x, y)} f_\theta(y|x)\right]$$

$$\begin{aligned} - \sum_{\mathbf{x}, \mathbf{y}} p(\mathbf{y} = 1|\mathbf{x}) \log f_\theta(\mathbf{x}) + p(\mathbf{y} = 0|\mathbf{x}) \log f_\theta(\mathbf{y} = 0|\mathbf{x}) = \\ - \sum_{\mathbf{x}, \mathbf{y}} b(\mathbf{y} = 1|\mathbf{x}) \frac{p(\mathbf{y} = 1|\mathbf{x})}{b(\mathbf{y} = 1|\mathbf{x})} \log f_\theta(\mathbf{x}) + \\ b(\mathbf{y} = 0|\mathbf{x}) \frac{p(\mathbf{y} = 0|\mathbf{x})}{b(\mathbf{y} = 0|\mathbf{x})} \log f_\theta(\mathbf{y} = 0|\mathbf{x}) \quad (7) \end{aligned}$$

$$b(\mathbf{y} = 1|x) = \frac{M}{M + N} = \frac{\frac{M}{N}}{1 + \frac{M}{N}} = \frac{p(\mathbf{y} = 1|x)}{1 + p(\mathbf{y} = 1|x)}$$

$$b(\mathbf{y} = 0|x) = 1 - b(\mathbf{y} = 1|x) = \frac{1}{1 + p(\mathbf{y} = 1|x)}$$

$$\begin{aligned} \mathcal{L}_{IS}(\theta) = - \sum_{\mathbf{x}, \mathbf{y}} b(\mathbf{y} = 1|\mathbf{x})(1 + p(\mathbf{y} = 1|\mathbf{x})) \log f_\theta(\mathbf{x}) + \\ b(\mathbf{y} = 0|\mathbf{x})p(\mathbf{y} = 0|\mathbf{x})(1 + p(\mathbf{y} = 1|\mathbf{x})) \log f_\theta(\mathbf{y} = 0|\mathbf{x}) \quad (10) \end{aligned}$$

Delay Feedback问题

纠偏方案

Fake Negative Calibration: 模型建模实际样本正负例，推导出真实正负例，简单易实现

serving 是输出的概率值应该是

$$\frac{P}{P+N} = \frac{1}{1+e^{-x}}$$

训练目标:

$$\frac{P}{r * (P+N) + P} = \frac{P/(P+N)}{r + P/(P+N)} = \frac{1/(1+e^{-x})}{r + 1/(1+e^{-x})}$$

调整训练logit:

$$\frac{1/(1+e^{-x})}{r + 1/(1+e^{-x})} = \frac{1}{1+e^{-x^*}}$$

$$x^* = -(\ln r + \ln(1 + e^{-x}))$$

Delay Feedback方案效果

参考文献

《Addressing Delayed Feedback for Continuous Training with Neural Networks in CTR prediction》

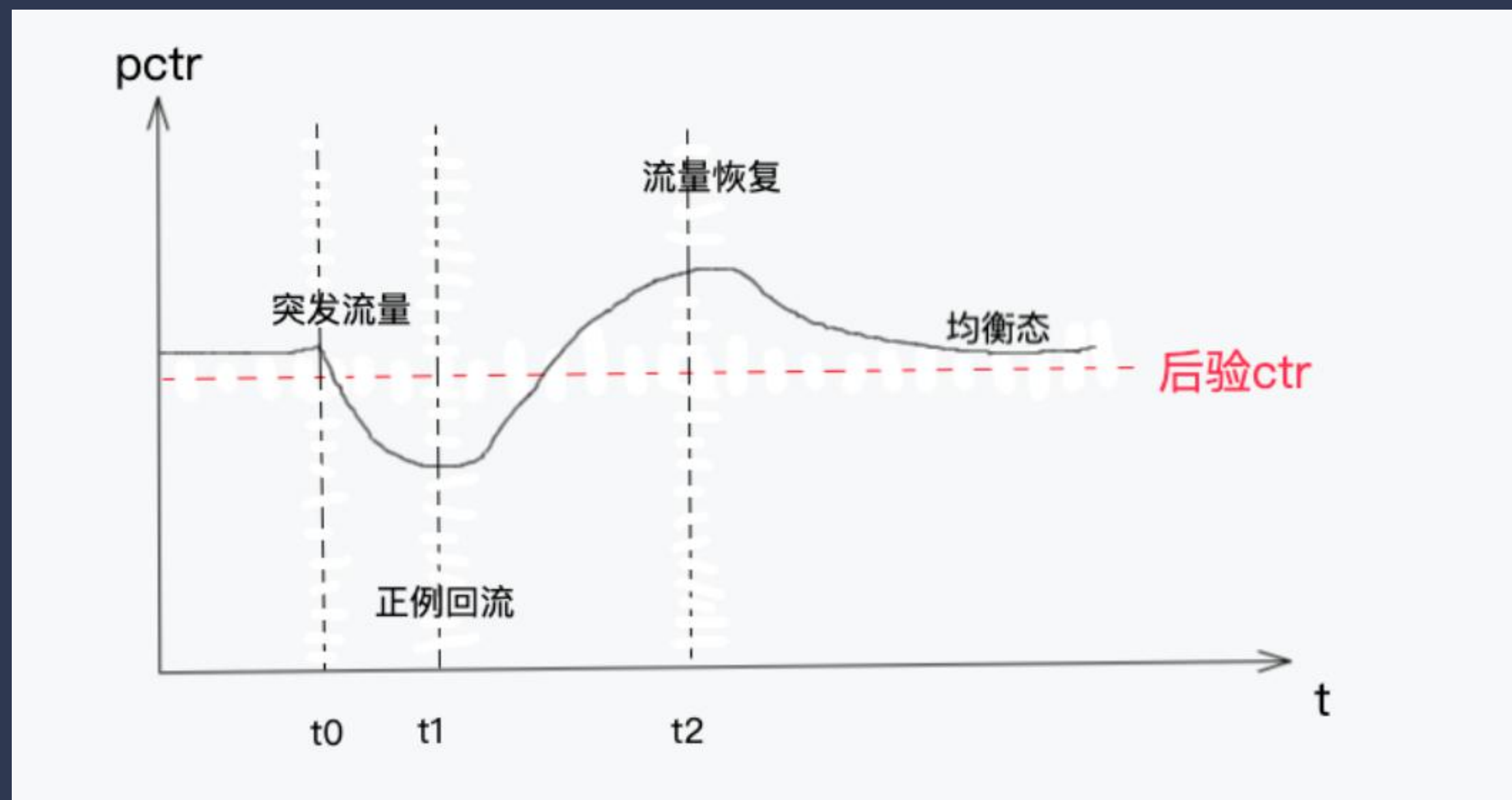
Wide & deep - Online experiment			
Loss function	Pooled RCE	RPMq	Monetized CTR
Log loss	7.68	100.00	100.00
PU loss	12.27	137.00	118.59
FN weighted	13.39	155.10	123.01
FN calibration	13.37	154.37	123.19

Table 4: Online results with the wide & deep model and the best performing loss functions (Twitter data). For RPMq and monetized CTR presented results correspond to relative improvements with respect to the log loss. Results for PU loss are prior to its divergence (within 2 days).

无等待窗口隐患

预估值波动

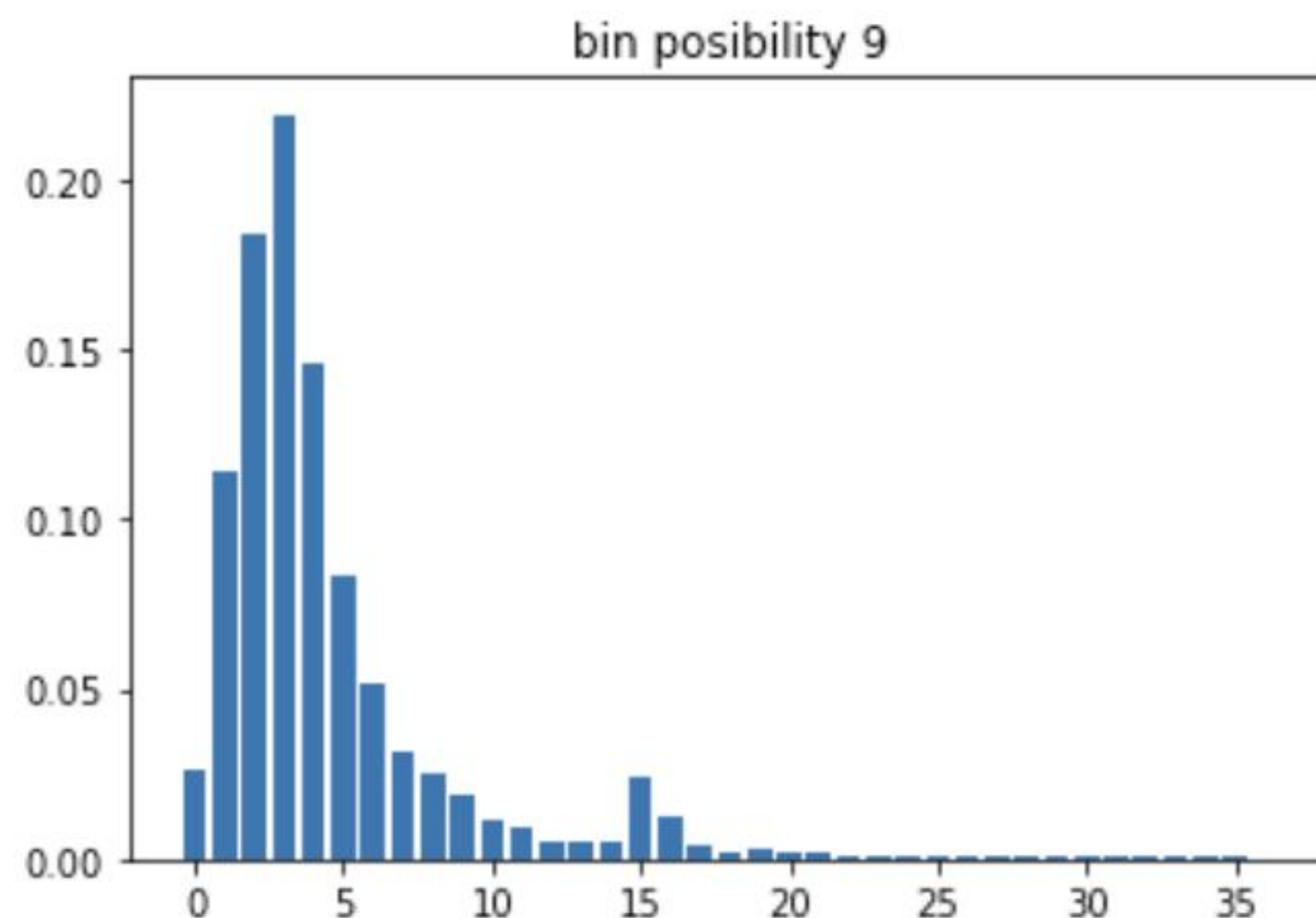
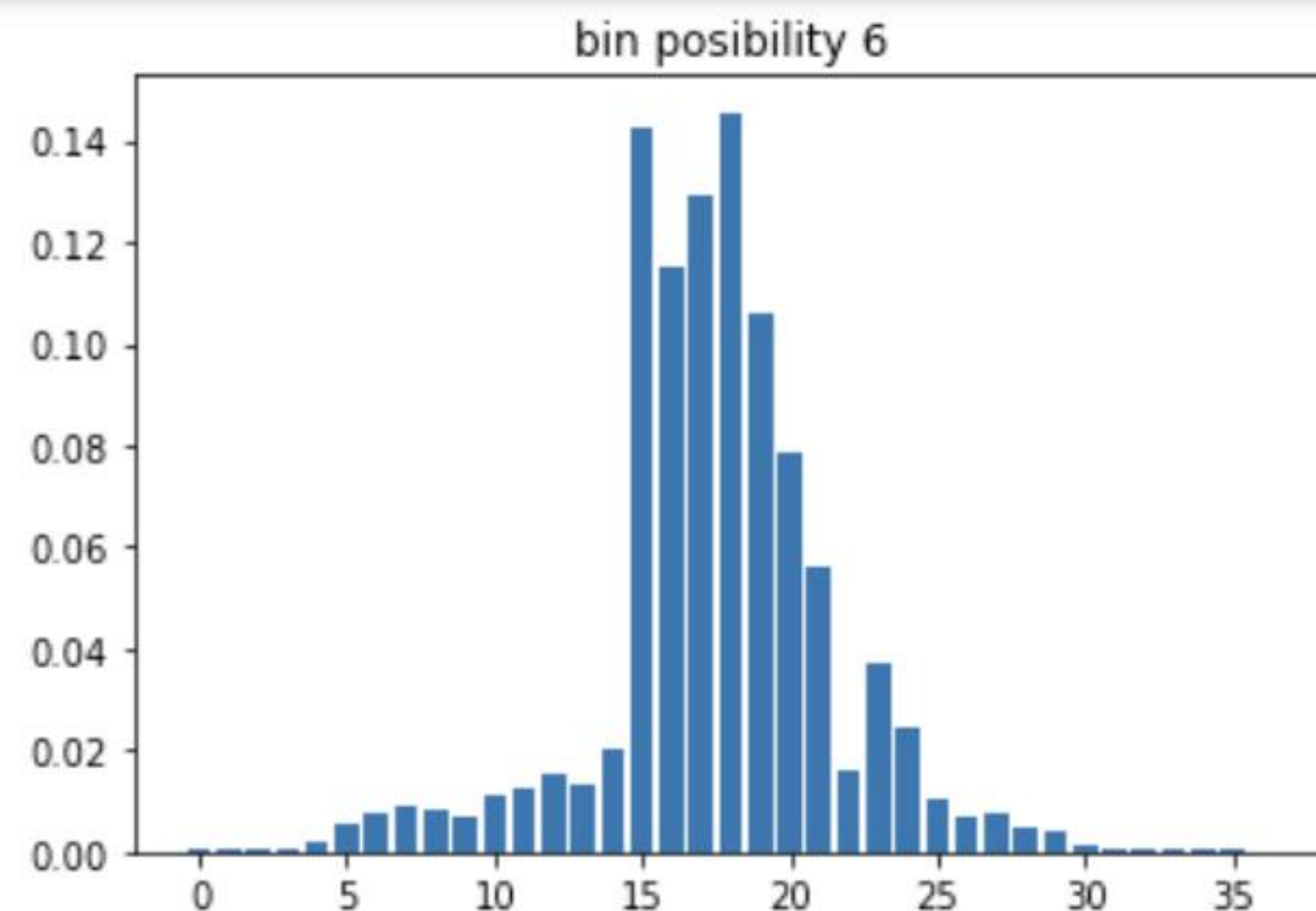
突发的流量和曝光，预估值在后验值上下波动



预估值波动解决方案

负例打散

负例样本产出后，缓存在joiner
之中，按照正例回流时间分布
PDF概率打散发送



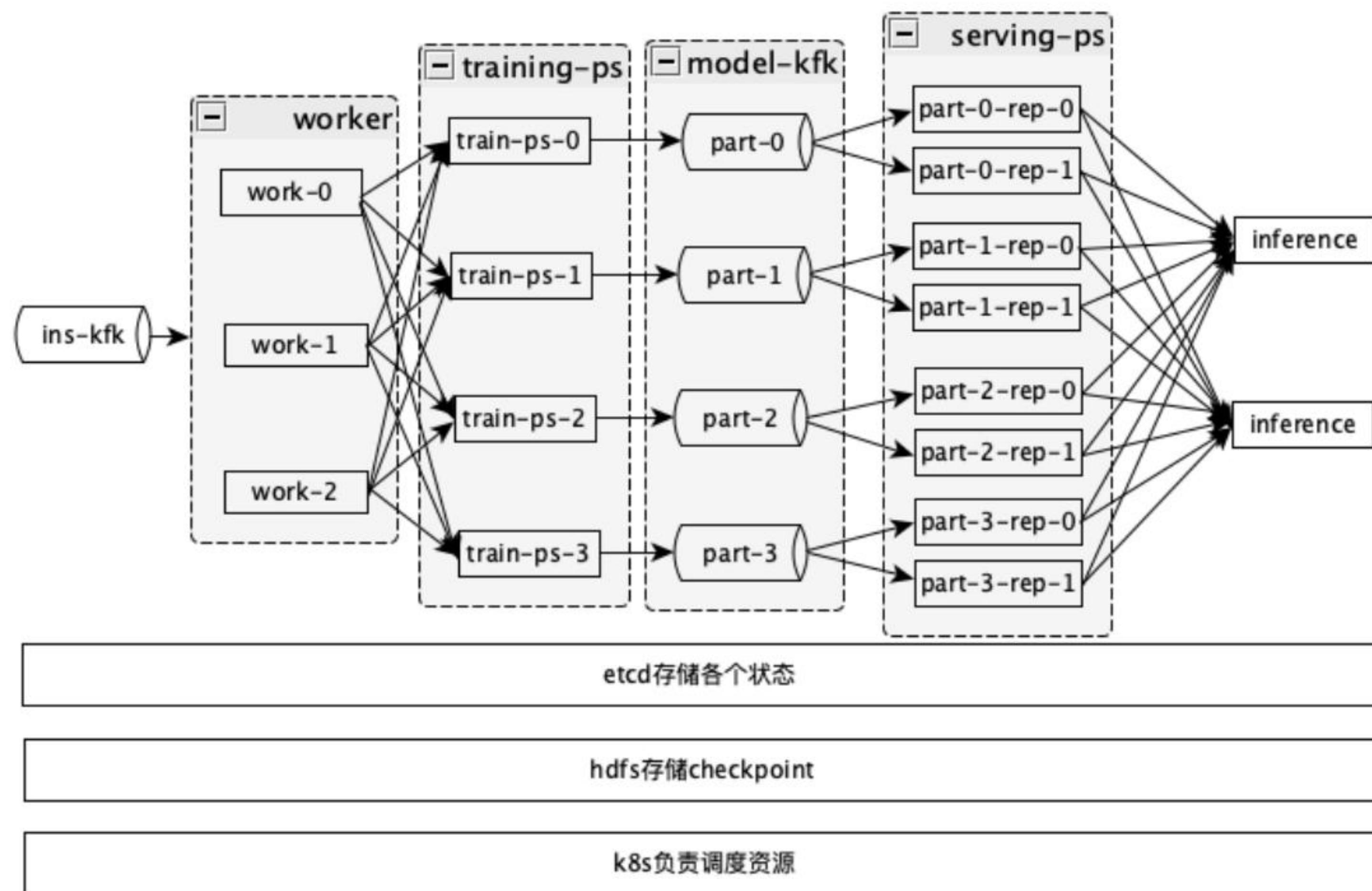
流式训练设计

设计目标

- 自研GPU同步训练，支持T级别的模型及推理
- 秒级流式更新，模型更新延时少于30S
- p99增加量小于15ms

流式训练设计

架构框图



系统开发与实践

工程问题

- 开发和设计的目标不一致
- 性能指标不达标
- 工期严重超时

解决办法

- 设计原理分享，理解后再开发
- 架构修改，换组件，读写分离等
- 后续制定方案时，留下更多冗余犯错空间

系统开发与实践

样本排查

- 事件和特征的join率排查
- 确定正负例样本数量与对齐率

AUC差距排查

- Reload收益：天级/流式各自的收益
- Shuffle收益：天级/流式样本各自shuffle后训练差距
- 不同纠偏方式，带来的auc收益

训练超参

- 学习率调小
- 去掉momenta

系统开发与实践

AUC评估

- 天级全天评估改batch内/间隔时间内评估
- 天级/流式样本有diff，选定相同的测试样本评估
- 离线评估实时性收益

编号	日期	超参	auc
0	天级模型测试 20220508	base组	auc:0.820882 pos:2775784, neg: 31892504
1	测试 20220508, 测试gap 1h	dense_lr: 0.00001 sparse_lr: 0.007 dense beta1:0	auc:0.809649 pos:2775784, neg: 31892504
2	测试 20220508测试gap 40min	dense_lr: 0.00001 sparse_lr: 0.007 dense beta1:0	auc:0.817282 pos:2775784, neg: 31892504
3	测试 20220508测试gap 30min	dense_lr: 0.00001 sparse_lr: 0.007 dense beta1:0	auc:0.822885 pos:2775784, neg: 31892504
3	测试 20220508测试gap 20min	dense_lr: 0.00001 sparse_lr: 0.007 dense beta1:0	auc:0.823649 pos:2775784, neg: 31892504
4	测试 20220508测试gap 10min	dense_lr: 0.00001 sparse_lr: 0.007 dense beta1:0	auc:0.826062 pos:2775784, neg: 31892504,

实验/上线

重点方向

- 流式模型常驻，模型稳定更新
- Inference时，emb更新策略
- 预估监控与报警
- 模型回滚策略
- 错误样本过滤，模型重训

总结

- 方案设计与代码实现——差距很大，见招拆招
- 流式样本对齐天级样本——繁琐零碎，耐心细致
- 流式模型AUC对齐与评估——控制变量，不放过任何diff
- 在线实验与模型上线——监控齐全，稳定可靠

Q&A

精彩继续！ 更多一线大厂前沿技术案例

📍 北京站

GIITC

全球大前端技术大会

时间：10月30-31日

地点：北京·国际会议中心

扫码查看大会
详情>>



📍 北京站

QCon

全球软件开发大会

时间：10月30-11月1日

地点：北京·国际会议中心

扫码查看大会
详情>>



📍 上海站

QCon

全球软件开发大会

时间：11月25-26日

地点：上海·宏安瑞士大酒店

扫码查看大会
详情>>



想一想，我该如何把这些
技术应用在工作实践中？

THANKS