

推荐-排序

姚凯飞

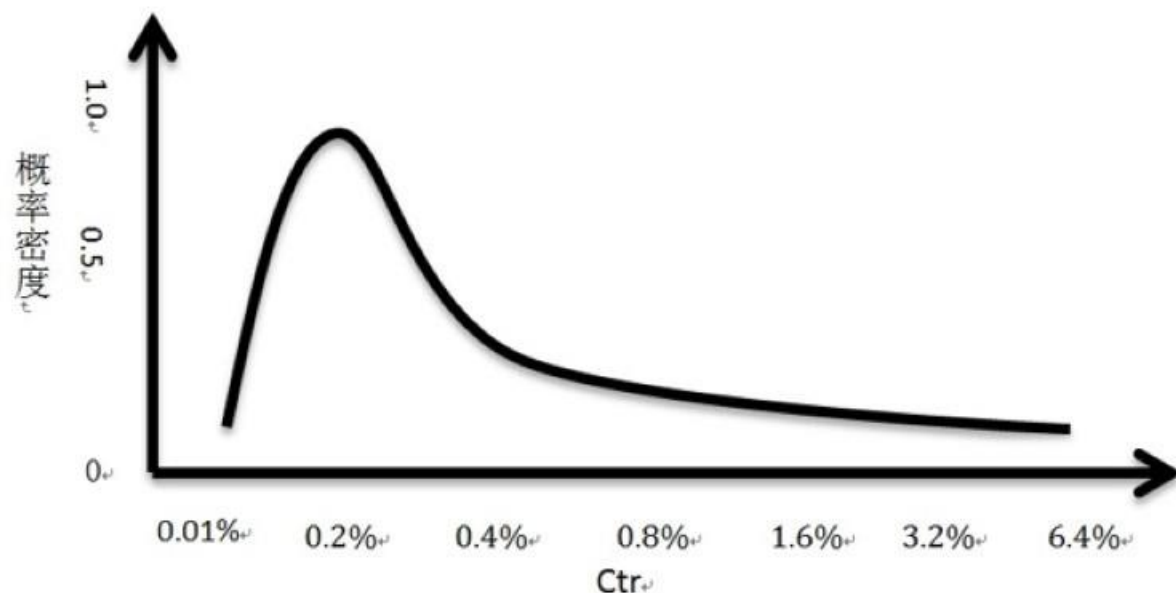
2017.10.22

什么是推荐排序

- 推荐系统的子模块
- 召回的后一阶段
- 以一个或多个指标为依据，进行打分

什么是推荐排序

对于一个 <user-商品> 展现后，该展现被用户“接受”的概率
通过历史用户的多维度多粒度行为特征去解释具有什么样特征的商品会被用户接受
个性化模型认为这些特征是造成商品被用户接受与否的差异



Machine learning goal: Estimate CTR for a given user/ad pair.

$$f(\text{User Image}, \text{Ad Image}) = \text{CTR}$$

\$50 off Your 1st
Bonobos
cim.meebo.com



"Best Mens Pants" - NY
Magazine. Get \$50 off
your first purchase of
Bonobos stylish men's
apparel.

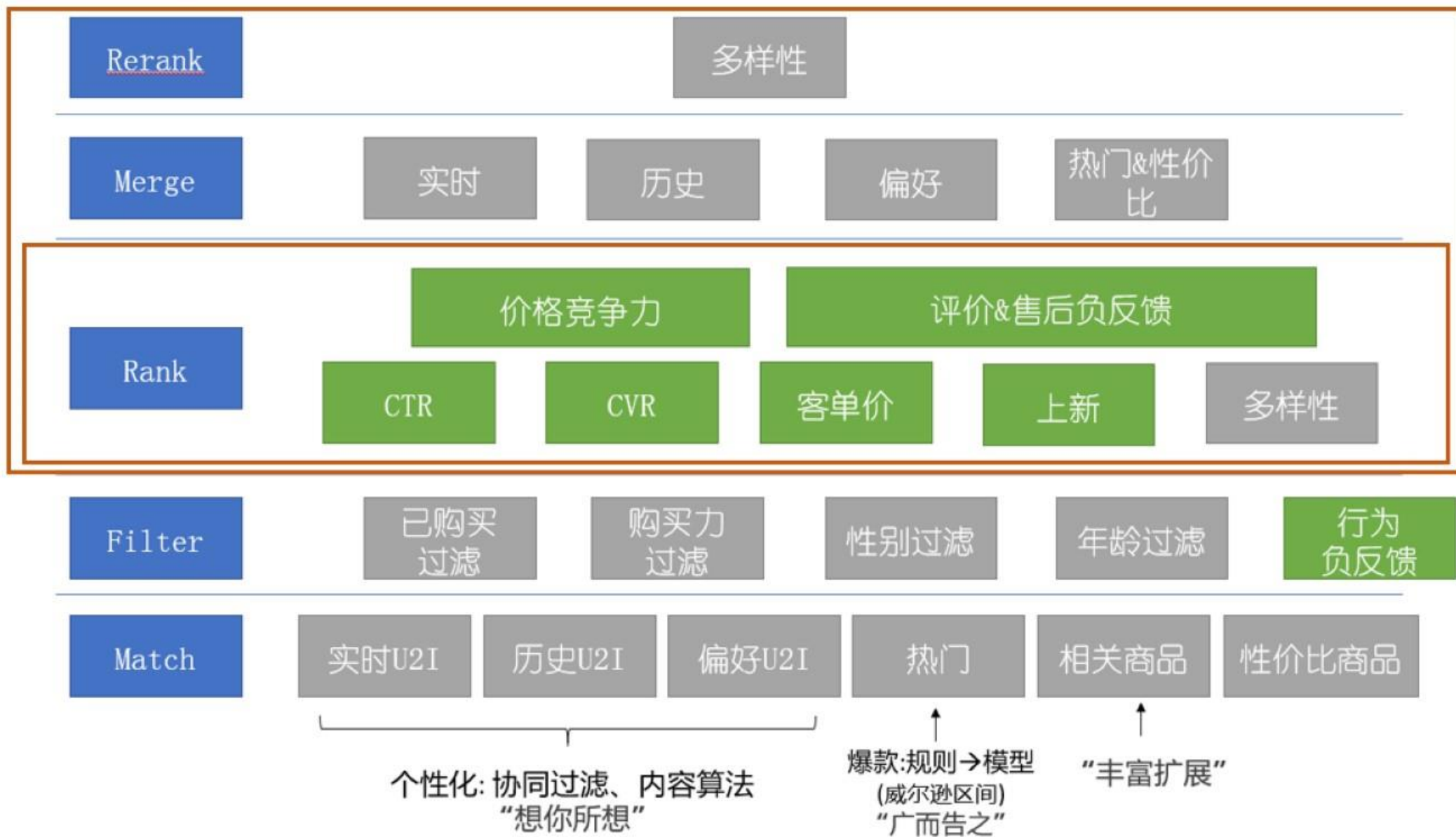
什么是推荐排序

- 用户体验好 → 被用户“接受”的概率高。
- 命题：成交的商品一定是用户接受的。
- 对<用户，商品>对不同的用户，商品被用户“接受”的概率是不同的：预测商品被用户接受的概率，将概率高的排在前面。

推荐排序的拆解

- 粗排与精排
- 单目标与多目标

推荐排序的拆解



排序阶段的实时化

特殊日期与日常

实时特征

离线构造线上实时特征样本

FTRL

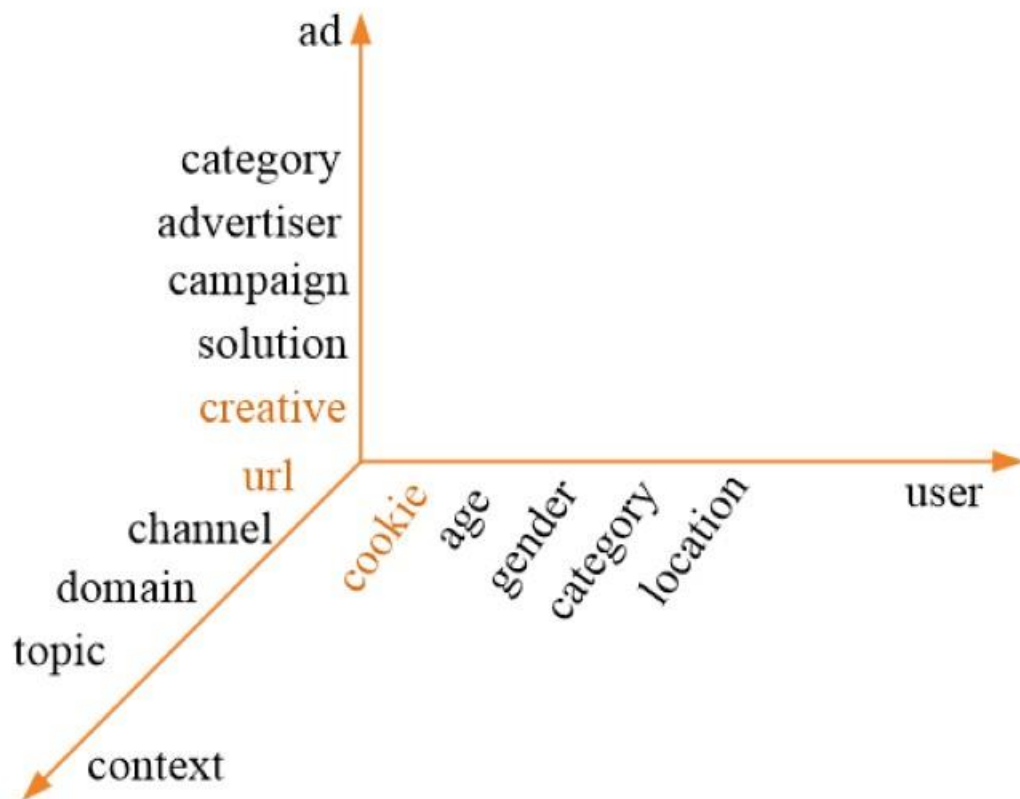
实时模型

Online Learning/Reinforce Learning

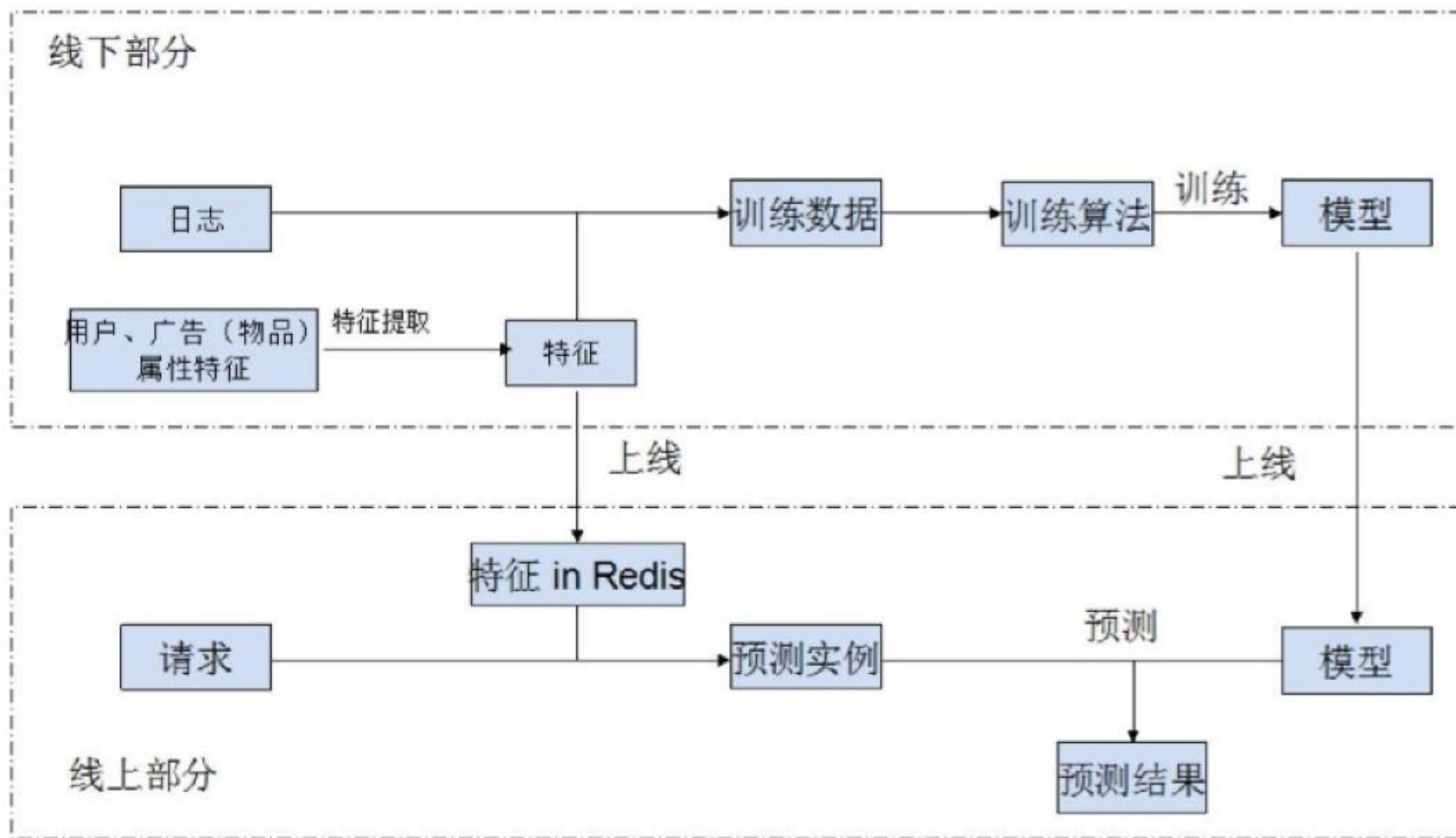
Q Learning

动态特征 - 多层次点击反馈

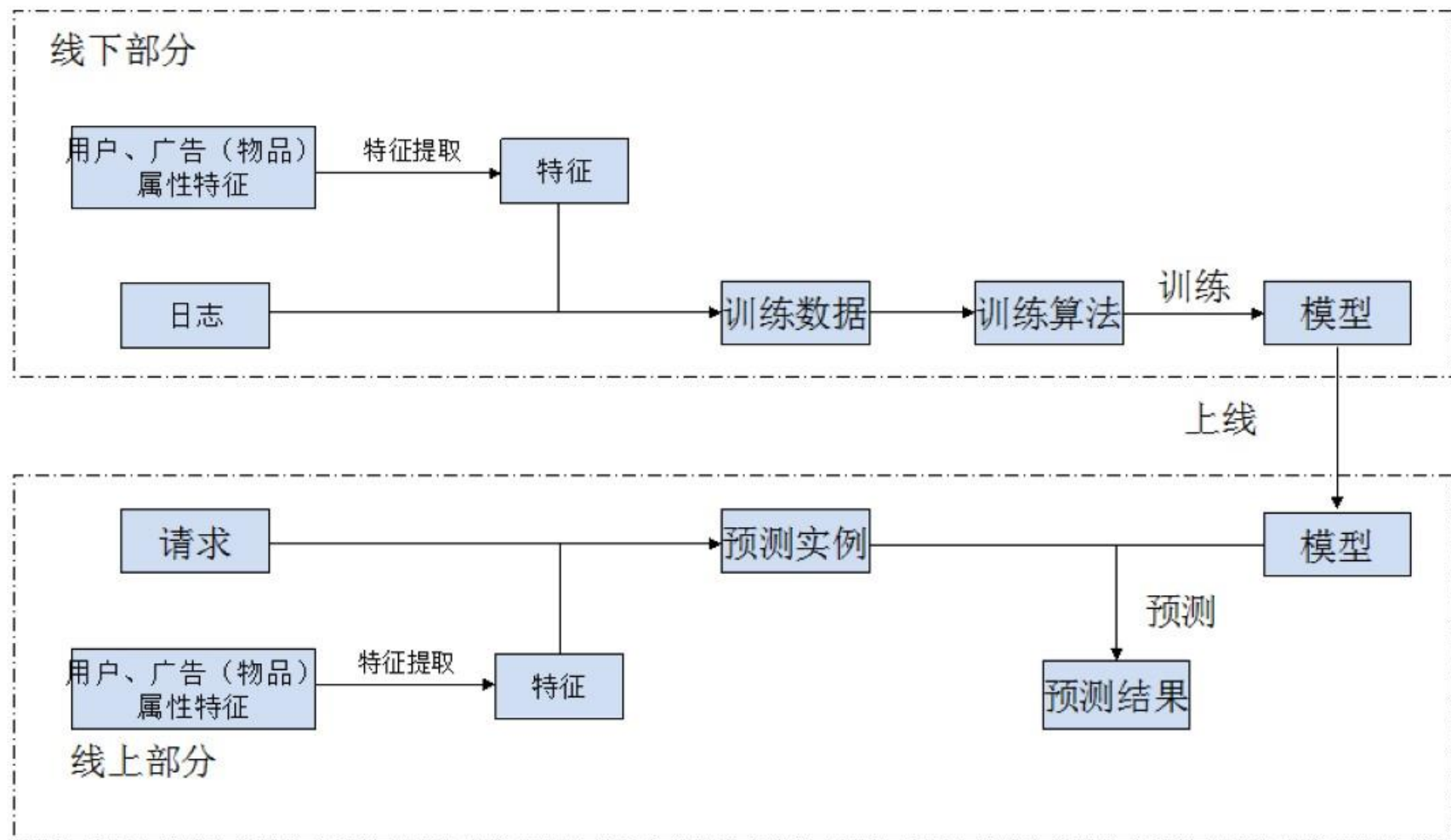
- 在标签组合维度上聚合点击反馈统计作为CTR预测的特征
- 优势:
 - 工程架构扩展性强(与在线学习相比)
 - 对新(a, u, c)组合有较强back-off能力
- 缺点:
 - 在线特征的存储量大, 更新要求高
- 组合维度举例:
 - $\text{cookie}(u)$ and $\text{creative}(a)$
 - $\text{gender}(u)$ and $\text{topic}(c)$
 - $\text{location}(u)$ and $\text{advertiser}(a)$
 - $\text{Category}(a)$ and $\text{category}(u)$
 - $\text{cookie}(u)$
 - $\text{creative}(a)$
 - $\text{gender}(u)$



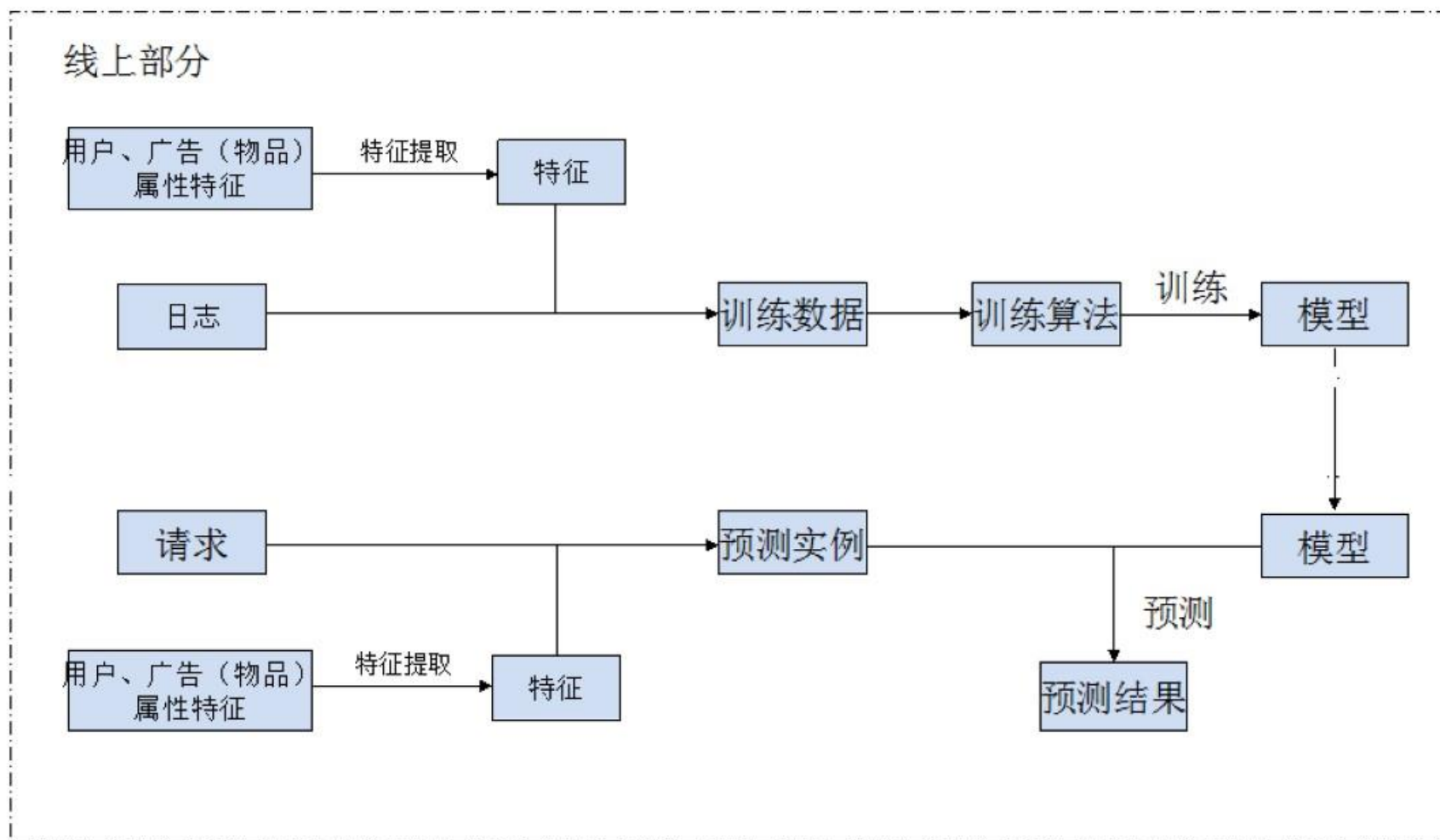
推荐排序的拆解



排序阶段的实时化



排序阶段的实时化



多模型级联

Offline GBDT

Online LR \rightarrow GBDT / GBDT \rightarrow LR

常见的排序算法

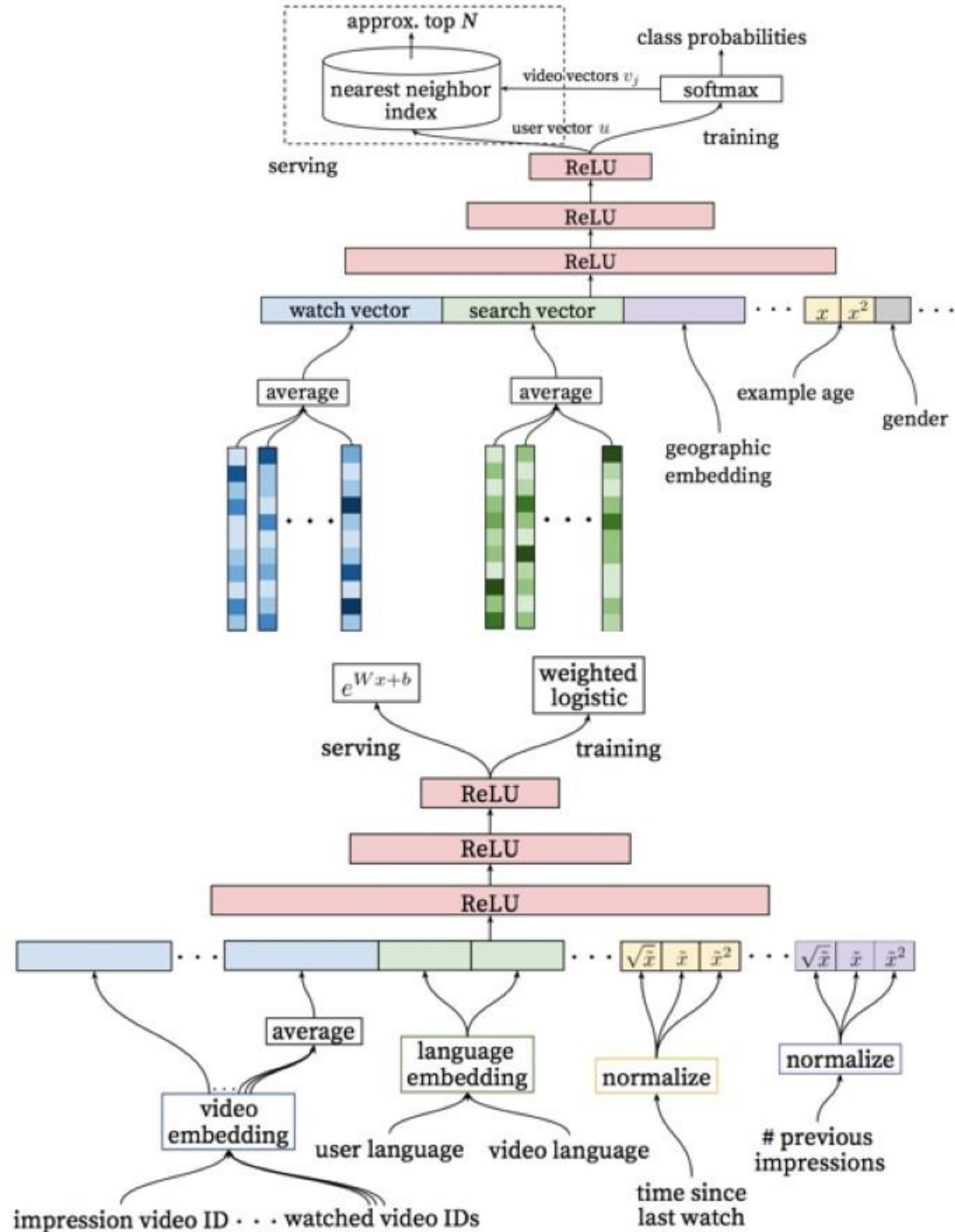
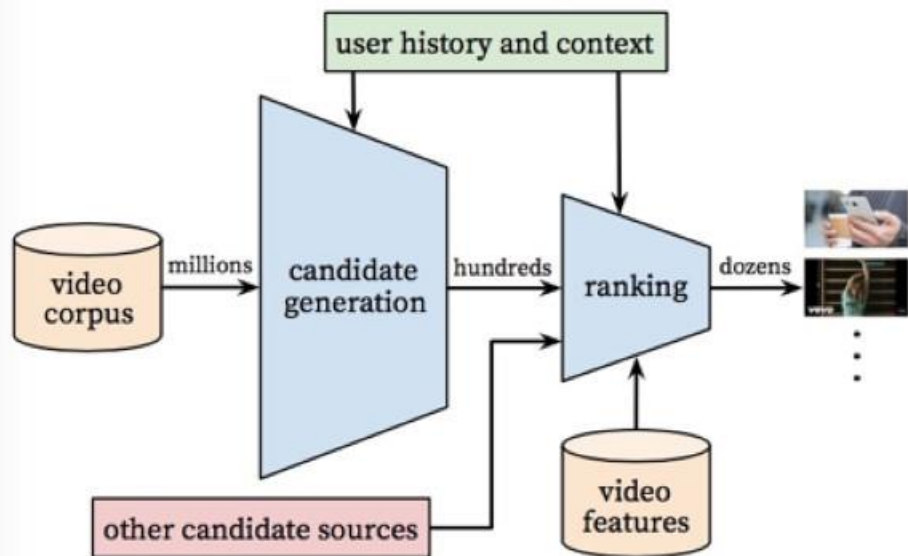
- 线性与非线性
- 宽与深

- LR
- GBDT
- DNN
- Wide&deep



- LSTM
- GRU

常见的排序算法



推荐排序与召回存在的意义

- 为何不全部rank，而需要match先召回？
- 为何不直接精排，而需要粗排？

- 效率
- 机器
- 成本
- 实时问题



• 粗选到精选



- 召回→排序
- 召回→粗排→精排

排序模型建模-日志清洗

- 黑名单
- 爬虫
- 非理性操作
- 有效曝光

排序模型建模-样本处理

- 本场景样本为主，其它场景为辅
- label随时间衰减
- 负样本采样
- 时间窗口选取，时间衰减加入

排序模型建模-样本处理

- 样本关联: 串联展示、点击、下单
- 样本选择: 时间窗口(近N月)、过滤(时长/访问次数限制)、样本迁移、skip above(点击模型 + 2)
- 样本采样: 访购模型: 减少负样本(非必要不做采样, 采样常常可能使实际数据分布发生变化, 但是如果数据太大无法训练或者正负比例严重失调(如超过100:1), 则需要采样解决)
- 样本权重: 支付 > 下单 > 点击
- 负样本: 热门但未被点击、主动删除的显示负反馈数据(高质量)

排序模型建模-位置偏倚

可见点击率=位置点击率+真实点击率

- 1.通过改变样本的分布,将每次点击归一化到首位上去,归一化的系数需要满足某个概率分布(可求得,如 COEC,E&E,Contextual Bandit)
- 2.将位置作为特征直接处理掉,训练时使用,预测时不使用

排序模型建模-特征工程

- Id特征
- 动态特征：偏好特征，统计特征
- 静态特征：属性特征

排序模型建模-特征工程

- 特征降维：特征维度高、稀疏、误差大
- 相似特征有相似的权重
- 特征的权重近似于后验概率



Low Level特征，High Level特征：

- Low Level 比较有针对性，单个特征覆盖面小（含有这个特征的数据不多），特征数量（维度）很大。
- High Level比较泛化，单个特征覆盖面大（含有这个特征的数据很多），特征数量（维度）不大。长尾样本的预测值主要受High Level特征影响。
- 高频样本的预测值主要受Low Level特征影响。
- 对于访购率问题，有大量的High Level或Low Level的特征

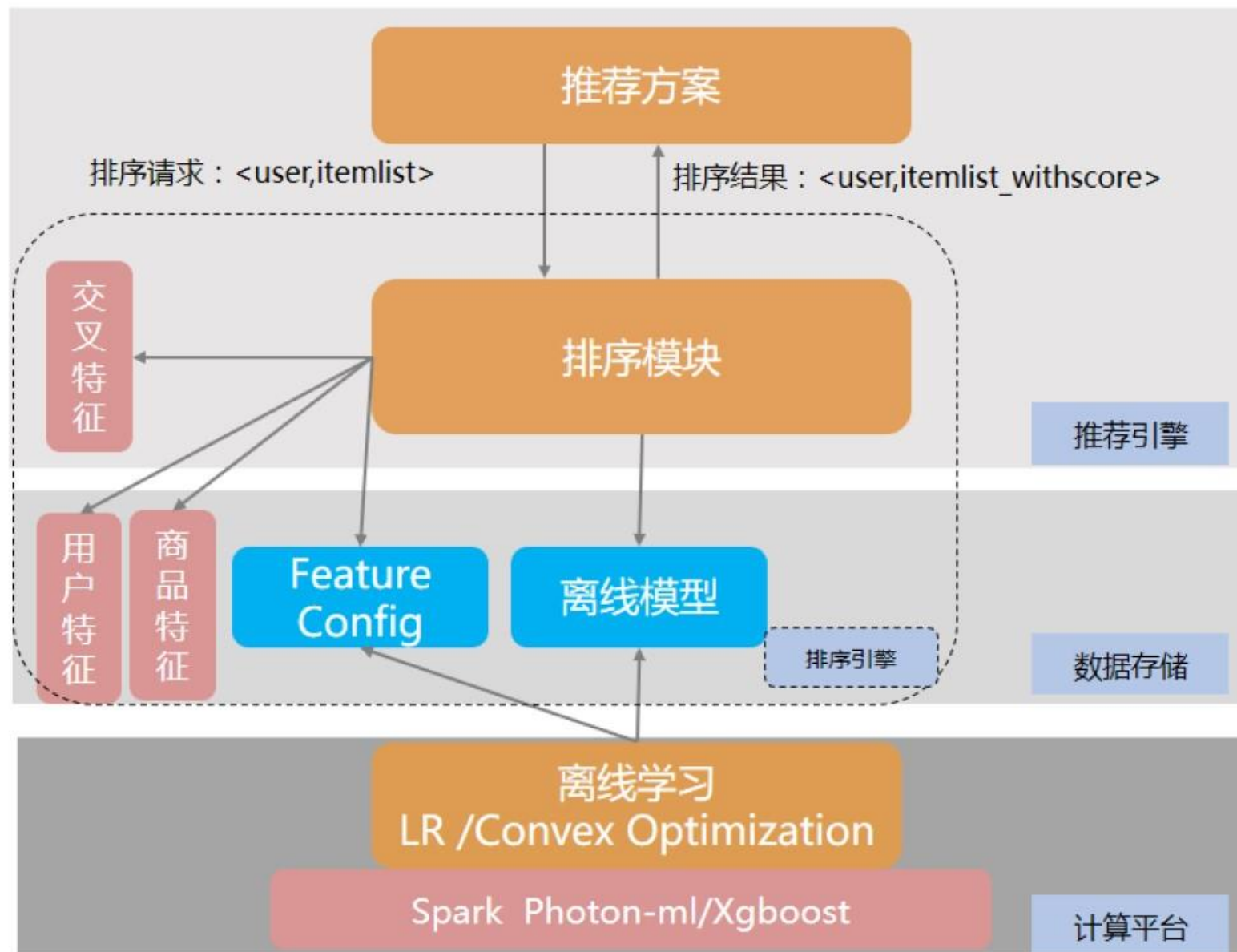
排序模型建模

点击率/转化率/观看时长预估

评估指标

- 线下
 - 特征有效性分析(相关系数、卡方检验、平均互信息、条件熵、后验概率、逻辑回归权重)
 - 特征权重分析
 - AUC (area under curve)
 - NDCG (Normalized Discounted Cumulative Gain)
 - 分query AUC
 - 分用户+类目AUC
- 线上：A/B test
 - 成交转化率，CTR业务指标
 - Case 查看

预估部署



排序模型建模-Rerank

- 从用户点击到购买整个流程来考虑
- Ctr Cvr Price
- Ctr Ts
- Match Score
- Trigger score time decay
- Match type ctr cvr
- Match type score
- 多样性
- 惊喜性
- 新颖性
- 准确率
- 召回率

排序模型功效

- 算法角度考虑，宝贝池越大，推荐可发挥的余地越多，会带来更好的效果

召回对排序影响

- match来源/相似分作为排序特征
- 漏斗的顶部好坏对漏斗底部的影响

精细化的排序

- 少量核心特征粗排
- 多过程建模：如直接购买还是收藏比价使用模型判断
- 粗粒度召回:离线产出不同类目下的精品数据
- 融合价格/人气分/质量分/兴趣分等
- E&E探索解决低置信度预估模型问题

工程的贡献

- 算法工具
- 离线训练速度与模型更新速度

推荐系统-打分-架构

