淘宝定向广告技术简介



定向广告算法负责人: 王勇睿





广告业务简介

后台系统





前台展示



- 几十个站内位置
- 无数站外位置:淘宝推广



定义



- 利用淘宝庞大的数据库,通过创新的多维度人群定向技术,锁定目标客户,将广告主推广的宝贝展现在目标客户 浏览的网页上。
- 通过用户的行为,构建用户模型,将用户兴趣和需要映射 到广告分类体系,将类目宝贝主动推荐给用户,实现精准 营销。

广告系统: 多方博弈



数据量



- 10亿商品
- 千万级别的广告
- 实时竞价数据
- N亿站内广告位流量
- 10亿级以上的站外广告位流量
- 用户所有行为数据:浏览、收藏、查询、Listing、购买、 长期行为、短期行为.....Tb级别

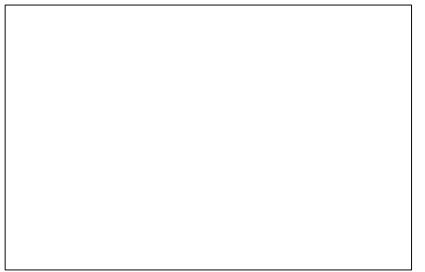


广告算法简介

主算法逻辑——两段式



• 两段式算法 vs 一段式算法?



离线部分是重头戏

www.etao.com

- 广告信息分析、聚类
- 广告点击率、点击转化率预估......
- 广告主信息分析
- 浏览者历史数据分析,兴趣预测.....

•

方向划分

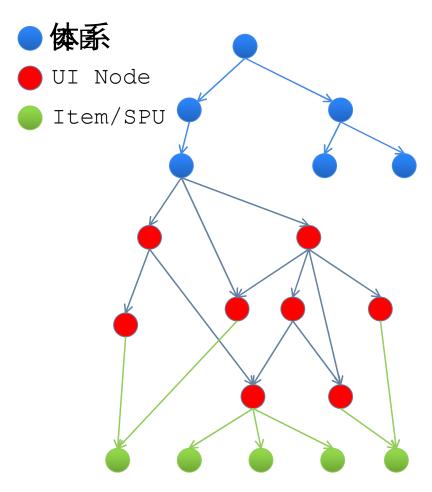


- 基础数据
 - 广告聚类
 - 用户聚类
- 相关性 (Match)
 - 用户兴趣模型
 - 推荐方法
- 排序 (Rank)
 - 大规模CTR预估
 - CVR**预估**
 - 展现概率

基础数据—广告聚类



• UIT (User Intension Tree): 全新的淘宝宝贝分类



- •淘宝类目体系的优缺点
- ·UI Node的必要性
- •Spu (Standard Product
 Unit)

UIT效果

www.etao.com

- 全部流量CTR平均提升37%
- 浏览者意图精准度大大提升
- 广告主覆盖度大大提升
- 广告流量分配更加合理
- 更多的行为模式挖掘

相关性 (Match) — 用户兴趣模型



- User → <UI Node>: 综合模型
 - 长期历史行为: 行为类型、item信息、时间、User Profile......
 - 混合逻辑回归模型
 - 训练数据的季节性问题
 - MPI
 - 短期行为 (RCC)融合
 - 购买降权、用户疲劳.....

相关性 (Match) —推荐方法

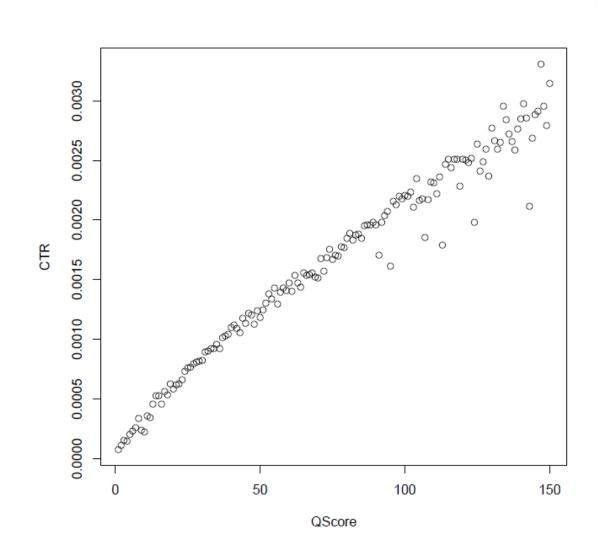


- User → <UI Node>: 特定模型
 - 看了又看
 - 买了又买
 - 稀疏问题:广告聚类、用户聚类.....
 - 关联推荐、协同过滤
 - item → <item>
 - item → <UI Node> → <item>
 - item \rightarrow <UI Node> \rightarrow <UI Node> \rightarrow <item>

排序 (Rank) —大规模CTR预估



- <user, ad>粒度
- 特征设计
 - 行为
 - 用户
 - 文本
 - 反馈

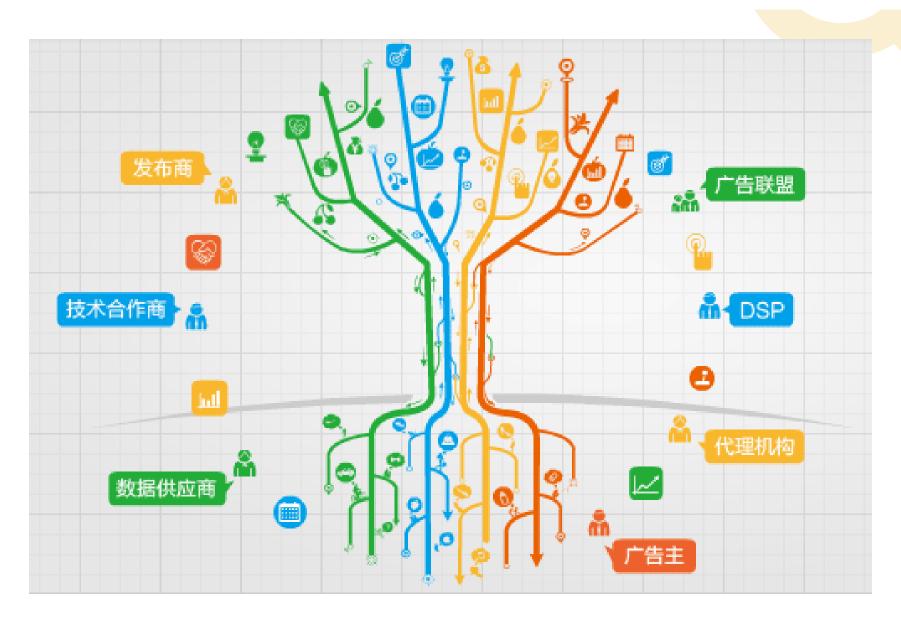




定向技术在RTB领域的应用

TANX (tanx.com)







• 角色:

- AdX (Ad eXchange): network $\grave{\boldsymbol{z}}$ network
- RTB: RealTime Bidding
- DSP: Demand/Buy Side Platform
- SSP: Supply/Sell Side Platform

• 淘宝DSP:

- 単品
- 图文

• 技术:

- 流量成分维度图
- 报价算法 (forecasting): 建立在true bidding
- CPC跟CPM的转换风险

定向技术被高估?



- <How Effective is Targeted Advertising>
- 没有定向技术体系,就没有定向广告市场的繁荣

• 谢谢 ☺