

智能推荐技术--案例分析: 天猫

潘微科

价值

- 2013年11月11日: **推荐引导的成交量**约占天猫总体的30%
- 手机端个性化会场: 如果没有个性化就很难展示(因为**屏幕大小受限**)
- ...

数据

- 用户行为
 - 点击,浏览
 - 购买
 - 收藏
 - 加入购物车
 - 关注
 - 时间
- 文本(描述信息)
- 基础推荐实体关系库: 商品关系库,品牌关系库,专辑关系库,活动关系库
- 天猫导购路径结构

张奇 (阿里巴巴天猫推荐算法团队). 个性化的天猫. 2014年4月20日.

张奇 (阿里巴巴天猫推荐算法团队). 天猫推荐业务与算法架构. 2013年7月14日.

评估维度

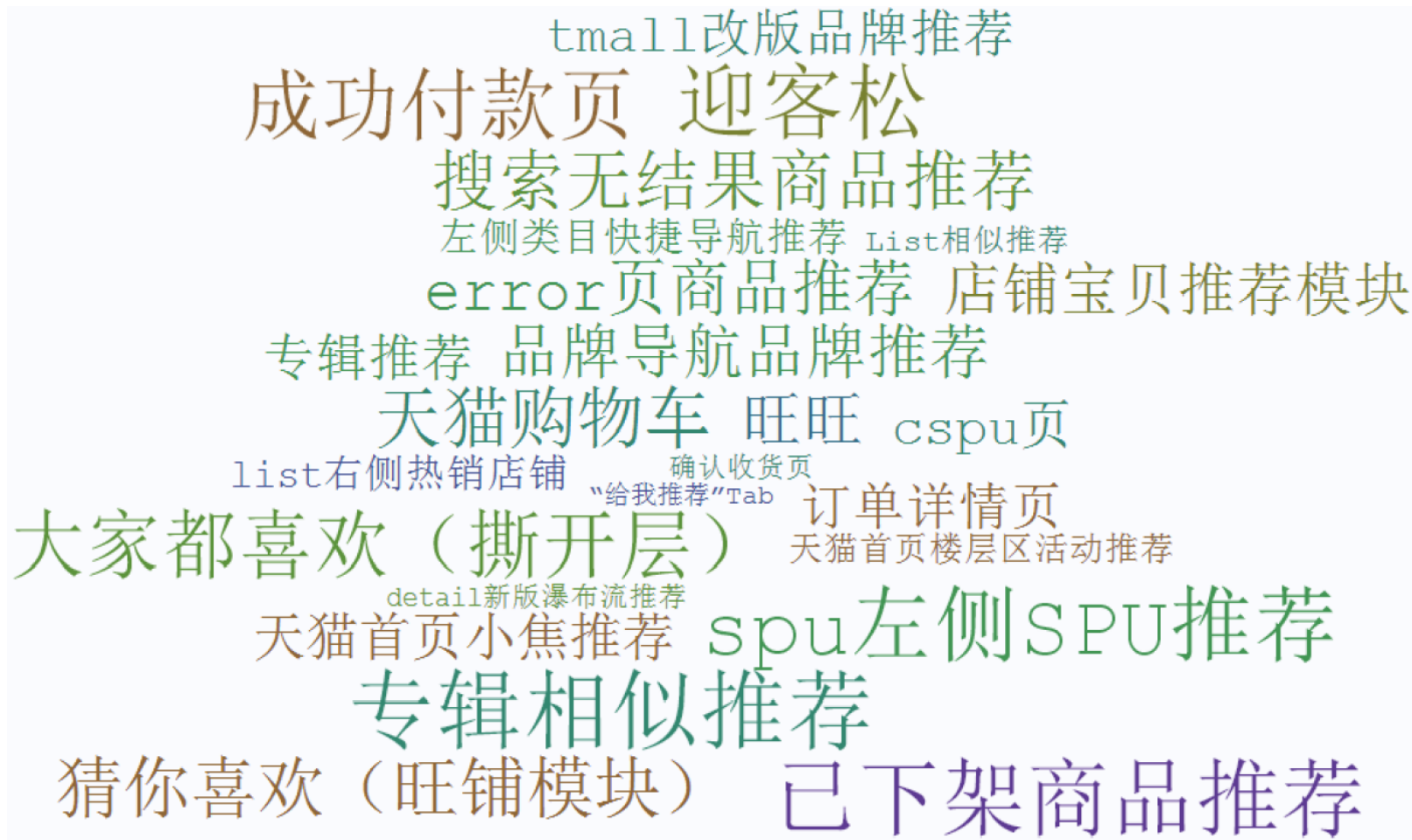
- 用户: UV (unique visitor)购买转化率, 人均点击商品数
- 卖家: 让流量的分配更公平
- 天猫: 通过推荐帮天猫带来多少成交量

天猫推荐算法团队的那些事儿. 2014年4月1日.

<http://www.infoq.com/cn/news/2014/04/tmall-recommendation-team/>

张奇 (阿里巴巴天猫推荐算法团队). 天猫推荐业务与算法架构. 2013年7月14日.

应用(PC端)



算法

- 任务: 作为一个CTR (Click-Through Rate)预估的问题
- 主要算法
 - Learning to rank (pointwise: LR, MLR, Random Forest, GBDT; pairwise: RankSVM), LDA, NMF
 - Online learning
 - Image processing and pattern recognition

张奇 (阿里巴巴天猫推荐算法团队). 个性化的天猫. 2014年4月20日.

张奇 (阿里巴巴天猫推荐算法团队). 天猫推荐业务与算法架构. 2013年7月14日.

技术(1/3)

- Feature engineering (特征工程)
- Fusion (模型融合)

张奇 (阿里巴巴天猫推荐算法团队). 个性化的天猫. 2014年4月20日.

张奇 (阿里巴巴天猫推荐算法团队). 天猫推荐业务与算法架构. 2013年7月14日.

技术(2/3)

- 场景引擎
 - 长期兴趣
 - 用户实时意图(intent)计算
 - 品牌意图(brand),类目意图(category),单品意图(product)
 - 逻辑层: 召回与重排序分离
 - 数据层: 数据与业务分离(通用性高,可支持多个业务,开发成本低)
- (注: 对比腾讯精准推荐中心的方式)

张奇 (阿里巴巴天猫推荐算法团队). 个性化的天猫. 2014年4月20日.

张奇 (阿里巴巴天猫推荐算法团队). 天猫推荐业务与算法架构. 2013年7月14日.

技术(3/3)

- 双11挑战
 - 系统: 流量暴增带来的系统压力
 - 非个性化推荐
 - 降低系统依赖
 - 算法数据本地Cache
 - 算法: 用户行为模式特征, 实时计算转离线计算, 效果压力

策略

- 服装: 基于风格、搭配来推荐
- 书籍: 基于话题、类别、知识水平等来推荐
- 音乐 ...

数据平台/系统架构

- 检索/索引(retrieval/indexing) -> 召回 -> 重排序 (CTR预估) -> 展现
- 计算框架: Hadoop/ODPS集群, MPI集群, Spark集群, MPI
- 支持: 在线实验, 离线实验
- 支撑非常规访问(流量暴增): 2013年11月11日, Detail访问次数**16亿次/当天**, 峰值吞吐率**6.9万次/秒**, 当天点击推荐产品的用户数**36M**

张奇 (阿里巴巴天猫推荐算法团队). 个性化的天猫. 2014年4月20日.

张奇 (阿里巴巴天猫推荐算法团队). 天猫推荐业务与算法架构. 2013年7月14日.

参考文献

- 张奇 (阿里巴巴天猫技术部算法组). 天猫推荐业务与算法架构. 2013年7月14日.
- 张奇 (阿里巴巴天猫技术部算法组). 个性化的双11-天猫推荐技术应用. 2013年11月23日.
- 张奇 (阿里巴巴天猫技术部算法组). 个性化的天猫. 2014年4月20日.
- 天猫推荐算法团队的那些事儿. 2014年4月1日.
<http://www.infoq.com/cn/news/2014/04/tmall-recommendation-team/>