

# ALS推荐的应用

赵海臣

# 背景

- ▶ ALS应用于
  - 对活跃用户进行商品推荐
  - 生成用户标签
    - 品牌喜好度
    - 类别喜好度

# 主要流程

- ▶ 参考codes中mainFunc.py
- ▶ 流程
  - 数据清洗
    - 活跃用户筛选: 只推荐活跃用户, 原因:
      - 1.活跃用户的推荐结果精确度高, 有推荐价值;
      - 2.构建更稠密矩阵, 有利于提高整体精确度
        - userId
    - 用户浏览数据
      - userId, productId, Time
    - 用户订单数据
      - userId, productId, Time
    - 用户购物车数据
      - userId, productId, Time
    - 用户收藏夹数据
      - userId, productId, Time

# 主要流程

- 召回商品数据：只推荐top5000热销商品，原因：
  - 1.Top5000商品的推荐结果精确度高，有推荐价值；
  - 2.构建更稠密矩阵，有利于提高整体精确度；
  - 3.top5000热销商品占据销量总量87%以上
    - productId
- \*Top热销占总销量的比例
  - # Recall products threshold
  - # Facts(Mar,16th,2017):
  - # <= 10000 - 93.37%
  - # <= 7000 - 90.52%
  - # <= 5000 - 87.00%
  - # <= 3000 - 80.18%
  - # <= 1000 - 62.19%
  - # <= 100 - 27.19%

# 主要流程

- 筛选活跃用户数据
  - 用户浏览数据
    - userId, productId, Time
  - 用户订单数据
    - userId, productId, Time
  - 用户购物车数据
    - userId, productId, Time
  - 用户收藏夹数据
    - userId, productId, Time

# 主要流程

- 聚合用户-商品评分

- 用户-商品评分=所有行为记录权值累加

- 用户-商品1次浏览记录 =  $1 * \text{浏览记录权重} * \text{时间衰减因素}$
- 用户-商品1次订单记录 =  $1 * \text{订单记录权重} * \text{时间衰减因素}$
- 用户-商品1次购物车记录 =  $1 * \text{购物车记录权重} * \text{时间衰减因素}$
- 用户-商品1次收藏夹记录 =  $1 * \text{收藏夹记录权重} * \text{时间衰减因素}$

- \*用户的行为分值随时间衰减，时间衰减方程 =  $0.3 + e^{(-\omega * |dateDiff|)}$

# 主要流程

- ALS训练
- 复购商品过滤
  - 对用户最近15天购买过的商品进行过滤，防止重复推荐
- 将推荐列表，每个用户截取分数最高的前100个商品，生成CSV文件

# 标签推荐

- ▶ 将流程中的productId换成对应的brandId, category\_v3\_3, 用户推荐的得分即用户对不同标签的喜好得分。



# ALS商品推荐AB测试

- ▶ 聚美优品APP购物车下“今日推荐”，AB测试结果，推荐组为ALS推荐，对照组为TOP100热销爆款



# ALS商品推荐AB测试

- 聚美优品APP购物车下“今日推荐”，AB测试结果，推荐组为ALS推荐，对照组为TOP100热销爆款



# 详细报告

## ▶ 邮件“首页猜你喜欢”

测试汇总数据如下：

数据类型	数据说明	推荐组 (A)	对照组(B)	差值比 (A/B-1)	结论备注
入口数据	入口曝光UV	1479203	1481406	-0.1%	入口流量基本均衡，差距在误差范围内
	入口点击UV	19334	19334	0.0%	
	入口用户CTR	1.31%	1.31%	0.1%	
流量数据	素材曝光PV	528073	480018	10.0%	流量数据有了比较大幅度的提升
	素材点击UV	6863	4677	46.7%	
	素材点击PV	17234	8007	115.2%	
销售数据	销售额	16099.7	6270.4	156.8%	订单数据有了比较明显的提升
	订单量	174	76	128.9%	
核心数据	素材点击PV/入口点击UV	89.1%	41.4%	115.2%	平均每用户点击比例大幅提升
	素材点击UV/入口点击UV	35.5%	24.2%	46.7%	发生点击的用户比例增加
	素材曝光PV/入口点击UV	27.31	24.83	10.0%	曝光素材数增加，用户阅读深度提升
	eCPM	832.7	324.3	156.8%	每UV价值大幅度提升
	销售额/订单量	92.5	82.5	12.1%	用户每次消费消费额有部分提升

The end