

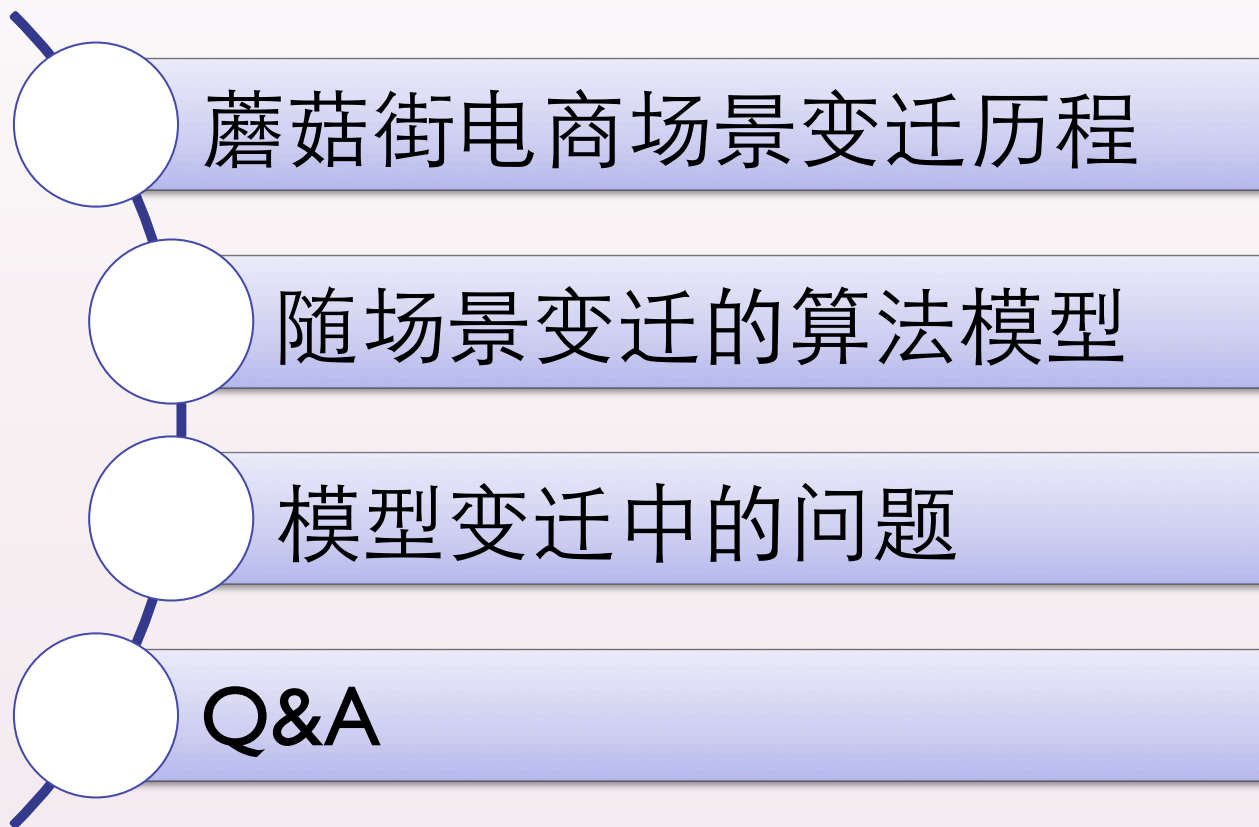
# 蘑菇街电商算法演化

从爆款模型到个性化模型

# 个人简介

- 2010-2015 百度凤巢
- 2015-今 蘑菇街
- 主要工作:
  - 数据挖掘

# 蘑菇街电商算法演化



# 蘑菇街电商算法演化



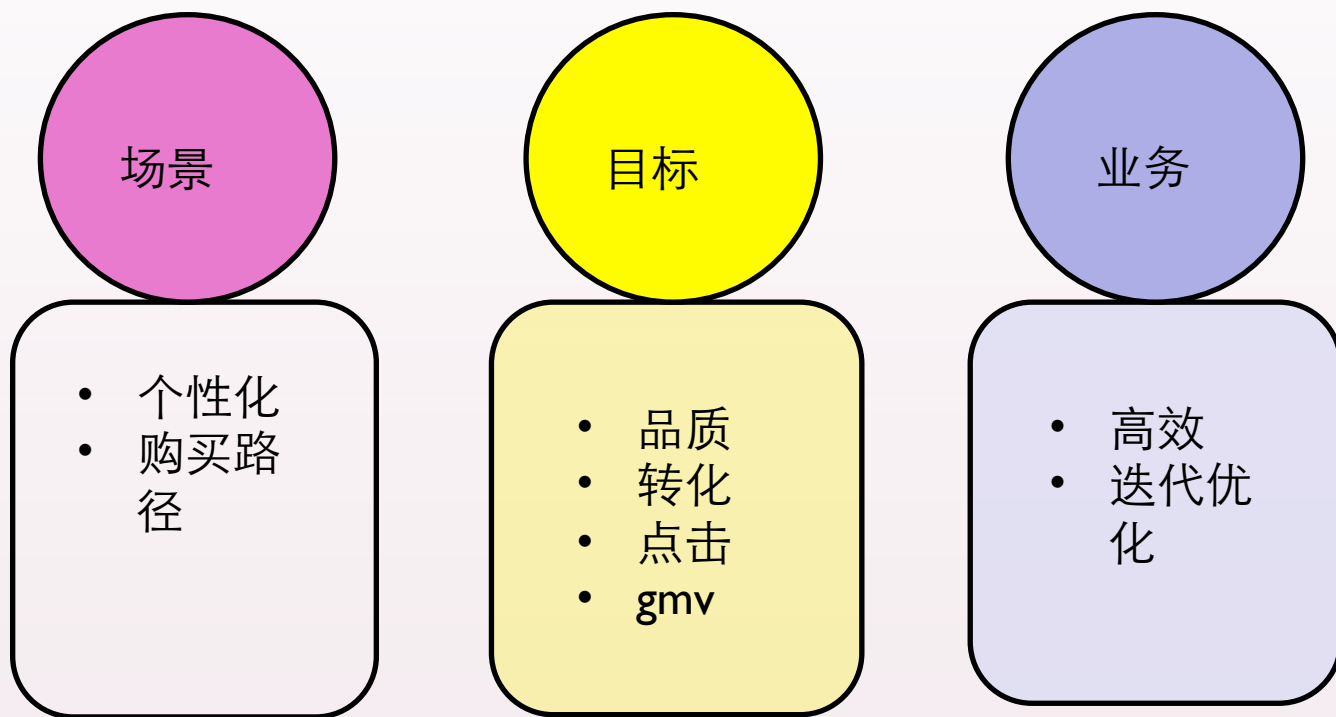
蘑菇街电商场景变迁历程

随场景变迁的算法模型

模型变迁中的问题

Q&A

# 蘑菇街电商场景变迁



# 蘑菇街电商算法演化



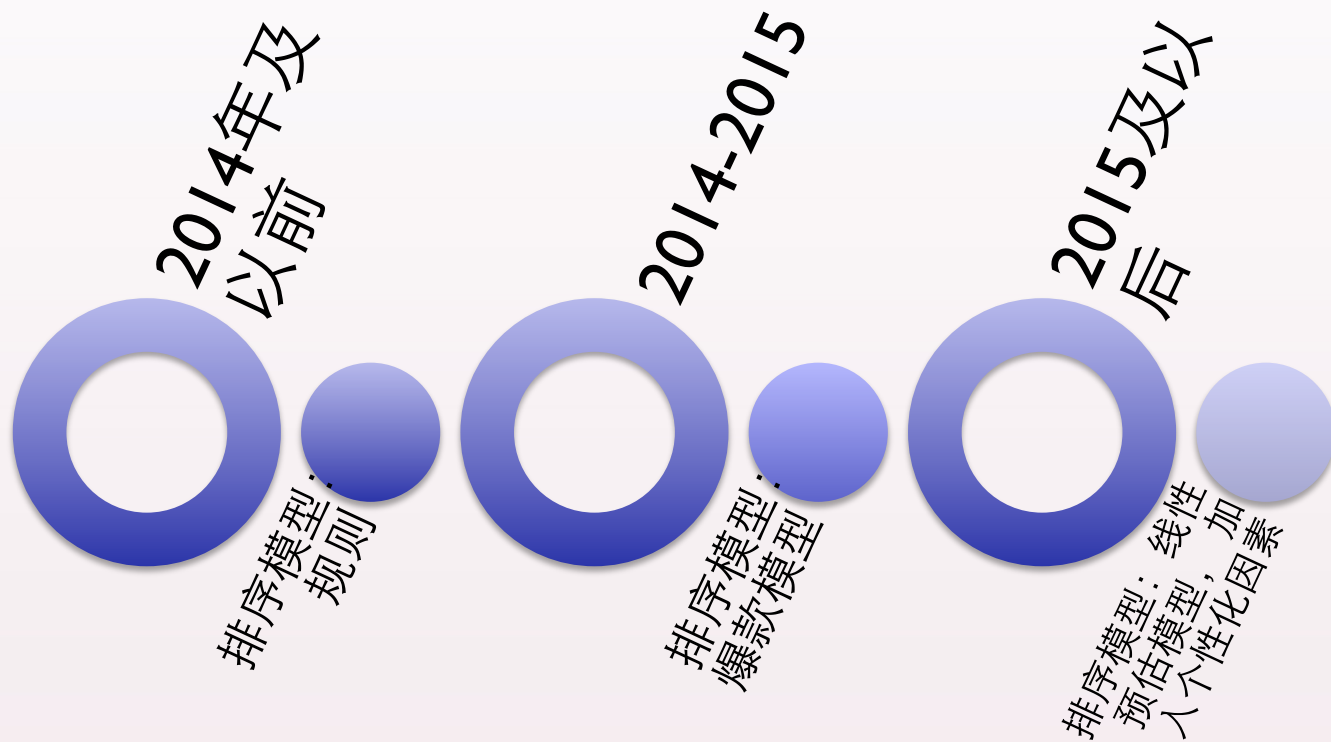
蘑菇街电商场景变迁历程

随场景变迁的算法模型

模型变迁中的问题

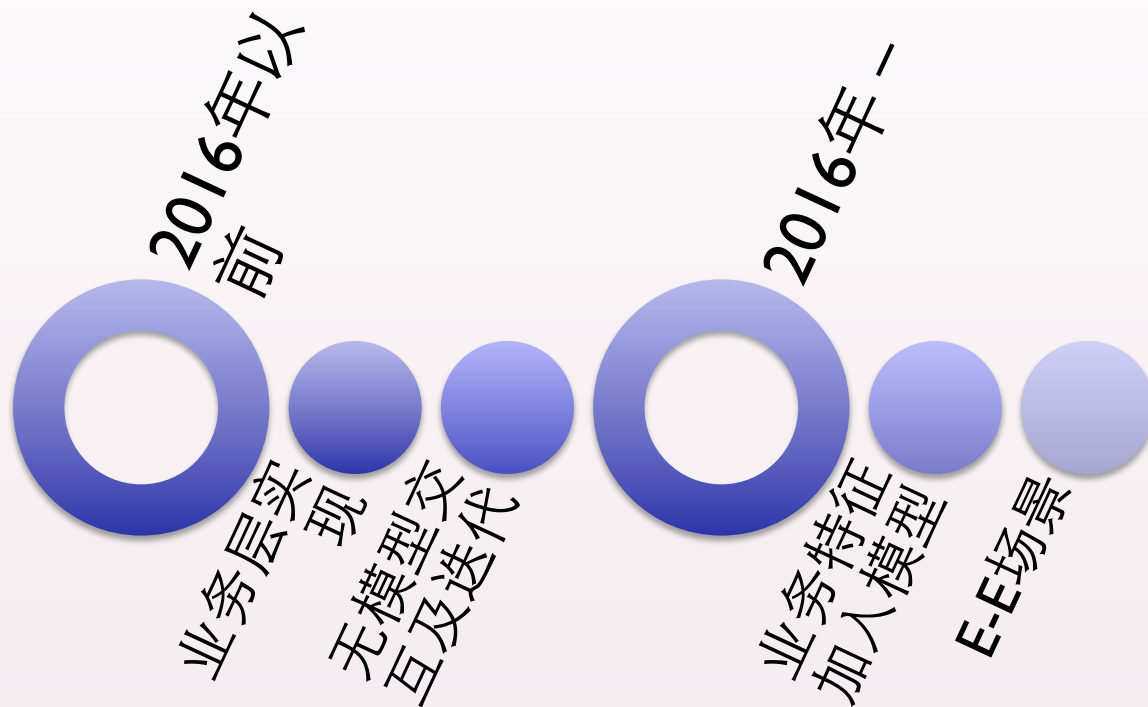
Q&A

# 初排层变迁过程



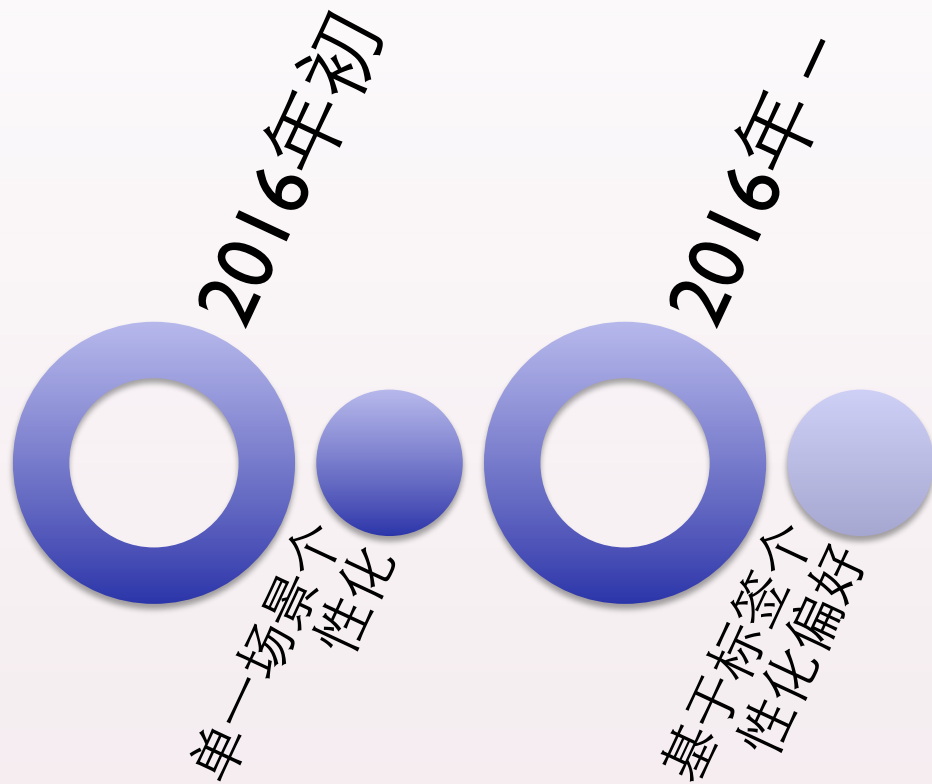


# 业务层变迁过程

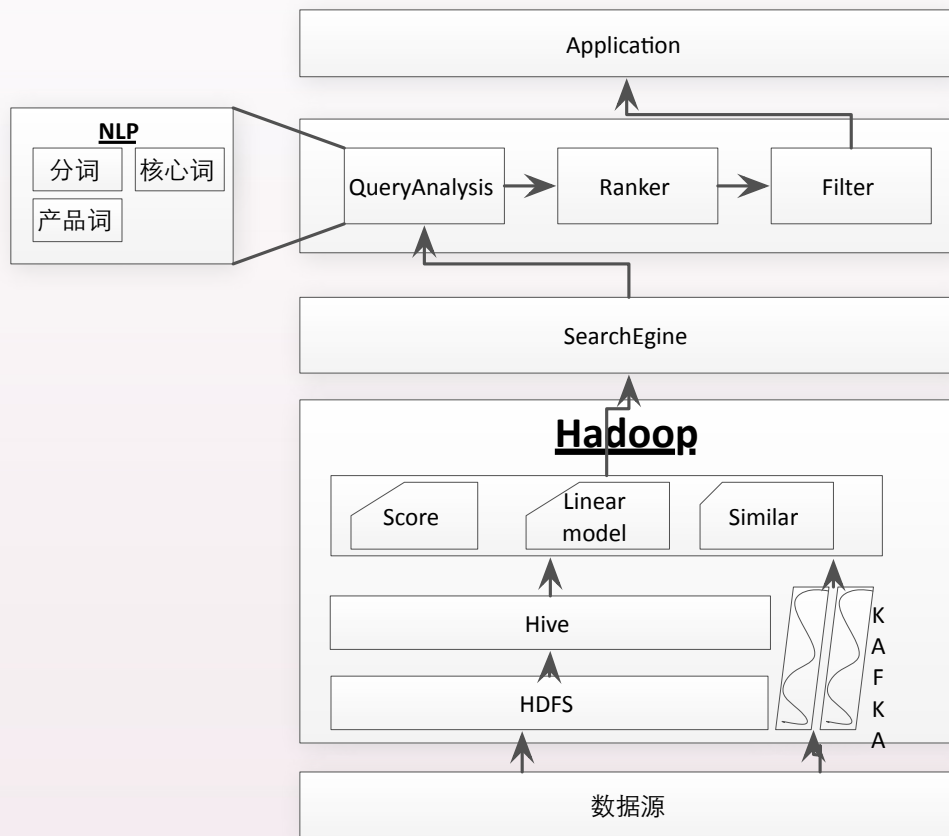




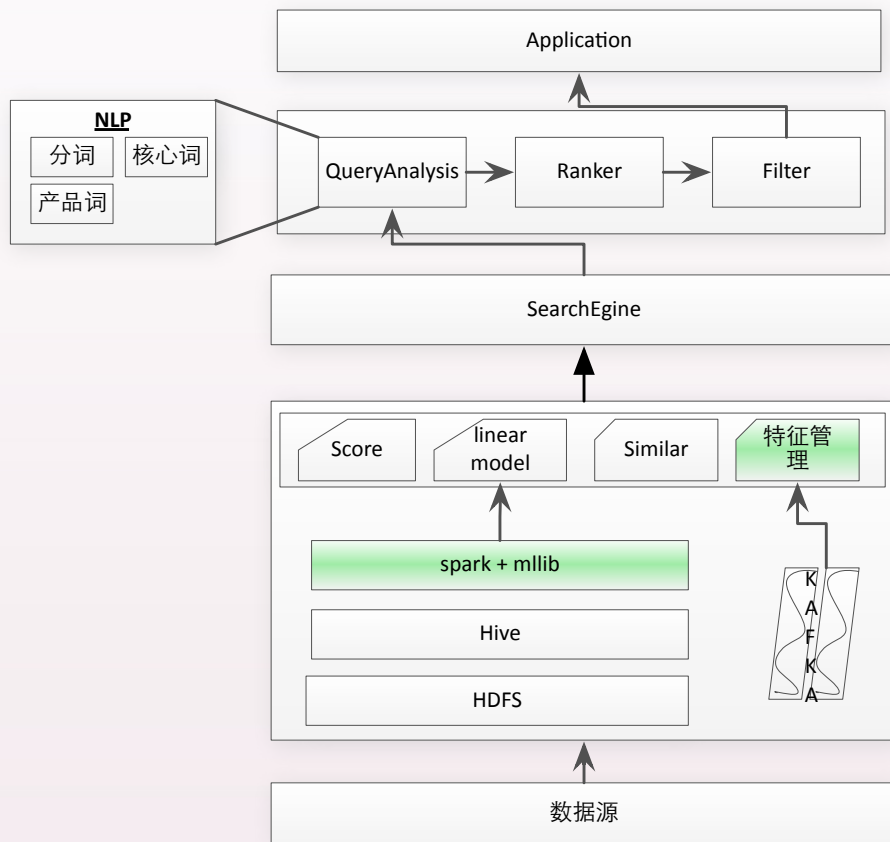
# 个性化层变迁过程



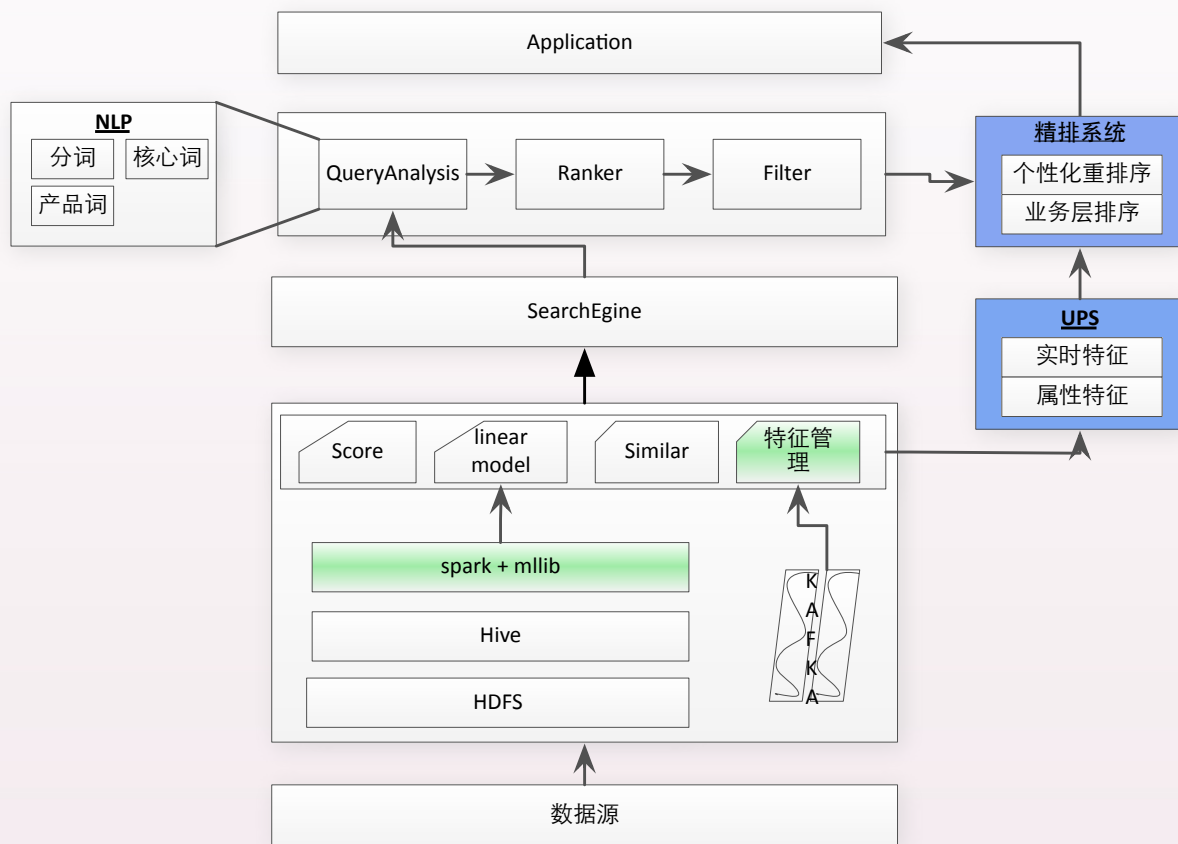
# 初始模型架构



# 初排模型架构



# 个性化后架构



# 蘑菇街电商算法演化



蘑菇街电商场景变迁历程

随场景变迁的算法模型

模型变迁中的问题

Q&A

# 模型变迁中的问题

●Q1：为什么要替换爆款模型？

	优势	劣势
爆款模型	预估速度快 与规则相比，预估效果更好	非对真实场景及数据分布的描述，对于中间商品缺乏排序区分度 个性化特征／业务特征难以加入
概率预估模型	对中间商品区分能力更强 个性化特征，业务类型特征可以低成本加入	预估速度慢 需要集群，分布式计算能力的支持 数据清洗，特征处理的工作更为复杂

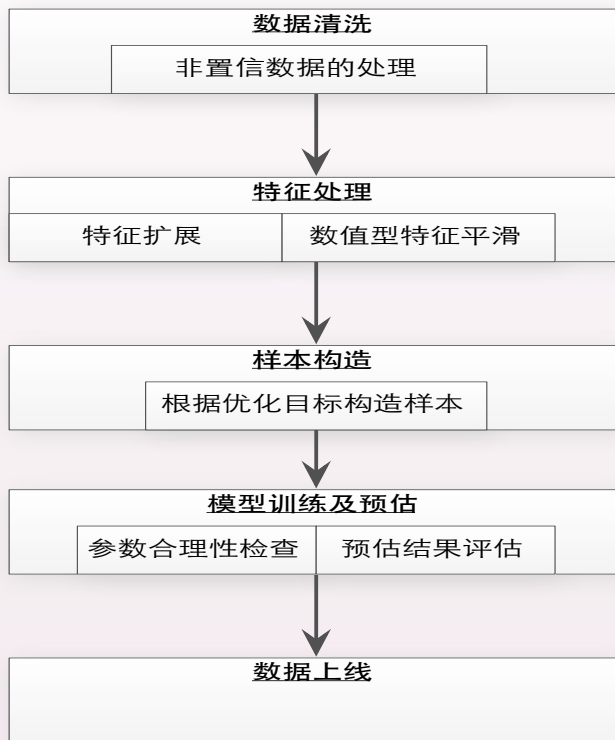
# 模型变迁中的问题

- 替换爆款模型带来的收益：
  - 业务收益
  - 系统收益
    - 为后续个性化/业务特征加入提供基础支持



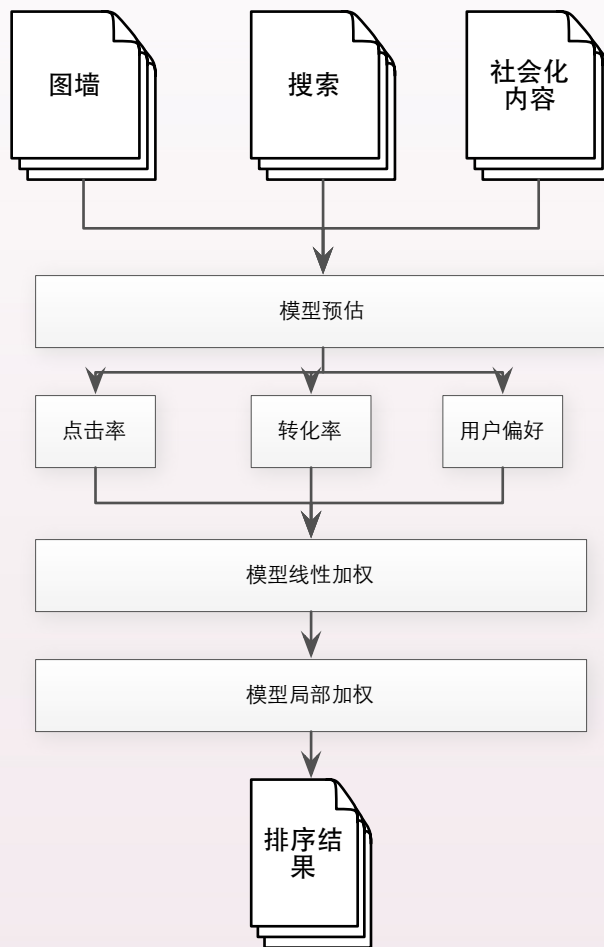
# 模型变迁中的问题

- Q2: 在模型切换过程中遇到的问题



# 模型变迁中的问题

- Q3: 多个优化目标如何融合



# 模型变迁中的问题

- Q4: 个性化因素加入

	优势	劣势
个性化因素作为初排之后的微调模型	单独模型训练，可控，可解释；不需要前期大量的数据	不能获得全局最优解，需要和explore模型做结合，效果有限
个性化因素作为特征加入模型训练	全局最优解，效率更高，可以自身完成数据迭代优化	需要前期有一定规模的数据，以支持特征获取更高权值

# 模型变迁中的问题

	线下实践	线上实践
个性化因素作为初排之后的微调模型	单独计算偏好auc较高	gmv和转化率均有收益， gmv提升2%
个性化因素作为特征加入模型训练	特征空间增加很大， 特征及其稀疏， 整体权值较低	暂未上线

# 模型变迁中的问题

- Q5:业务因素加入排序模型
  - 业务特征进入模型训练
  - E-E场景加入
  - 选品系统

# 模型变迁中的问题

- Q6: Explore-Exploit在业务上的使用
  - 统一业务场景下的e-e
    - 小流量explore，通过样本处理影响模型迭代
  - 不同业务场景下的e-e
    - Explore业务场景
    - Exploit业务场景

Thanks!

Q&A