

数据分析的职责（从业务中来，到业务中去）：

- 基于业务理解，把 **商业问题** 抽象成 **数据问题**
- 利用 **模型** 和 **工具** 来解答这个抽象后的问题，产生 **业务洞见**

常见误区：

- 堆数据：做报表，跑数据，excel达人
- 猜结论：有结论没数据支撑，分析结果靠猜
- 不落地：看似有数据、逻辑、结论落不了地

典型的“失职”

数据说明了什么？

——商业问题缺失

[illegible]

结论怎么来的？

——理论建模技术不足



数据分析能力三要素

- 数据分析能力模型

- 业务理解

- 工程、统计理论

- 呈现与说服

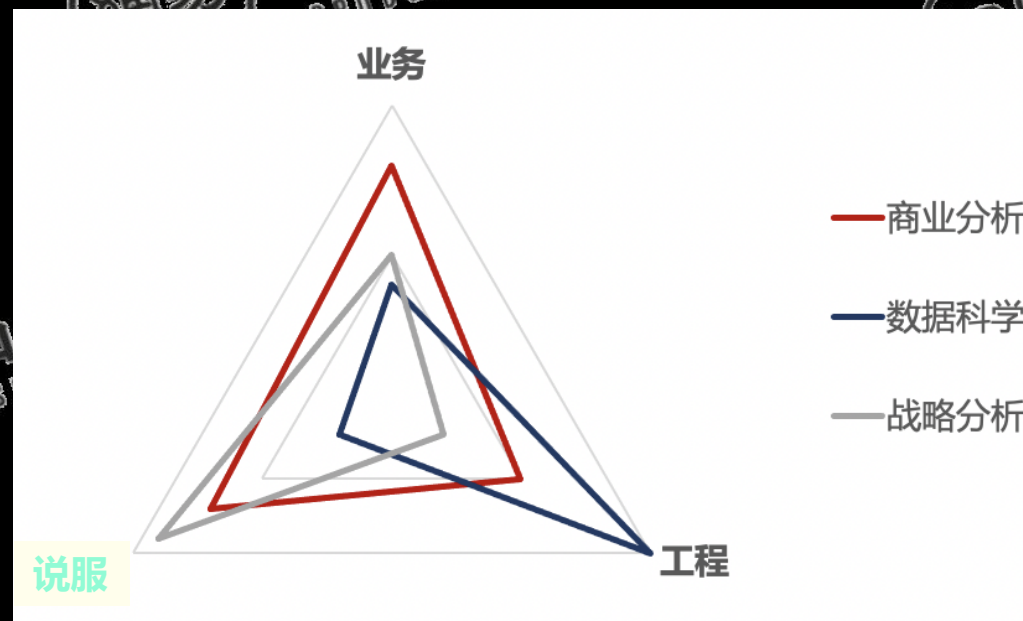
- 数据分析要交付什么价值

- 是什么（60分）：描述业务现状 – 基础技能

- 怎么样（80分）：现状是好是坏 – 把业务问题抽象成数据问题

- 怎么做（90分）：有什么建议

思考：之所以做了详细的报告，仍然被说“堆数据”，根本是缺乏业务视角



思考：不懂业务会有什么问题

视频平台首页改版：使用时长，观看次数.....

如果用到在线支付首页分析，那么.....

- 首页强化信息流
- 隐藏扫一扫和付款
- 付款链路越复杂越好

视频平台



在线支付

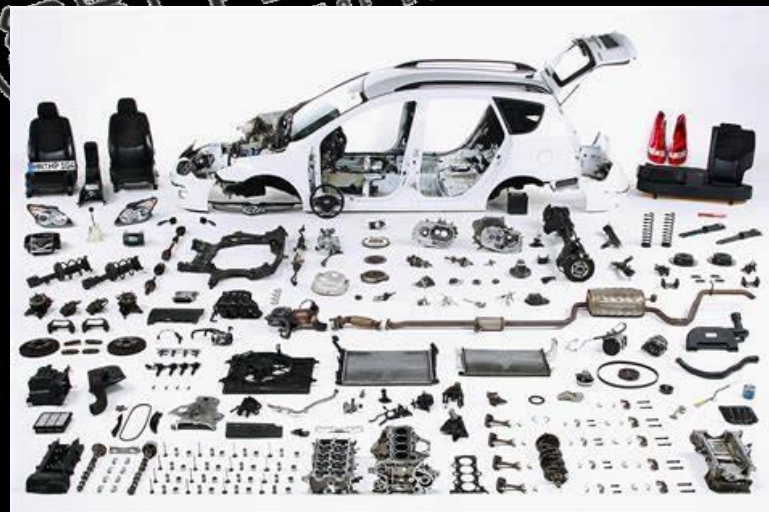


相同
指标?

不正确的业务理解，导致错误的分析方向，最终让工具产品失去了效率价值

业务理解：三个检验标准

技术层，分析通用能力



商业层，分析业务场景



理解业务的关键

- 全局视角，避免轻重颠倒、以偏概全
- 避免常识性结论
- 给建议接地气，避免太“虚”不落地

资深用户 & 技术大拿 → 有价值的分析？

- 技术优化通用能力，商业优化业务场景
- “我” << 用户

目录

- 商业分析的职责
- 内容分发业务简介
 - 内容分发的基本原理
 - 多种分发机制和产品形态
- 推荐产品分析基础

内容分发的基本原理

内容分发产品 **相同的价值**

- 帮人 获取 优质的内容和服务
- 帮 优质的内容和服务 找 人

内容/服务

用户

消费反馈

平台

内容分发产品 **不同的产品结构**

内容类型

音视频

商品

图文

.....

内容生产者

权威机构

明星大V

用户

.....

分发机制

运营精选

社交分发

算法推荐

.....

内容交互形式

实时互动

AR/VR

沉浸弹幕

.....

从分发机制看内容分发场景

主动获取

需求：用户有明确需求

分发：基于需求，按照相关度

推荐内容

分发价值：效率

被动推荐

需求：没有明确需求

分发：基于兴趣、热度、位置

等推荐内容

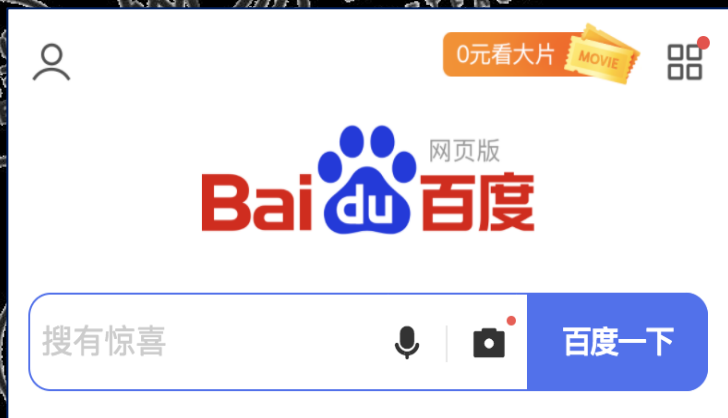
分发价值：发现兴趣

分类检索

需求：有明确需求

分发：基于分类和需求

分发价值：效率



目录

1. 商业分析的职责
2. 内容分发业务简介
3. 推荐产品分析基础
 - 3.1 用户价值
 - 3.2 算法机制
 - 3.3 数据标签
 - 3.4 核心指标

用户价值

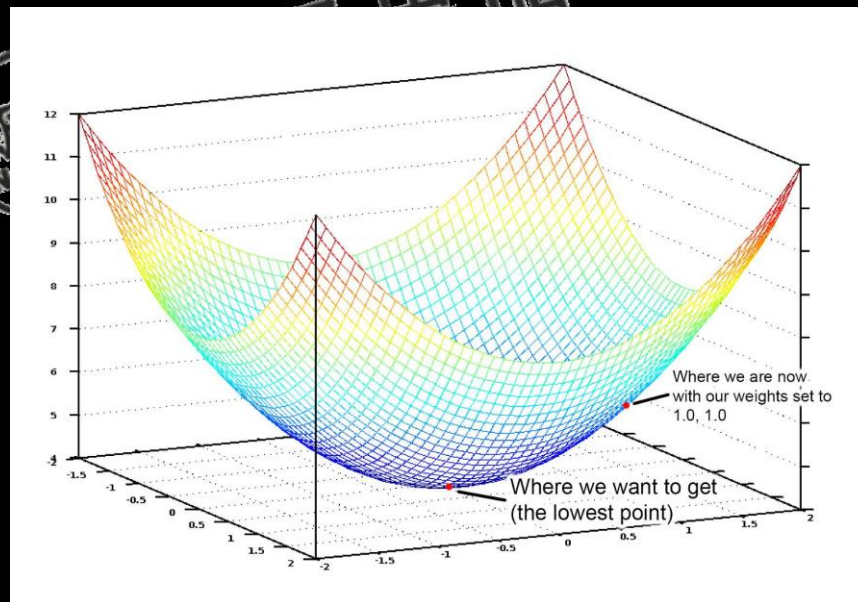
• 用户价值 如何实现

- 降低用户的消费决策, 挖掘用户的消费需求
- 通过**用户反馈、算法模型、优化目标**提升帮内容找人的效率
- **优化目标**包含模型的效率和用户的价值

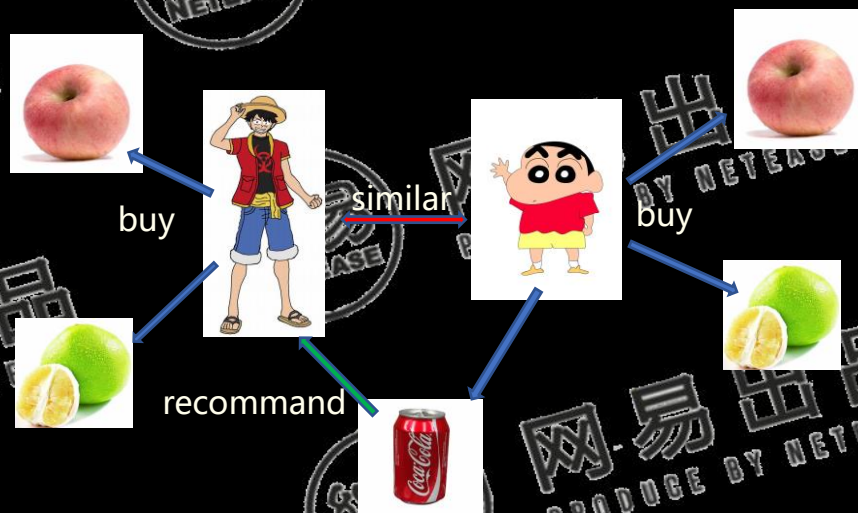
• 用户价值 如何衡量

- 内容消费: **点展比、停留时长**
- 关注引导: **关注UV**
- 造热点: **传播量**

基于用户反馈、优化目标, 模型寻找最优匹配



基于兴趣, 找到“志同”之人



算法机制

基于内容推荐

- 基于历史行为，按内容之间相似性推荐
- 典型场景：猜你喜欢
- 局限：信息茧房、冷启动

协同推荐

- 基于兴趣相似群体，进行推荐
- 典型场景：电影推荐 (Netflix)
- 局限：冷启动

关联规则推荐

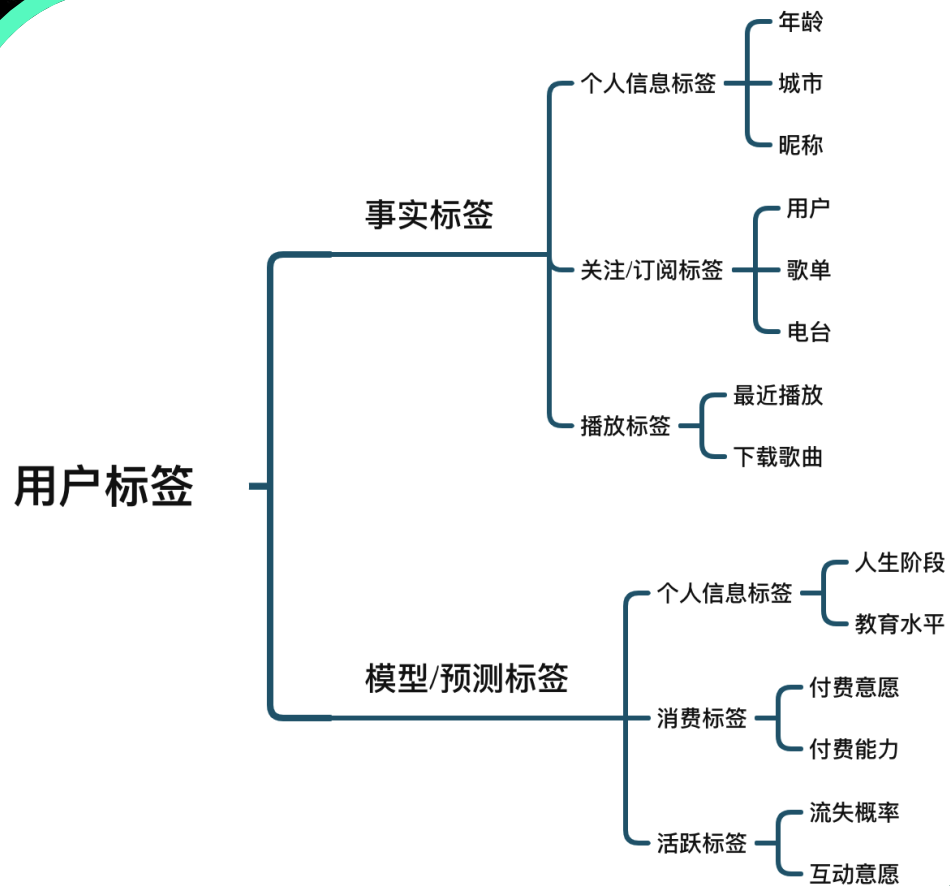
- 按置信度找事件之间强关联规则
- 典型场景：电商（啤酒和尿布）
- 局限：过度推荐、冷启动

基于POI推荐

- 基于地理位置信息，按照距离等推荐
- 典型场景：辅助性，如位置广告
- 局限：粒度粗，个性化程度低

数据标签：用户标签、Item标签

类型：事实标签、模型/预测标签； 来源：用户输入、专家建议、算法预测...



用户标签案例



item标签案例

数据标签的应用场景

① 精细化运营，比如基于「用户活跃度」来制定促销策略

参与活动的老用户数

生命周期标签

预算有限，奖金如何设置

价值分层标签

如何召回沉默用户

兴趣偏好标签

推送时间如何设定

营销时机标签

【案例】标签业务使用场景

② 个性化推荐

- 基于标签构建模型
- 召回策略

③ 业务分析

用户兴趣标签

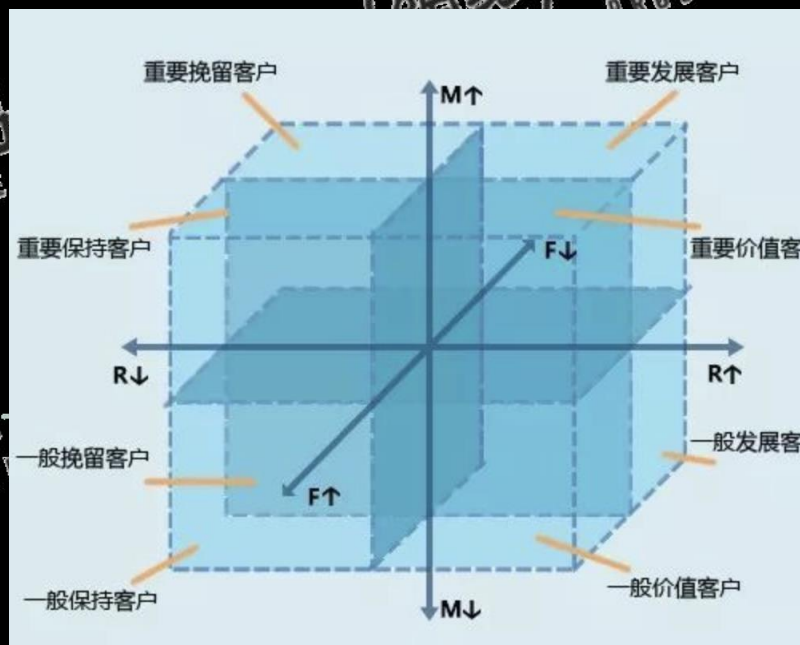
球鞋	0.5
键盘	0.2
耳机	0.1
手办	0.1

对应商品
TOP N结果

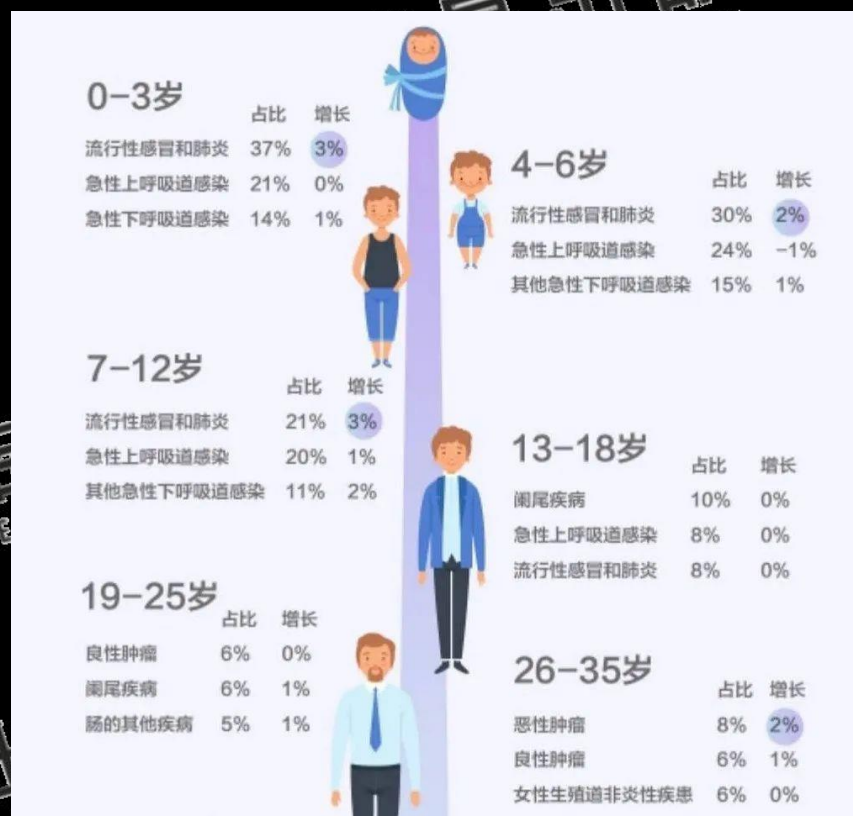
推荐模型

数据标签：分析使用场景

分层：以RFM为例



分群对比：以各年龄层用户为例



用户反馈

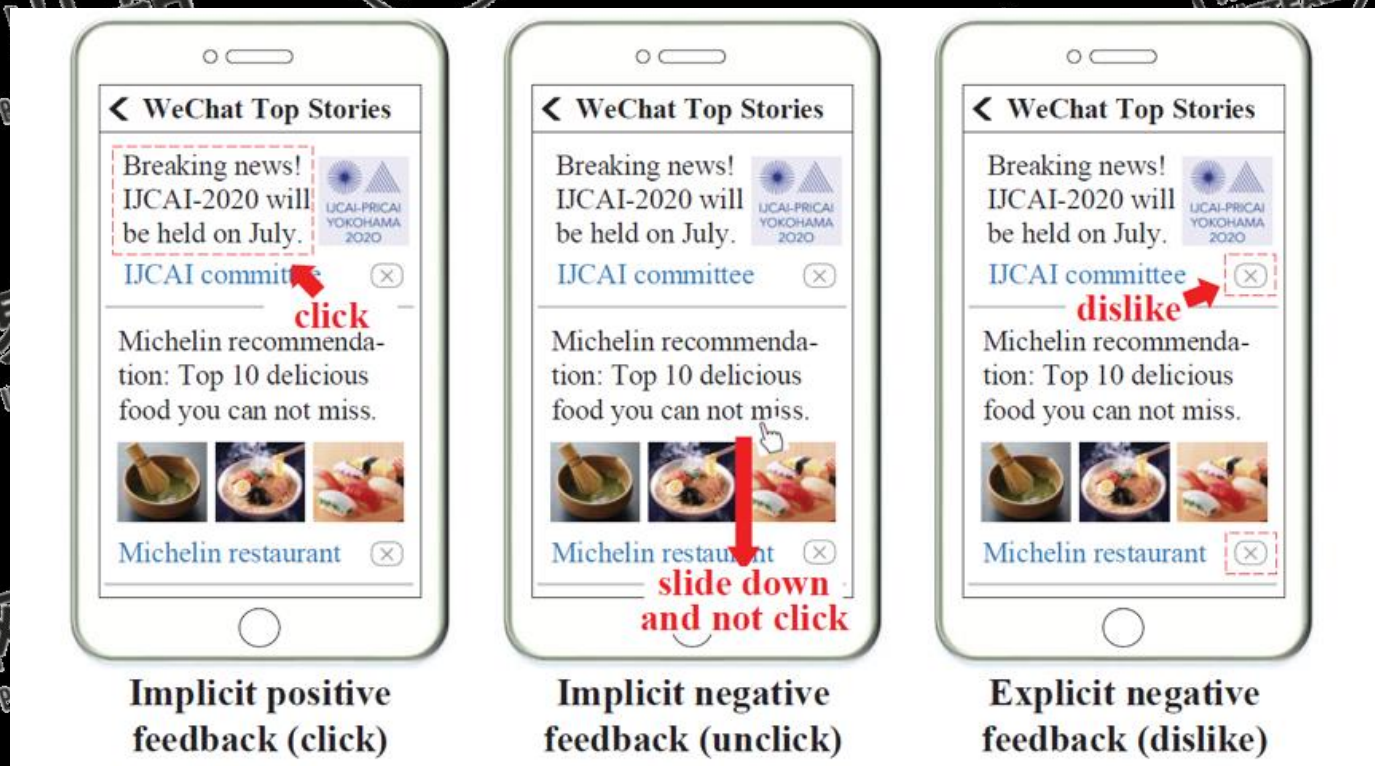
反馈类型

➤ **显性反馈行为**：包括用户明确表示对物品喜好的行为，主要方式就是评分和喜欢/不喜欢；

➤ **隐性反馈行为**：指的是那些不能明确反应用户喜好的行为，最具代表性的隐性反馈行为就是页面浏览行为；

➤ **正反馈行为**

➤ **负反馈行为**



基于用户交互反馈，发现兴趣，准实时更新推荐策略

核心指标：从推荐系统的三个节点评估

通用基础指标

内容

- 数量：存量、增量
- 质量：
 - 优质内容占比
 - 独有内容占比
- 丰富度：垂类内容占比

优质的定义：

- 平台调性、内容形态

丰富度的重要性：

- 通用型、垂类内容品台

分发

- 分发效率：
 - 点展比
 - 有效消费率
 - 互动率
- 流量分配：gini系数

分发效率的定义：

- 不同内容类型 有区别
- 用户导向、商业化导向

流量分配策略：是否中心化

用户

- 正向体验：
 - 留存、粘性
 - 使用时长
- 负向体验：
 - 跳出率

用户体验重点：

- 与内容类型
- 商业模式有关

业务差异定义

模型的优化目标：本质是算法背后的“价值观”

- 不同产品定位 导致 不同“价值观”（由业务目标、用户结构等决定）
- “价值观” 决定 推荐策略



微博：推荐话题越“引战”，
越激发评论欲

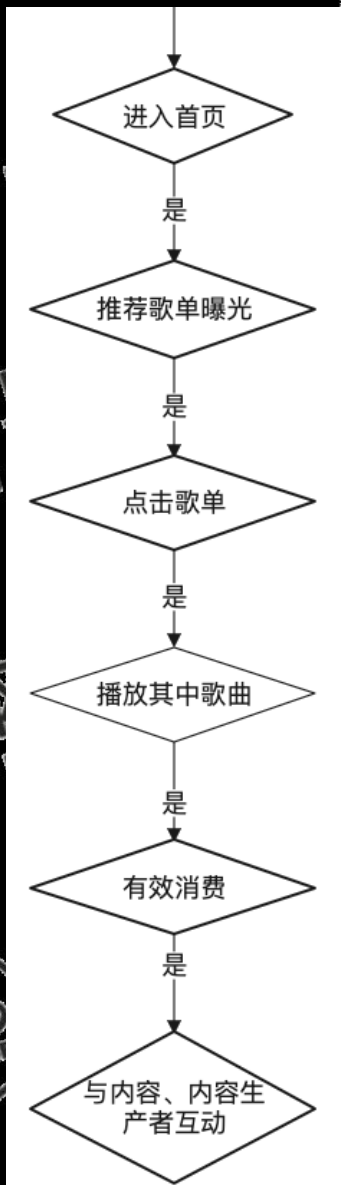


UC：推荐内容越“震惊”，用
户越愿意点击

算法没有价值观

张一鸣

推荐业务场景-云音乐推荐歌单消费链路



内容（相应的歌单数量）：

- 歌单库是否充足？
- 用户兴趣相关的歌单库是否足够？
- 推荐的歌单质量是否好？ 包含封面、文案

分发效率（点击率、有效播放率）：

- 推荐的内容是否有吸引力：有点击
- 点击的内容是否有效：有效消费
- 消费的内容是否产生粘性：赞评传、关注