

# DATAFUNCON

2020大数据 AI的最新技术实践

## 面向用户增长的信息流分发机制

阿里文娱-人工智能部-信息流推荐

天师



# 目录

## 1. 内容信息流--用户增长

- 问题定义
- 问题分析
- 增长要素

## 2. 内容信息流--推荐算法

- 推荐系统回顾
- 信息流推荐的增长目标

## 3. 核心增长机制

- 双边冷启动与流转机制
- 消偏与因果推断
- 面向增长的用户画像
- 效用理论应用

# DATAFUNCON

2020大数据 AI的最新技术实践

## 内容信息流 用户增长



# 问题定义

- 本文的问题域
  - 内容信息流app用户增长问题
  - 内容信息流：图文信息流，短视频信息流、内容+电商信息流
- 旨在解决的问题
  - 基于推荐算法视角的信息流产品用户增长问题
    - 留存问题
    - 幸存者偏差
- 本文受众
  - 信息流产品的
    - 推荐算法专家
    - 数据科学家
    - 产品专家





# 问题分析

## ➤ 背景

- 在移动互联网进入下半场的大趋势下，过去粗放式的买量、厂商合作等模式越来越会受到掣肘，将更加依赖精细化的用户增长策略和产品用户体验的细致打磨；经典的AARRR模式会逐步转向RARRA模式，提升产品留存、拉活、分享传播等方式是构建增长的主要战场。而在此之中，对于一个内容型产品，个性化算法对于用户留存、拉活将起到决定性的作用。

## ➤ 信息流产品增长的成功模式

- 头部内容模式
  - 该产品利用精准的内容采买眼光、引入优质的头部内容创作者，利用头部内容的流量聚集效应，迅速圈定大批用户，并形成内容app特有用户心智。由于内容头部化，个性化算法在其中发挥的空间和作用较小，产品、模式趋于同质化。
- 下沉/激励模式
  - 该产品参考了网络游戏模式，从各个环节设计用户里程碑和激励，不断引导新用户一步步完成点击、下刷、完整阅读、分享、关注等目标里程碑，并给予虚拟货币和真实货币的激励，在短时间内可以获取大量下沉用户。
- 生态构建模式
  - 该产品构建了完善的内容生产和消费生态，旨在通过推荐系统同时刺激生产和消费，实现两端的同时增长。

## ➤ 个性化核心问题

- 用户状态建模：深度建模用户状态和行为，从大数据集中找到使用户从低阶状态到高阶状态转化的干预因子。
- 个性化分发的升级：将用户行为建模后，在多个场景将这些干预动作落地为个性化推荐算法和营销算法，满足用户的消费需求。



# 增长要素

## 增长要素

- 优质内容 / 时效性
- 个性化体验
- 多渠道获客
- CPC vs LTV

## 算法如何助力增长?

- 精细化买量/外投
- 提升留存
- 衡量Action的效用
- 消除幸存者偏差

AARRR



RARRA



## STAGES OF RARRA



Acquisition: 获客  
Activation: 激活  
Retention: 留存  
Referral: 分享  
Revenue: 收入

# DATAFUNCON

2020大数据 AI的最新技术实践

## 内容信息流 推荐算法



# 推荐系统回顾

## ➤ 劣质系统

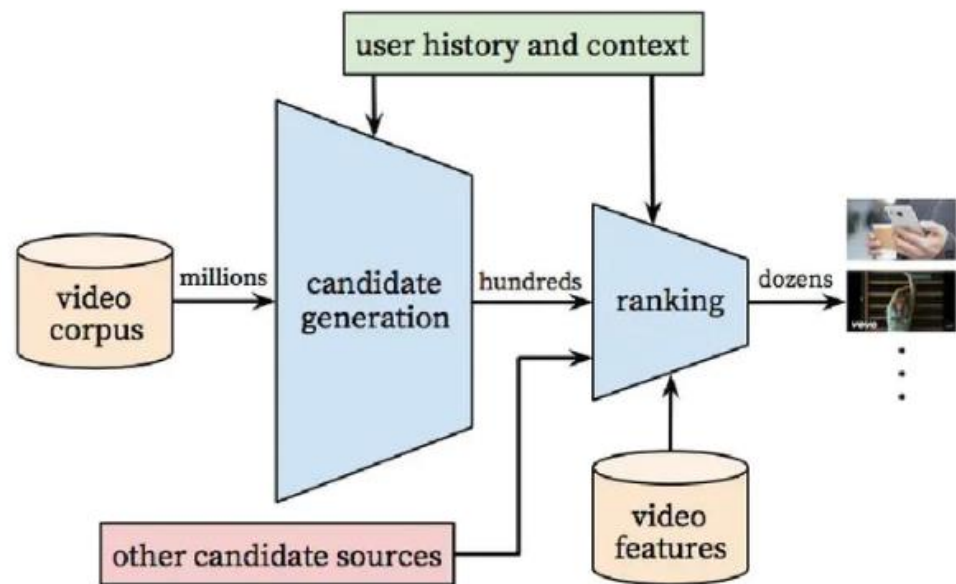
- 低俗（只满足低层次需求）
- 兴趣收窄
- 搬运内容（无稀缺性）
- 陈旧内容（无稀缺性）

## ➤ 良性系统：在各个环节在不断增加信息量/多样性

- 不同层次的用户引入（用户多样性）
- 各类高质内容的引入（内容多样性）
- 时效性（提升媒体属性）
- 探索出用户中长尾兴趣：当头部内容过期时更合理的承接

## ➤ 问题在哪儿？

- 统计机器学习模型的缺陷
- 指标观测体系的匮乏/业务短视
- 缺乏合理的机制设计/产品视角



经典视频推荐架构 (Youtube, 2016)

### 传统系统

标签召回

Item cf

User cf

统计机器学习

### NN base系统

表征学习

MTL

强化学习

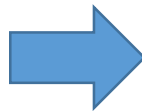




# 信息流推荐的增长目标

## 用户满意度的衡量

- 指标很多，但主要归结于三大类
  - 相关性、内容质量、内容时效性
  - 时效性一定程度代表内容稀缺性
- ctr代表什么：一定需要ctr吗？
  - 对内容无认知：对列表页素材的满意度
  - 对内容有认知：部分包含对内容本身的满意度
- 用户对内容的真正认可：点赞、分享、评论



## 信息流推荐的增长目标

- 提升用户留存，LTV
- 通过分发筛选出优质内容、优质生产者
  - 传统的认知：喜欢什么就给什么
  - 更新的认知：创造需求和玩法，持续优质内容生产
- 构建内容生态
  - 分发时效性
  - upgc主的激励：曝光、点击、粉丝、分成
  - 准入、扶持打压、激励

# DATAFUNCON

2020大数据 AI的最新技术实践

## 核心增长机制

# ● 双边冷启动与流转机制：概述

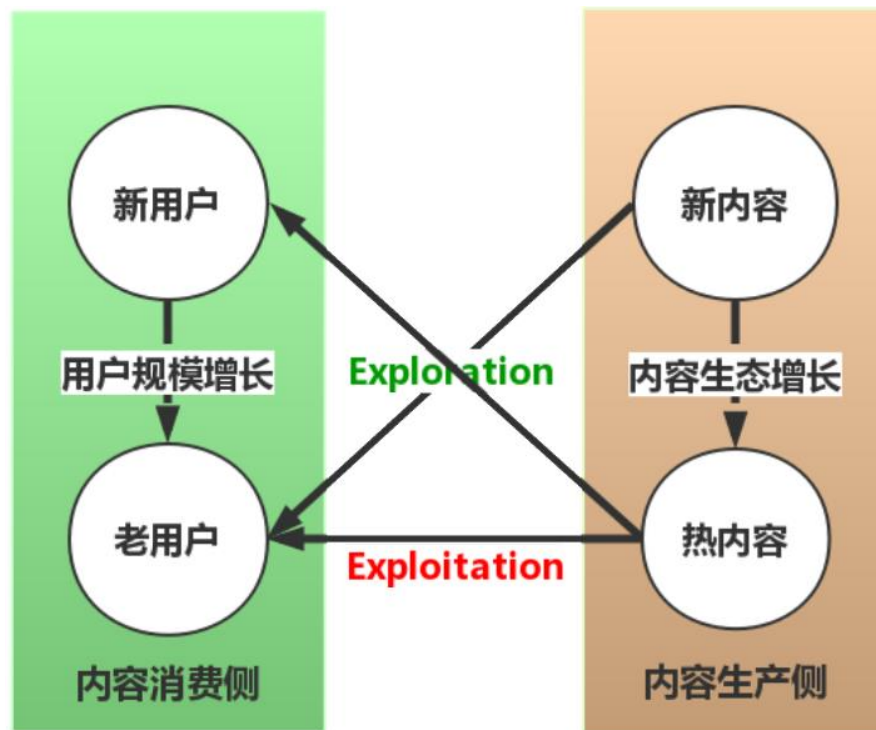
## ➤ 流转机制思想

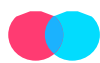
- $new\ item / new\ publisher \rightarrow old\ user$ : 利用相关性进行新内容冷启
- $old\ item / old\ publisher \rightarrow new\ user$ : 用置信的 $item$ 进行（新）用户兴趣探索

- $new/old$ 的定义通过置信度对应到多峰兴趣级别
- $new/old$ 的准确定义：经过某些流量的分发验证
- 新内容冷启机制也叫爬坡机制

## ➤ 问题

- 基于表征学习的排序技术，难以表达置信度  $\rightarrow$   $uncertainty$ 预估问题
- 新内容冷启技术选型：随机保量 +  $Bandit$ 类算法 +  $uncertainty$ 预估
  - $Bandit$ 类算法：短程收敛性有一定劣势
  - 随机保量分发
    - 短期降效
    - 对生态 $fairness$ 问题相对友好
    - 构建宝贵的无偏数据集：消偏应用
- 新用户冷启技术选型：强化学习、联邦学习





# 双边冷启动与流转机制：详细实现

## ➤ 早期以显式标签+统计连续值为主的推荐系统

- $Rank = pRelevance(topic / user)^{cu} * pCTR(item / topic)^{ci}$
- 新item冷启探索
  - $ci$ 低 $\rightarrow$  $cu$ 变高, 侧重 $pRelevance(topic / user)$ 的值 + 预估稳定性
- 新用户冷启/兴趣探索
  - $cu$ 低 $\rightarrow$  $ci$ 变高, 侧重 $pCTR(item / topic)$ 的值 + 预估稳定性
- 纯利用
  - $cu$ 高,  $ci$ 高, 推荐的短期效率指标最高

## ➤ 以表征学习为主的推荐系统

- 一种构建排序模型uncertainty的方式
  - $user\ interest\ uncertainty(u) = \sigma ( pdf \{ pCTR(\bar{U}(u), i) \} )$
  - $item\ efficiency\ uncertainty(i) = \sigma ( pdf \{ pCTR(\bar{U}(i), u) \} )$



构建Risk-aware Recommendation

## ➤ 超参学习/调控的目标 $\rightarrow$ 系统全局E&E

- $User = [(t1, cu1, w1), (t2, cu2, w2), ...]$
- $Item = [(t1, ci1, s1), (t2, ci2, s2), ...]$
- $t$ : 兴趣topic
- $w$ : 兴趣权重
- $s$ : 分发效率
- $cu$ : 兴趣置信度/活跃度
- $ci$ : 内容置信度
- $pRelevance$ : 相关性打分
- $pCTR$ : 预估ctr

- $\sigma$ : 风险衡量
- $\bar{U}$ : embedding空间相似领域
- $pdf$ : 概率密度函数





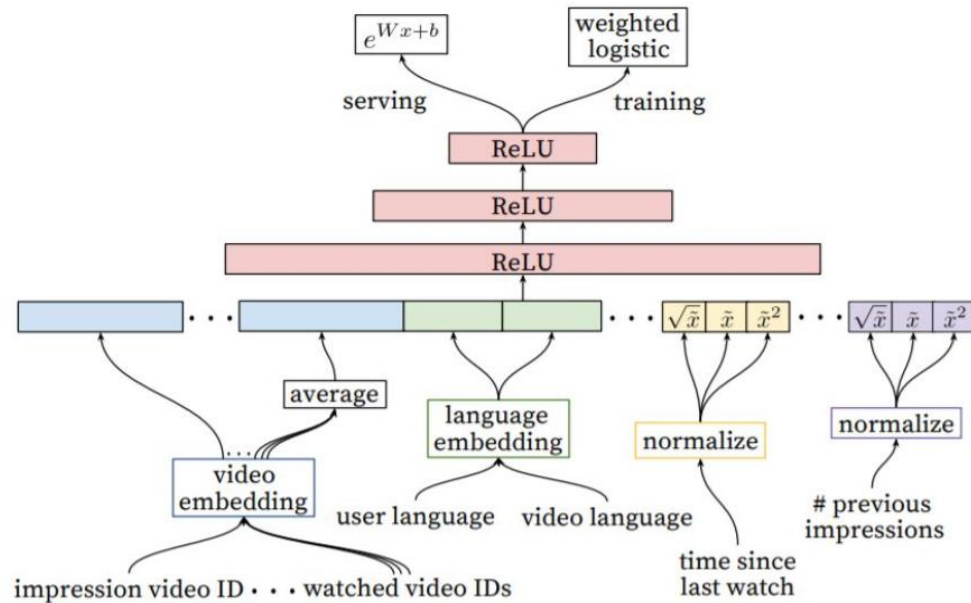
# 消偏与因果推断：背景介绍

## ➤ 推荐系统其实是一个因果推断问题

- *user emedding*: 用户是什么样的人(*cause*), 从而推测用户会喜欢的*item*
- 推荐的*causal effect*: 我这么推用户会不会喜欢? 效用有多大?

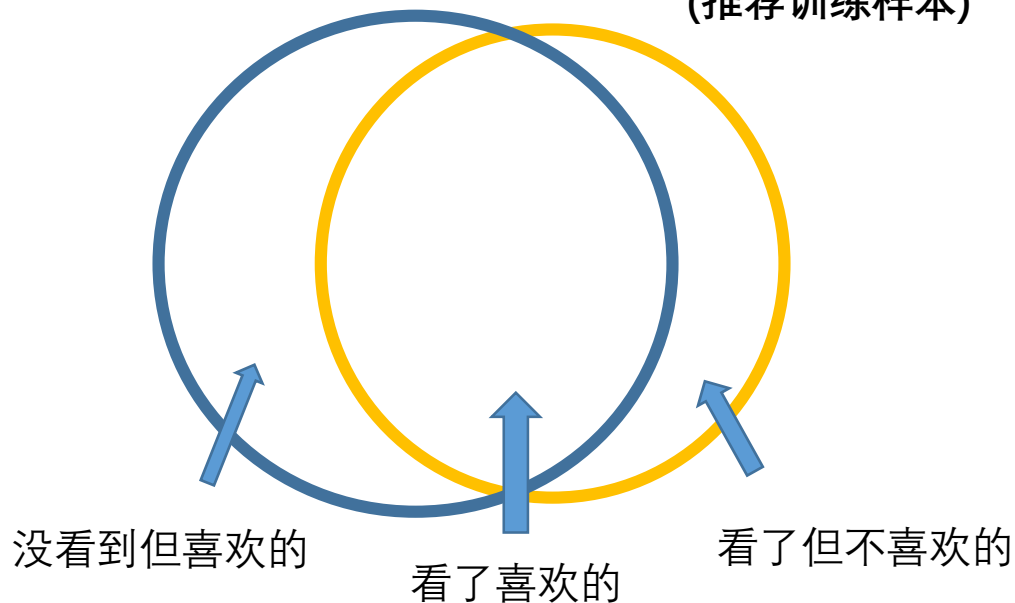
## ➤ 为什么会有幸存者偏差

- 推荐模型训练的样本有偏 (*user*、*item*偏同时存在)
- 对应到*item*, 就是*selection-bias*和*fairness*问题

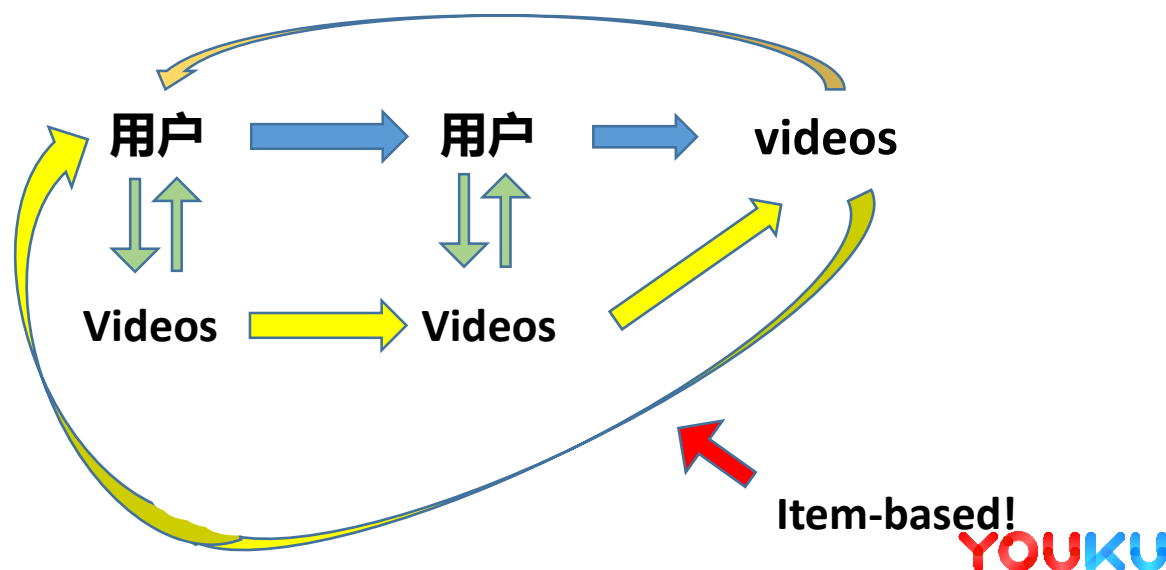


用户A喜欢的

用户A看了的  
(推荐训练样本)



Youtube net的问题: *user embedding average pooling*导致本质上其实依然是*item-based*。后续的诸多改进没有本质解决消偏问题。



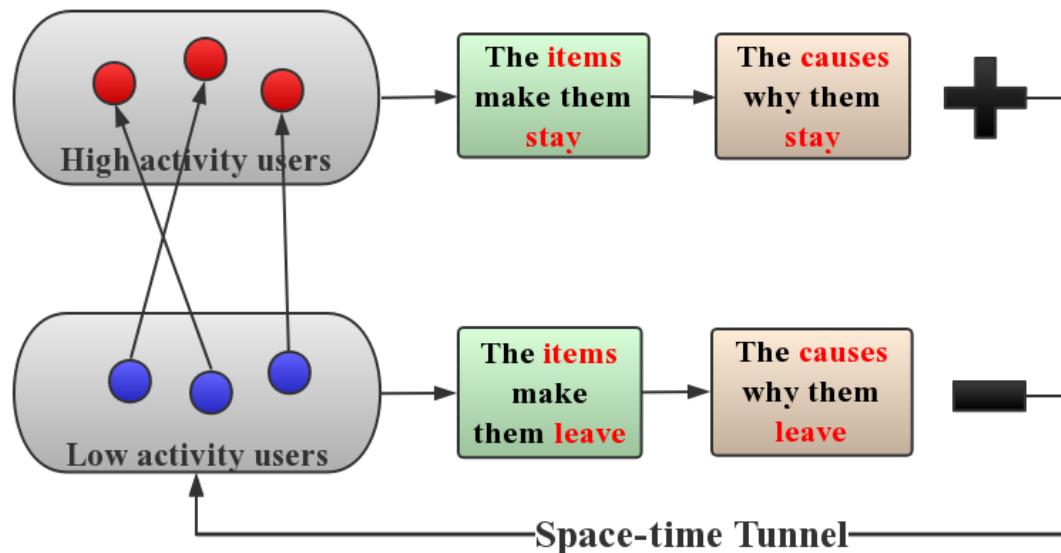
# 消除幸存者偏差：Causal Inference推荐框架

## 假设

- 用户变成低活、沉默的原因主要是用户对之前推荐的内容不满意

## 方法

- 构建反事实镜像人：利用无偏信息构建相似度量，构造低活User到高活User的匹配
  - Matching / Propensity Score / IPW
  - Causal Embedding
- 去除低活、沉默用户的leave causes，推荐高活镜像人的stay causes



**Causal Inference框架将成为用户增长方向的理论基石**  
消偏/效用衡量/归因/生态公平性问题

**Causal Inference为经典推荐算法难题提供解法**  
selection-bias/低活用户画像/推荐可解释性

# ● 面向增长的用户画像：状态里程碑表示法

## 经典用户画像

- 服务于运营可解释性
- 服务于推荐或广告系统的模型预估



## 通常建模成向量

- 高维离散:  $U=[(t1,w1),(t2,w2),...,(tn,wn)]$
- 低维稠密embedding:  $U = e$
- 多峰embedding:  $U = (e1, e2, ..., em)$



- 信息流产品具备连续型消费的特点（同主题下的消费，upgc关注/直播），用户的消费行为可以在连续的时间上进行切分，状态表示法是对向量表示法的有力补充。
- 将难度较大的“促留存”问题拆分为“目标达成”问题，通过策略不断使得用户完成高阶里程碑。



## 状态表示法

- 活跃度/置信度：高中低活
  - 多峰兴趣的置信度/活跃度
- $$U=[(t1,cu1,w1),(t2,cu2,w2),...]$$
- $$H=[S1, S2, ...]$$
- $$\text{跃迁概率 } P=(Hi \rightarrow Sj)$$

“促留存”的抓手 → 用户状态跃迁速度提升

# 面向增长的用户画像：全生命周期因果推断

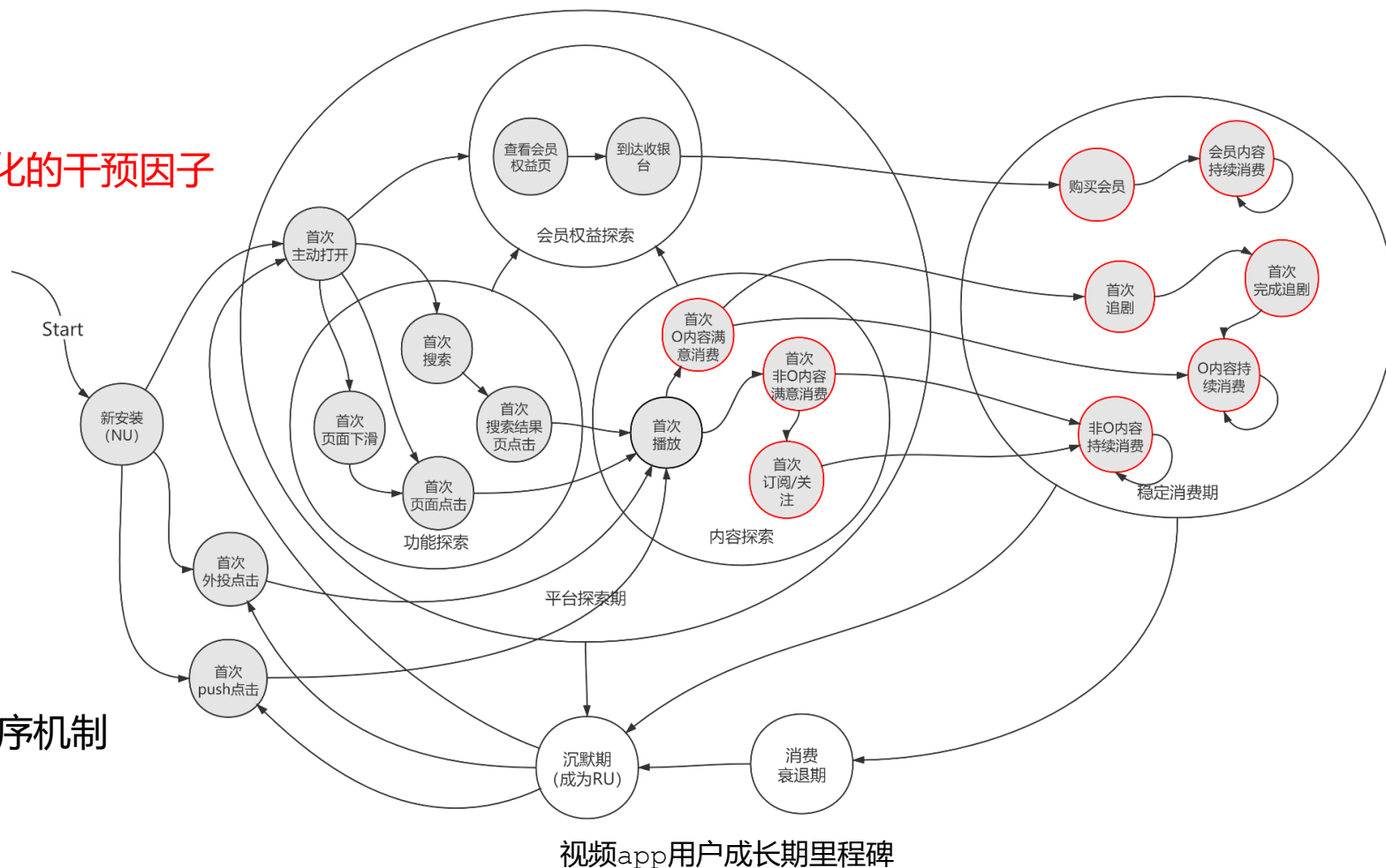
## ➤ 全生命周期的因果推断

- 1) 状态跃迁效用衡量
- 2) 找到使用户从低阶状态到高阶状态转化的干预因子

- 内容变化
  - 新内容上下架
  - 热点事件的产生/消逝
- 捕捉兴趣的变化
- 分发幸存者偏差问题

## ➤ 根据Causes的Actions

- 推断后的数据统计 → 内容采买/生产
- 干预场的设计、页面组织
- 内容供给指导 (2B供应链)
- 构建时间线上的推荐系统 → 个性化排序机制



### Causal Effect

单用户不可重复实验(Uplift model)

红包/券/激励 特定时期的头部内容/热点

### Causal Inference

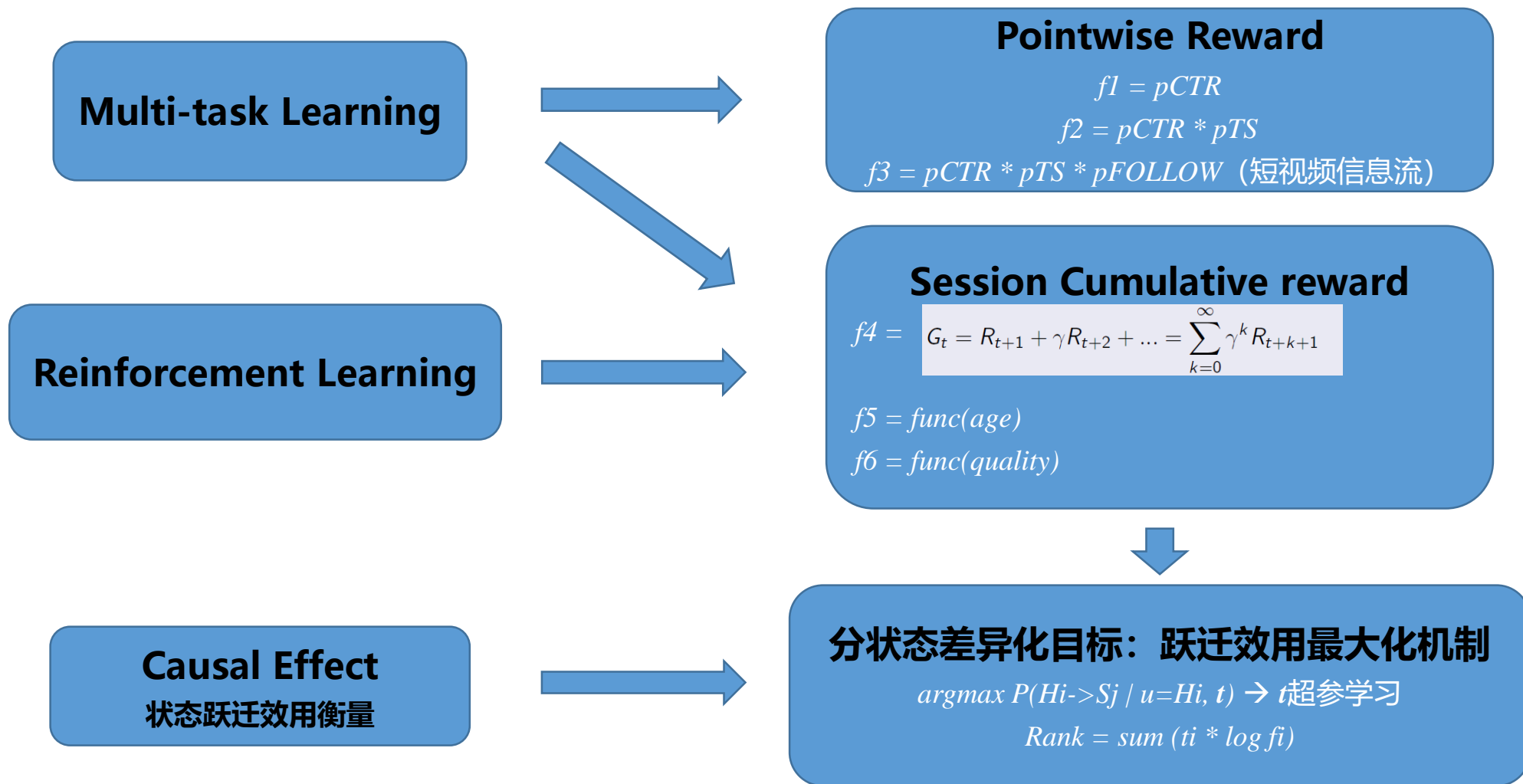
单用户可重复实验

用户长期兴趣推断





# 效用理论应用：个性化排序机制



## 效用的普遍认知

新用户 → 首次满意消费/互动，内容质量控制  
低活用户 → 浏览深度和ctr，内容质量控制  
高活用户 → 新颖性和多样性



# 效用理论应用：生态效用技术

## ➤ 供给归因

- 兴趣覆盖度、兴趣点下的内容质量、时效性、竞争激烈程度
- 内容供给指导（2B供应链）→ 内容采买/生产/激励机制

## ➤ up主激励机制

- up主质量体系：关注量曲线/生产效率/生产质量：内容正负反馈
- 输入：历史分发数据（曝光量、点击量、关注量、up主质量体系）
- 计算：
  - 曝光量 → 爆款成功率
  - 最大化效用：激励函数 → （内容质量，生产效率）的拟合
- 输出：一定时间内up主内容的保量、限量range → 在线排序的调控约束机制

## ➤ 未来的方向

- 流量货币化
- 更多经济学、机制设计理论的引入
  - 演化博弈论分析
  - 竞对分析

# DATAFUNCON

2020大数据 AI的最新技术实践

# THANKS! & QA