



www.leadleo.com

2022年

# 中国推荐算法应用市场研究

报告标签：智能推荐、推荐算法、机器学习、深度学习

报告主要作者：彭承玺

2022/07

报告提供的任何内容（包括但不限于数据、文字、图表、图像等）均系头豹研究院独有的高度机密性文件（在报告中另行标明出处者除外）。未经头豹研究院事先书面许可，任何人不得以任何方式擅自复制、再造、传播、出版、引用、改编、汇编本报告内容，若有违反上述约定的行为发生，头豹研究院保留采取法律措施，追究相关人员责任的权利。头豹研究院开展的所有商业活动均使用“头豹研究院”或“头豹”的商号、商标，头豹研究院无任何前述名称之外的其他分支机构，也未授权或聘用其他任何第三方代表头豹研究院开展商业活动。



# 摘要

**01 推荐算法经过多年的发展已较为成熟，融合数学、计算机等多学科，进行分类与标签匹配，再通过海量运算后进行精准推荐。**

- 2012年至今，推荐算法进入了快速的技术革新阶段，也逐步依托技术带来的领先性有效的实现商业化落地，与互联网领域进行了深度的结合。2010年提出的FM技术在2012年成为主流的推荐算法，2014年的GBDT+LR更是带来了技术上的突破，2015年推荐算法由机器学习也正式转为深度学习。推荐算法也自此进入到各行各业，目前广泛应用于图书、音乐、视频、新闻、电影、地图、网购等等领域；

**02 推荐算法与不同领域融合时，考虑到场景的运行逻辑不同也会采用不同的运行机制，以适应场景的特殊性**

- 推荐算法是通过合理的逻辑运算，为用户推荐最适合的内容，在模型角度上讲，是拟合用户对于内容满意程度的预测函数。新闻资讯领域的推荐算法着重考虑三方面因素，对应模型中的三大要素，第一要素为内容，第二要素为用户特征，第三要素为环境特征；以抖音为代表的短视频领域，多以内容发布为起点，通过智能算法匹配到合适的用户，再根据反馈决定是否扩大内容的传播范围；生活类更多围绕用户的兴趣标签，以及用户的历史搜索标签进行关联推荐与深入推荐；

**03 平台与机构运用推荐算法，商业运行效率提高、用户快速获取所需、内容与产品供应商实现高效供应**

- 推荐算法作为桥梁快速匹配用户与内容或产品，提高整个链条的运行效率；(用户角度，效率明显提高，算法通过分析精准推荐用户感兴趣的内容或产品，省去了搜索与查找得时间，快速定位到用户需要；内容产出者与商家角度，对于内容或产品对于目标用户/客户群体的精准投放得以快速实现，加快了匹配效率，使得内容与产品可以快速传播。

## 推荐算法不仅要广泛应用，更需要广泛善用

推荐算法的应用已经涉及到国民生活的方方面面，确实为广大用户与内容或产品供应方提供了巨大的便利，但随着应用推荐算法的机构与平台影响力逐步扩大，对于算法的善用成为更值得探讨的话题；

推荐算法应用伴生的个人信息安全问题、个人隐私泄露问题、泛娱乐化趋势、内容传播低质化趋势等，结合各类平台与机构的海量流量，影响范围巨大。平台与机构在获得大量收益的同时，是否也需要采取措施，肩负起一定的社会责任，由重视企业利益，转变为用户利益为主，在社会或是国家层面起到更加积极的作用。

# 目录

## CONTENTS

◆ 推荐算法总览		
• 推荐算法的定义与分类	-----	6
• 推荐算法的发展历程	-----	7
• 推荐算法的相关政策环境分析	-----	8
• 推荐算法的技术分析（例举）	-----	9
◆ 推荐算法的产业图谱		
• 推荐算法的应用产业图谱总览	-----	11
• 资讯类 - 常用推荐算法逻辑	-----	12
• 短视频类 - 常用推荐算法逻辑	-----	13
• 生活类 - 常用推荐算法逻辑	-----	14
◆ 智能推荐算法的优劣势分析		
• 推荐算法以企业利益为主的负面影响	-----	16
• 个人信息乱用与泄露的隐患	-----	17
• 推荐算法有效提高用户的匹配效率	-----	18
• 推荐算法具备引导优质内容的巨大潜质	-----	19
• 推荐算法拥有深度帮助用户的广阔可能性	-----	21
◆ 企业研究		
• 字节跳动 – 抖音	-----	23
• 阿里巴巴 – 淘宝	-----	26
◆ 法律声明、方法论	-----	28
◆ 研究院介绍	-----	30

# 图表目录

## List of Figures and Tables

图表1: 小红书搜索端流量来源占比情况 (2020年)	-----	14
图表2: 遭遇信息泄露的网民比例 (截止2021年12月)	-----	17
图表3: 信息泄露中电商渠道扩散的比例 (截止2021年06月)	-----	17
图表4: 拼多多各类通道流量占比情况 (2020年-大致)	-----	18
图表5: 抖音与快手月活量人数 (2020年1月-2021年4月)	-----	19
图表6: 抖音与快手月使用时长 (2020年1月-2021年4月)	-----	19
图表7: 主流新闻门户的月活量情况 (2020年1月与2021年1月)	-----	24



# 中国推荐算法应用市场背景分析

- 推荐算法的定义与分类
- 推荐算法的发展历程
- 推荐算法相关政策环境分析
- 推荐算法相关技术模型分析





# 定义与分类

• 推荐算法根据算法逻辑不同分为多个类别，随着智能化的提升对于组合推荐的需求明显更高

## 推荐算法定义与分类

智能推荐算法是计算机专业中的一种算法，通过智能化的数学算法，分析用户的行为数据，判断用户的喜好并进行精准的推荐，目前多应用于互联网领域

### 基于内容

以项目的内容信息为评判主体，不考虑用户对于项目的评价意见，基于对用户的特征、兴趣范围的判断，评估用户与项目内容的关联性，并进行关联推荐

### 基于效用

结合用户对于项目的使用效果反馈，综合评定供应商的可靠性与产品的可得性，在推荐的项目范围中，更优先推荐用户可靠性高的项目供应商

### 基于协同

目前应用最早与最成功的推荐算法技术，采用临近技术，基于同群体用户爱好重合度高的特性，评估用户关联度高的其他用户的兴趣范围，并推荐给目标用户

### 基于知识

结合推理技术，与用户需求与偏好分析不同，以项目本身的知识性为研究主体，分析其是否满足某一特定用户群体，并进行精准推荐

### 基于关联规则

以用户已购项目为基点，分析此类项目在销售过程中与其他项目销售的相关性，借助用户对于个别项目的需求，挖掘用户的潜在需求，进行关联推荐

### 组合推荐

实际应用中最为常被采用，结合上述两个或多个算法进行有效结合，避免与弥补各单项技术的弱点或局限性，共同计算产生更加精准的推荐预测结果

### 组合推荐

加权组合法

变换组合法

元级别组合法

混合组合法

特征扩充组合法

特征组合法

叠层组合法

## 算法优劣势分析

### 内容推荐

#### 优势

- 推荐结果直观，容易解释；
- 无需领域知识

#### 劣势

- 新用户问题；
- 复杂属性难处理；
- 需要足够数据构造分类器

### 协同推荐

#### 优势

- 新异兴趣发现、无需领域知识；
- 性能随时间的积累而提高；
- 自动化程度高；
- 能处理复杂的非结构化对象

#### 劣势

- 稀疏问题；
- 可扩展性问题；
- 新用户问题；
- 质量取决于历史数据集

### 规则推荐

#### 优势

- 能发现新兴趣点；
- 无需领域知识

#### 劣势

- 规则抽取难、耗时；
- 产品名同义性问题；
- 个性化程度低；

### 效用推荐

#### 优势

- 无冷开始和稀疏问题；
- 对用户偏好变化敏感；
- 可考虑非产品特性

#### 劣势

- 用户必须输入效用函数；
- 推荐是静态的，灵活性差；
- 属性重叠问题；

### 知识推荐

#### 优势

- 可把用户需求映射到产品上；
- 可考虑非产品属性

#### 劣势

- 知识难获得；
- 推荐是静态的

## 头豹洞察

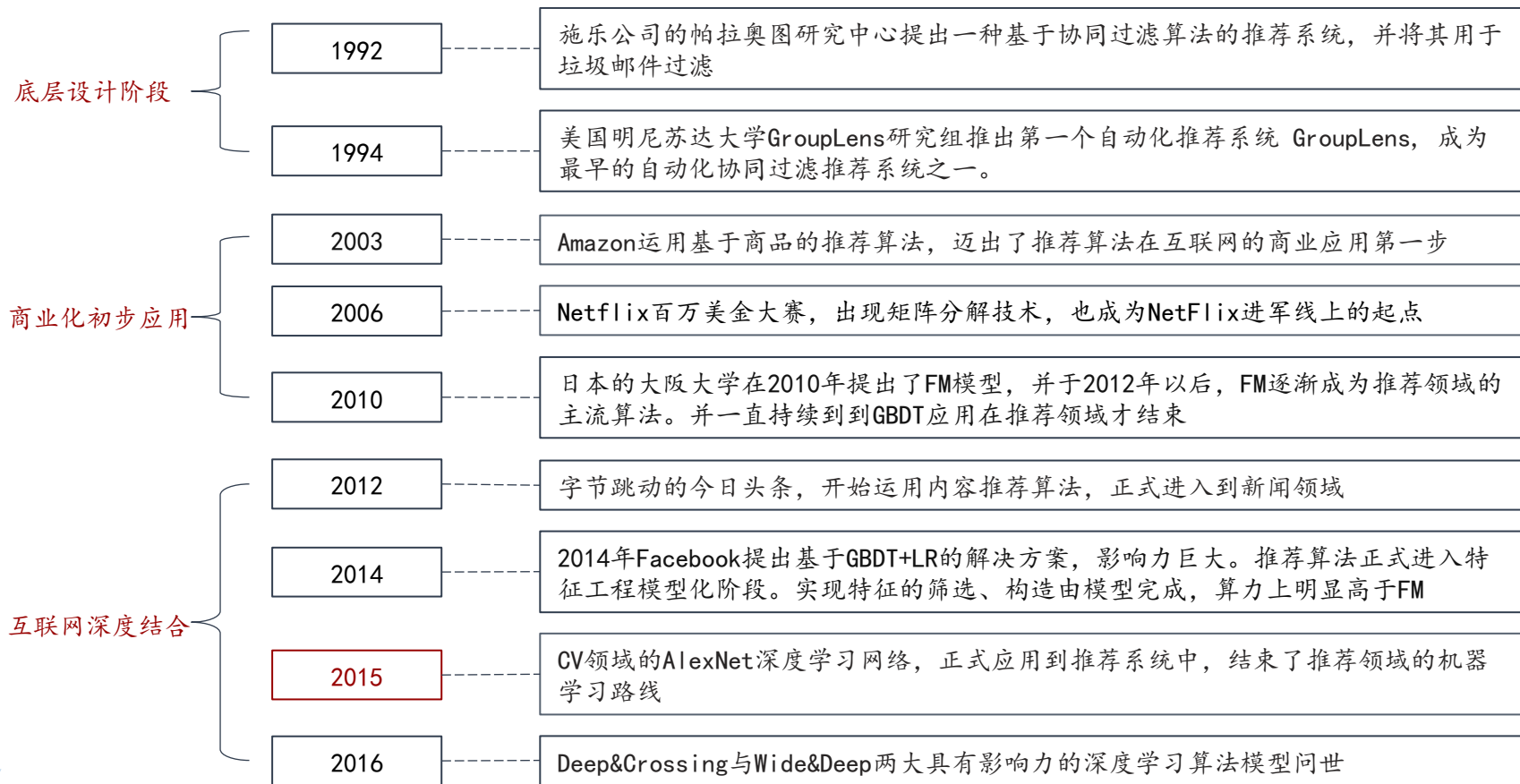
- ❑ 智能推荐算法的不同应用逻辑，在具备优越性的同时也存在着自身的局限性；
- ❑ 组合推荐法通过七种主流的融合方式，在保留各自优势的同时，对于局限性进行互补，通常测算与推荐的结果会有明显的提升，是当前智能推荐的主要应用逻辑

来源：头豹研究院编辑整理

## 发展历程

- 推荐算法的发展经历了底层设计阶段、商业化初步应用阶段于互联网深度结合阶段

### 推荐算法的发展历程



### 头豹洞察

- 推荐算法的发展已有约30年的历史，其技术的发展史也逐步推动着整个算法的在不同行业的应用；
- 1992年至2003年为底层设计阶段，最初的协同过滤算法出现，起到推荐作用，最早应用于垃圾邮件的过滤，并在原有基础上不断进步；
- 2003年至2012年为商业化初步应用阶段，协同过滤技术结合矩阵分解技术，实现了最初级的推荐算法。Amazon与NestFlix也有效利用推荐算法实现了其商业化价值；
- 2012年至今，推荐算法进入了快速的技术革新阶段，也逐步依托技术带来的领先性有效的实现商业化落地，与互联网领域进行了深度的结合。2010年提出的FM技术在2012年成为主流的推荐算法，2014年的GBDT+LR更是带来了技术上的突破，2015年推荐算法由机器学习也正式转为深度学习。推荐算法也自此进入到各行各业，包括字节跳动，由最初的今日头条所在的新闻领域，也迅速扩张到抖音所在的短视频等领域。目前广泛应用于图书、音乐、视频、新闻、电影、地图、网购等等领域

## 政策环境

- 推荐算法在互联网领域的逐步广泛，影响群体庞大，政策上的监管也逐步完善且明确

### 相关法规与政策列举

①缓解推荐算法限制用户选择范围

②保护用户信息与隐私

③严禁非法利用个人数据进行杀熟或垄断

颁布时间	颁布部门	政策名称	重点内容	政策解读
2016年11月	十二届全国人大常委会第二十四次会议	《网络安全法》	第十条建设、运营网络或者通过网络提供服务，应当依照法律、行政法规的规定和国家标准的强制性要求，采取技术措施和其他必要措施，保障网络安全、稳定运行，有效应对网络安全事件，防范网络违法犯罪活动，维护网络数据的完整性、保密性和可用性。	立法保护用户的信息，维护网络安全与健康运行，对推荐内容进行严格审查
2019年1月	十三届全国人大常委会第五次会议	《电子商务法》	第十八条 电子商务经营者根据消费者的兴趣爱好、消费习惯等特征向其提供商品或者服务的搜索结果的，应当同时向该消费者提供不针对其个人特征的选项，尊重和公平保护消费者合法权益	立法限制推荐算法的过分应用，防止消费者或用户的选择范围被算法框定
2021年2月	国务院	《国务院反垄断委员会关于平台经济领域的反垄断指南》	平台经济领域垄断协议是指经营者排除、限制竞争的协议、决定或者其他协同行为。协议、决定可以是书面、口头等形式。其他协同行为是指经营者虽未明确订立协议或者决定，但通过数据、算法、平台规则或者其他方式实质上存在协调一致的行为，有关经营者基于独立意思表示所作出的价格跟随等平行行为除外。	禁止平台通过个人信息进行不正当竞争行为
2021年9月	十三届全国人大常委会第二十九次会议	《数据安全法》	第八条 开展数据处理活动，应当遵守法律、法规，尊重社会公德和伦理，遵守商业道德和职业道德，诚实守信，履行数据安全保护义务，承担社会责任，不得危害国家安全、公共利益，不得损害个人、组织的合法权益。	立法保护数据安全，严谨算法底层数据库中的数据用于非法用途
2021年11月	十三届全国人大常委会第三十次会议	《个人信息保护法》	第二十四条个人信息处理者利用个人信息进行自动化决策，应当保证决策的透明度和结果公平、公正，不得对个人在交易价格等交易条件上实行不合理的差别待遇。通过自动化决策方式向个人进行信息推送、商业营销，应当同时提供不针对其个人特征的选项，或者向个人提供便捷的拒绝方式。	立法保护用户数据，保护个人信息不被泄露或用于非法用途
2022年3月	网信办、工信部、公安部、市场监督管理总局	《互联网信息服务算法推荐管理规定》	针对服务者：算法推荐服务提供者应当坚持主流价值导向，积极传播正能量 针对用户权益：明确了对于算法推荐服务提供者的用户权益保护要求①算法知情权②算法选择权③未成年、老年人、劳动者、消费者的合法权益保障	针对推荐算法与各类应用场景，全方位进行明确的规范

来源：各政府部门官网、头豹研究院编辑整理

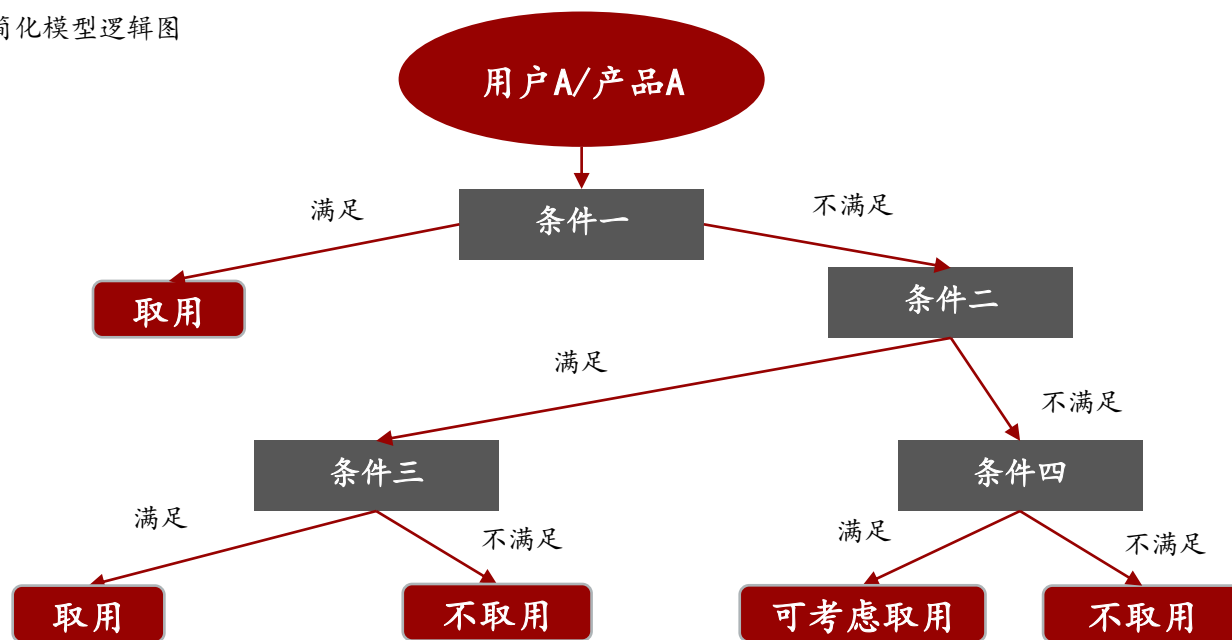


## 技术分析

- 推荐算法经历了机器学习与深度学习的两大阶段，融合了多学科与众多的底层算法模型

### 决策树（例）

简化模型逻辑图



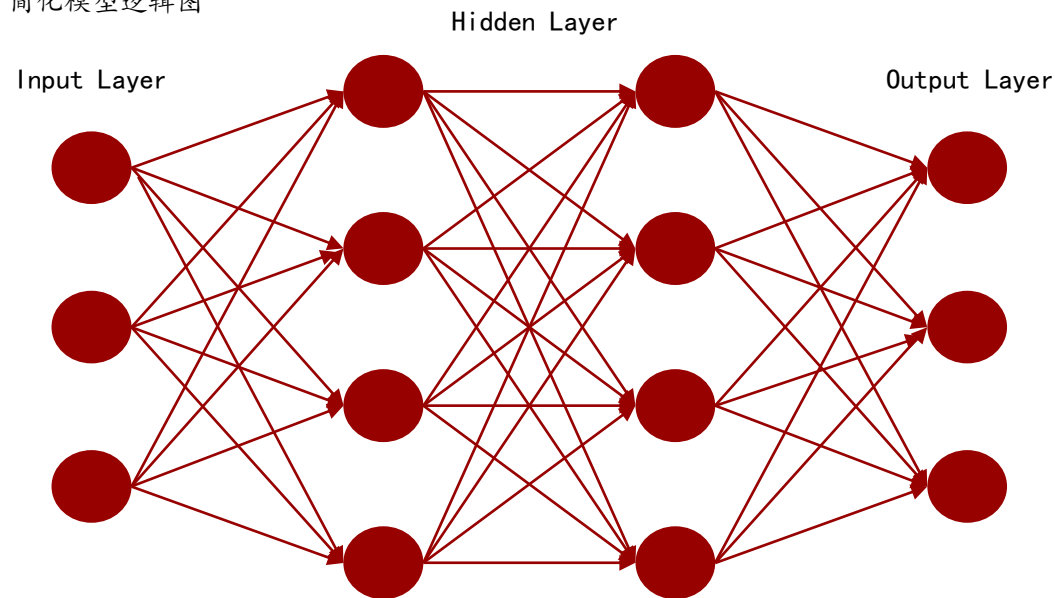
### 头豹洞察

- 决策树是机器学习中较为重要的预测模型之一，用于判断可行性的决策分析方法；
- 决策树的模型逻辑为树形结构，其中每个节点表示在某个属性或特性上的测试，每一个分支代表一个测试输出，最终评估可行性；
- 决策树属于监督学习的常见模型，需要在模型搭建时给定足够的样本，并事先确定样本的分类方式与类别范围，后期机器可以通过学习得到分类器，对新样本给出正确的分类；
- 决策树的可以满足用户与需求的推荐匹配，通过不同的节点测试，最终筛选出最为合适的用户或是产品，进行有效推荐。

来源：头豹研究院编辑整理

### 神经网络（例）

简化模型逻辑图



### 头豹洞察

- 神经网络是深度学习的基础，其中又可分为多种不同的应用算法；
- 神经网络算法模拟人体的神经系统，传输由不同的节点向下传递，作为初始的输出层，经历每一隐藏层的算法分析，向下传递结果，并最终在输出层汇总，显示结果；
- 神经网络算法是更贴近如甚至高于深度学习的算法范围，对于运算复杂度的阈值较高，且计算结果更加准确且智能；
- 神经网络算法应用于推荐领域将推动整个领域的智能化进一步提升。



# 推荐算法应用场景与逻辑分析

- 推荐算法应用场景图谱总览
- 资讯类推荐算法应用分析 - 今日头条
- 短视频类推荐算法应用分析 - 抖音
- 生活类推荐算法应用分析 - 小红书





## 应用范围

- 推荐算法的应用范围广泛，以互联网行业为基础，涵盖众多领域，涉及生活的方方面面

### 算法应用图谱



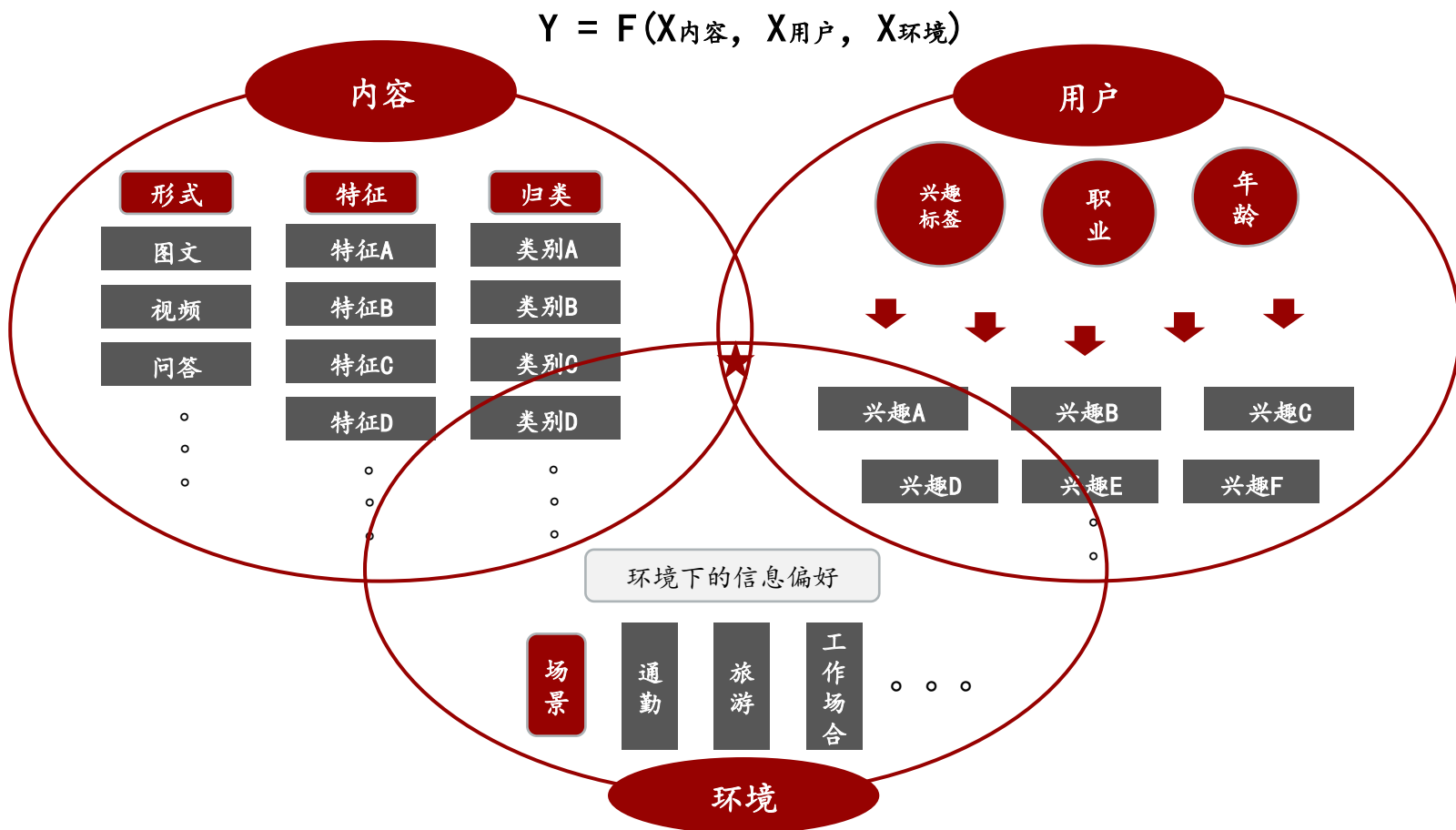
来源：头豹研究院编辑整理

## 资讯类算法逻辑

- 资讯类App与平台均大量采用推荐算法，虽采用逻辑或有不同，但均为实现对用户进行个性化内容的推荐

### 今日头条（例：概述）

资讯推荐系统的本质是解决用户、咨询以及环境的匹配



### 头豹洞察

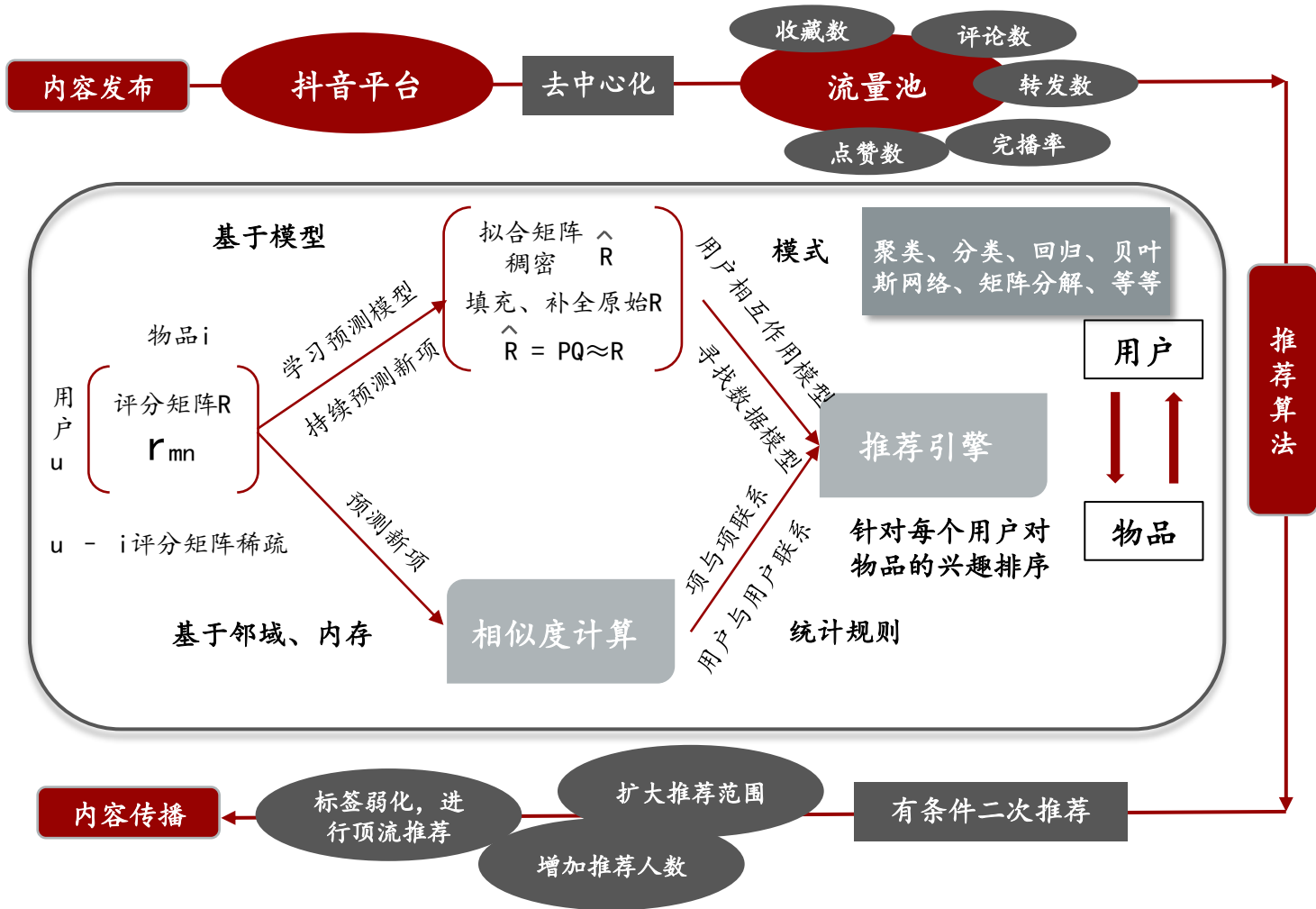
- 推荐算法是通过合理的逻辑运算，为用户推荐最适合的内容，在模型角度上讲，是拟合用户对于内容满意程度的预测函数；
- 资讯类的推荐算法着重考虑三方面因素，对应模型中的三大要素，第一要素为内容，第二要素为用户特征、第三要素为环境特征；
- 内容要素的考虑逻辑，作为资讯类平台，内容无论是形式或是话题都异常丰富，则需要初步划分内容展现形式的基础上，再对内容的特征进行不断的分类与标签化，明确每个内容的特征，方便后续的匹配；
- 用户特征的算法逻辑，在充分考虑到用户主动输入的兴趣标签，以及各客观方面包括职业、年龄、性别等因素，结合后续的个人订阅与查看行为，不断对用户的特征进行完整且深入地描述。明确用户的对于内容的兴趣范围，甚至挖掘用户的隐式兴趣，进行内容的精准匹配
- 环境特征的应用逻辑，结合现实中用户在不同环境下对于内容的需求差异化的特征，例如工作、通勤、旅行的不同环境场景，用户对于内容会具有一定规律性的兴趣转移，把环境因素加入到算法中，使得推荐算法更加的准确；
- 推荐算法结合三大重要要素，在合适的环境下向用户推荐其感兴趣的内容。当然是简化后的逻辑，其真实的算法模型要复杂的多



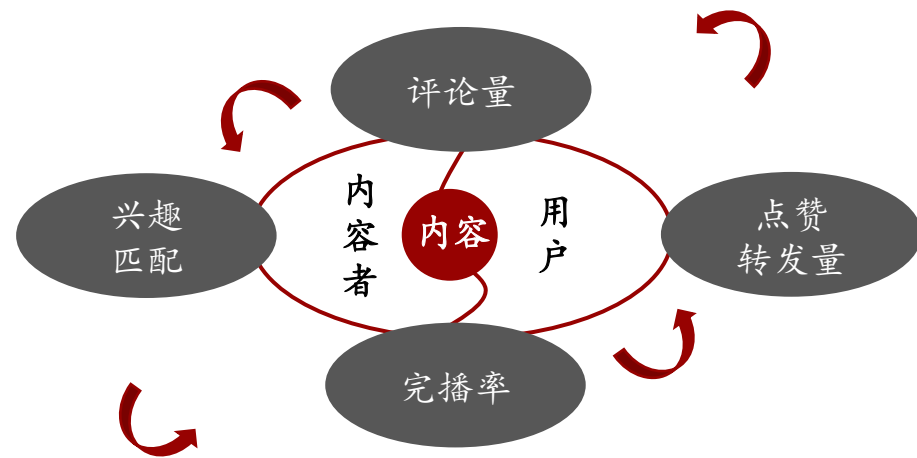
# 短视频类算法逻辑

- 短视频类算法相对复杂，对内容、内容生产者以及用户三方进行综合评估后，给出推荐与流量分配意见

抖音（例：概述）



头豹洞察



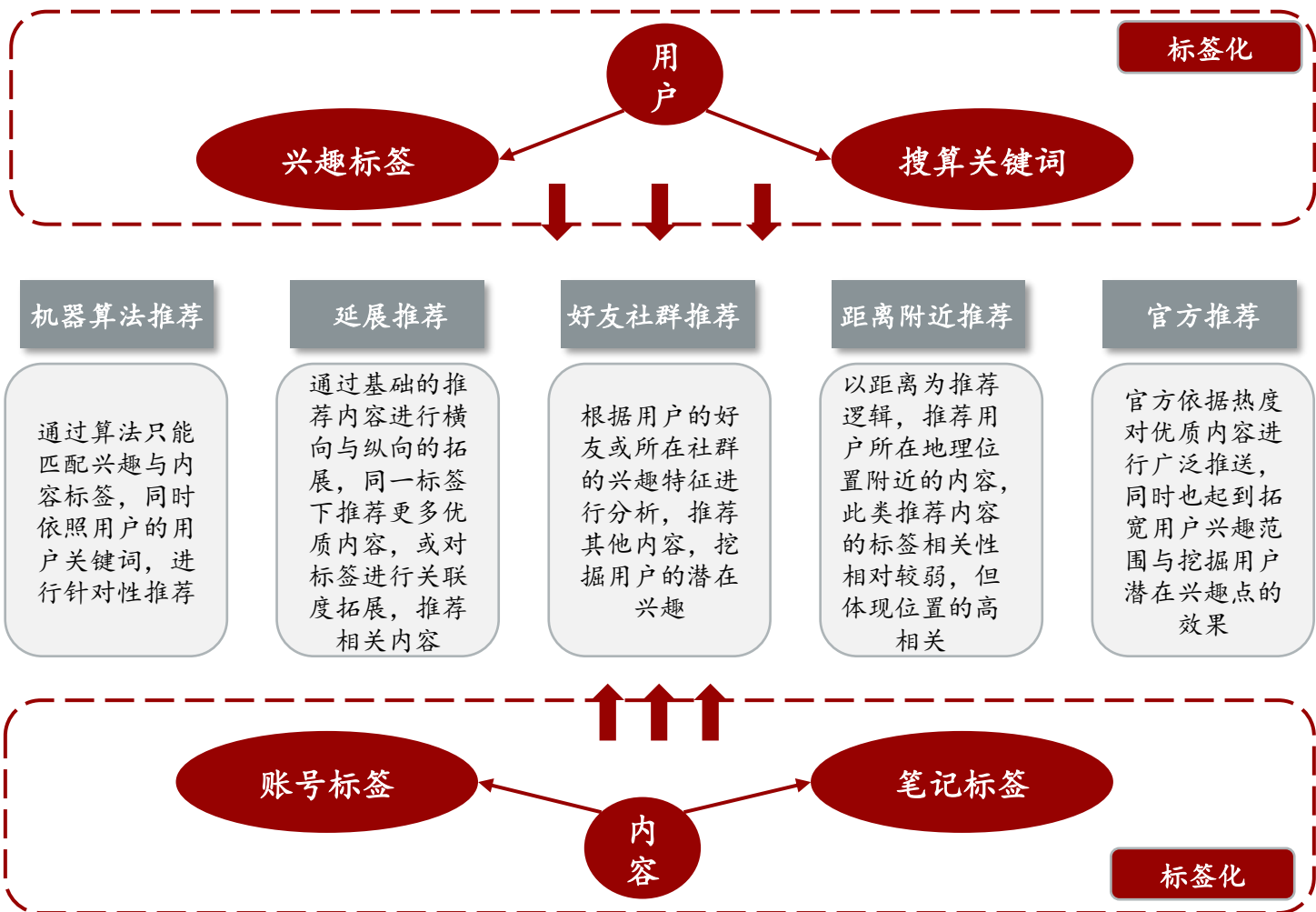
- 抖音推荐的模式，主要以内容发布为起点，进入抖音平台内容流量池，通过智能算法匹配到合适的用户，再通过点赞、转发等的用户反馈行为，评判内容质量；质量高且认可度高的内容会被二次分配到更多的用户，以此类推，若内容成为热点，则会淡化标签的作用，而进行大范围推荐，扩大内容的传播范围；
- 作为日活量与用户基数最大的短视频平台，抖音的推荐算法相对更加复杂且智能化，推荐体系不仅可以推给用户其感兴趣的内容，同时也会更容易发掘潜在的爆点内容，通过其推荐体制，扩大传播范围，也更易成就爆点内容。

来源：头豹研究院编辑整理

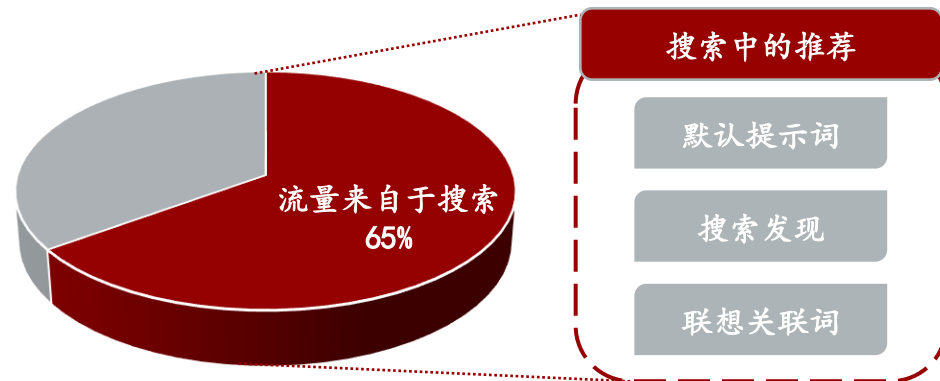
## 生活类算法逻辑

- 生活类平台用户更偏向于主动搜索内容，因此相关应用中智能算法的核心更多以搜索行为为原点进行展开

### 小红书（例：概述）



### 小红书用户习惯



### 头豹洞察

- 小红书的用户习惯，具有明显特征，区别于抖音的主动推荐，小红书的流量更多的来自于用户的主动搜索，占比约65%。说明用户具有较为明确的需求与关注点；
- 小红书的推荐算法主要围绕用户的兴趣标签，以及用户的历史搜索标签。以相关标签为原点，通过智能推荐算法进行匹配，路径分为深入推荐，就搜索标签本身推荐更多的优质内容；以及延展推荐，基于搜索标签进行拓展，扩大推荐内容的范围；
- 同时结合多个其他推荐逻辑，包括附近推荐、社群好友兴趣推荐、官方热点推荐等；
- 小红书在此类推荐算法逻辑的应用下，用户粘性相对较高，但因具有较为明确的兴趣内容，通常平台单次使用时长相对较短。



# 推荐算法应用优劣势分析

- 推荐算法以企业利益为主的负面影响
- 个人信息乱用与泄露的隐患
- 推荐算法有效提高用户的匹配效率
- 推荐算法具备引导优质内容的巨大潜质
- 推荐算法拥有深度帮助用户广阔可能性



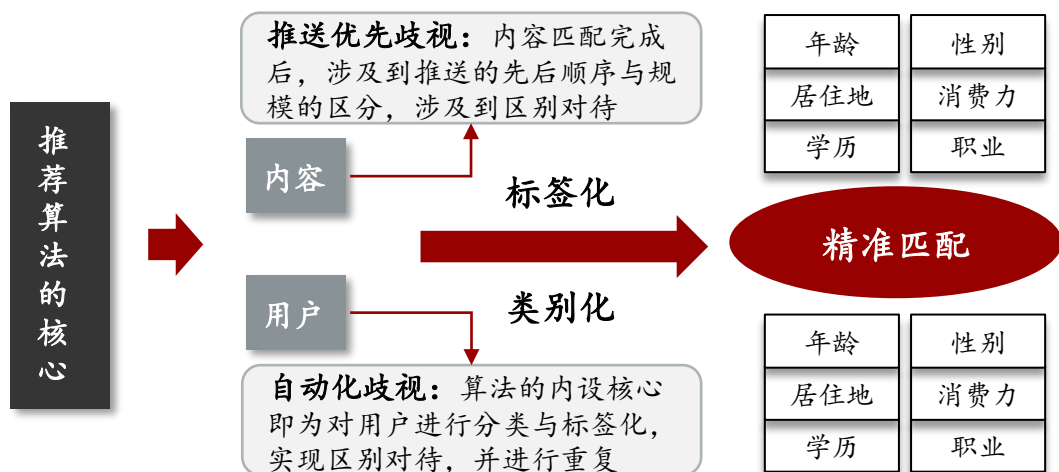
## 企业利益为主的巨大隐患

- 推荐算法引导的商业行为，若是不以消费者利益为出发点，而以企业利益为核心，则会产生较大的社会隐患

### 算法歧视问题

歧视定义（概括）：

基于外在环境和自身情况的差异，使得人们对他人或者特定群体因种族、性别、民族、宗教信仰、社会出身等存在的不同而产生的损害其机会或者待遇平等的概括性固定看法



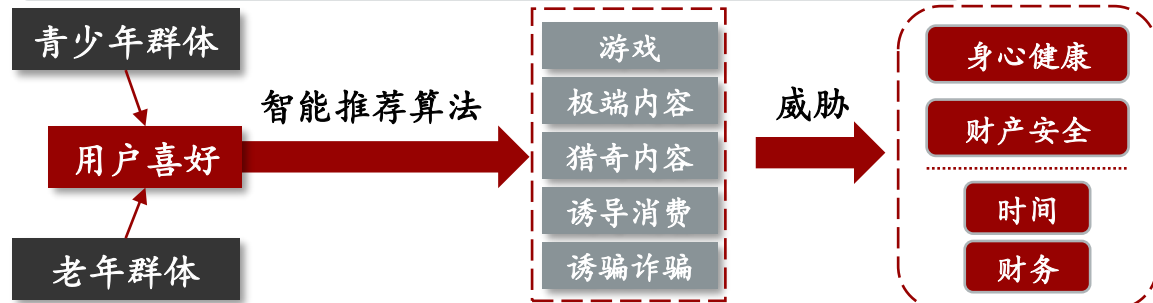
### 头豹洞察

- 算法歧视一定程度上代表推荐算法的内核逻辑，无论对内容或用户进行标签化与分类的行为都涉嫌歧视，但在实际操作中尽量保证公正，使得标签与分类的使用尽可能脱离歧视的范畴；
- 算法通常是平台或其他使用方不会公开的部分，所以在算法中嵌入歧视机制的隐蔽性较强，用户一般难以察觉，或许借助企业自身或监管机构采取相关措施进行监管；
- 推荐算法的应用广泛，若放任歧视，影响巨大，需要多方共同努力，杜绝对于性别、年龄、职业、收入等各类标签的歧视对待，维护社会的公平公正。

来源：头豹研究院编辑整理

### 诱导与沉迷问题

青少年等的未成年群体：价值观与社会观还处于塑造阶段，判断能力还尚不成熟，时间管理与自控能力有待提高，易被猎奇或极端的内容诱导、易被玩乐的内容诱导花费大量时间或是金钱



老年群体：对于诱导性强的内容通常较难以判断，极易成为不法分子诱导的目标群体，导致老年群体发生财物损失的情况

国家网信办、工信部、公安部、国家市场监督管理总局联合发布《互联网信息服务算法推荐管理规定》，明确规定不得利用算法推荐服务诱导未成年人沉迷网络，以及保障老年人依法享有的权益，依法开展涉电信网络诈骗信息的监测、识别和处置

### 头豹洞察

- 算法逻辑即是针对用户的特点与兴趣进行内容匹配。若不加以规范与引导，对于老年人群体与未成年群体的诱导与沉迷问题将不断出现；
- 未成年群体由于自制力与判断力尚未足够成熟，很容易被游戏、猎奇等娱乐或偏移主流价值观的内容吸引，并花费大量时间，而影响自身的成长与学业知识的获取；
- 老年群体对于诱导消费或骗局的识别能力相对较弱，若是算法不加以识别，很容易被不法分子抓住漏洞，而大量的触及到老年人群，使其产生财物的损失，并影响老人的身心健康。

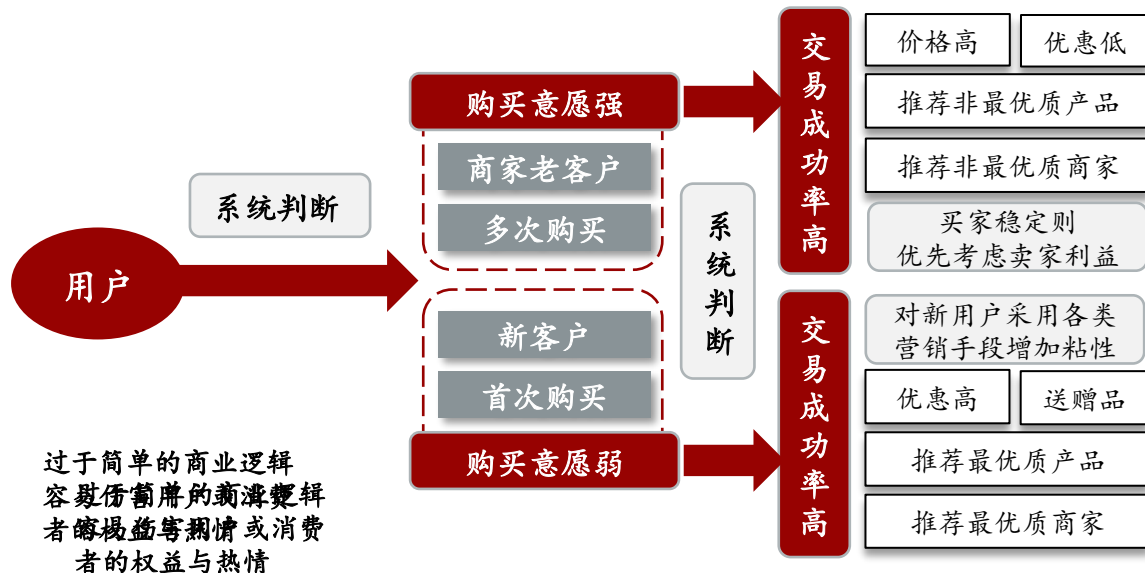


## 个人信息乱用与泄露的隐患

- 应用推荐算法的平台与主体拥有海量的用户信息，而用户个人信息的合理使用与保护需要更加的规范

### 用户个人信息乱用 - 大数据杀熟

特别涉及到购买行为中产品或商家推荐，以及营销策略的智能选用



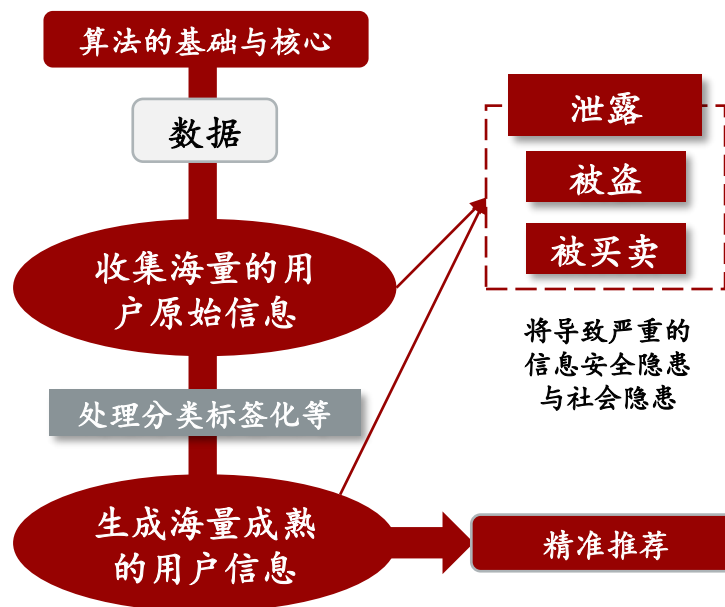
过于简单的商业逻辑  
容易伤害消费者的权益与热情  
者或消费者  
者的权益与热情

#### 头豹洞察

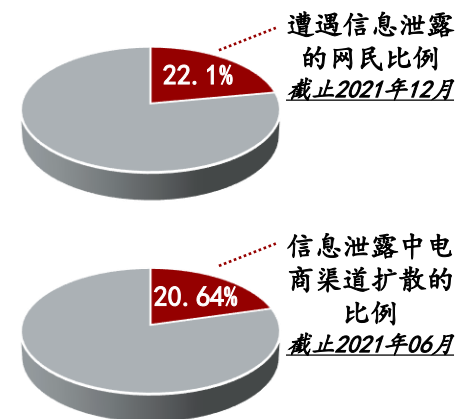
- 应用推荐算法的各类平台，特别是涉及到商品销售与消费行为的，近年来出现大量的大数据杀熟事件，损害消费者的合法权益与消费热情，广受诟病；
- 大数据杀熟的简单逻辑即为，通过判断购买者的消费意愿来评判交易成功率，若是成功率高，则相应保证卖家的权益，隐藏优惠政策，隐蔽调高价格，或是推荐非最高品质的商品或商家，让卖方获得更多的利益，然而对于老客户甚至商家VIP而言，此类行为不仅损害其消费权益，更伤害其消费热情，严重的导致对商家或平台失去信心。此类行为也与消费领域的常规做法相违背，在吸引新客户的同时，更要用心维护老用户，在保证客户存量的同时扩大增量；
- 大数据杀熟已被国家立法禁止，共同维护线上/平台消费的健康环境与可持续发展。

来源：中国互联网络信息中心、头豹研究院编辑整理

### 用户个人信息泄露



不完全统计，我国已知个人信息泄露数量达到55.3亿左右，平均每人泄露4条个人信息



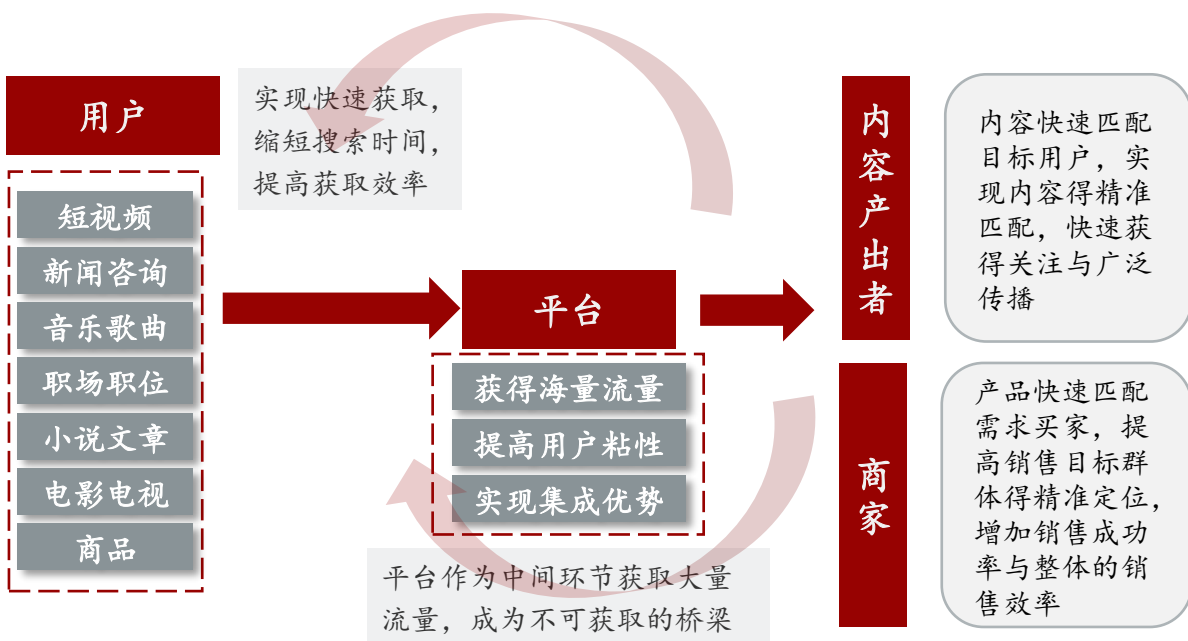
#### 头豹洞察

- 推荐算法的核心技术即为数据，汇总海量的原始个人信息，并通过算法进行分类与标签化区分，无论是原始数据或是处理过的数据，发生泄漏都将会造成巨大的社会隐患；
- 不完全统计，我国已知个人信息泄露数量高达55.3亿左右，据猎聘网统计，中国网络黑产从业者超过40万人，从事网络诈骗的行业人数超过160万，黑色产业年产值超过1000亿元；
- 电商平台作为信息泄露的主要源头之一，依托智能推荐算法，拥有海量的处理后成熟数据，此类数据对于不法分子更容易使用，带来的隐患与可造成的危害程度也将更高。

## 内容与用户的匹配效率明显提高

- 各类平台或机构运用推荐算法，商业运行效率提高，用户快速获取所需，内容与产品供应商实现高效供应

### 效率提高



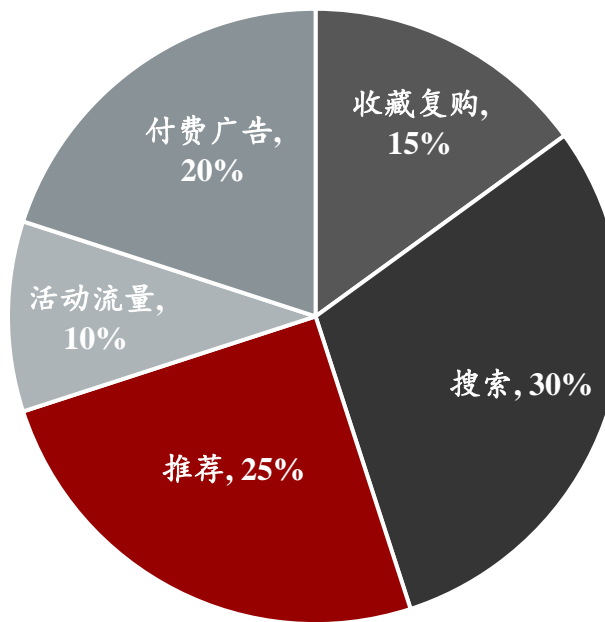
### 头豹洞察

- 推荐算法作为桥梁快速匹配用户与内容或产品，提高整个链条的运行效率；
- 用户角度，效率明显提高，算法通过分析精准推荐用户感兴趣的内容或产品，省去了搜索与查找得时间，快速定位到用户需要；
- 内容产出者与商家角度，对于内容或产品对于目标用户/客户群体的精准投放得以快速实现，加快了匹配效率，使得内容与产品可以快速传播。

来源：广发证券、头豹研究院编辑整理

### 推荐算法的使用率高

2020年拼多多流量大致占比情况



作为主流电商平台之一的拼多多，来自于推荐算法的流量占到平台总量的25%，与搜索流量占比的30%，并没有显著的差距。说明随着算法的智能化、推荐的精准化，用户已经大量使用推荐功能区获取购物所需。

### 头豹洞察

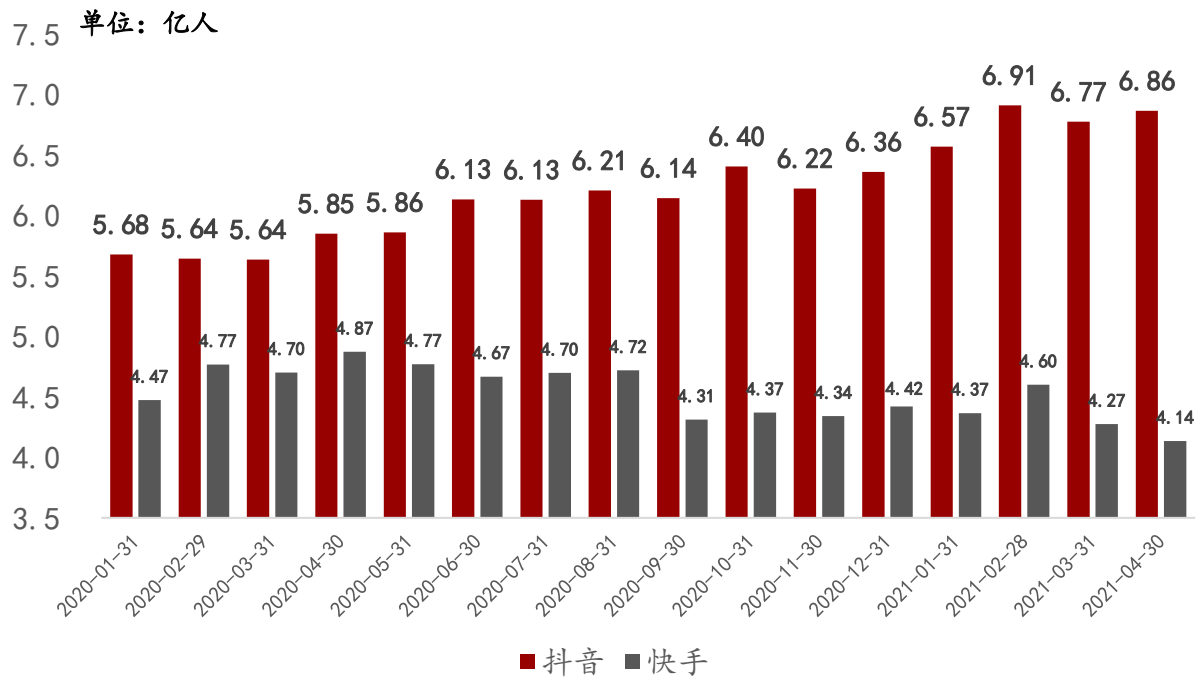
- 各类应用推荐算法的平台，推荐入口获得的流量比例虽有不同，但拥有共同点，便是推荐流量占比显著，在个别平台甚至远超过主动搜索的流量体量；
- 这一现象首先可以说明，推荐算法为底层逻辑的推荐功能，已经获得用户的认可，被广泛使用，或是因为便捷性或是因为精准性；
- 商家与内容制造者也同样收益，在推荐专区被用户或客户快速查阅并选择，加速了传播与经营效率，从而获得积极的正面效果。



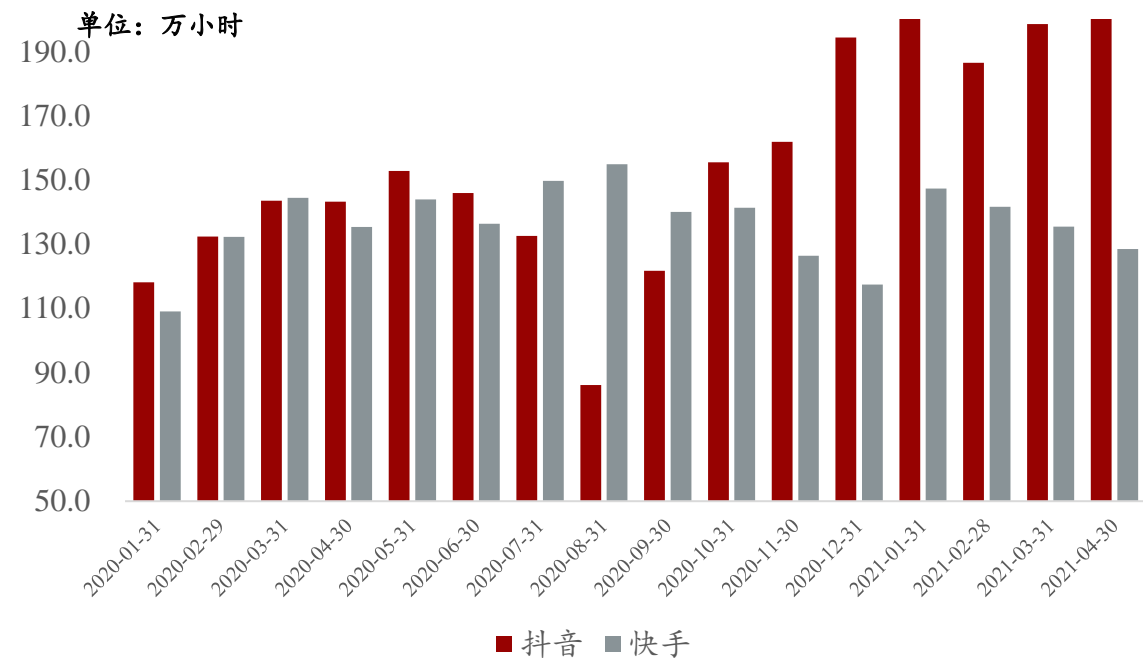
## 优质内容引导的潜力 (1/2)

- 推荐算法所应用各类平台掌握巨大流量，大量的用户基础与累计使用时间，对于社会的影响深远

### 抖音与快手月活量人数 (2020年1月 - 2021年4月)



### 抖音与快手月使用时长 (2020年1月 - 2021年4月)



#### 头豹洞察

- 以抖音与快手两大短视频平台巨头为例，抖音的月活量人数维持在6亿以上，快手的月活量人数维持在4亿以上，流量体积庞大；
- 数据清晰显示两大平台的影响力，拥有庞大的用户群体，且皆具备一定的活跃度，总结来讲有大量的国民在频繁使用两大短视频平台获取内容；
- 抖音与快手等平台以及其他应用推荐算法的资讯类平台，日常向整个社会输出大量的内容，若在内容输出上提高标准，将提高大量国民获取的知识质量

来源：国家统计局、头豹研究院编辑整理

#### 头豹洞察

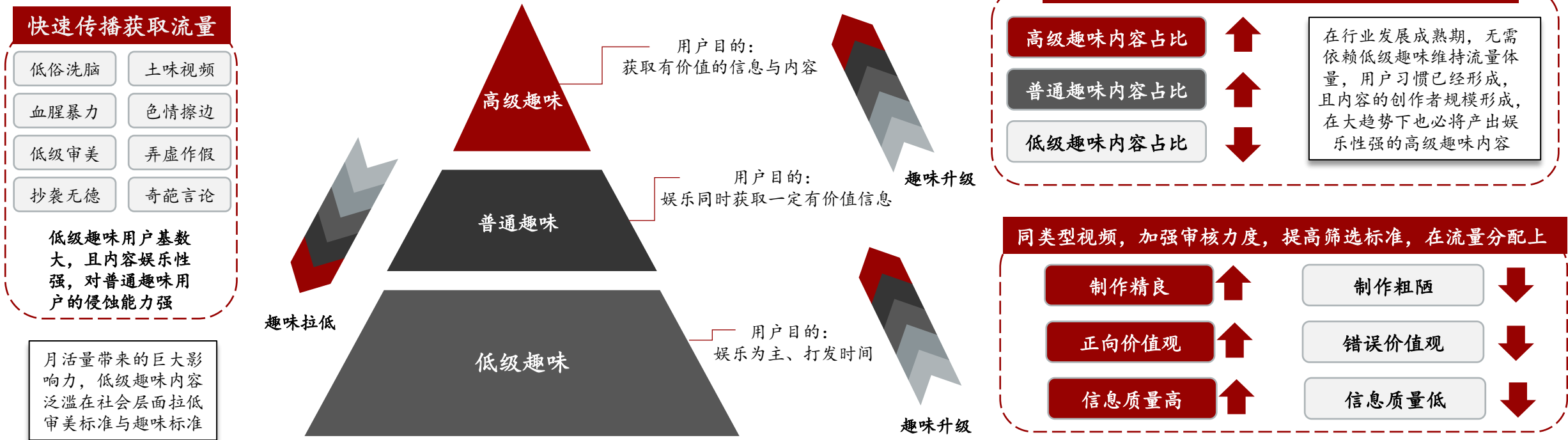
- 抖音月使用时长最高接近200万小时，快手最高也超过150万小时，用户的使用时长相当可观，大量用户花费大量时间使用，因此平台提供的内容对用户甚至是社会都具有深远的影响；
- 短视频平台目前已荣升为使用时长最多的App或平台类型，200万小时约等于228年，整个社会在抖音上花费了大量的时间，同时也获得了海量的信息与内容；
- 作为流量如此巨大的内容传播平台，对于内容的选取与审核格外重要，避免低质的内容与错误的价值观大量传播。平台在获得大量收益的同时，也需要肩负起传播优质内容的责任。

## 优质内容引导的潜力 (2/2)

- 推荐算法若主观加以规范, 可加速实现对于优质内容、高级趣味、正确价值观的广泛传播

### 算法机制改变内容侧重的重要社会意义

单以流量判断的算法, 会导致低级内容得到快速传播



**玩物丧志** 大量时间用于单纯娱乐而毫无进步, 是整个社会生产力的浪费

**扭曲审美** 大量内容通过丑化而增添娱乐性, 传播错误信息

**价值观偏离** 大量低级趣味传播错误价值观, 青少年极易受到影响

**提高素质** 在娱乐的同时获得知识与技能, 提高全民素质与社会生产力

**提高审美** 优质的高级趣味内容提高审美理解, 根本上杜绝低级审美传播

**正确价值观** 帮助梳理正确的价值观, 特别是青少年群体, 提高社会稳定性

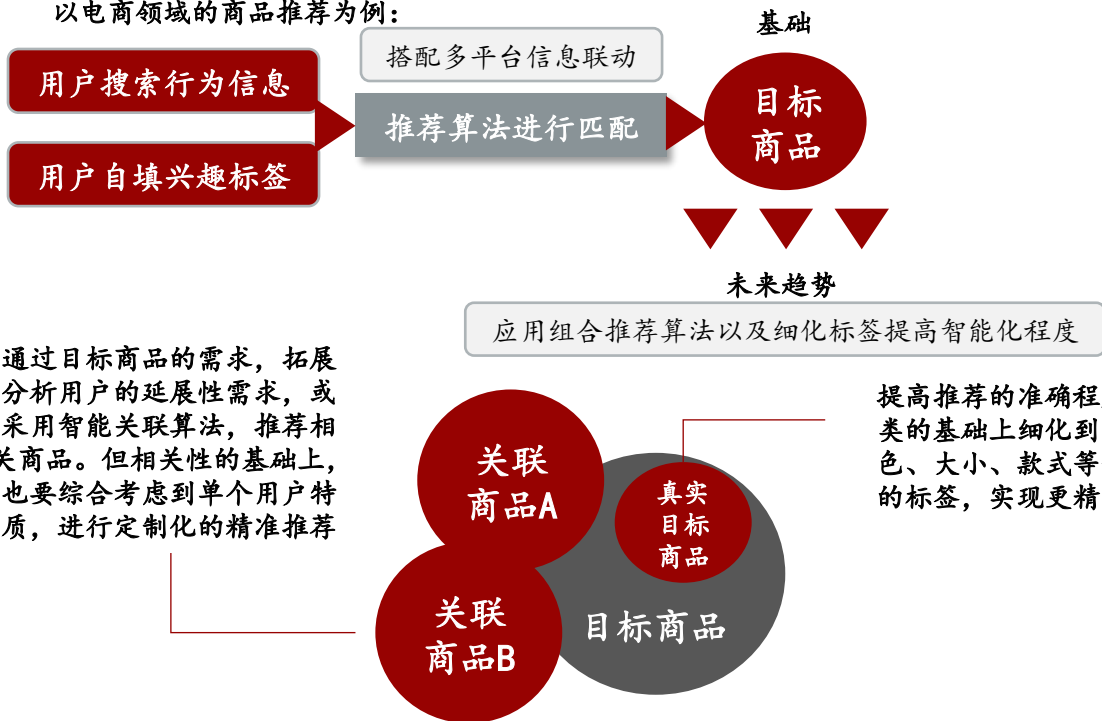


## 深度帮助用户进步的可能性

- 推荐算法基于海量的用户信息，可以通过智能化提高满足用户需求满足效率，甚至挖掘用户潜在需求

### 推荐精准度提高用户需求定位效率

以电商领域的商品推荐为例：



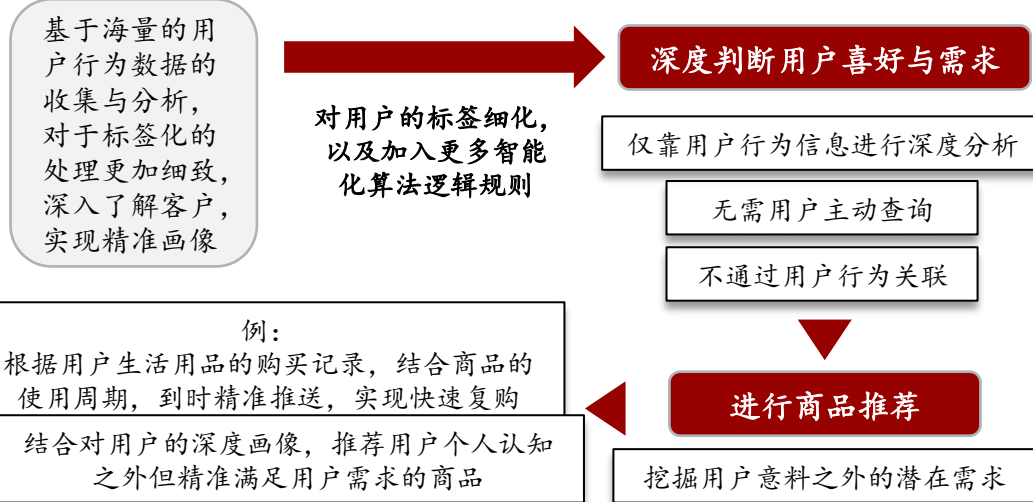
### 头豹洞察

- 电商平台常通过用户的搜索历史以及自填兴趣标签，进行商品的推荐，一定程度上提高了用户对商品查找的效率，但是未来依旧拥有很大的进步空间；
- 采用组合推荐算法或细化标签等措施提高推荐的智能化程度，实现对于用户商品的精准的定位，考虑因素除了价位、品类等常用指标外，提前考虑品牌、款式、大小、颜色等等细化指标
- 关联推荐算法同理可以做到更加精准，切实提高用户的需求定位效率。

来源：头豹研究院编辑整理

### 算法基于数据挖掘用户潜在需求

不基于用户搜索以及自填兴趣标签



### 头豹洞察

- 推荐算法需要提高智能化，做到比用户更了解用户自身的程度；
- 算法可以实现通过对用户信息的海量收集与分析，完全模拟出个人用户的具体画像，除了主观因素的年龄、消费能力、职业、爱好范围等，更可以分析出相对隐蔽的包括消费习惯、个人偏好、行为习惯等，用于更加精准的商品推送；
- 商品推送的智能化可以体现在挖掘用户的潜在兴趣与需求，对当前以搜索为主的重复推送进行升级，延伸出对用户推送非主观选择、但是有购买需求的商品，提高用户的购物效率，一定程度上可以提高用户的生活质量；同时也包括算法帮助用户合理规划购买频率，对于生活必需品，结合使用周期，进行提前推荐。



# 知名企业分析

## ➤ 字节跳动

- 企业背景分析
- 字节跳动成绩分析
- 新闻领域推荐算法的局限性分析

## ➤ 阿里巴巴

- 企业背景分析
- 淘宝推荐算法逻辑与优劣势分析





• 字节跳动发展迅猛，结合智能推荐算法的应用，覆盖多个行业领域，拥有众多知名品牌，企业价值不断提高

### 融资大事件

融资日期	融资轮次	企业估值	融资金额	融资比例	投资方
2012年3月9日	天使轮	X	X	X	晨兴资本、顺为资本、以及天使投资人
2012年7月1日	A轮	X	100万美元	X	海纳亚洲
2013年9月1日	B轮	X	数千万美元	X	DST Global、奇虎360
2014年6月1日	C轮	5亿美元	1亿美元	20%	微创投、红杉资本
2016年12月30日	D轮	110亿美元	10亿美元	9%	红杉资本、建银国际
2017年8月1日	E轮	220亿美元	20亿美元	9%	General Atlantic
2018年10月20日	Pre-IPO	750亿美元	40亿美元	5%	软银中国资本、春华资本、KKR、云锋基金
2020年3月30日	战略投资	1000亿美元	X	X	老虎全球管理基金

### 头豹洞察

- 字节跳动自2012年成立以来，经历了多轮的融资事件，深受资本市场的信赖，屡次获得业内知名机构领头，公司估值稳步上涨；
- 字节跳动2020年的战略投资阶段，公司估值已经达到1000亿美元，正式迈入估值超千亿美金的巨头行列，发展十分迅猛；
- 公司近期虽因为国际环境与经济环境影响，上市暂缓，但可预料资本市场对其抱有的期待态度。

来源：公司官网与公开资料、头豹研究院编辑整理

### 旗下品牌及上线（投资）时间线（例举）



### 头豹洞察

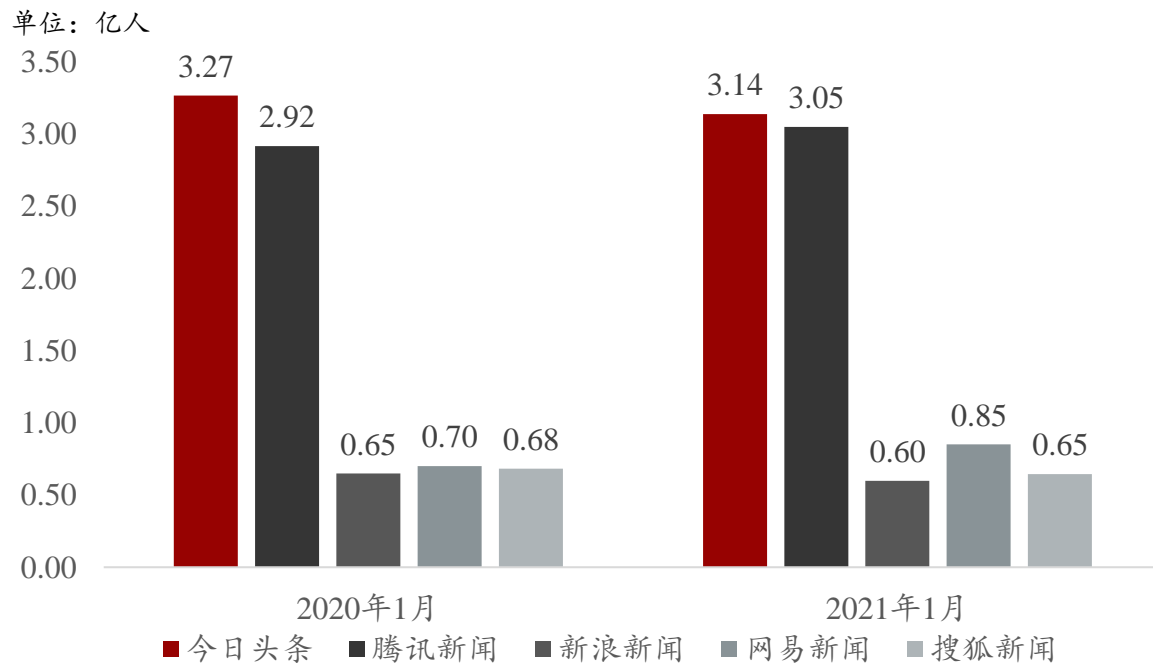
- 字节跳动在近十年的时间内，以今日头条为起点，充分发挥推荐算法优势，建立起庞大的商业帝国，涉及众多不同领域，16年上线的抖音也已成为最具商业价值的App之一；
- 字节跳动、较为主打的依然是内容与用户的匹配，包括短视频平台（抖音）、内容分发平台（今日头条）、PUGC视频平台（西瓜视频）、汽车领域的信息交流平台（懂车帝）、体育类社区（虎扑），推荐算法的商业价值得以凸显。

- 字节跳动通过应用了推荐算法的今日头条，在新闻资讯领域打破固有模式，突破重围飞速扩充市场份额

### 今日头条打破壁垒



### 新闻门户的月活量情况 (2020年1月与2021年1月)



#### 头豹洞察

- 2012年之前的互联网新闻领域基本被四大门户所瓜分，网易新闻、新浪新闻、搜狐新闻、腾讯新闻，市场集中度极高，也形成了坚实的行业壁垒，但今日头条融合算法与新闻资讯，打破了市场原有模式，并获得了市场的认可；
- 互联网初期的新闻门户依然采用过去的编辑制模式，与报纸、杂志的展现形式类似，门户中集合大量新闻资讯，并按照板块分类，用户需要查找自己所需或感兴趣的信息；
- 今日头条融合推荐算法，采用算法制模式，使得用户可以快速地获取感兴趣的内容，也促使新闻资讯门户由制式化时代转型进入到定制化的时代，具有众多优势也深受用户追捧。

来源：国家统计局、头豹研究院编辑整理

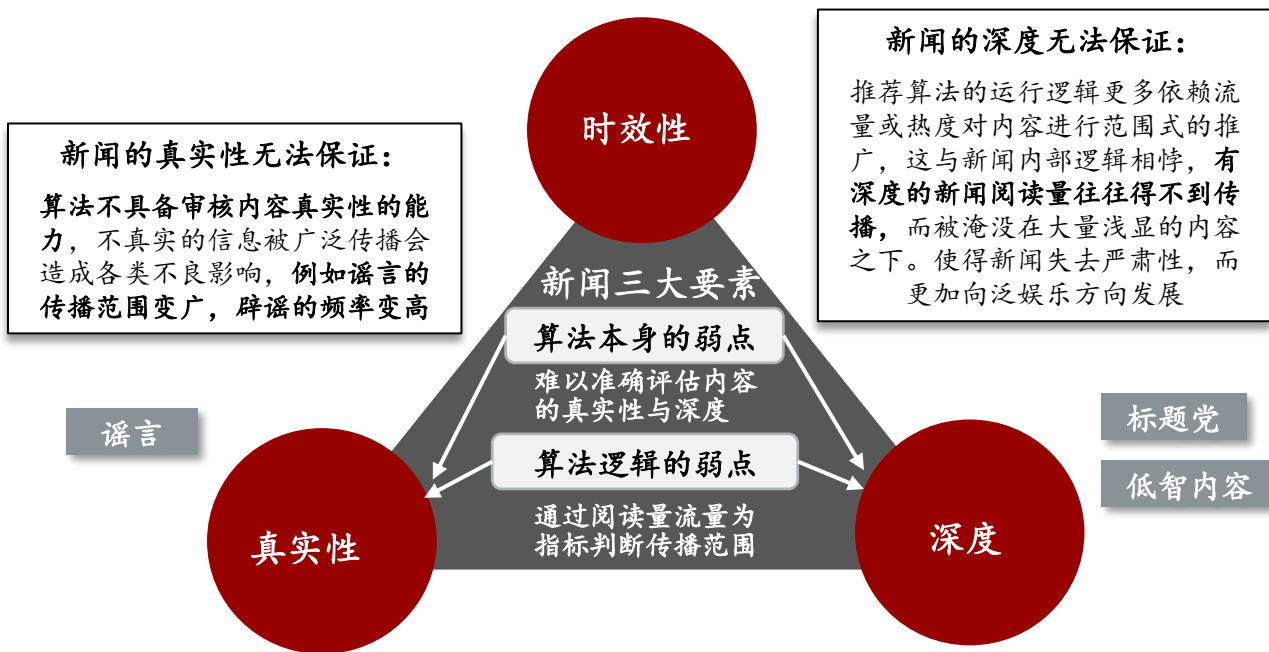
#### 头豹洞察

- 今日头条已然成为目前月活量最高的新闻门户，月活量基本维持在3亿左右，不仅彻底的打破了过去四大门户的时代，更是站在了行业的领先地位；
- 今日头条的发展离不开推荐算法，算法与新闻咨询的有效结合，不仅可以帮助用户快速获取所需或感兴趣的内容与资讯，更是可以有效地规划平台收集的海量资讯，合理利用门户，有效提高用户以及平台双方的效率；
- 今日头条的成功也说明推荐算法因其具备十分可观实际意义，得到了用户与社会的广泛认可。



今日头条结合算法颠覆了传统的新闻领域，取得显著成就，但模式更改带来的娱乐化趋势却引人深思

### 算法在新闻领域应用的局限性

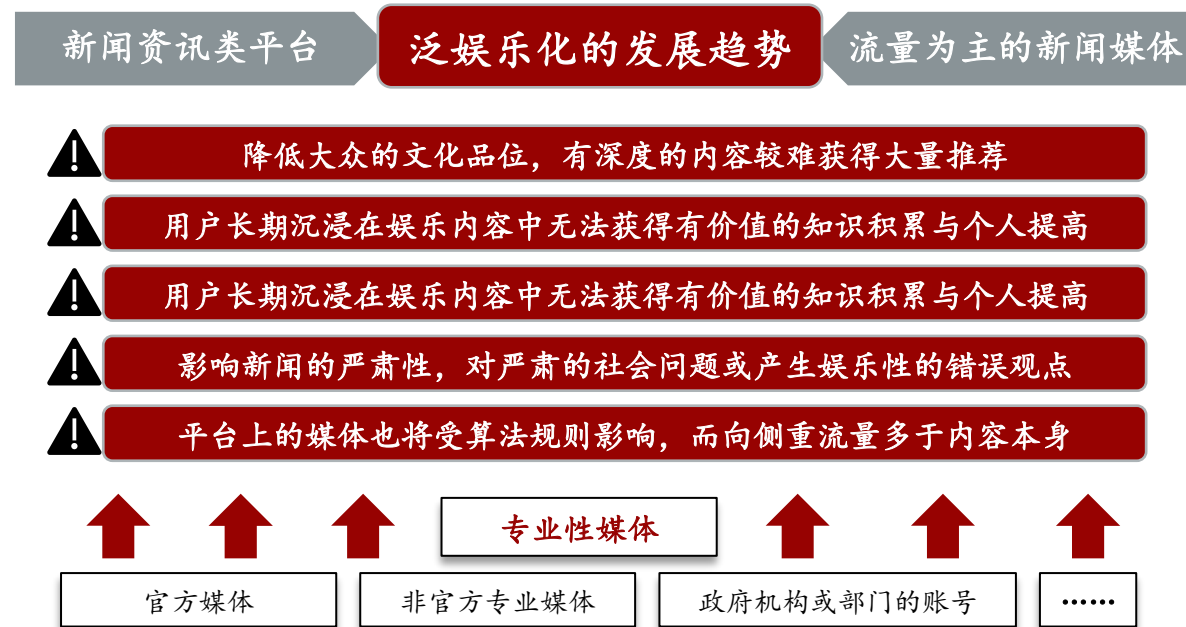


### 头豹洞察

- 推荐算法自身的局限性导致与新闻资讯领域融合时，会带来一定的负面影响；
- 算法作为机器工具，对于内容的审核，包括真实性与深度的判断能力有限，没有办法在技术层面去除谣言被传播的风险，同时算法对于内容深度的判断难以做到精确，极易导致同一话题下高质内容与低质内容的优先级相同；
- 推荐算法的应用逻辑中，较多采用以流量为标准调整传播范围的模式，深度的内容往往相对严肃且枯燥，导致阅读、转发等流量指标相对较低；而低智、搞笑、内涵等泛娱乐的内容更容易获得流量。此消彼长劣币驱逐良币，存在优质的专业内容被大量娱乐化内容淹没的风险。

来源：头豹研究院编辑整理

### 新闻门户泛娱乐化的趋势

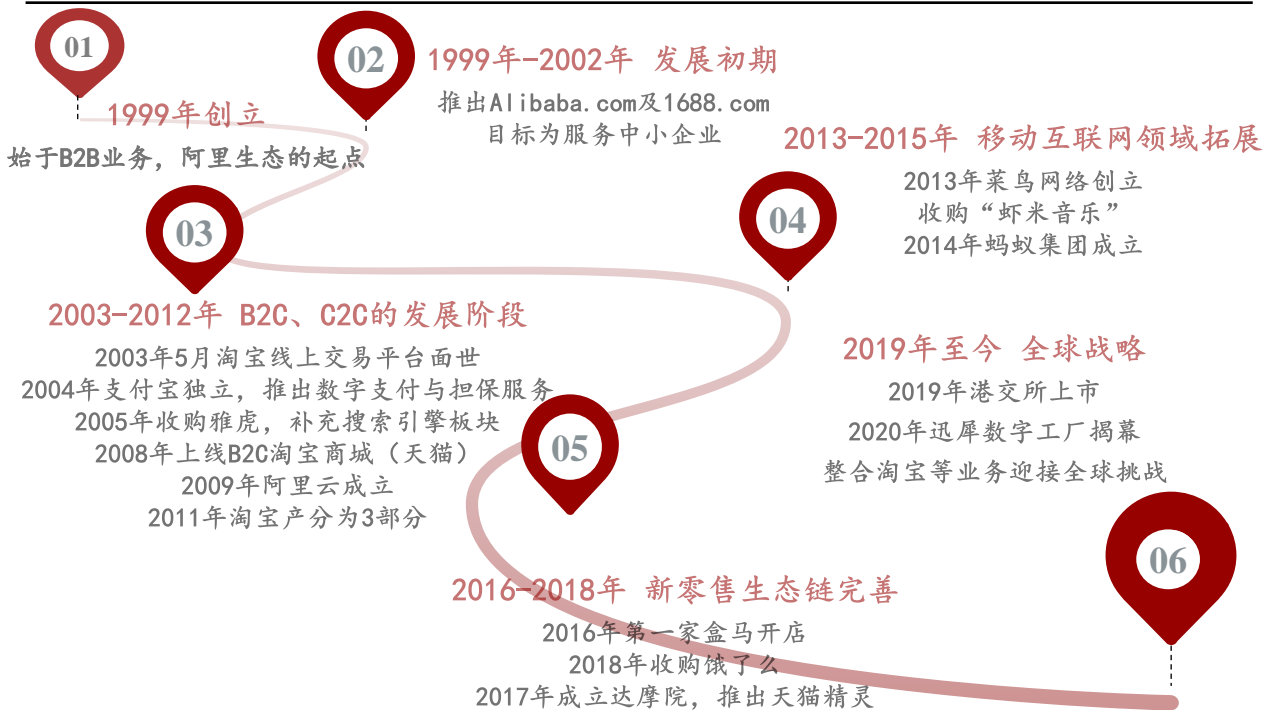


### 头豹洞察

- 推荐算法融合进入新闻资讯领域，难免会产生泛娱乐化趋势。并影响部分媒体亲近流量而弱化内容的专业性，与新闻挖掘产出的投入；
- 新闻泛娱乐化趋势在影响内容专业性的同时，会产生对用户与社会的一定负面影响，或将出现娱乐性强内容覆盖专业性强的现象发生，导致用户的信息获取量大但低质；或会影响部分社会话题的严肃性，以及导致谣言等不真实内容得到广泛且快速的传播；
- 或将形成由专业性媒体作支撑，提供专业严肃内容，流量媒体自媒体提供娱乐内容的格局。

• 阿里巴巴作为国内最知名的互联网企业之一，其产品生态逐步完善，与推荐算法的融合也在逐渐加深

## 发展历程与大事件



## 商业版图



### 头豹洞察

- 阿里巴巴在20多年的发展中，逐渐优化战略方向，不断充实完善商业版图，现已成为全球化、多板块、高科技的行业最领先企业之一；
- 阿里巴巴由最初的B2B业务方向逐步拓展到B2C、C2C领域；由交易平台、零售消费领域逐步拓展成为覆盖零售、金融、制造、技术、能源等领域的庞大商业帝国；
- 推荐算法无论是在最初的淘宝，还是目前的淘宝、天猫等系列平台，都起到重大的积极作用。

### 头豹洞察

- 阿里巴巴的商业版图庞大，涉及到零售、批发、数字媒体娱乐、生活服务、物流、金融与支付、营销服务与数据管理、云计算以及其他创新领域；
- 最核心的业务板块还是以淘宝、天猫所在的消费零售领域，中国作为主市场分为零售与批发两块业务，同时在全球化布局的战略下，配合海外与全球化的跨境零售业务；
- 阿里巴巴基本核心消费领域与生活服务领域对于推荐算法的应用与融合逐步加深，也为用户提供了巨大便利。

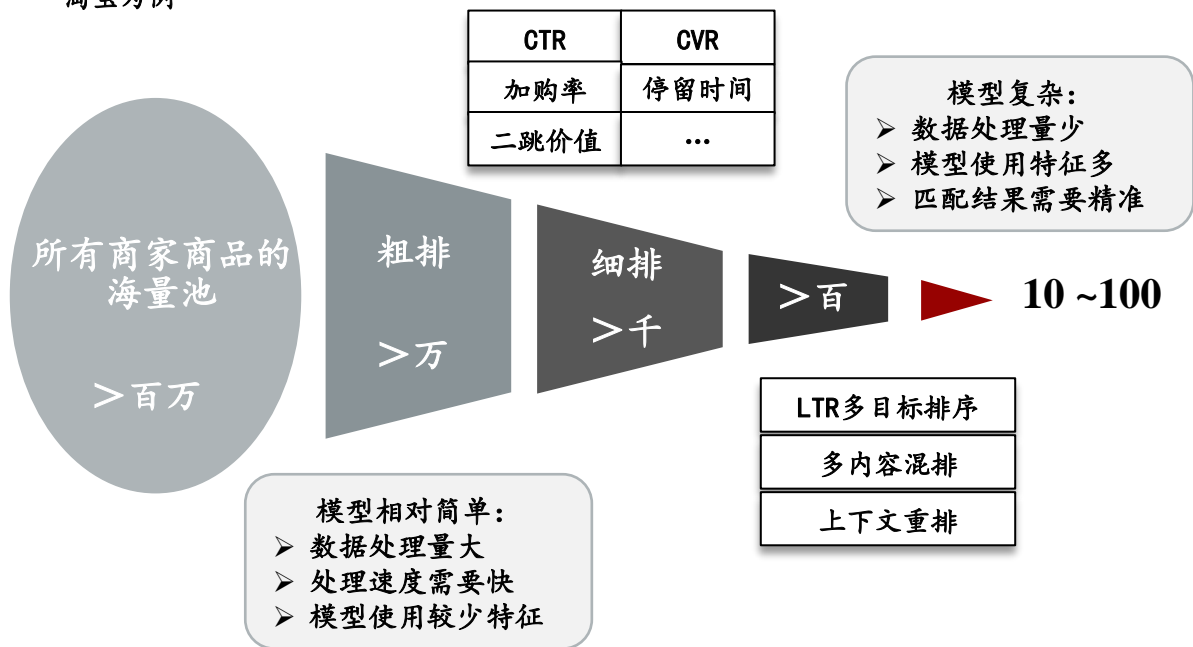
来源：公司官网与公开资料、头豹研究院编辑整理



- 淘宝作为阿里巴巴的核心品牌，GMV领域在同领域持续领先，其中推荐算法的应用也为用户提供实际的便利

## 算法逻辑 (概述)

淘宝为例

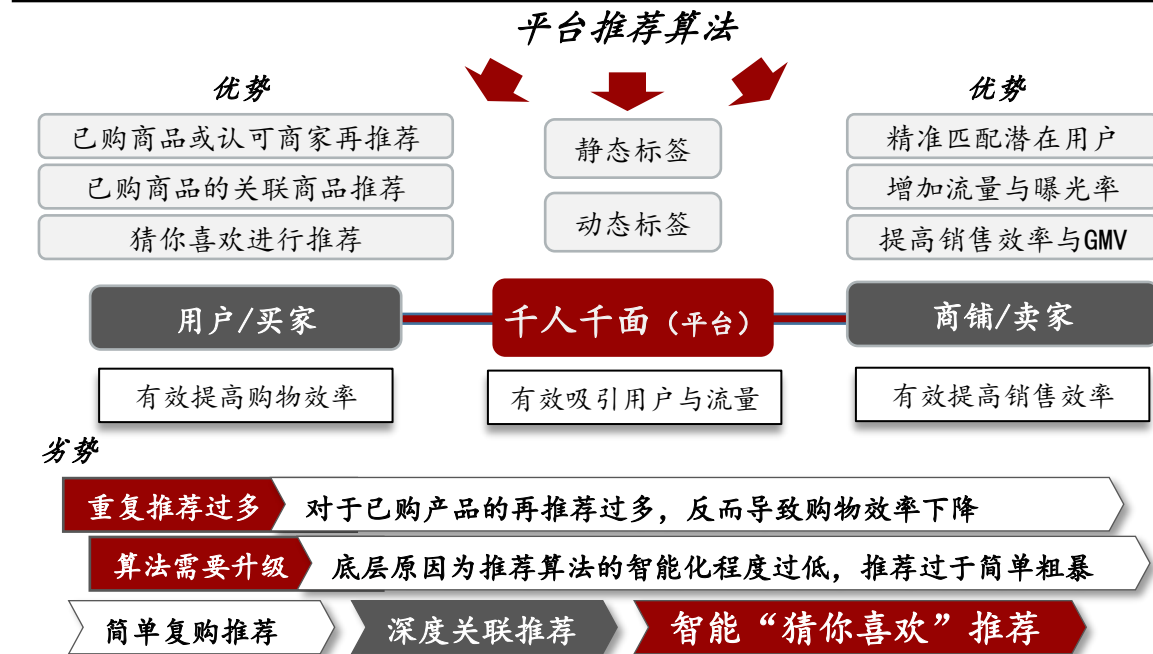


## 头豹洞察

- ❑ 以淘宝为首的电商平台，大量依托推荐算法的应用，提高商品与用户的匹配效率；
- ❑ 推荐算法的核心逻辑就是在数以百万计的商品中，针对用户信息的分析结果进行匹配，首先经过采用相对简单的模型进行召回，可以保证在海量的信息中快速匹配出不需过于精准的目标，再通过相对复杂的模型逐步进行更高精度的排序与筛选；
- ❑ 推荐算法的应用最终呈现就是在平台购物首页上可以罗列10-100个用户最有可能购买的产品。

来源：头豹研究院编辑整理

## 淘宝推荐算法的优劣势



## 头豹洞察

- ❑ 推荐算法在淘宝以及各大电商平台的运用优劣势均较为明显；
- ❑ 优势在于突破过去商店式或黄页式浏览方式，正式开启了千人千面的网购时代，通过标签与行为的分析进行匹配与推荐，有效提高买家的购物效率与卖家的销售效率，而此优势也为帮助平台大量的吸引客户与流量，可谓三方皆赢的局面；
- ❑ 劣势在于算法的智能化程度有待提高，特别针对已购产品的同类产品过度推荐的现象，导致推荐界面千篇一律，平台需要改进算法，进行更深层次的关联推荐，以及更智能化的“猜你喜欢”推荐，使得推荐功能更加真实有效，满足用户对于购物体验的优化需求。

# 方法论

- ◆ 头豹研究院布局中国市场，深入研究19大行业，持续跟踪532个垂直行业的市场变化，已沉淀超过100万行业研究价值数据元素，完成超过1万个独立的研究咨询项目。
- ◆ 头豹研究院依托中国活跃的经济环境，研究内容覆盖整个行业发展周期，伴随着行业内企业的创立，发展，扩张，到企业上市及上市后的成熟期，头豹各行业研究员积极探索和评估行业中多变的产业模式，企业的商业模式和运营模式，以专业视野解读行业的沿革。
- ◆ 头豹研究院融合传统与新型的研究方法论，采用自主研发算法，结合行业交叉大数据，通过多元化调研方法，挖掘定量数据背后根因，剖析定性内容背后的逻辑，客观真实地阐述行业现状，前瞻性地预测行业未来发展趋势，在研究院的每一份研究报告中，完整地呈现行业的过去，现在和未来。
- ◆ 头豹研究院密切关注行业发展最新动向，报告内容及数据会随着行业发展、技术革新、竞争格局变化、政策法规颁布、市场调研深入，保持不断更新与优化。
- ◆ 头豹研究院秉承匠心研究，砥砺前行的宗旨，以战略发展的视角分析行业，从执行落地的层面阐述观点，为每一位读者提供有深度有价值的研究报告。

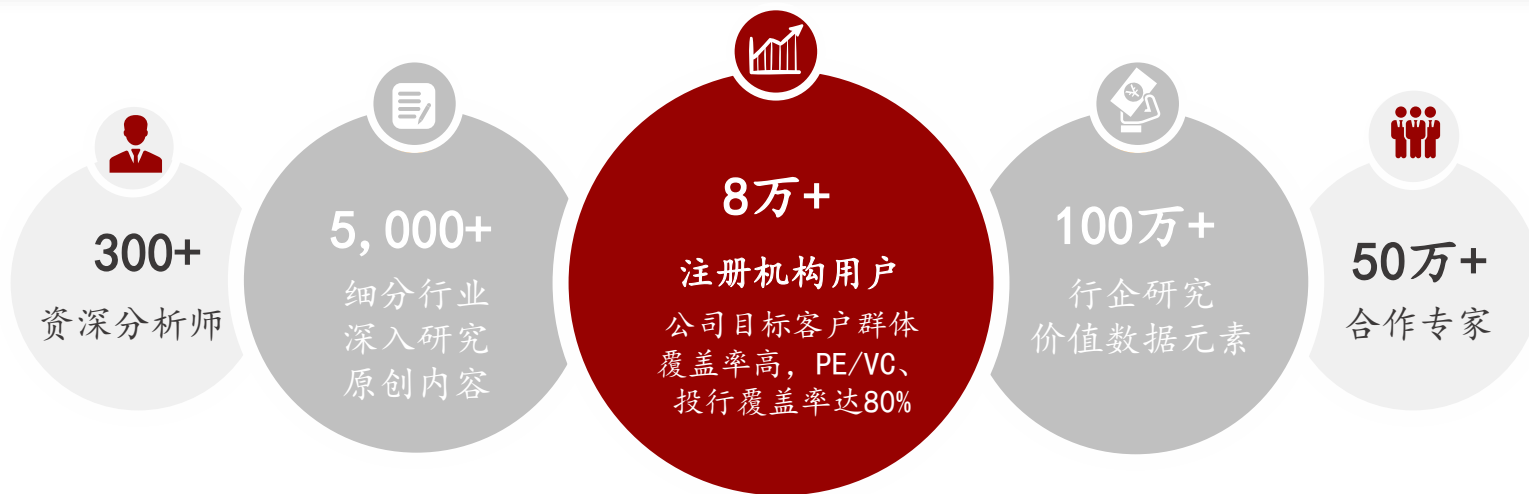


# 法律声明

- ◆ 本报告著作权归头豹所有，未经书面许可，任何机构或个人不得以任何形式翻版、复刻、发表或引用。若征得头豹同意进行引用、刊发的，需在允许的范围内使用，并注明出处为“头豹研究院”，且不得对本报告进行任何有悖原意的引用、删节或修改。
- ◆ 本报告分析师具有专业研究能力，保证报告数据均来自合法合规渠道，观点产出及数据分析基于分析师对行业的客观理解，本报告不受任何第三方授意或影响。
- ◆ 本报告所涉及的观点或信息仅供参考，不构成任何证券或基金投资建议。本报告仅在相关法律许可的情况下发放，并仅为提供信息而发放，概不构成任何广告或证券研究报告。在法律许可的情况下，头豹可能会为报告中提及的企业提供或争取提供投融资或咨询等相关服务。
- ◆ 本报告的部分信息来源于公开资料，头豹对该等信息的准确性、完整性或可靠性不做任何保证。本报告所载的资料、意见及推测仅反映头豹于发布本报告当日的判断，过往报告中的描述不应作为日后的表现依据。在不同时期，头豹可发出与本报告所载资料、意见及推测不一致的报告或文章。头豹均不保证本报告所含信息保持在最新状态。同时，头豹对本报告所含信息可在不发出通知的情形下做出修改，读者应当自行关注相应的更新或修改。任何机构或个人应对其利用本报告的数据、分析、研究、部分或者全部内容所进行的一切活动负责并承担该等活动所导致的任何损失或伤害。

# 头豹研究院简介

- ◆ 头豹是中国领先的原创行企研究内容平台和新型企业服务提供商。围绕“协助企业加速资本价值的挖掘、提升、传播”这一核心目标，头豹打造了一系列产品及解决方案，包括：**报告/数据库服务、行企研报服务、微估值及微尽调自动化产品、财务顾问服务、PR及IR服务**，以及其他以企业为基础，利用大数据、区块链和人工智能等技术，围绕产业焦点、热点问题，基于丰富案例和海量数据，通过开放合作的增长咨询服务等
- ◆ 头豹致力于以优质商业资源共享研究平台，汇集各界智慧，推动产业健康、有序、可持续发展



备注：数据截止2022.6

## 四大核心服务

### 研究咨询服务

为企业提供定制化报告服务、管理咨询、战略调整等服务

### 企业价值增长服务

为处于不同发展阶段的企业，提供与之推广需求相对应的“内容+渠道投放”一站式服务

### 行业排名、展会宣传

行业峰会策划、奖项评选、行业白皮书等服务

### 园区规划、产业规划

地方产业规划，园区企业孵化服务



# 研报阅读渠道

◆ 头豹官网：登录 [www.leadleo.com](http://www.leadleo.com) 阅读更多研报

◆ 头豹小程序/微信小程序：搜索“头豹”手机可便捷阅读研报

◆ 头豹交流群：可添加企业微信13080197867，身份认证后邀您进群

## 详情咨询



客服电话

400-072-5588



上海

王先生：13611634866

李女士：13061967127



深圳

李先生：13080197867

李女士：18049912451



南京

杨先生：13120628075

唐先生：18014813521