

# 淘系千人千面 分发体系全拆解

搜索&推荐技术应用系列



- 搜索和推荐如何支撑分发体系关键环节
- 上千个搜索业务背后的实践经验总结
- 个性化推荐的典型问题和经典打法





钉钉扫码加入  
“推荐与搜索技术交流群”



微信扫码关注  
“搜索与推荐技术”  
公众号



扫码关注订阅  
“搜索与推荐技术”社区



阿里云开发者“藏经阁”  
海量免费电子书下载

# | 目录

1.首篇	4
1.1 搜索推荐广告三位一体的在线服务体系怎样支撑千人千面的分发体系	4
2.搜索篇	14
2.1 “业务指标” 衡量电商搜索引擎的优劣	14
2.2 商品搜索“无结果率高，体验差” 怎么办？	17
2.3 导购电商怎样“提升 CTR、CVR 业务指标”	20
2.4 “减少输入，更快找到用户所想”	24
2.5 “随便逛逛，看大家都在搜什么？追踪潮流和热门”	27
3.推荐篇	29
3.1 推荐场景化构建与业务价值	29
3.2 智能推荐在电商的经典打法与实践	35
3.3 商品推荐的核心问题、架构和应用效果	41
3.4 常规和大促，怎样转变思路应对不同场景	48
4.视频课程	52
4.1 搜索篇	52
4.2 推荐篇	54

# 1. 首篇

## 1.1 搜索推荐广告三位一体的在线服务体系怎样支撑千人千面的分发体系

作者：阿里巴巴搜索推荐事业部高级研究员 沈加翔

### 一、三位一体的在线服务体系 AI · OS 介绍

AI · OS ( Artificial Intelligence Online Serving ) 是由阿里巴巴搜索工程团队打造, 集个性化搜索、推荐、广告三位一体的在线服务平台。AI · OS 引擎体系所支撑的业务场景包括: 手机淘宝所有搜索页面、手机淘宝上的信息流(猜你喜欢)、大促重点活动会场、淘宝首页商品推荐、分类行业个性化推荐以及商品选品场景, 覆盖了手机淘宝上 80% 以上用户群体。使用一套技术来支撑搜索、推荐乃至广告, 这在大型互联网公司里是比较少见的。阿里的平台技术战略, 本质上是两样东西, 一是电商技术, 另一个是和电商配合的大数据 AI 技术的应用。电商 OS, 包括商品管理、类目管理、运营管理、交易链路。而在大数据、深度学习的时代, AI 的投放、搜索推荐, 甚至广告的投放, 已经是独立于传统电商的技术场景。除了手淘上的场景之外, AI · OS 还支持阿里电商集团内能够想到的所有场景, 典型的包括东南亚的 Lazada、聚划算、飞猪、优酷、钉钉、菜鸟、盒马、本地生活——饿了么、口碑, 甚至和经济体内部的兄弟公司支付宝在合作。

淘宝搜索、推荐与广告



在深度学习的时代，AI·OS 引擎体系的架构演进的变化是相当大的。但是我们并没有像其他的互联网公司一样，在搜索和推荐的系统之外单独做深度学习的技术。之所以成为今天的 AI·OS 体系，是因为我们把搜索、推荐、信息流、广告、深度学习这几项，从技术到业务场景都不同的东西，非常好地结合到了一起，得以形成了有实质内涵，能相互借助合体发力的基础引擎平台。

从另一个角度纵观 AI·OS 上众多的业务场景，整个体系内会涵盖以下技术或概念(如图)：



第一层，分布式引擎系统需要具备的能力。无论是搜索、推荐、广告，都需要召回、排序，系统做大之后，需要分布式通信，高性能索引的存储，要有比较高效率的支持索引灵活构建更新的能力，以上是基础能力。

中间这一层，是在深度学习时代沉淀出来的技术要求和场景要求。比如深度学习需要做样本处理，做训练，在线预测，同时，与之相应的，是个性化投放，这在搜索、推荐、广告都有所体现。最后，索引的部分还应该支持实时更新，这个概念在电商体系尤为重要。

再下面，就是整个体系内有资源管理，高可用性，相应的计算引擎的支撑，运维的管控和插件的支持。

以下是 AI·OS 工程体系技术的概念图。



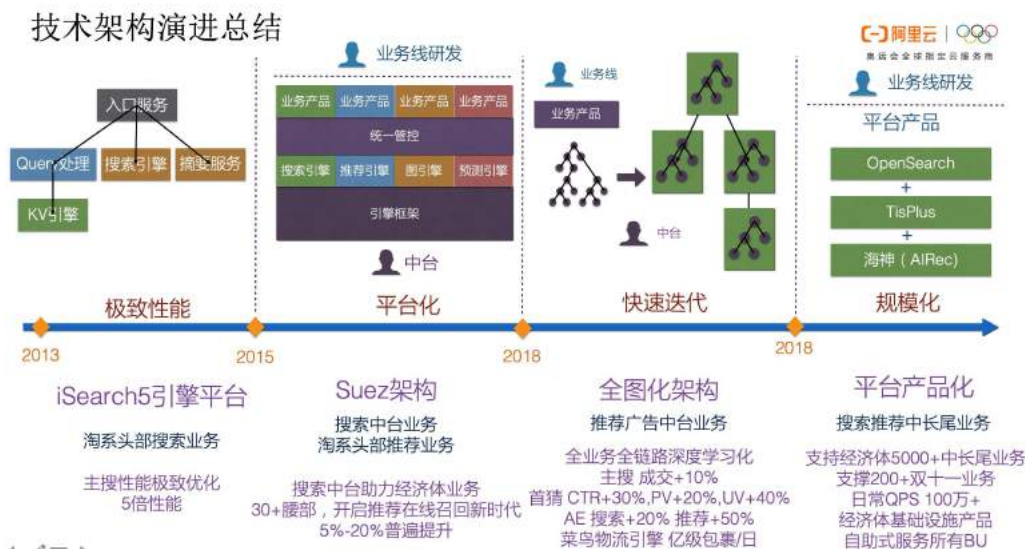


推荐改变和结果混排的变化，而是真正在端上做模型的训练，深度模型的预测，在这方面，手淘信息流是全世界规模最大的深度学习、训练和预测的应用场景，这是我们比较有特色的探索。

HA3 搜索引擎的服务是我们最经典的具有全文检索能力的引擎。商业化引擎，是与 HA3 相对应的，支持广告业务、关键词匹配带广告，或者支持定向场景投放的召回引擎。iGraph 图引擎，是具备在线图计算和图检索能力，在业界规模较大的图检索引擎，它里面具备的用户个性化关系、知识图谱的在线推导能力都是很大规模的。这些引擎都是支持数据实时更新的，这个能力的来源就是 AI · OS Framework，支持对数据的管理，对更新的管理，是对右侧整套技术的依赖，也有延伸出来的深度学习的能力。

## 二、AI · OS 在线服务技术架构演进

AI · OS 在线服务架构，对逐步走向规模化的初创公司，具备一定的参考意义。



AI · OS 虽然有十年时间，但由于之前一直围绕淘宝搜索业务深耕细作，在 2013 年至 2015 年期间主要在搜索引擎上做性能优化，在阿里体系内部做搜索引擎平台化。搜索搭建主要是运用比较经典的架构——Query 的处理+搜索引擎+摘要服务，Query 的部分，会有一些个性化存储，当时是用简单的 KV 来完成。这个架构是很多初创型企业选用的架构，也是我们云产品上提供的一个经典解决方案。

2015 年到 2018 年，随着信息流业务的进入，我们将搜索底层（Suez 或 AI · OS framework）数据抽象沉淀出来，在沉淀的数据上面，衍生了图引擎、预测引擎、搜索引擎、推荐引擎，形成了今天 AI · OS 的主体框架。在这个过程中，我们也统一了全集团的搜索和信息流基础框架，但这个过程是很依赖于阿里集团内部自下而上的搜索平台推进，这是在业界获得证明，在集团获得认同后，纳入到集团战略里的。

2018 年~2019 年，我们推进了全图化的架构，是从深度学习开源框架 TensorFlow 学来的。在深度神经网络迭代过程中，全图化架构这种图化 DAG 的表达，对业务描述更为标准通用化，我们把这个全图化架构学来后，推广到所有业务线，不仅是深度学习上，还包括业务逻辑调整、场景迭代、功能调整，如粗排、精排、统计、过滤这些功能，都用图化的方式来表达，在业务迭代效率上，有了质的飞越。

由于之前在定制业务逻辑时，难以避免有些代码级的，像 C++、JAVA、插件开发，虽然能解决业务需求，但维护升级成本较高。解决方案就是用 DAG 算子化的表达，完成算子图后，随着版本升级，不需要改变算子图，仅需改变算子的实现，把业务迭代、平台升级的耦合度大大降低，非常好地解决了这个问题，这是近两年重要的技术突破。

在这个过程中，我们也把搜索和推荐的技术，用到了非常有趣的场景上面，比如菜鸟物流引擎，本质上是图检索和图计算的表达，跟我们的引擎体系，与 iGraph 图引擎一起成长，成就了这个典型的场景，在这个场景上每天有亿级别的包裹在图中流动，寻找最优化的路径，都是通过这个引擎来支持的。还有像钉钉消息搜索，消息是加密的，我们是无法查看的，而这个加密是从上到下贯彻到每一个环节的，在常规的搜索引擎中，是无法实现的。这个加密能力，是靠引擎的迭代实现的，除了深度学习外，我们在逐步引入 SQL 的能力。

随着集团内部进一步贯彻中台战略，不管是在软件的抽象，还是在能力的衍生，都需要在云上发挥。我们已经在集团内部已与蚂蚁金服业务拉通，并开辟了新的突破口。我们也真正在践行平台化、通用化的思路，用最高效的方式解决问题，并打造出了海神（一站式选品投放系统）、开放搜索（一站式内容智能搜索服务）等经典应用产品。其中，海神平台支撑了阿里集团 1000+ 个性化场景，业务覆盖淘宝、天猫、聚划算、双 11 大促等业务线，支持了超过 10 亿次的选品投放。开放搜索作为高搜索质量的一站式内容智能搜索服务，以产品方式规模化支持搜索业务，覆盖了集团绝大部分搜索业务，业务应用数量超过 1 万个。并在双 11 大促时，担负了重要角色，支撑了高达 100 万的 QPS 峰值。

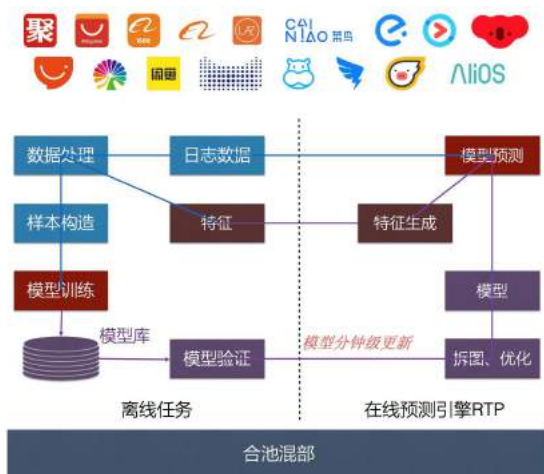
### 三、AI · OS 技术概览



AI · OS 主要部件包括：

### 端到端深度学习平台：星云&AOP

- 预测引擎
  - 全图化架构
  - 支持TB级大模型
  - 全量、增量、实时更新
  - 单次打分文档超过6000
- 样本处理，训练
  - 批流一体
  - 实时样本，流式训练
  - 更新：小时级->分钟级
- 样本、训练、预测混部
  - GPU分时复用
  - 资源按需使用
- 支持集团算法中台
  - 算法能力沉淀
  - 29个BU，750个实例，3万核



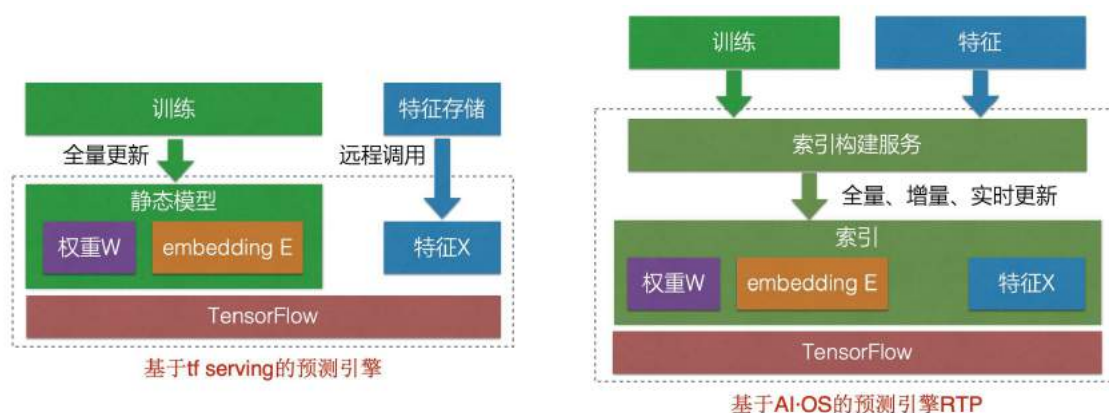
深度学习极大加速了模型工程的发展，模型迭代越来越频繁，网络结构越来越复杂多样，因此给算法迭代效率、数据计算效能以及模型交付可靠性等都带来了巨大挑战。为此我们提出并构建了面向大规模商业化场景的一站式深度学习建模平台——星云。借助星云，用户可以快速完成从特征引入、样本特征变换，到模型训练及评估，再到模型交付整个算法建模闭环。在此基础上，星云提供了完善的数据模型校验体系，确保用户的离线建模和模型交付具备生产级别的可靠性。星云支持全量学习、增量学习和在线学习，并通过高层抽象实现了各种学习模式之间的低成本切换。

### 大规模分布式深度学习框架：XDL



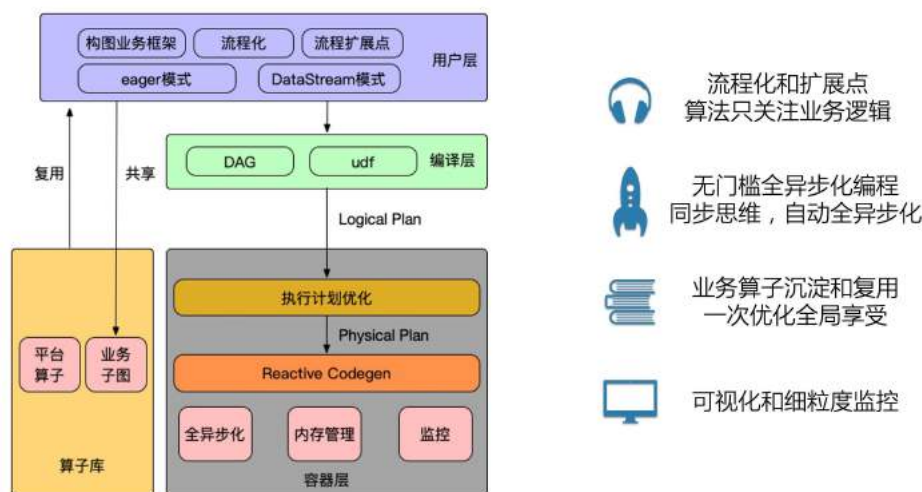
XDL 是以开源框架为基础，面向广告、搜索、推荐等场景打造的分布式深度学习框架，针对高维稀疏特征、互联网结构化数据和结构化模型进行了专门设计与优化。XDL 支撑着阿里妈妈直通车、钻展、超级推荐、品牌等多个营销产品对用户的深度理解与智能投放，是智能营销 AI 的核心驱动力。

## 预测引擎：RTP



传统的模型增量学习通过 restore 当前模型，持续训练，天级更新，或者实时流式训练，小时级更新，新模型需要经过全量切换才真正生效。基于 AI · OS 的预测引擎 RTP，把 TensorFlow 的能力集成进来，让深度大模型支持实时更新，充分利用实时数据分布，提升 CTR/CVR 预估精度，获取业务效果。模型特征支持实时更新，模型可以增量训练，通过对 RTP 在线 graph 分解，提取可更新模型 weights 成最大不连通可执行子图，实时发送模型数据消息，真正实现了模型离线流式训练，在线实时更新的效果。最终将在线模型更新周期从之前的小时级缩短到分钟级别，模型生效时间从之前的分钟级缩短到秒级别。

## 全图化推荐引擎：TPP



TPP (The Personalization Platform, 阿里个性化平台) 为集团众多的个性化业务提供开放、一致的解决方案，让搜索、推荐技术轻松服务于业务发展，业务也能快速得在平台找到需要的技术，是 AI · OS (Online Serving) 大数据深度学习在线服务体系中的入口之一。用户在 TPP 平台上编写方案代码，通过场景的形式对外提供服务。用户不用关心机器资源申请，应用部署结构，不需要编写服务框架，只需要实现自己的推荐逻辑函数，在 TPP 产品页面管理方案的生命周期，从编译，调试到发布上线。

#### 四、AI · OS 云原生产品与实践

我们将集团内沉淀出的技术能力，从 2014 年开始逐渐向外部推出。目前，依托于 AI · OS 体系构建的产品矩阵如下所示。



## 一站式内容智能搜索服务：开放搜索OpenSearch



开放搜索（OpenSearch）源于阿里巴巴 AI·OS 体系技术，是免运维、一键式的云上平台，将搜索平台服务化、产品化，完全屏蔽了搜索系统的底层复杂度，以标准化产品的方式支持业务。开放搜索拥有比较好的搜索质量，效果可以在线调优。只需要提交内容、配置，就可以直接得到搜索体验。搜索体验中相关性可灵活配置选择，可以改写 query、定制分词、提交行业词典。

在阿里巴巴集团内部，我们依靠开放搜索统一了各个搜索中长尾业务，自助接入的业务数量达上千个，基本覆盖了集团各个 BU 的业务。经过数年的耕耘，我们在阿里云也拥有数千家用户，包含内容、电商、视频等行业的典型应用。

## 一站式内容智能推荐服务：智能推荐AIRec



## 行业以及场景垂直化的一站式内容智能推荐服务





智能推荐（AIRec）延申自阿里集团内部的海神系统，提供一站式的个性化推荐服务，服务使用者在提交内容和用户行为，在确保数据保密的前提下，按照不同行业的模板，应用个性化推荐算法，实时调整效果，为用户提供服务。除了阿里巴巴深耕的电商领域外，我们在内容、视频等其他行业算法的投入是相当充沛的。我们要把内部的技术用好之后，很自然地更加抽象、通用地拓展新的场景。

在这一过程中，数据源的埋点规范、使用是有些门槛的，但我们也致力于让产品更易用，同时效果提升的监控、调试、运营干预的策略，都已经在产品里提供了。在中小企业的创业期间，这可能将成为一种刚需配置。



还有围绕开源生态的 Elasticsearch，这款服务由我们和 Elastic 公司合作推出。我们秉承开放的生态理念，将 Elasticsearch 的灵活易用和我们调度管控系统的稳定高效相结合，并根据用户需求，不断迭代自研的创新性功能。用户需要什么，我们就会提供什么，Elasticsearch 在我们体系上做适配，是一个相互扶持、相互助推的局面。

在这三个垂直产品之下，是我们基础的云上技术积累，比如 ElasticFlow，在数据进入引擎之前，需要对数据做打平处理，再灌到搜索引擎里，例如 Elasticsearch 就需要 ElasticFlow。同时，它还实现了开放搜索开箱即用的离线能力。在这一层，我们需要的是具化的计算引擎能力，这个计算引擎是在集团的计算平台之上，为搜索、推荐定制了一个数据采集、开发、共享以及模型训练的能力，这个能力提供出来之后，可以让其上的产品横向地拉通共享，这也是 AI·OS 体系沉淀的重要标志。

再下层是管控平台，搜索的基础平台以及阿里云的基础产品，在这上面我们会有一些生态化的产品。

## 2. 搜索篇

### | 2.1 “业务指标” 衡量电商搜索引擎的优劣

一般电商搜索的核心是搜索精度和搜索广度，精度就是搜索的精确性，广度就是搜索结果的范围，其关键结果肯定是“为用户找到想要的商品”，但过于追求搜索的精确度就会导致出现搜索的结果比较少或结果为 0 的情况，用户搜不到商品势必会引发流失，因此在搜索服务里面还可以做的就是给用户提供一些相关性搜索结果。那么搜索做的好不好，其实就是在搜索精度和搜索广度二者之间做一个比较好的平衡点。那如何评定我们的搜索是否可以满足业务需求呢？又可以通过哪些场景的优化实践提升我们搜索的整体性能呢？

#### 1. 驱动搜索业务价值提升的原因

先和大家回顾一下你是不是也是遇到以下几个情况后，开始思考搜索到底做的怎么样，有什么方法可以评估呢？

##### 1.1 随机性发现的 Bad case

领导或运营人员工作中发现搜索某商品时，无结果率较高，搜索排序体验不好的情况，认为不能完全满足实际用户搜索需求，直接影响整体业务发展，建议开发同学提升搜索效果和搜索体验；

##### 1.2 KPI 考核

KPI 考核驱动优化提升搜索业务，比如要求 CTR 提升 10%；

##### 1.3. 业务方诉求

合作某品牌旗舰店要求优化搜索排序效果；

那接下来我们剖析一下，从哪些方法和角度去评估搜索质量、体验与业务价值。

#### 2. 业务指标衡量搜索体验与价值

##### 2.1 完整体系化的衡量方法



阿里云开放搜索经过千锤百炼的实践与总结形成了一套体系化的、完整的、持续的评估方法。通过周期性的监控与评估，制定相应的迭代与专项跟进，做 A/B Test 或灰度后才去放全量，然后再持续进行监控。

## 2.2 核心业务指标

【指标维度】：流量指标、点击类指标、用户分析类指标、Query 分析类指标、成交指标。



### 【核心指标】

“搜索 PV”：指访问搜索页面的次数；

“搜索 UV”：访问过搜索结果页的用户数；

“无结果率”：空结果 PV/搜索 PV，无结果率越低，代表客户搜索需求解决情况越好；

“TOP5 PV—CTR”：指该 query search 结果中，排在前五位的 item 有被点击的搜索 PV/该 query 搜索 PV 该指标能一定程度反应排序效果；

“人均搜索 PV”：搜索 PV/搜索 UV；该指标的含义比较复杂，一方面人均 pv 大的话可能代表用户对搜索比较感兴趣，但另一方面人均 pv 大也可能代表搜索召回的结果较差，导致用户无法使用较少的点击找到满足需求的结果；

“有点击搜索 PV 占比”：有点击搜索 PV/搜索 PV 数；

“PV-CTR”：搜索结果页 item 点击数/搜索 PV 数；

“UV-CTR”：点击的 uv / 曝光的 uv；

“Item-CTR”：搜索结果页 item 点击数/搜索结果页 item 总曝光 PV 数；

通过体系化的业务指标和报表可以清晰反映搜索的质量和用户体验效果，再通过系统的评估服务，找到对应的问题原因和解决方案。

3. 搜索技术等级

参考业界对搜索等级的不同定义：

等级	名称	定义	数据处理	在线召回	算法能力	数据运维	产品化
L0	基础检索	基础的大数据检索能力	单表全量、实时增量	简单召回	文本相关性	系统级监控	人工接入和运维
L1	标准检索	通用的搜索功能	多表join、实时全量	下拉提示、可定制分词、query分析导航	rank formula	业务监控和报警	运维托管
L2	高级检索	利用用户行为数据和离线机器学习算法提高搜索体验和效果	复杂数据处理逻辑、离线机器学习	语义理解和智能导航	大规模机器学习、LTR	AB Test和天级别的数据报表	接入定制化
L3	个性化检索	提供实时个性化搜索体验、千人千面	复杂数据处理逻辑、实时数据计算	个性化召回、个性化排序	实时计算和在线学习	分层AB test、实时报表、MAB	AIOPS、自动诊断
L4	智能检索	实时场景感知和智能决策	复杂数据处理逻辑、在线实时机器学习	意图识别、智能交互	深度学习、强化学习、智能决策	OLAP分析	全链路管控、SaaS端对端体验

市面上大部分产品搜索还处 L0 等级和 L1 等级的位置。想要实现高效、可持续的高质量搜索价值需要投入大量的人力成本才能达到，这也是很多开源自建用户所面临的问题与挑战。只有一套高效、完整落地的搜索体系才会不断驱动电商业务更好的持续发展。



## 2.2 商品搜索“无结果率高，体验差”怎么办？

垂直电商与综合类电商相比，具有更精准的市场定位、更深化的产品与服务质量、更强的客户粘性和更独特的品牌附加度的优势，所以搜索性能的好坏直接影响着业务最终结果。

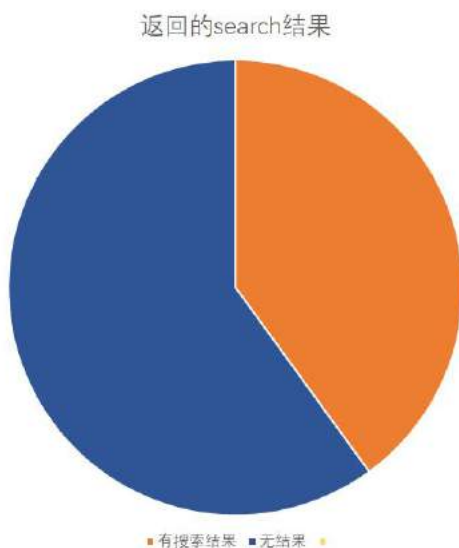
### 案例背景

某日活千万级的垂直类电商平台，业务以社区+商城形式开展，商城业务是商业收入的主要来源，大部分用户有明确的购买指向性，其中商品搜索天级 PV3000 万+，搜索引导的成交占比全部成交的 60%以上，是站内最重要的功能，在用户满意度调研中发现对搜索体验吐槽连连，反馈的主要问题是站内商品搜不到，个人卖家发布的商品排序靠后。

### 搜索问题反馈

(1) 用户：搜索不到想要的商品，体验差；

(2) 运营：站内搜索的无结果率接近 60%，说明每天有 1800 万的 PV 转换为 0，流量白白浪费；



(3) 个人卖家：个人卖家发布的商品排序靠后；打击发布商品积极性，影响平台价值定位和圈层生态，从而直接影响平台收益；

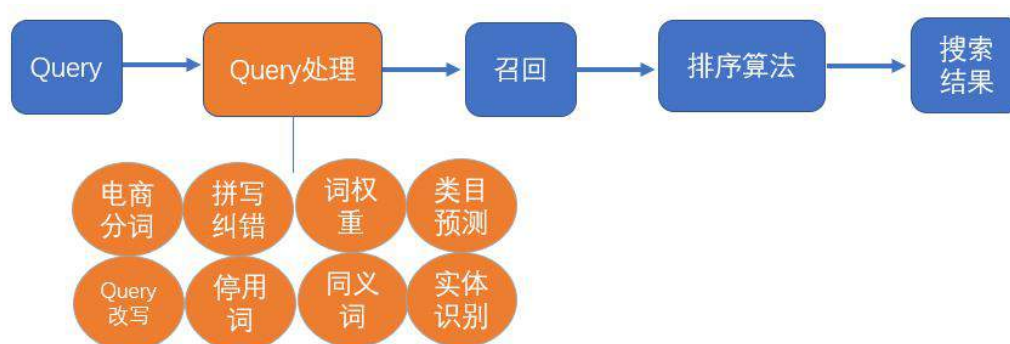
### 搜索问题成因：

- (1) **垂直小众的圈子，对于商品的叫法非常多样**，并形成主流，用户搜索中不一定按照实际商品名称进行查询（例如：用户会搜“喷泡”其实想找的商品是 Air Jordan AirFoamposite 系列的鞋）；
- (2) **用户搜索表述错误**（例如：搜“连衣群”其实是想搜“连衣裙”）；
- (3) 站内的搜索结果分 3 个 tab 呈现，分别为“销量”、“价格”、“新品”，用户搜索后默认展现的是“销量”tab 下的结果，因此个人卖家发布产品由于销量少或无销量自然导致排序靠后，曝光量小，销量难增长，恶性循环；

### 问题分析:

- (1) 针对召回结果不理想情况，经分析发现自建 ES 服务没有对搜索关键词做智能的语义理解，甚至有些实体名词分词还是错误的；
- (2) 针对排序问题，经分析需增加“综合”搜索结果呈现，根据核心索引优化排序算法；

### 开放搜索解决方案:



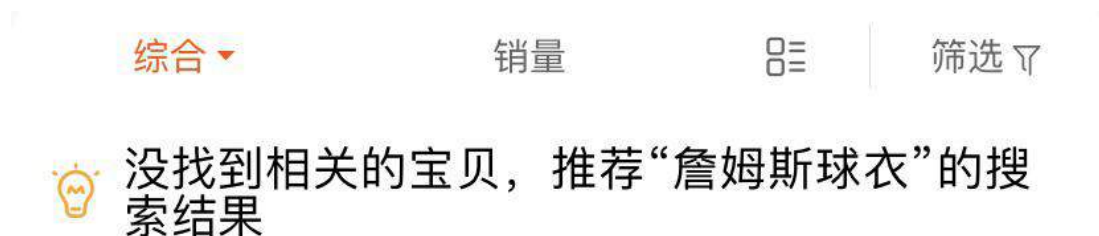
- (1) 核心索引上配置使用了**电商行业的查询语义理解**，包括**同义词、停用词、电商拼写纠错、电商实体识别等**，就是这些功能将搜索关键词进行了系统可识别的智能改写，扩大召回相关结果；
- (2) 针对商品别称问题，运营同学通过平时运营积累的**专业词汇可视化同步到开放搜索做查询语义理解功能的补丁**，通过灵活干预得以解决；
- (3) 创建核心索引“商品标题、颜色、类目名称、品牌名称、运营优化文案、系列名称等”，将它们引入到排序表达式中，**通过多个维度构建出更精细化的排序模型**；

(4) 增加“综合”搜索 tab，并默认展示“综合”搜索结果；

### 实践后的搜索性能对比：

(1) 搜索“詹姆斯球衣”输入成“詹慕斯球衣”

- Before: 服务无法召回相关结果；
- After: 纠错改写为“詹姆斯”进行查询，并且前端会提示“以下的结果是查询：詹姆斯球衣，仍然搜索詹慕斯球衣”；



(2) 搜索“喷泡”

- Before: 无法召回相关结果；
- After: 召回到 Air Jordan AirFoamposite 系列的鞋；

(3) 排序效果

- Before: 以销量默认排名，个人卖家排序靠后；
- After: 提高搜索相关性增加更多商品曝光机会；

## | 2.3 导购电商怎样 “提升 CTR、CVR 业务指标”

### 案例背景

某导购类电商 APP，与淘宝天猫等一线商家合作，亿级商品索引量，类目和子类目多层嵌套，商品有不同子款式和尺码，搜索和筛选需求复杂。通过采用分销+券模式，优惠券帮助普通 C 用户降低了单品价格，分销模式帮助推广者 B 增加了用户数量，平台获得分成，用户下单后还可以返现金和优惠券，从而提升复购率，其中搜索的流量占比站内流量 60%以上，所以对于搜索结果的召回和排序有极高的要求，衡量搜索效果的直接指标就是成单转化率。

### 需求反馈

业务方希望迅速扩大市场提高用户体验，需求开发团队针对产品功能快速优化，进一步提升 CTR、CVR 的业务指标

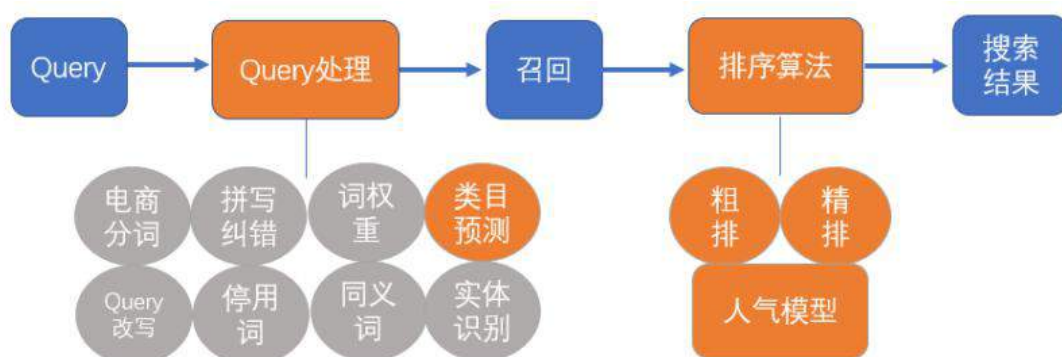
### 需求分析

- 电商行业中，“搜索”是帮助用户定位自己想要的商品提升转化的重要渠道；搜索引擎的效果优化是一个很大的话题，在查询意图理解阶段可以有语义理解、命名实体识别、词权重分析、拼写纠错等优化手段，在排序阶段可有文本相关度、人气模型、类目预测等优化手段，通过配置查询分析策略和调整排序公式，我们对于效果优化可以有很大的发挥空间，再通过 AB 测试来对比不同优化策略的效果表现，我们可以做到效果优化心中有数。
- 业务指标数据中“点击率和转化率”则直观反映了搜索结果页的商品结果是否满足用户的需求。
- 优化召回和排序结果可以帮助用户快速找到心中所想的内容，是改善用户体验，降低跳出率，促进用户转化率的最好方法。

### 阿里云开放搜索解决方案



Query 在开放搜索的执行流程：



### 1. 开放搜索的优化方案:

查询意图理解优化方案可以参考：[商品搜索“无结果率高，体验差”怎么办？](#)

### 2. 开放搜索-类目预测功能

类目预测是开放搜索里基于物品的类目信息改善搜索效果的算法功能,类目预测根据用户的查询词来预测用户想要查询哪个类目的结果，结合排序表达式，可以使得更符合搜索意图的结果排序更靠前。

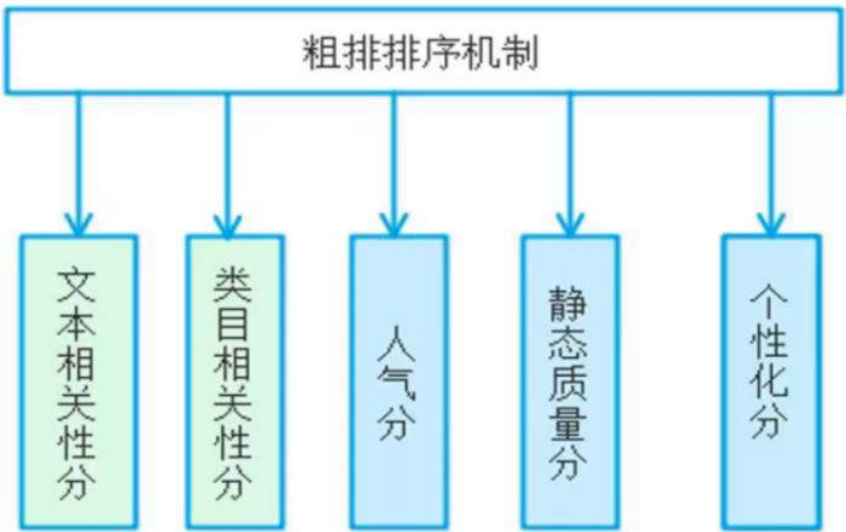
例如：用户搜索“华为”

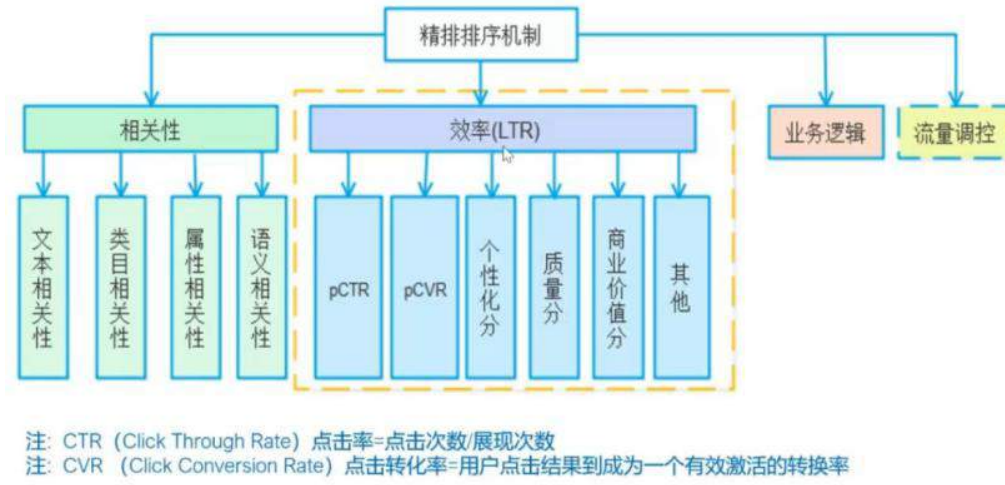
- 大部分人意图其实是想获得“华为手机”，但因为销量大小、价格高低、店铺等级等各种原因，存在“华为手表”等配件商品排在“华为手机”更前面。
- 当我们训练“类目预测模型”，模型就会表达出一个信息，根据行为数据统计发现点击“手机”类目要比点击“配件”类目的人多很多，那么模型会给出这样的预测结果，对于“华为”这个 query 来说，“手机”类目与“华为”的相关度，比“配件”类目与“华为”的相关度高，所以在计算每个物品的排序分的过程中，“手机”类目下的物品所获得的得分要比“配件”类目下的物品得分高，从而“手机”类目下的物品会排在更前面。
- 这样的排序结果才是一个比较符合用户预期的结果，用户才更有可能点进去了解详情，从而提升搜索的业务价值,提升 CVR 的业务指标；



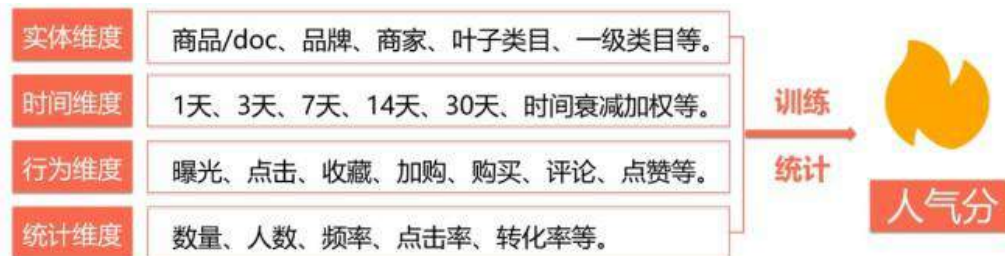
3. 排序算法优化

- 支持两轮相关性排序定制，搜索结果相关性排序是影响用户体验最关键的一环，开放搜索支持开发者定制两轮相关性排序规则来准确控制搜索结果的排序。第一轮为粗排，从命中的文档集合里海选出相关文档。第二轮为精排，对粗排的结果做更精细筛选，支持任意复杂的表达式和语法。方便开发者能更准确控制排序效果，优化系统性能，提高搜索响应速度。





- **引用智能排序人气模型:** 离线计算的模型，淘宝搜索最基础的排序算法模型。人气模型会计算量化出每个商品的静态质量及受欢迎的程度的值，不断训练统计形成人气分，构建更精细化的排序模型，精准命中搜索需求，将人气模型 involve 成为排序的一个因子，搜索结果的转化率还会有质的飞跃。



## 2.4 “减少输入，更快找到用户所想”

### 案例背景

某测评 APP，通过对化妆品成分和肤质的分析，由于涵盖全球各国化妆品，有 300 万+化妆品 SKU，所以用户在搜索框输入查询词的过程时，经常遇到中英文切换、名称过长等导致的效率低的问题。

### 问题反馈

希望提高用户搜索输入效率，更快速找到所需要的商品或内容,从而提高业务转化

### 问题分析

1. 通过搜索框输入查询词的过程中，**智能推荐高质量候选 Query**，可以帮助用户更快速找到搜索内容；
2. **优化下拉提示的内容排序**，让用户搜索意图商品排序更前。

### 阿里云开放搜索解决方案

扩展功能：



#### 1. 开放搜索下拉提示功能：

起到减少用户输入的作用，自动补全搜索关键字，提升用户使用搜索引擎的体验，同时减少用户的查询次数，减小我们服务端的压力。

#### 2. 支持智能抽取候选词，电商行业模板

- **文档字段原值保留**(内容较短含义明确适用于不需要分词的字段，例如：店铺名，应用



名等)；

- **文档字段智能抽取** (阿里海量自热语言训练的分词器，进行分词，抽取有意义的 term 进行组合)；
- **历史查询词** (引用用户近期的历史查询词)；

### 3. 支持多种干预项

- 对于一些突增热词，例如当“iphone Xs”发布时，它的 query 频次肯定不及发布了一年的“iphone X”，这种 case 下，客户肯定希望当他的用户敲入“iphone”时，“iphone Xs”排在下拉提示的最前面，这时我们提供的**推荐名单功能**就可以派上用场了，客户可以手动输入“iphone Xs”词条，**这样它就可以排在最前面了**。相反，有些 query 客户是不希望被展示的，例如一些**法律敏感的词汇**，或 query 返回的结果集较少的词汇，此时客户可以使用我们提供的**黑名单功能**，运用该功能后，用户输入的 query 如果部分匹配或全部匹配黑名单词条时，**相关的下拉提示结果会被屏蔽**。

### 4. 下拉提示请求与搜索请求关联

基于关联请求数据，得到下拉提示点击结果等数据，可以优化下拉提示排序模型，进而提升下拉提示搜索的效果。

## 实现效果

### 1. 搜索框输入“elix”,下拉提示结果



2. “保湿霜大瓶”这个 query，可以通过如下方式查询得到：

- 中文前缀：保，保湿，…
- 全拼前缀：B, Ba, Bao, Baoshi, Baoshishaung, …
- 简拼前缀：B, Bs, Bss, …
- 汉字加拼音：保 shi, 保湿 shuang;
- 分词后前缀：大瓶，大瓶保湿，保湿大，…
- 中文同音别字：保湿双, 宝石霜，…

## 2.5 “随便逛逛，看大家都在搜什么？追踪潮流和热门”

### 案例背景

某销量领先的互联网食品品牌同时也有电商平台，主打休闲零食，自有品牌 SKU 超 500 个，应用程序主要为 APP 端和微信小程序端。那如何促使用户产生更多消费那？在现有搜索方案的基础上上寻求更多扩展功能提升用户购买转化率。

### 需求反馈

- 最终用户：随便逛逛，想知道大家都在搜什么？现在什么最热门？
- 业务运营：吃货们典型“喜新厌旧”的特质，追逐新奇和个性化，需要不断的用新鲜的零食爆款去冲击他们的钱包，产生更多消费。

### 需求分析

#### 从最终用户的角度：

1. 零食消费者主力军为 80-95 后女性，对流行事物更加的关注，追求快速消费，并且猎奇心较重，对榜单、热门、新品比较敏感；
2. 用户只是随便逛逛，没有明确的搜索意图，搜索框可以推荐优质的查询词，引导其搜索点击，从而实现转化；
3. 大家都在搜些什么？大家都喜欢的内容，用户也可能会感兴趣；
4. 不想输入查询词，或输入不便，可以减少用户输入，优化体验；
5. 多样化推荐查询词，既兼顾热门结果，又可以探索兴趣之外的其它内容；

#### 从业务运营的角度：

1. 在用户没有任何输入的时候，也可以推荐优质查询词，对用户意图进行引导；
2. 推荐查询词既要关注热门程度，又要兼顾多样化，给部分次热门查询词一些曝光机会；
3. 可以结合用户兴趣来推荐查询词，既兼顾用户体验，又可以有效的提升业务目标；

### 阿里云开放搜索解决方案

扩展功能：



热词和底纹处于搜索引擎整个工作流程的最上游，通过推荐热门、优质、多样化的查询词，对用户搜索意图起到重要的引导作用，既可以减少用户输入，优化体验，又可以优化查询词质量，大大降低后续查询意图理解、相关性、排序、运营干预等环节的调优难度，对提升整体业务目标可以起到非常好的铺垫作用；

## 实现效果



- **最终用户：**多样化推荐优质查询词，简化了输入，体验更佳，通过热词和底纹可以探索兴趣之外更多更新更热门的商品；
- **业务运营：**
  - a: 提高了搜索 UV 转化率和点击率；
  - b: 通过分析热门查询词可以把握用户兴趣走向，制定相应的运营策略，还可以给部分次热门查询词增加曝光机会，即兼顾了用户体验，又可以有效提升业务目标；

## 3.推荐篇

### 3.1 推荐场景化构建与业务价值

#### 内容要点

本文从推荐的业务价值、推荐领域难点问题、可实现的场景功能来阐述推荐的价值和意义。

#### 推荐的起源和价值实现

具有“推荐系统之王”称号的电子商务网站-亚马逊，在推荐的应用上走在了行业的最前列。

在初期的设计中，亚马逊主要将推荐应用于 3 个板块：

第一种是最常见的基于 item 的协同过滤算法，利用过去的历史行为进行推荐；

第二种是利用 facebook 间好友关系的推荐；

第三种是相关推荐-打包销售场景，这个和后续中场景部分的相关推荐差别较大。打包销售主要的应用场景类似于买了手机买手机壳的场景。

以上三种，在当时是一个较为全面的推荐管理，串联了消费者浏览、购买以及购买后的场景。从已有的数据材料中统计，当时 20%-30%的销售来源于推荐系统。

对于每一个消费者而言，推荐系统使自己拥有了一个在线商店，能够不虚此次访问，在商店中找到自己感兴趣的商品。

由推荐的引入，快速促进业务增长的案例中，值得一提的还有 Netflix。与前者类似，Netflix 应用的是基于用户历史行为的协同过滤，为用户推荐感兴趣的影视作品。据统计，Netflix 有 60%的用户，是通过其推荐系统找到自己感兴趣的电影和视频的，可见推荐成为了承载流量的主要入口。从 2006 年起，Netflix 开始举办推荐系统比赛，通过技术底层驱动的形式来促进业务增长，对推荐应用的重视程度可见一斑。





回到现在，市场上很多火爆的 app 已经将推荐应用得更加多元化，比如典型的电商场景下，通过信息流的混排打造逛起来的体感，利用消费者评价或 app 领袖级买家引导促成消费。纵观各行业，不论是内容论坛、广告分发，还是教育培训、书籍小说，个性化推荐都成为了必不可少的功能。

之所以能够引发这种现象，就是因为推荐这种智能分发能够带来的业务价值提升绝不是一星半点。客观而言，通过千人千面进行推荐，可以大幅提升分发效率，促进点击率指标的提升。

对于老用户，能够在推荐板块找到自己的兴趣变化、惊喜发现，提升粘性；新用户能够迅速在 app 中获取最喜好最关心的内容，从而促进新、老用户的留存粘性提升。

利用内容分发促进转化，一方面可能由当下主打下沉市场的 app 应用中的返利机制促成，另一方面是为了培养粘性用户，提升用户活跃度，从而为后期变现、转化打下良好的基础。

推荐并不止完成长短兴趣的探测与精准触达就可达成目标，也需要给到用户多样化的体验，使其能发现从未获取的新颖好物，不断发现惊喜，真正喜欢上一个 app。有时候推荐要比用户更懂在当下他需要什么，通过什么能够给他带来欢喜，这也是人们对推荐的心里预期。

从客观指标、主观体验通盘来看，推荐的价值体现在，作为一股强大的动力，贯穿于从获客到粉丝型用户的整个培养流程，推动用户的转化与粘性提升。

## 推荐领域难点问题

将一个个性化推荐的场景搭建好并不是一件易事，通常需要解决几个推荐应用中比较棘手的问题。

阿里云 | 智能推荐

### 难点问题



首先是新用户冷启动，有时候我们花了很大力气做了活动、营销、广告或是买入一些渠道的用户，但最后的结果并不理想。新用户留存难，对于推荐而言难在缺少新用户的信息，而怎么解决这个问题，后期我们将进行剖析。

第二是回音壁问题，指的是当用户的兴趣比较集中时，我们往往在内容推荐过程中陷入局限，从而导致用户走不出兴趣圈，逐渐疲惫。

第三是马太效应，往往 80% 的流量会集中在 20% 的内容之中，剩下的 80% 并不是内容不好，而是比较小众，或因为流量分发的限制就此埋没。因此，当存在一些小众爱好的用户时，很难做到准确分发。

第四，上下文联系、访问心智、行为变化也是一个先进的推荐系统应该解决的问题。

综上所述，推荐想要做好仍面临着很多挑战。

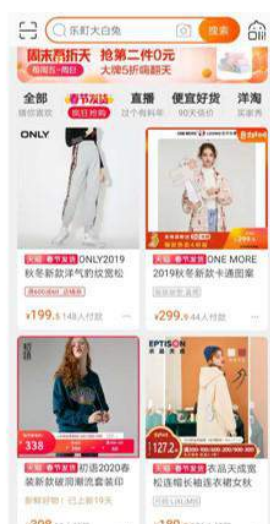
## 可实现的场景功能

当我们想要把推荐部署应用在 app 前，需结合具体的场景构建合理的应用方案。

以阿里云智能推荐主打的场景为例，如以下几个常见的场景，就可以拉近 app 的资深体验用户与推荐的距离。

阿里云 | 智能推荐

### 推荐方案-心智



#### 首页推荐

#### 猜你喜欢

- 长短期兴趣观测
- 多元兴趣探索发现
- 用户偏好画像分析

常见应用位置：  
首页推荐  
购物车页推荐  
订单页推荐



#### 详情页推荐

#### 相关推荐

- 相关性计算分析
- 多维度向量分析
- 用户偏好画像分析

常见应用位置：  
详情页推荐

首先，我们进入一个消费场景，可看见猜你喜欢和相关推荐这 2 个场景。

猜你喜欢，承载的是当消费者的购买意向不明确时，更多的是让我们能够逛起来。看了很久没找到合适的商品、一直感兴趣没下狠心下单的商品、很有意思看了就想要的小物件，都被我们期待着出现，或者换个角度看，这是电商平台让我们逛起来的同时，提升转化率的意向商品集。

阿里云 | 智能推荐

@猜你喜欢

推荐方案-心智



相关推荐的差异主要在于，当我们愿意花心思研究某个商品的参数、详情，说明购买意向已经相对明确。看了又看就是推荐在商品性质上和当前驻足商品类似的，并且结合用户口味容易引导购买的商品。

其次，例如一个新闻阅读的场景。

阿里云 | 智能推荐

推荐方案-分栏



新闻的展示往往会区分几个频道，称之为专栏。针对每个专栏的推荐，都代表用户的兴趣相对集中的分布。在不同的专栏涉及不同内容的选举、分发，基于智能推荐，快速实现场景化专栏搭建是一个很方便的操作。作为一个智能的反馈，这些专栏也能够根据不同用户的兴趣覆盖点，进行优先级展示的重组。就像在很多 app 里看到的多功能板块、多业务板块，在推荐中的应用就会体现在一个基于 banner 热点图的推荐，这样会展示给用户更加定制化的体验。

而实际看到的结果，应该是具备视觉与类目丰富覆盖的体验。在智能推荐内部，提供了 2 种功能，分别是打散和混排，这些业务干预也是为了满足提升用户体验的诉求。这一步的设置最好先完成一个用户行为分析的分层，比如，从电商领域分析的数据可知，大部分的男性消费者更倾向于使用搜索的功能，发现自己想要的商品，推荐直击搜索内容会具备更高的转化率，再多元的展现都不及准确的推荐有效。而女性则不具备这样的特点。

以上也可见，推荐和用户行为倾向也具备很强的关联，推荐这个板块指引着我们更加全面、细粒度地了解平台用户，已发展成为业务反哺的参考。



### 3.2 智能推荐在电商的经典打法与实践

## 内容要点

本文从电商行业经典个性化场景、会场搭建实战分析来分享个性化推荐在电商的经典应用和平台化搭建。

## 电商行业经典个性化场景

在早期，个性化推荐技术，首次应用到淘宝双十一主会场，形成了一个从楼层顺序排列个性化、楼层内坑位个性化再到坑位素材个性化的三层结构，为每位消费者打造了一个专属的类目会场。



第一层，通过楼层顺序个性化，女神看到的楼层顺序可能是女装、美妆、天猫国际等，男神看到的楼层顺序可能是男装、旅行、数码等。

第二层，通过楼层内坑位内容个性化，使得在同一个楼层内，不同用户看到的商品或店铺不同，比如同样都是美食控，喜欢吃辣的用户可能看见麻辣牛肉干，喜欢甜口儿的用户可能看见巧克力。

第三层，通过坑位内容素材个性化，使得同一个楼层的同一个坑位，即便算法预测两个用

户都需要巧克力，但一个喜欢费列罗，而另一个喜欢德芙，也会在入口图上展示不同的品牌。

通过这三层的个性化层层递进，将海量的商品按照消费者从面到线再到点这样一个深入的过程进行喜好挖掘，从而达到精准迅速的信息分发、触达。

早期的双十一类目会场包括了 51 个天猫会场和 18 个淘宝会场，在底层技术层面，应用了基于人群特征的选品技术、对实时行为的迅速响应乃至自动合图技术等，在给予用户畅快体验的同时，极大提升了 uv 的引流效率，对比非个性化相对提升了 2.56 倍。



可以说，那个阶段是个性化推荐的探索元年，在斩获大幅数据增长后，个性化推荐的场景开始在淘宝逐渐铺开。之后，主要打造的个性化场景有首焦推荐、banner 推荐、店铺内个性化推荐等，在早期的基础上，进行技术内核升级的同时也加入了商家赋能的个性化推荐系统，从而能够结合会场流量的增长、每个场景的差异点。比如说推荐各场景大小不一、定位差异大，有导购类场景、有成交类场景等，需要根据场景本身的特性来进行流量智能调控。与之前相比，之后的双十一个性化场景也融入了更多平台业务策略、流量调控逻辑的因素。

阿里云 | 智能推荐

## 手机淘宝的个性化演进简介



再之后，个性化的场景不论在广度上还是深度上都有了更多的延伸、拓展。例如近年的双十一，从10月20号的预售/造势期开始，到11月1号预热期开始再到双十一当天，个性化的推荐场景都由始至终贯穿，并成为促成交易额提升的重要手段。除之前的业务以外，还扩展出了入口图个性化、多楼层入口图个性化、多楼层店铺个性化、店铺内推荐等业务形态。

阿里云 | 智能推荐

## 手机淘宝的个性化演进简介



到近两年，瀑布流首次加入到众多商品推荐之中，较之前双十一手淘首页个性化推荐区块，在延续往年大促传统的基础上又有了新意，每个区块各有侧重，能够覆盖到不同用户的多层次需求。

首页焦点图以滑动帧的形式展示商品图片和宣传语，以品牌活动为主；主会场入口是导流最集中的区域，满足用户最迫切的 shopping 需求；六宫格历史悠久，已经培养起了一定的用户心智；氛围会场入口为各行业会场分发流量，也是最需要使用流量调控手段的区域；瀑布流会场入口主要面向手淘深度用户，从一定程度上迎合长尾需求和发现性需求。



到去年，淘宝双十一推出了互动城个性化和云主题，互动城和品牌墙首次合二为一，具有很强商业化目标的互动城首次尝试了个性化。互动城作为之前双十一重点项目，因为领喵币盖楼玩法带来了大量用户流量，通过用户做任务领喵币/兑换红包等互动方式，成功引领了网络潮流，创造了流量新阵地。商家通过分层级付费/参与店铺投放获取了对应的层级流量，实现了平台、商家及用户的三方共赢，是一次重大的创新和突破。

随着之前手淘大改版，信息流已经成为手淘首页的用户个性化推荐主阵地。云主题，作为信息流的重要组成部分，是一个服务于用户多元化需求的轻量导购体系。云主题推荐是结合了电商认知图谱，基于对用户本身及其每个行为背后隐含的深层次需求，再与场景建立映射，最终通过实时的推理、计算，然后得到当前推荐的商品和场景组合。

这些年与推荐、个性化相关的双十一，很多玩法都是基于大数据，基于业务诉求，包括平台的愿景等，如果针对每一个场景都需要定制化开发，对于阿里来说无疑是一个不可收敛的人力投入。从早期推荐展露头脚，到如今端智能应用到推荐板块，除了淘宝、天猫，咸鱼、1688 等也已具备了非常丰富的推荐应用场景。面临整个集团推荐场景的广泛铺开，从很早开始，集团就开始了平台式推荐业务支持的探索，不论是算法开发、还是运营，都能



使用适配的平台工具快速搭建场景，进行推荐实验与实时的效果观察、流量调控等。

## 会场搭建实战分析



以阿里小二们经常使用的平台为例，这个平台提供了很多能力，是一个高效、实时、支持多种投放类型、提供高阶算子的开放的个性化推荐平台，可以一站式算法选品、智能搭建、个性化投放。在这个平台上，小二们只要明确会场的搭建诉求、面向对象和业务指标，即使没有推荐的算法基础，对埋点和大数据没有经验，也可以很快搭建出个性化的推荐场景，从选品到投放的每一步都有数据赋能，具备业务经验加成引导。



除一些非常核心的需要算法实时跟进，通常，小二们从搭建场景到部署到上线只需小时级



别的时间，也不需要固定的算法支持开发。目前来说，推荐的场景非常多，淘系的大部分个性化场景都是通过平台来搭建上线的，平台化已是一种必须的业务支撑方向。



如上图，这是两种截然不同的场景，但是不论从业务类型的多样性，还是业务目标的多样性，小二都可以在这个平台上进行选配。每一个场景都有着搭建的目标，有的是转化，有的是留存，场景化搭建的贯穿也是平台的一大特色。



因此，基于淘系算法、业务经验的沉淀，我们打造了一款推荐的 Saas 化服务——智能推荐 AIRec，希望以标准化服务的形式赋能客户，通过行业化、场景化，下钻找到适配行业、效果优异的算法模板，通过丰富的数据侧服务为数据采集、上报减负，配合多样的运营策略，匹配上层的特色业务诉求。目前，智能推荐已进行了电商行业算法模型的升级，电商模板 2.0 已发布。

## 3.3 商品推荐的核心问题、架构和应用效果

### 内容要点

本文主要围绕电商推荐，从个性化推荐的核心问题、架构、推荐效果阐述智能推荐的技术演进应用。

### 电商推荐问题和价值



在看推荐系统要解决哪些问题之前，首要了解它能带来什么价值。个性化推荐之所以能如此迅速地占领各大 APP 的首要版块，取代了之前很多人工运营的坑位，最核心的原因是它可以更拉近用户与商品的距离，用最短的时间和路径让用户找到想要的商品，也就是常说的流量分发效率。

这就好比一个人逛超市，他喜欢吃肉，而且一般只逛 3 个货架，虽然还没确定这次要买什么，但是如果货架上摆的都是蔬菜，那么他大概会扫兴而归，对于超市来说，最现实的就是收入少了，归因就是流量分发效率低。所以由此引出的就是，推荐要解决的最基本问题，就是保证流量效率，也就是点击率、转化率这些指标。

电商首页的版块，一般都代表着产品的门面，它所承载的，不仅仅是怎么赚更多的钱。所以这也需要推荐系统再往前走一步，不仅仅是流量，还要能承接住用户的心智和产品的调性。这些可以概括为两个方面：用户体验和运营干预。

运营干预，就是推荐结果要可以被一些外部的输入所影响，这些输入主要来自于人的主观判断。比如能够做一些流量调控，侧重一些行业的倾斜，甚至对新店铺做一些扶持。首先这些判断对产品的发展是很重要的，其次其背后的信息是系统无法量化计算的，因此只能交给人来处理，系统只需要留好接口，尽可能柔和地接住这些硬性的规则。

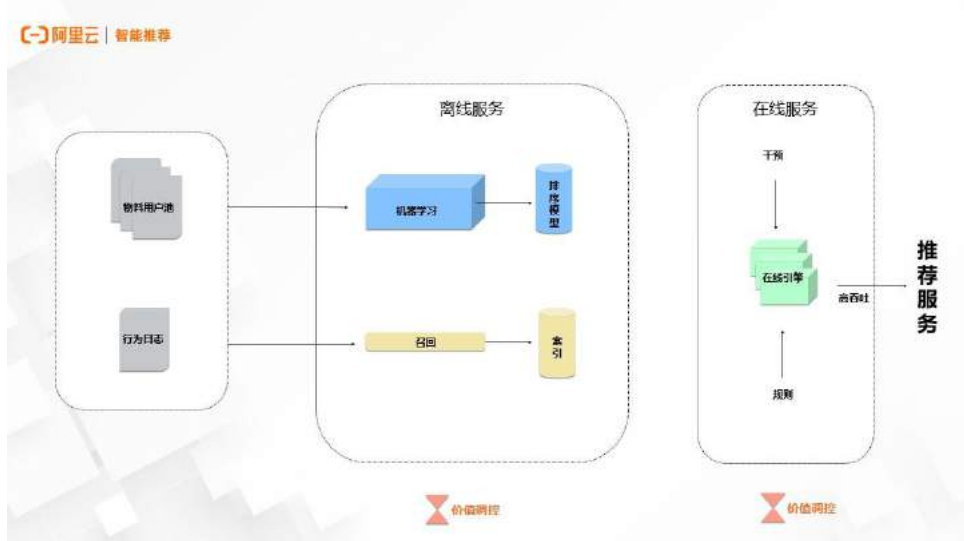
怎么才算是好的用户体验？这是一个综合的衡量指标，很难从单一维度回答。换个角度，当一个用户逛完某个平台，有了一定的精神满足感，那么就基本可以判断用户体验还不错。仍以逛超市举例，我们不仅希望用户在 3 个货架内成交，还希望他能逛更多的货架，希望他明天能来，逛完后不仅仅喜欢吃肉了，还喜欢吃水果了。所以具象化到推荐系统，就有了更多的衡量指标，如发现性、停留时长、浏览深度、复购率等，也就是说，一个好的推荐系统，不仅是让用户快速找到喜欢的商品，还能让用户逛起来。

总结起来，推荐需要解决这三方面的问题，但单靠推荐系统肯定是不够的，还需要商品、流量、运营这些因素共同来支撑。

由左侧的问题，可以推断出，判定一套推荐系统的价值，一定是多元化的。不同的业务系统，会有各自侧重的标准，所以推荐做到最后，不仅仅是算法模型的问题，它更是一个业务系统。

在此有一点需要强调，这种多元的价值判断，不同指标看上去都是相辅相成的，但是仔细思考，里面还是有很多相悖的。就比如前面说的，一个用户满足感高了，他自然买的就多了，可买的多并不仅仅是因为流量效率高了，而是因为他浏览了更多的货架，感觉越逛越开心。所以怎么去衡量这些指标之间的平衡，也是推荐系统需要考量的问题。

## 推荐系统架构



对于一个推荐系统来说，它的架构已经比较明确了。

第一部分是商品池、用户池的构建。数据是算法的基础，底层数据的丰富程度决定了推荐能力的天花板。更多的用户录入系统，使请求推荐的用户未登录占比更低。更多的特征，比如用户的年龄、性别、地域、商品的店铺、价格、类目等。更丰富的行为，曝光和点击是目前智能推荐的基本要求，其他还包括点赞、停留时长、下载等等。算法系统本质上是拿过去预测未来，如果过去什么都没有发生，那么预测出来的结果，就可想而知。

第二部分是怎么从海量的商品池选择推荐的对象。这方面主要是依靠检索的技术，包括通过协同过滤、内容过滤、用户偏好建立索引，并根据索引生成第一步的结果。这些涉及到的工程技术主要是性能，怎样一次性查得更多、更快。以阿里云智能推荐来说，主要是复用了阿里内部的系统架构，单次单链路的召回至少是几千级别的，这是这款产品的固本之策。

在算法上，主要就是各个召回链路。以最基础且最通用的协同过滤为例，ItemCF 可算是推荐系统发展过程中的一个里程碑节点，它第一次解决了大规模的计算问题，并且能在效果上碾压其它普通的基于信息检索，语义检索的推荐系统。它效果好的原因，第一就是它本质上是一种迁移学习，收集用户在全场景上的行为、推荐、搜索等打造出来的信息，是一种半监督式的。另一个就是，它是通过用户的近期行为去做相关性的匹配，所以也是一个核函数的方法，就是越近的结果越容易被推荐出来。但是，i2i 效果太好也会有挑战性的一面，很难去突破，而且对数据质量要求很高。以上问题的关键就在于心智的体现，落地

到系统上就是需要扩召回，需要更丰富的，能体现心智的链路进来。

第三部分是在之前的基础上，需要一个精选排序。不管是 LR 或者 GBDT，还是 WDL 在神经网络上 4 层、7 层甚至 10 层的尝试，都是希望 rank model 学得更准。再配合推荐价值，不仅仅能优化例如停留时长之类的某一个目标。

最后一部分是运营干预，也就是对最后结果的一次规则性的过滤。这里其实有一个主要的矛盾点，对于运营同学来说，仅仅对最后结果的干预，还是太弱了。对于算法同学来说，强行对最后的结果进行人工规则的过滤、重排序，是非常影响效果的。所以，从系统落地的角度，我们在做的一个事情就是把运营规则前置化，可以更柔性地应用到选品、召回、排序中。

## 推荐效果

第三部分内容聚焦在怎样提升推荐效果上，也就是流量分发效率。

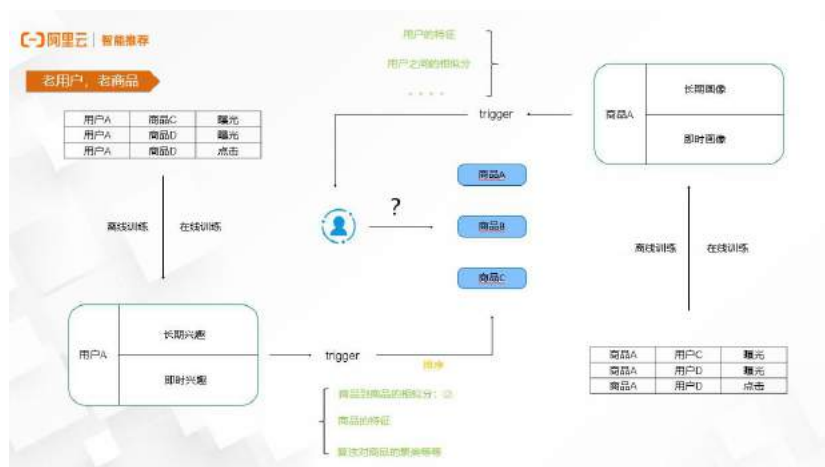


现阶段，机器还是一个把人的思维落地成自动化的过程。所以试想一下，如果面临一个用户要选择三个商品，你会怎么去推荐，需要什么样的信息，怎样去做决策。

第一步，需要什么样的信息。我们需要知道这 3 个商品是什么。是美妆，还是食品？是大品牌，还是性价比高？还需要知道用户的信息。是男是女？地域怎么样？年龄多大？最重要的是，我们得知道用户喜欢什么，他肯定不会直接告诉你，有时候他自己也不知道，那



么，我们就用他过去的行为来推测，点过什么商品，买过什么商品，什么商品看了看就走了，什么商品看了很久很久。所以你需要这些信息判断，推荐系统也需要，把这些信息抽象一下，就是以上图右边的这三类，对于智能推荐这个产品来说，就是三张基础数据表。



第二步，需要把这些信息抽象一下，得出这个用户喜欢什么的结论。对于推荐系统，就是一个用户画像构建的过程。简单的做法，就是把用户过去的行为做个聚类。也可以做的很复杂，如需要考虑时间的衰减，需要把即时兴趣抽出来，需要考虑同特征的人群画像的输入等等，这也是各个推荐系统的核心竞争力之一。当然，还是那句话，没有最好的，只有最符合你的业务的。

拿到用户画像之后，还有很关键的一步，就是怎么把这个画像和要推荐的商品关联起来。

这里就可以抽象出一个 trigger 的概念，是一个桥梁，以之前提到的 i2i 举例，图中左边这个 item 就是一个 trigger，可通过画像知道用户更喜欢哪个 item，然后进而去寻找和它最相近的右边 item，这样就完成了一次召回的查询。

有了用户画像，还可以构建物品画像。举个例子，一个商品，北京的人更喜欢购买，那么它的画像中的一个因素就是北京人，一般物品画像更多地会用在新用户冷启动上。

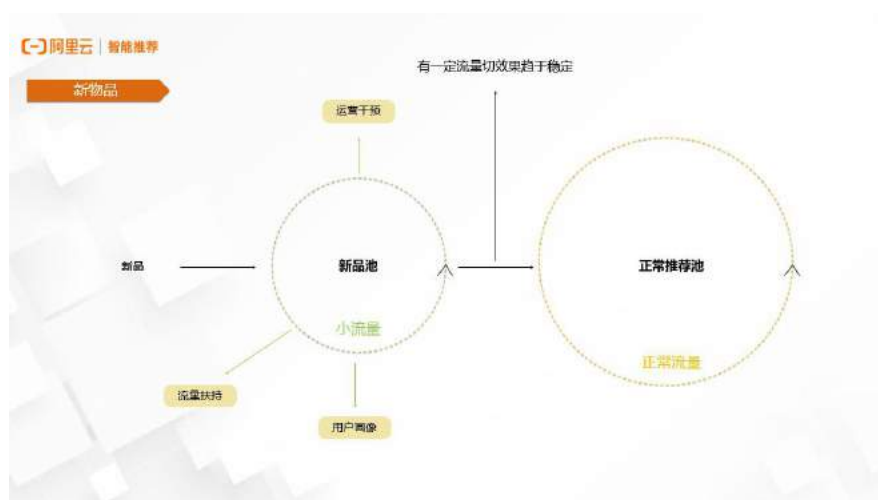
第三步，通过这种相关性，可以拿到多商品，之后就是推荐的最后一步——精排。还是用之前的这些信息量，换个角度，重新再模型计算打分。这样，就完成了一次推荐。

在这个过程中，其实还有一些更细致的考虑，比如当拿到的信息缺失了一部分怎么办，映射到推荐系统，也就是用户冷启动和物品冷启动。



用户冷启动，需要明确一个前提，就是可以拿到这个用户的部分特征（每一个平台用户注册的基本信息），但是没有这个用户过去在这个平台的行为。仅仅从分发效率上说，能想的办法不多。一方面，可通过物品画像进行推荐，比如这个平台上，男生都喜欢什么，那么就给新来的男性推送这些商品。另一方面，也可以想办法扩充用户画像和特征，他在这个平台上是新用户，但并不表示在全网其它平台上也是新用户，这是一个思路。具体的操作需要第三方数据授权，在此也提醒大家，一定要注意隐私保护，因为这些信息对推荐系统来说，确实很重要。最后，如果前两个都不起作用，那就只能用热门推荐了，这是考虑效率后的最优选择。

其实，新用户对一个平台来说，留存是最重要的，所以对于推荐系统来说，就更需要特殊考虑了。以阿里云智能推荐来说，除了分发效果，也会考虑用户心智的探索，比如适当地推荐一些新鲜的东西，让用户能快速建立平台的心智，同时还会结合一些运营的需求。



其次，是新品冷启动，这个问题的核心在于怎么让一个未知的商品在兼顾效率的情况下，快速地热起来。对于我们智能推荐的设计，主要就是两个流量环，我们会给新品单独建立一个索引池，小流量地灰度预热这些商品，它的推荐策略主要是兼容效果和运营需求。当这个商品热起来之后，再放到正常的推荐流量环中。这样做主要的考量因素就是新品的曝光可控性。考虑到如果新品直接放入正常的推荐池中，那么大概率是得不到流量的，如果人工运营强行加权，又会干扰到正常商品池的推荐，还会导致新品自身的流量不均衡。

另外有一点也值得深思，就是怎么判断一个商品热起来了。这里有很多办法，简单点，PV 达到一定量就可以了，再进一步，需要这个商品的 CTR 稳定下来不再波动，这个问题我们会持续迭代探讨下去。

## 智能推荐 2.0 及未来



基于以上产品技术的思考，本次阿里云智能推荐 2.0 发布了。相比之前，智能推荐加入了前文提到的价值判断的环节，目标就是为了优化用户体验。在新品的策略上进行了一些调整，希望做到可控，可调整，有分发效率。同时，2.0 版本也提升了整体推荐效果，模型持续迭代，扩了很多召回，很多基础链路也进行了重构，主要是为了更适宜云上中小客户的行为规模。以上优化仅仅是个开始，我们的目标是推荐赋能用户，这条路还很长，期待我们一起探索。

## 3.4 常规和大促，怎样转变思路应对不同场景

从大促实战，复盘如何利用大促最大程度提升转化与销售增长？

在线上商城，主要有如下 2 大方向：

### 1. 恰到好处的优惠券下发、玩法搭建

从最传统的线下店铺所谓“买一送一”、“返季清仓”，到线上平台的万券齐发、组合优惠，玩法也越来越丰富。优惠券就像是一次购买行为发生的临门一脚，往往能够覆盖到消费者关注的店铺、品牌、品类时发挥出临门一脚的作用。

通过多样化的玩法，比如之前很火的“盖楼”达到消费者驻留时长与活跃度的提升。不容小觑的新起之秀“直播带货”也成为各大平台 GMV 引导的重要火力。对于不同的平台，在直播 KOL 的引进、合作商铺的拓展、业务发展阶段各不相同，往往容易受到多方资源的限制，很难在大促中大放异彩。

### 2. 高转化率商品的分发提效、智能预测

抛开资源等因素限制，我们能够做的常规操作，或者说必备功能，就是在消费者具备消费倾向时，像导购员一样推荐各类商品，在消费者没有明显的消费倾向时，像推销员一样展示可能会转化的商品，来促进一笔笔订单的发生。

那么，如何快速搭建个性化会场，促进转化？

我们从 2 个维度展开：

#### 1、常规操作

想要搭建个性化会场，促进消费转化。“个性化”的前提是，对平台的消费者的消费意向具备合理预测的把握，而消费者意向将会在他的行为中表达，如一次搜索、一次点击、一次收藏、一次下单等等，所以在常规的操作中我们需要记录下每一位消费者的行为。

那我们不由会发问：

存多久的行为？

哪些行为需要着重对待？

哪些行为对于消费者来说只是看看而已，哪些又表达了真实的购买意向呢？

每天产生数以亿计的行为，就这么存着留到大促采用，未免 ROI 不高吧？

那这些行为究竟表达了消费者何种意向，是需要不断的互动、预测、更新发现的，在时间周期上，也同时具备长期兴趣、短期兴趣的区分。这些行为我们可以用来做什么，或者说能够带来什么收益？

- 1、将传统的基于热品、新品、运营精品的推荐功能升级为智能算法、业务驱动的个性化推荐。
- 2、通过个性化的关怀能力，促进新老消费者在平台的粘性、活跃度提升。

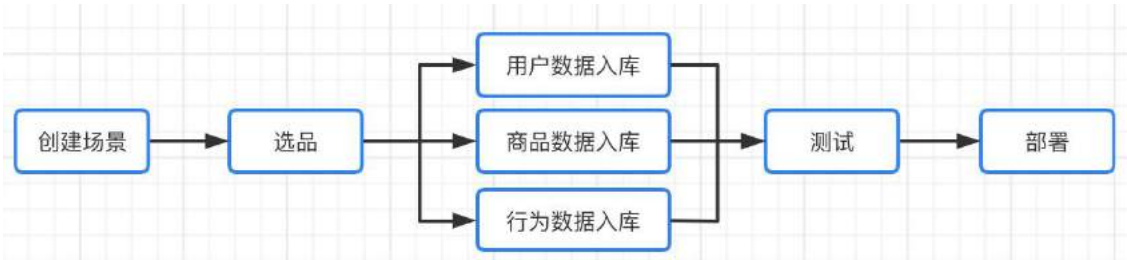
说到智能算法，听上去门槛就很高。而如何快速获取算法能力，匹配业务，加速转型？

通过阿里云智能推荐（<https://www.aliyun.com/product/bigdata/airec>）这个 SaaS 类型的服务即可快速实现。



(背景：智能推荐底层算法是手淘猜你喜欢、相关推荐算法的升级输出，于2019年8月商业化后对外赋能，目前已服务超过20家电商平台，提供千人千面的个性化推荐服务。)

如何利用智能推荐搭建一个常规的商品个性化瀑布流？



仅需 5 步，即可搭建功能。

2、大促操作

参加大促的商品曾积累部分行为（无行为则视为新品），与其他商品不同的是，大促的商品具备优惠、大促生命周期、推广力度等等。

所以，我们可以创建一个“大促”的场景，在这个场景中，只推荐参加活动的商品。

在商品维度

需要与其他场景的商品存储形态保持一致。

更新商品的生命周期与上下架的信息，尤其是下架（当商品库存不足时，及时下架商品，避免流量浪费）。

维护商品的特征信息如类目、标题、品牌、店铺等。



在业务维度

需要扶持部分优质的商品，可以在商品层面进行加权的推广，从而实现其流量分发的增长，促进转化。



基于平台售卖商品品类、店家等级、高利润商品等因素，商家、平台需要**扶持商品**。而对商品的扶持在一定程度上容易与用户兴趣表达相悖，容易导致整体推荐效果下跌，AIRec基于此情况封装了同时兼顾个性化与商品扶持的方案。



### 在消费者维度

能够在活动商品中进行互动，不断增强购买兴趣的刻画与推荐的反馈。

实时交互是促进消费者沉浸式浏览的**必备基础功能**。AIRec可实时学习消费者当前兴趣表达、变化，并更新在下次生成的推荐结果中，从而实现实时的互动式推荐功能。



### 在平台维度

减缓大促中常见的二八效应，促进平台流量的合理化分配，实现用户转化率提升的最大化。

## 4. 视频课程

---

### | 4.1 搜索篇

#### 1. 从阿里巴巴搜索业务中台到行业智能搜索云服务的实践



观看地址: <https://developer.aliyun.com/live/43906>

#### 2. 智能搜索技术在电商行业的最佳实践



观看地址: <https://developer.aliyun.com/live/43907>

#### 3. OpenSearch 知名客户业务实践分享——企查查



网址: <https://developer.aliyun.com/live/43913>

#### 4. OpenSearch 知名客户业务实践分享——趣店



观看地址: <https://developer.aliyun.com/live/43914>

#### 5. 企业智能推荐&Opensearch 案例分享——安利



观看地址: <https://developer.aliyun.com/live/43915>

## | 4.2 推荐篇

### 6. 0-1 激活推荐系统，解锁冷启动难题



观看地址: <https://developer.aliyun.com/live/2614>

### 7. 7 天搭建电商个性化推荐场景和最佳实践



观看地址: <https://developer.aliyun.com/live/2506>

### 8. 推荐场景与用户心智



观看地址: <https://developer.aliyun.com/live/43737>

## 9. 智能推荐 AIRec 功能使用指南



观看地址: <https://developer.aliyun.com/live/43387>

## 10. 炼数成金：智能推荐大数据和 AI 实践



观看地址: <https://developer.aliyun.com/live/43916>

了解更多:

智能推荐系列公开课



观看地址: <https://developer.aliyun.com/article/746419>



钉钉扫码加入  
“推荐与搜索技术交流群”



微信扫码关注  
“搜索与推荐技术”  
公众号



扫码关注订阅  
搜索与推荐技术”社区



阿里云开发者“藏经阁”  
海量免费电子书下载