




推荐系统的工作流程

 第四范式...
已认证的官方帐号

关注他

7 人赞同了该文章

在互联网飞速发展的现代社会，人们每天都要受到成百上千条信息的轰炸，APP推送、新闻热点、信息流广告.....一个有效的“信息过滤器”已经成为了人们日常生活的刚需，也是信息供应商在激烈的市场环境中脱颖而出的必杀技。

推荐系统正扮演着这样一个角色，它如同筛子一般，帮我们找到最想要的内容。但是，推荐系统过高的技术门槛和研发成本把很多企业挡在了门外。第四范式基于机器学习技术推出的智能推荐产品先荐，专注于内容行业的个性化推荐，凭借自身的技术优势有效解决着这一难题，已经服务了36氪、花瓣、果壳等知名媒体，不断受到行业内的广泛好评。

在接下来的文章中，先荐将系统地讲解推荐系统的相关知识，希望各位技术爱好者能对推荐系统有更多、更多的了解。首先，我们将从推荐系统的工作流程讲起。

1. 信息收集阶段

知乎

推荐系统依赖于不同类型的输入，例如最直接的显式反馈，即用户直接输入感兴趣的内容，或隐式反馈，即通过观察用户行为间接地推断用户偏好，还可以通过显式和隐式反馈的组合来获得混合反馈。

以网络学习平台为例，用户画像是与特定用户相关联的个人信息的集合。这些信息包括该用户的认知技能、智力水平、学习方式、兴趣爱好以及交互行为等。用户画像通常用于用户模型构建时信息检索所需。换句话说，用户画像粗略地反映了用户模型。要想做成功一个推荐系统，很大程度上取决于其对用户兴趣的表征能力。要想获得准确的推荐结果，准确的用户模型必不可少。

1.1 显式反馈

网站一般会在用户操作界面上提示用户对内容做出评价，以便构建和改进该用户的用户模型。推荐结果的准确性取决于用户提供的评级数量。用户的评级数量越多，推荐结果越准确。显式反馈的唯一缺点是，非常依赖用户评级的积极性，而且，用户不是时时刻刻愿意做出评级。不过，相比之下，显示反馈不涉及到从用户行为中获取用户偏好这一步，因此提供的数据更可靠，整个推荐过程也更透明，能够更好地感知推荐系统的质量，从而提高用户满意度。

1.2 隐式反馈

网站后台通过监测用户的不同行为，自动推测用户的兴趣偏好，例如购买历史、导航历史，在某些网页上停留的时间、用户点击的链接、按钮、以及电子邮件内容等。隐式反馈从用户行为中推断用户的偏好，减轻了用户的评级负担。隐式反馈对用户评级的积极性要求不高，准确性也较低。

也有一些人认为，用户隐性反馈的数据实际上更客观，在隐式反馈的情况下，用户不需要以社会大众期望的方式做出反应，也没有任何维护自我形象的需求，因此提供的数据更真实。

1.3 混合反馈

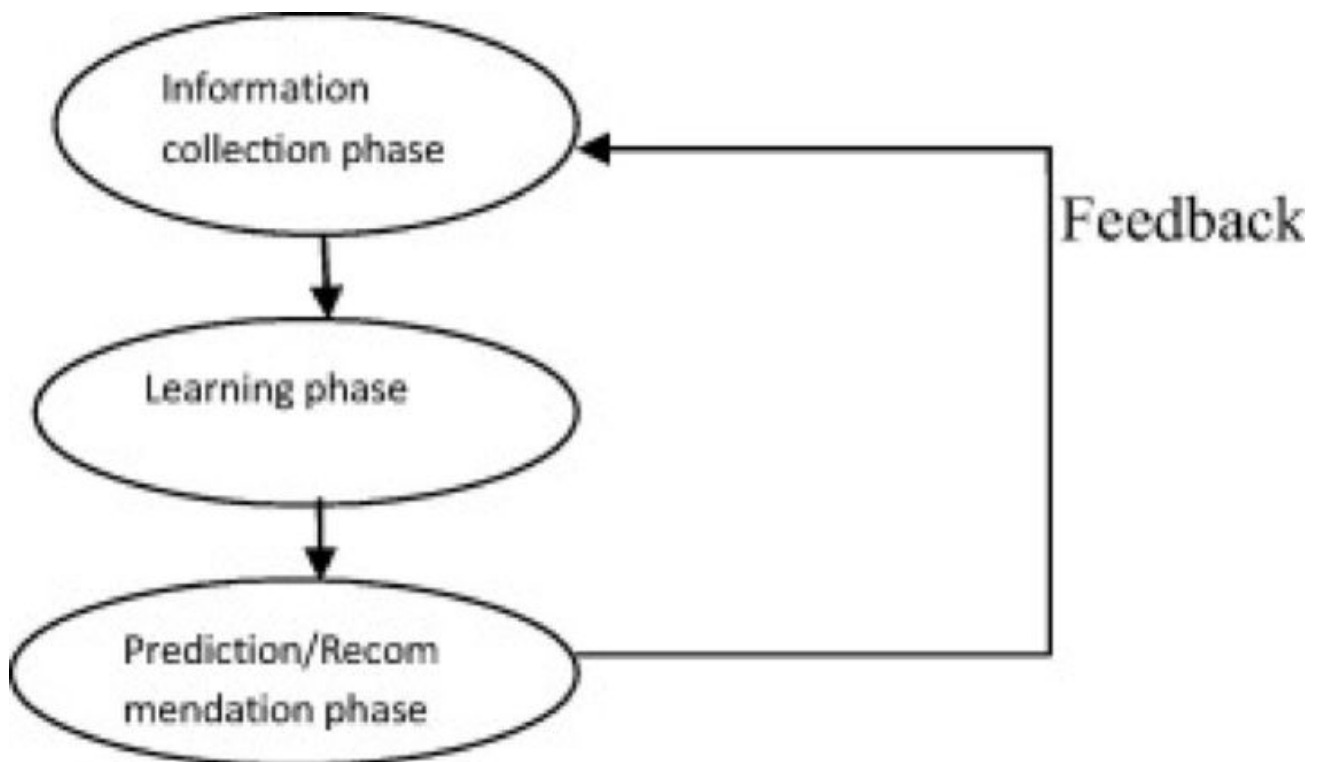
隐式和显式反馈的优势可以在混合系统中结合，以最大限度地降低二者的不足并实现性能最佳的推荐系统。具体来看，用隐式反馈的数据来校验显式反馈的数据，或仅允许用户在表达明确兴趣时给出显式反馈。

2. 算法学习阶段

在这一阶段，系统会通过学习算法，过滤上一阶段得到的用户反馈，并提取用户特征。关于这一部

知乎

在这一阶段，系统会预测用户可能喜欢的内容类型。这一步可以直接基于在信息收集阶段收集的数据集（基于存储器或基于模型）来实现，也可以通过后台监测到的用户行为来实现。



推荐系统的工作流程

在下一篇文章中，我们将会详细介绍推荐的过滤技术，敬请期待。

相关阅读：

[用Python搭建推荐系统的最佳开源包](#)

[如何用Python搭建一个简单的推荐系统？](#)

[想要了解推荐系统？看这里！（2）——神经网络方法](#)

[想要了解推荐系统？看这里！（1）——协同过滤与奇异值分解](#)

[入门推荐系统，你不应该错过的知识清单](#)