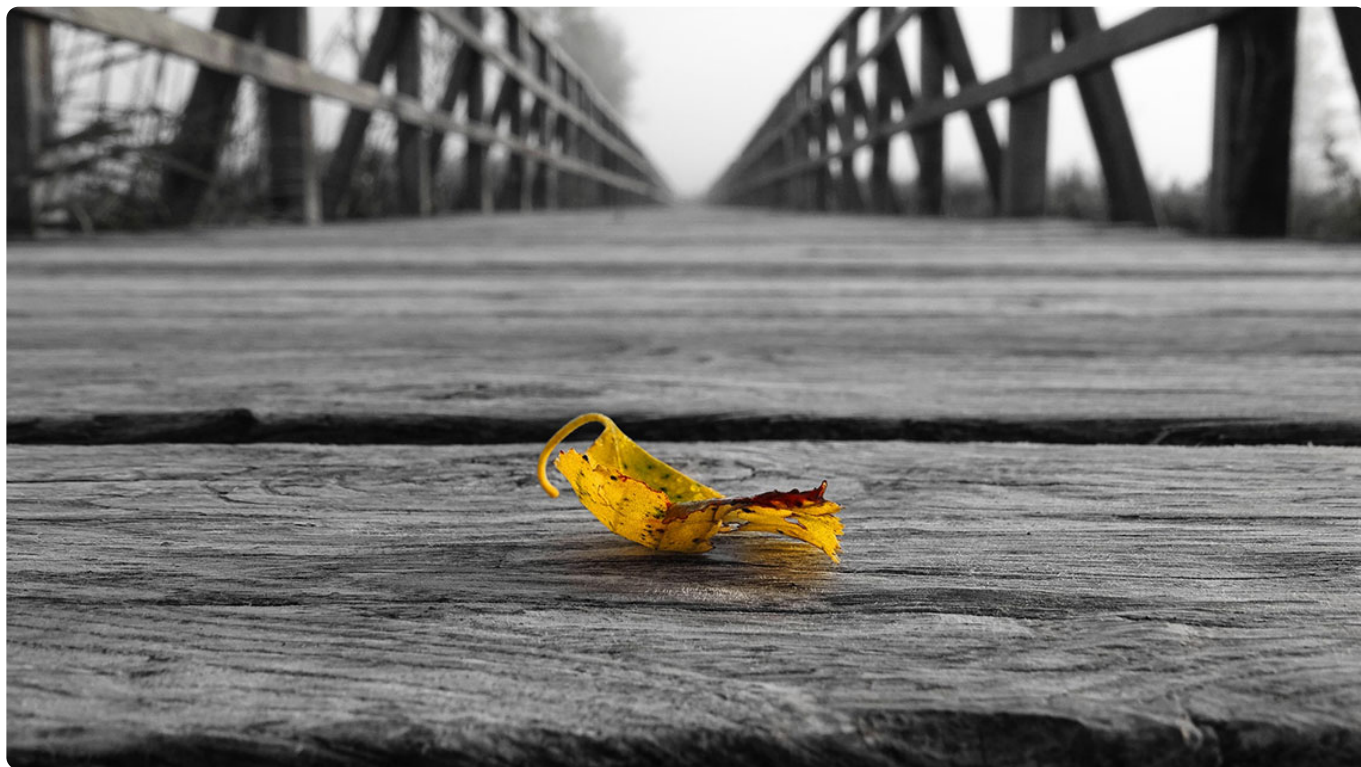


072 | 推荐的Exploit和Explore算法之一：EE算法综述

2018-03-19 洪亮劼

AI技术内参

[进入课程 >](#)



讲述：初明明

时长 07:55 大小 3.63M



上周，我们聊了一些比较高级的模型，包括张量分解和协同矩阵分解，讨论这些模型如何能够抓住更多的用户和物品之间的关系。最后，我们还讨论了如何优化更加复杂的目标函数。

这周，我们来看一个完全不同的话题，那就是 **Exploitation**（利用）和 **Exploration**（探索）的策略，俗称“**EE 策略**”。

一个推荐系统，如果片面优化用户的喜好，很可能导致千篇一律的推荐结果。其实，EE 策略是推荐系统里很有意思，但也非常有争议的一个话题。一方面，大家都明白这类算法的目的，每年有很多相关论文发表。但另一方面，工业界对于部署这类算法非常谨慎，有的产品经理甚至视之为“洪水猛兽”。我们今天就来分析一下导致这个现象的一些因素。

走一步看一步的策略

这里再简单阐述一下什么是 EE。简单来说，就是我们在优化某些目标函数的时候，从一个时间维度来看，当信息不足或者决策不确定性（Uncertainty）很大的时候，我们需要平衡两类决策：

选择现在可能是最佳的方案；

选择现在不确定，但未来可能会有高收益的方案。

在做这两类决策的过程中，我们也逐渐更新对所有决策不确定性的认识。最终，从时间的维度上来看，我们在不确定性的干扰下，依然能够去优化目标函数。

也就是说，**EE 可以看作是一个优化过程，需要多次迭代才能找到比较好的方案。**

EE 的应用历史

早期把 EE 应用于新闻推荐系统的文章，主要关注在雅虎的今日新闻（Today Module）这一产品上，这也基本上是 EE 最早在互联网应用的尝试，目的是为了优化点击率（CTR）。而更早的一些奠基性的文章，则是在广告的数据集上展示实验结果。

雅虎的今日新闻其实为 EE 提供了一些被很多学者和工业界人士忽视了的条件和成功因素。如果不考虑这些因素，鲁莽地在其他场景中使用这些文献中相似的算法，很可能会产生相当差的效果。那么是哪些因素呢？主要有两点。

1. **相对少量的优质资源。** 今日新闻每天的内容池（Content Pool）其实并不大。这里面都是网站编辑精选了的大概 100 篇文章。这些文章原本的质量就非常高，无论是这里面的任何一组，用户体验都没有明显变差。内容池每天都人为更换。
2. **非常大的用户量。** 有亿万级的用户，最终可能是从算法随机产生的文章排序中选择了阅读的文章。因为用户数量巨大，所以算法就相对比较容易收敛（Converge）到稳定的方案，也就是前面讲的，优化 CTR 的状态。

正因为有了以上两个因素，在这个模块上，工程师和科学家们享受了后来学者们所无法想象的“奢侈”，比如运行 Epsilon-Greedy 这种简单粗暴的 EE 算法；甚至是完全随机显示新闻，收集到了大量无偏（Unbiased）的数据，都为很多学术工作奠定了数据基础。时至今日，依然有不少后续学者，在今日新闻随机数据的基础上进行算法改进。

如果没有了这两条因素，最早的解决方案可能都没法在当时的雅虎施行。原因很简单，如果资源良莠不齐，资源数量非常大，那么在仅有的几个展示位置，优质资源显示的可能性在短期内就会比较小（因为系统对于大多数的资源还有很高的不确定性，需要 Explore）。

由于优质资源显示得少了，用户就会明显感受到体验下降，最直接的可能就是倾向于不点击甚至放弃使用产品。用户不和系统交互这样的行为，又进一步减缓了系统学习资源不确定性的速度。这时，也许亿万级用户数都没法满足学习所有资源的用户数量（毕竟所有用户只有一部分会落入 Exploration）。

关于这个解决方案有一个很有意思的点值得一提，在雅虎的研究人员跳槽到了 LinkedIn 以后，LinkedIn 也推了相似的方案。为了模拟雅虎今日新闻的这些条件，就对用户内容流里的数据进行了大规模的过滤。这样就有少数的信息符合高质量的要求，并且能够在用户数允许的情况下探索到合适的解决方案。

EE 的产品部署难点

我们来讲一下 EE 的产品部署难点，这些难点普遍高于具体的 EE 算法选择（比如选某一个 UCB 或者某一个 Thompson Sampling）在产品工程解决方案上的抉择。

为了便于讨论，我们把所有 EE 算法整体叫作“Random”，简称“**R 算法**”，而把不做 EE 的算法叫作“Deterministic”，简称“**D 算法**”。这里面的假设是，D 算法比较静态，能够产生高质量、一致性的内容。这里的一致性是指用户在短时间内的体验比较稳定，不会有大幅度的界面和内容变化。相反，整体来说，R 算法的不确定性比较大，用户体验和产生的内容可能会有比较大的波动。

第一个难点是如何上线测试。这看上去不应该是难点，但实际上需要格外小心。传统 EE 文献，只是把问题抽象为每一次访问需要做一个决策的选择。然而，文献却没有说，这些访问是否来自同一个用户。

那么，理论上，EE 应该对所有的访问不加区别，不管其是否来自同一个用户。用我们这篇文章的术语来说，就是所有的流量都做 R 算法。虽然从理论上讲这样没有问题，但实际上，用户体验会有很大的差别。特别是一些推荐网站，用户希望自己前后两次对网站的访问保持一致性。如果不加区分地做 R，对于同一个用户来说，很可能两次看见的内容完全迥异。这对用户熟悉产品界面，寻找喜爱的内容会产生非常大的障碍。

那么，我们对绝大部分用户做 D，对另外一小部分用户做 R，这样会好一些吗？这其实就是“牺牲”少部分用户体验，来换取绝大多数用户体验的一致性。这样实现也是最直观的，因为很多在线系统的 A/B 测试系统是根据用户来进行逻辑分割的。也就是说，某一部分用户会进入一个 Bucket，而另一批用户会进入另外一个 Bucket。按用户来做 D & R 可以很容易和 Bucket System 一致起来，方便部署。当然，这样做也是有潜在风险的。那就是，这部分老做 R 的用户，在当了别人的小白鼠以后，很可能永远放弃使用产品。

另外一个难点就是如何平衡产品。EE 几乎是一定会导致用户对产品的体验下降，至少是在短期内会这样。如何弥补这一点，技术上其实比较困难。比如做过滤是一种思路，那就是只在优质内容里探索。当然，有人会说，这样其实也没有多大的意义。然而，一旦把质量的闸门打开了，那就会对用户体验带来很大的影响。

这也是很多产品经理对于 EE 非常谨慎的原因，能不做就不做。而且，牺牲了用户体验后，EE 所带来的好处其实是很难评测的，这主要是由线上产品的评测机制和评测原理决定的。目前还没有比较统一的解决方案。

如何能够做到“用户友好型”的 EE 呢？这里面可以有两种思路。

思路一，不是所有人群的所有访问都适合做 R。和传统 EE 不同的是，做“反向 EE”。也就是说，我们只针对常用产品的用户做探索，而并不是针对新用户或者是还没有那么熟悉系统的人群。这个思路和 EE 完全相反，但是更加人性化。

思路二，夹带“私货”。也就是更改 EE 的算法，使得高质量的内容和低质量的内容能够相伴产生，并且高质量的内容更有几率排在前面。这样用户体验的损失是可控的。

其实，EE 和产品的结合点应该是工程和研究的重点，但遗憾的是，碍于数据和其他方面的因素，这方面的研究工作几乎没有。

小结

今天我为你讲了推荐系统的一个重要的问题，就是如何持续挖掘用户可能的喜好，也就是做 EE。

一起来回顾下要点：第一，我们讲解了什么是 EE；第二，我们介绍了 EE 的一些简要历史；第三，我们讨论了 EE 的部署难点。

最后，给你留一个思考题，EE 策略是不是一定会导致用户看到多样不同的东西呢？

欢迎你给我留言，和我一起讨论。



AI 技术内参

你的360度人工智能信息助理

洪亮劼

Etsy 数据科学主管
前雅虎研究院资深科学家



新版升级：点击「👤 请朋友读」，10位好友免费读，邀请订阅更有**现金**奖励。

© 版权归极客邦科技所有，未经许可不得传播售卖。页面已增加防盗追踪，如有侵权极客邦将依法追究其法律责任。

上一篇 071 | 高级推荐模型之三：优化复杂目标函数

下一篇 073 | 推荐的Exploit和Explore算法之二：UCB算法

精选留言 (2)

写留言



吴文敏

2018-04-03



工业界用EE主要是一部分是为了解决冷启动问题，反向EE有助于解决新物品的冷启动问题。对于新用户可以从topN的物品中使用R算法探索用户兴趣



林彦

2018-03-19



EE策略需要用户花足够时间探索，同时保证总有新的高质量/高匹配度的内容的产生。此外如果对受欢迎的内容没有什么惩罚(有些产品形态会导致受欢迎的内容倍增地受用户关注)，对内容的新旧没有什么系数调整，则用户看到不同的东西的几率更小。

也和用户本身的做事风格有关，有的用户更倾向于看自己熟悉的东西，有的则愿意经常...
展开 ∨

拼课微信：171614366!