

聊一聊推荐系统中Exploit&Explore算法

原创 天渡 天渡漫谈 5天前

Exploit: 利用 Explore: 探索

推荐系统中，如果片面优化用户的喜好，推荐系统可能会造成信息茧房问题。也就是推荐的东西都是用户喜欢的东西，千篇一律。这时候不仅仅需要Exploit，还需要Explore新内容。另外呢，EE还可以通过Explore信息不足的物品（如很少曝光），提高对其信息的掌握程度（如ctr等）。随着时间的推移，推荐系统对item的信息掌握越来越多，也就可以更好的做出决策。

但是在工业界，EE算法其实是一个很矛盾的东西。上吧，确实可以提高新颖度，但是谁知道是正向影响还是负向影响。若不上EE，之前提到的种种好处也就无从谈起。但是，万一这些新东西很严重的损害了用户体验，造成用户流失，公司得哭死。所以，工业界对EE都有着很谨慎的态度。

EE算法这种思路是很好的，就是在实践中需要仔细斟酌使用方法，现在来看看工业界常用的EE算法：

首先是**EG(Epsilon-Greedy)算法**，这种算法思路很简单。以推荐场景为例，设定一个参数 P ，其中 $p\%$ 的人按照ctr大小顺序进行商品推荐，剩下的 $(1-p)\%$ 用户随机推荐商品。 $p\%$ 用户的策略为Exploit， $(1-p)\%$ 用户的策略为Explore。这样造成的后果是这 $(1-p)\%$ 用户的体验可能会遭到很严重的损害，以至于流失。而另外 $p\%$ 的用户体验反而因为这 $(1-p)\%$ 的用户而变好了，因为推荐系统牺牲了他们的用户体验来换取商品的信息(ctr等)，让系统对商品的推荐更加准确，从而提升了 $p\%$ 用户按照ctr推荐的用户体验。EG算法的随机性太强了，会对用户体验造成很不好的影响，有没有什么办法可以减小这种影响呢？答案肯定是有的。比如筛选出商品的优选池，每次Explore的商品都在优选池中选。有的小伙伴可能会说既然规定优选池了，里面的商品都是经过人工筛选过了的，那么Explore还有意义吗？好，那我们在Explore时，一部分从优选池中选，另外一部分完全随机。如此， $(1-p)\%$ 的用户即使有部分商品是随机呈现的，但因为有优选池商品的加持，体验损害有限。这些EG中减小用户体验损失的措施可以根据具体的业务场景再进行尝试。

第二个为**UCB(upper confidence bound)算法**。UCB算法是引入了置信度这个概念。这个置信度可以宽泛的去理解，大致逻辑是：该商品被选中的次数越多，其置信度

就应该越高，越接近真实值。最终UCB是根据均值+置信度的和进行排序的。UCB算法舍弃了随机推荐的概念，其实就是一个确定性(deterministic)算法了，根据置信度产生的随机是伪随机。极端来说，若每个商品的均值就是真实值了，那么每个人看到的都是一样的。这可不就是确定性算法了嘛。

第三个为**汤普森采样(Thompson Sampling)**。汤普森采样先给商品的信息(ctr)定义了先验分布，然后利用每次的观察结果去计算后验分布，从每个商品的后验分布中采样生成随机数，取这些随机数中最大的进行推荐展示，依次循环。由于每一轮汤普森采样中，都有根据分布采样随机数的过程，所以汤普森采样是个随机的过程。为了每一轮迭代，方便先验概率与后验概率转化，可以使用共轭先验。即先验概率根据观察结果更新后验概率时，分布形式不变，只有参数发生变化。对于伯努利分布来说，共轭先验是Beta分布。在后续迭代中，根据实验结果更新Beta分布的a、b参数即可。这也就是为什么在ctr预估时贝叶斯平滑采用Beta先验分布的原因。

喜欢此内容的人还喜欢

【AI100问(28)】什么是模拟退火算法？

清语赋

主流机器学习算法概述

网优小谈

【AI100问(25)】什么是遗传算法？

清语赋