

5.6 推荐系统评价

CSDN学院
2017年11月

- 推荐系统出现的背景
- 基于协同过滤的推荐
- 基于内容的推荐
- **推荐系统的评价**
- 案例分析

- 一个完整的推荐系统一般涉及3个参与方：用户、物品、网站
- 评价指标
 - 用户满意度(调查问卷、购买率、点击率、转化率、停留时间...)
 - 预测准确率(离线系统)
 - 覆盖率(对long tail物品的发掘能力)
 - 多样性(用户的兴趣本身是多样的，推荐不同的商品，用户选择更多)
 - 新颖性(将用户有过历史行为的物品从推荐列表中去掉)
 - 惊喜度(surprise，推荐用户不知道的商品)
 - 信任度(提供可靠的推荐理由)
 - 实时性(实时更新程度)
 - 健壮性(抗击作弊的能力)
 - 商业目标

- 打分系统:如豆瓣电影的电影评分，淘宝的商品打分
- 设 r_{ui} 为用户 u 对物品 i 的实际评分， \hat{r} 为预测分。 $|T|$ 是测试样本集，用 $RMSE$ 和 MAE 来评判预测打分和真实打分之间的差异性：

$$RMSE = \sqrt{\frac{\sum_{u,i \in T} (r_{ui} - \hat{r}_{ui})^2}{|T|}}$$
$$MAE = \frac{\sum_{u,i \in T} |r_{ui} - \hat{r}_{ui}|}{|T|}$$

► 准确度

- Top K 推荐：用准确率和召回率评定用户对推荐的喜欢程度
- 设 $R(u)$ 为模型对测试集中用户 u 的推荐， $T(u)$ 为测试集上用户的选择，准确率和召回率分别为：

$$Precision = \frac{\sum_{u \in U} |R(u) \cap T(u)|}{\sum_{u \in U} |R(u)|}$$

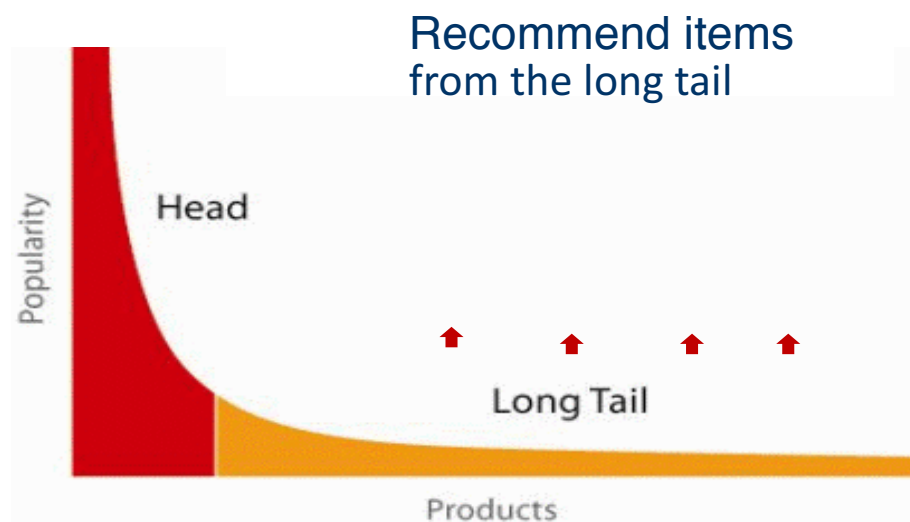
推荐的商品中用户感兴趣的比例

$$Recall = \frac{\sum_{u \in U} |R(u) \cap T(u)|}{\sum_{u \in U} |T(u)|}$$

用户感兴趣的商品中被推荐到的比例

► 覆盖率

- 推荐系统对物品长尾的发掘能力(消除马太效应)



推荐商品的集合占所有需要推荐商品集合I的比例

$$Coverage = \frac{|U_{u \in U} R(u)|}{|I|}$$

$$H = - \sum_{i=1}^n p(i) \log p(i)$$

$p(i)$: 推荐商品 i 的概率

I : 表示商品的总数

► 推荐系统面临的挑战

- 冷启动
- 多优化目标
- 异构的、多源数据

THANK YOU



AI100