

本章重难点梳理

学而需要时习之，下面让我们来快速回顾一下本章的内容，看看你都掌握了多少呢。

- 推荐系统模块

- 数据生产
 - 将有用的信息汇总、清洗，为后续分析和推荐做准备
- 数据储存
 - 持久化存储收集到的数据，方便后续不同阶段和模块灵活取用
- 算法召回
 - 将海量候选集进行初步筛选，把数据集规模降低至数百量级
- 结果排序
 - 对召回的结果进行进一步筛选并精确排序
- 结果应用
 - 根据不同的推荐场景给用户展示最终的推荐结果

- 推荐系统常用特征

- 用户特征
 - 自然属性：姓名、性别等
 - 画像特征：兴趣、行为
 - 关系特征：人群关系、关注关系
- 物品特征
 - 静态特征：分类、标签、价格
 - 动态特征：热度、标签热度
 - 相关性特征：主题相似度、分类匹配度

- 推荐系统评估标准

- 满意度：准确率、停留时长、转化率等
- 覆盖率：长尾商品能否被推荐
- 多样性：推荐的物品是否两两不相似
- 新颖性：是否能推荐用户之前没见过的东西
- 实时性：能否根据用户最新的喜好实时更新推荐结果

- 课程项目模块

- 数据处理模块
 - 处理、清洗以及可视化原始数据
- 召回层模块
 - 提供猜你喜欢和相似推荐的召回结果
- 排序层模块
 - 提供猜你喜欢场景的排序服务
- API接口服务模块
 - 组织上游模块，并对外暴露RESTful API
- 前端页面
 - 供用户访问推荐系统的网页