|  |  |
| --- | --- |
| 3 | 推荐系统 |
| 主要目的 | 对于前10名用户（具有更多checkins）构建：  - 一篮子推荐：场地（地方）  - 用户将根据他们的朋友访问的可能场馆（地点）列表。  - 他们将在哪里（with probability scores）。  对于前10名“社交”用户（与更多朋友）  - 将路径（带地图）绘制一周/月的用户checkins。  - 列出您的朋友以及他们在“品味”方面的近距离（基于场地访问和排名） |
| 参考数据 | 使用以下表格给出了一个sqlite（fsdata.db）文件：  •用户 - 包括一组用户，使得每个用户具有唯一的ID和代表用户home town位置的地理空间位置（纬度和经度）。  •场地 - 由一系列场地（例如餐馆）组成，以便每个场地都有一个唯一的ID和地理空间位置（纬度和经度）。  •Checkins - 标志着用户在场地的访问。每次访问都有一个唯一的ID以及用户ID和Venue ID。  •Sockoggh - 包含用户之间存在的社交图形边缘（连接）。每个社交连接由两个唯一ID（first\_user\_id和second\_user\_id）表示的两个用户（朋友）组成。  •评级 - 包括隐含的额定值，这些评级量化了用户喜欢特定场地的多少。  You can download the data from: <https://mltestpublicdata.s3-ap-southeast-1.amazonaws.com/fsdata.db.zip> |
| 参考工具 | 1. Pandas/numpy 2. RecQ, DeepRec, supervise, etc |
| 建议人数 | 4-5人； |
| 难度分 | 30/30 |
| 备注 | 1. 注意大数据集合的处理 2. 可以使用任何可行推荐算法；不限于上述参考 3. 使用jupyter notebook保存结果 |