

Kaggle event 推荐比赛及解答思路

七月在线 寒老师
2016年8月14日

问题简单分析

■ 首先这是一个推荐的问题

■ 我们有下面这样几类数据

1. 用户的历史数据 => 对event是否感兴趣/是否参加

2. 用户社交数据 => 朋友圈

3. event相关的数据 => event相关的信息

■ 简单思考

1. 要把更多维度的信息纳入考量

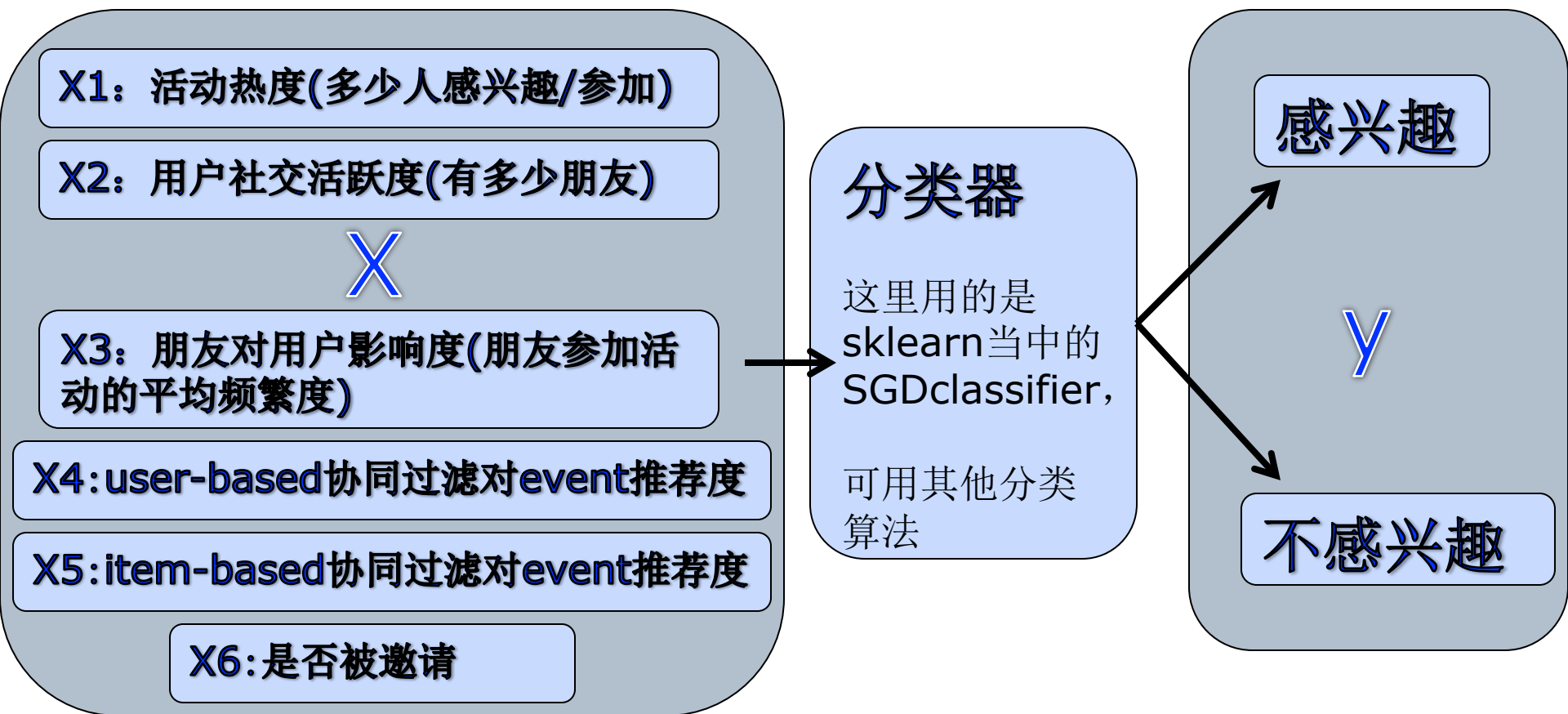
2. 协同过滤是基于 用户-event 历史交互数据

3. 需要把社交数据 和 event相关信息 作为影响最后结果的因素纳入考量

4. 视作分类模型，每一个人 感兴趣/不感兴趣 是target，其他影响结果的是feature

5. 影响结果的feature包括由协同过滤产出的推荐度

简图



更多细节

请参见
ipython notebook
与
对应注释

感谢大家！

恳请大家批评指正！