# 推荐系统考试内容3

Distance with similarity selection

实验步骤与要求：

1. 选择MovieLens small 数据集；
2. 按照试题1的数据划分步骤划分训练集和测试集；
3. 计算用户Pearson间相似性PCC；
4. 当用户间PCC大于阈值th，则计算用户间距离distance并保存；
5. 注意u到v的距离为d时，v到u的距离为-d都需要保存；
6. 针对测试集中预测目标，按照slope one预测用户的可能评分；
7. 基于考试题目2，画出不同邻居数下DSS与Pearson CF和CosineCF的各种指标对比；

------------------------------------分割线------------------------------

上交材料要求：

1. 打印版纸质的作业报告；首页、目录单面打印，报告正文双面打印；
2. 实验报告文档电子版，pdf或word文件均可；目录名：学号姓名
3. 实验代码和数据电子版，.py和.ipynb均可，实验代码与所需数据放在一个子目录中，这个子目录放置在实验报告电子档的目录中；
4. 自评报告纸质版：一张打印，按照评分要求对自己上交的材料进行评分，分析优势和扣分原因；

------------------------------------完结线------------------------------