第二章全部算法使用的数据集均为movie-lens 100k

获得参数说明，使用python XXX.py --help

如python userCF\_rating.py –help

## 各算法对应文件

|  |  |
| --- | --- |
| 算法名 | 对应文件 |
| 针对TopN推荐的userCF | userCF\_TopN.py |
| 针对评分预测的userCF | userCF\_rating.py |
| 针对TopN推荐的itemCF | itemCF\_TopN.py |
| 针对评分预测的itemCF | itemCF\_rating.py |
| 基于距离的相似度度量 | itemCF\_dis.py |
| slopeOne算法 | SlopeOne.py |
| 激活扩散模型 | spreadingActivation.py |
| 物质扩散模型 | spreadingSubstance.py |
| 热传导模型 | thermalConduction.py |

## TopN推荐算法运行结果：

|  |  |  |  |
| --- | --- | --- | --- |
| 算法名 | Precision | Recall | Coverage |
| userCF-余弦相似度 | 0.1862 | 0.1862 | 0.2598 |
| userCF-杰卡德相似度 | 0.1881 | 0.1921 | 0.2331 |
| itemCF-余弦相似度 | 0.1779 | 0.1796 | 0.1272 |
| itemCF-条件概率 | 0.1466 | 0.1483 | 0.0642 |
| itemCF-距离 | 0.1547 | 0.1558 | 0.0660 |
| 激活扩散模型 | 0.1027 | 0.1071 | 0.0434 |
| 物质扩散模型 | 0.1244 | 0.1283 | 0.0488 |
| 热传导模型 | 0.0042 | 0.0044 | 0.2122 |

## 评分预测算法运行结果

|  |  |  |
| --- | --- | --- |
| 算法名 | MAE | MSE |
| userCF-皮尔逊相似度 | 0.9615194990982957 | 1.52177173419827 |
| userCF-余弦相似度 | 0.9795184841124885 | 1.530409007014136 |
| itemCF-皮尔逊相似度 | 1.1449563800161946 | 2.22600813251409 |
| itemCF-余弦相似度 | 1.1399839529656934 | 2.219480403341986 |
| SlopeOne | 0.739291913465804 | 0.9122661047832548 |

## jupyter代码

jupyter代码位于RSBook\lec2-jupyter，对应教材上的示例，使用教材上的数据。

|  |  |  |
| --- | --- | --- |
| 算法名 | 对应文件 | 数据集 |
| 针对TopN推荐的userCF | userCF-TopN.ipynb | /data/TopN-jupyter.txt |
| 针对评分预测的userCF | userCF-rating.ipynb | /data/Rate-jupyter.txt |
| 针对TopN推荐的itemCF | itemCF-TopN.ipynb | /data/TopN-jupyter.txt |
| 针对评分预测的itemCF | itemCF-rating.ipynb | /data/Rate-jupyter.txt |
| slopeOne算法 | slopeOne.ipynb | /data/slopeOne-jupyter.txt |
| 激活扩散模型 | spreadingActivation.ipynb | /data/TopN-jupyter.txt |
| 物质扩散模型 | spreadingSubstance.ipynb | /data/TopN-jupyter.txt |
| 热传导模型 | thermalConduction.ipynb | /data/TopN-jupyter.txt |