|  |  |
| --- | --- |
| 1，模式挖掘 | **完成**  **成果**：  形成一种新的基于模式挖掘的评分填充算法。并在填充后的矩阵的基础上提出了改进的相似度计算方法。  **新问题**：  由于原始矩阵的稀疏，传统的计算相似度的复杂度比较小。  如果直接在填充后的相似度矩阵上使用传统的在线计算，则**复杂度大大增加。**  **解决办法**：填充后的矩阵，对于填充的部分使用**离线计算，（填充评分属于代替评分，由于无法做到太高的精确性，因此不需要太高的实时性，离线计算比较合适）** |
| 2，（模糊）聚类 | 在论文中加入聚类的方法（创新点，k-means算法初始选中心点的方法） |
| Idea1  1,选取打分普遍高的用户Uo（积极用户），打分普遍低的用户Up（消极用户），计算出中立用户Un（是计算出来的，实际上不存在）。得到K-means算法初始的三个中心点。  2，执行k-means算法，最终收敛后，得出三个分组。  3，执行模糊k-means算法，得出三个模糊分组  4，计算MAE |
| Idea2  1，计算每个用户的熵，根据熵的分布，选择初始的几个用户作为中心点。  2，进行k-means聚类。  3，计算MAE |
| 3，用户分组 |  |
| 4，信息熵（暂不研究） | 参考：基于加权信息熵相似度的协同过滤算法研究--刘文龙 |

# ML100K

## PCC

PCC\_MAE

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | PCC | HybirdSimilarity3  (len=2) | HybirdSimilarity3  (len=3) | HybirdSimilarity3  (len=4) | HybirdSimilarity3  (len=5) | HybirdSimilarity3  (len=6) | HybirdSimilarity3  (len=7) | HybirdSimilarity3  (len=8) | HybirdSimilarity3  (len=9) | HybirdSimilarity3  (len=10) | HybirdSimilarity3  (len=11) | HybirdSimilarity3  (len=12) | HybirdSimilarity3  (len=13) | HybirdSimilarity3  (len=14) | HybirdSimilarity3  (len=16) | HybirdSimilarity3  (len=18) | HybirdSimilarity3  (len=20) |
| 5 | 0.81401 | 0.80308 | 0.80367 | 0.80392 | 0.80564 | 0.80784 | 0.80765 | 0.80697 | 0.80596 | 0.80628 | 0.80504 | 0.80507 | 0.80542 | 0.80479 |  |  |  |
| 10 | 0.78190 | 0.77380 | 0.77432 | 0.77536 | 0.77591 | 0.77575 | 0.77614 | 0.77580 | 0.77599 | 0.77542 | 0.77489 | 0.77472 | 0.77482 | 0.77478 |  |  |  |
| 15 | 0.77053 | 0.76367 | 0.76375 | 0.76346 | 0.76375 | 0.76477 | 0.76493 | 0.76454 | 0.76426 | 0.76382 | 0.76362 | 0.76362 | 0.76341 | 0.76349 |  |  |  |
| 20 | 0.76463 | 0.75874 | 0.75876 | 0.75906 | 0.75973 | 0.76021 | 0.76015 | 0.75961 | 0.75929 | 0.75903 | 0.75888 | 0.75894 | 0.75885 | 0.75874 |  |  |  |
| 25 | 0.76180 | 0.75637 | 0.75639 | 0.75677 | 0.75711 | 0.75777 | 0.75772 | 0.75741 | 0.75712 | 0.75697 | 0.75696 | 0.75699 | 0.75681 | 0.75688 |  |  |  |
| 30 | 0.76045 | 0.75554 | 0.75563 | 0.75622 | 0.75676 | 0.75722 | 0.75713 | 0.75691 | 0.75653 | 0.75636 | 0.75617 | 0.75617 | 0.75612 | 0.75613 |  |  |  |
| 35 | 0.75921 | 0.75496 | 0.75530 | 0.75574 | 0.75612 | 0.75636 | 0.75660 | 0.75622 | 0.75606 | 0.75595 | 0.75580 | 0.75581 | 0.75568 | 0.75565 |  |  |  |
| 40 | 0.75847 | 0.75445 | 0.75468 | 0.75489 | 0.75515 | 0.75557 | 0.75567 | 0.75541 | 0.75531 | 0.75507 | 0.75495 | 0.75495 | 0.75494 | 0.75492 |  |  |  |
| 45 | 0.75798 | 0.75473 | 0.75476 | 0.75492 | 0.75501 | 0.75526 | 0.75542 | 0.75509 | 0.75500 | 0.75478 | 0.75468 | 0.75474 | 0.75461 | 0.75452 |  |  |  |
| 50 | 0.75774 | 0.75477 | 0.75485 | 0.75498 | 0.75517 | 0.75544 | 0.75548 | 0.75528 | 0.75516 | 0.75496 | 0.75486 | 0.75492 | 0.75484 | 0.75479 |  |  |  |

PCC\_Precision

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | PCC | HybirdSimilarity3  (len=2) | HybirdSimilarity3  (len=3) | HybirdSimilarity3  (len=4) | HybirdSimilarity3  (len=5) | HybirdSimilarity3  (len=6) | HybirdSimilarity3  (len=7) | HybirdSimilarity3  (len=8) | HybirdSimilarity3  (len=9) | HybirdSimilarity3  (len=10) | HybirdSimilarity3  (len=11) | HybirdSimilarity3  (len=12) | HybirdSimilarity3  (len=13) | HybirdSimilarity3  (len=14) | HybirdSimilarity3  (len=16) | HybirdSimilarity3  (len=18) | HybirdSimilarity3  (len=20) |
| 5 | 0.10806 | 0.17756 | 0.17756 | 0.17375 | 0.17081 | 0.16623 | 0.16569 | 0.16765 | 0.16743 | 0.16895 | 0.16917 | 0.17015 | 0.17070 | 0.16950 |  |  |  |
| 10 | 0.11209 | 0.17418 | 0.17353 | 0.17048 | 0.16819 | 0.16275 | 0.16209 | 0.16275 | 0.16383 | 0.16492 | 0.16623 | 0.16732 | 0.16656 | 0.16601 |  |  |  |
| 15 | 0.12484 | 0.17887 | 0.17832 | 0.17320 | 0.16993 | 0.16318 | 0.16264 | 0.16285 | 0.16383 | 0.16405 | 0.16460 | 0.16601 | 0.16525 | 0.16460 |  |  |  |
| 20 | 0.13780 | 0.18322 | 0.18268 | 0.17832 | 0.17407 | 0.16852 | 0.16645 | 0.16558 | 0.16645 | 0.16754 | 0.16786 | 0.16885 | 0.16852 | 0.16819 |  |  |  |
| 25 | 0.14597 | 0.18987 | 0.18911 | 0.18246 | 0.17952 | 0.17211 | 0.17048 | 0.17070 | 0.16993 | 0.17070 | 0.17157 | 0.17266 | 0.17244 | 0.17211 |  |  |  |
| 30 | 0.15174 | 0.19564 | 0.19521 | 0.18911 | 0.18420 | 0.17691 | 0.17549 | 0.17473 | 0.17407 | 0.17527 | 0.17538 | 0.17593 | 0.17582 | 0.17582 |  |  |  |
| 35 | 0.15773 | 0.20370 | 0.20327 | 0.19542 | 0.18911 | 0.18235 | 0.18017 | 0.18017 | 0.18007 | 0.18083 | 0.18126 | 0.18224 | 0.18203 | 0.18126 |  |  |  |
| 40 | 0.16296 | 0.20948 | 0.20839 | 0.20120 | 0.19684 | 0.18769 | 0.18606 | 0.18627 | 0.18649 | 0.18725 | 0.18813 | 0.18900 | 0.18867 | 0.18856 |  |  |  |
| 45 | 0.16754 | 0.21503 | 0.21351 | 0.20621 | 0.20076 | 0.19444 | 0.19150 | 0.19139 | 0.19205 | 0.19314 | 0.19379 | 0.19477 | 0.19444 | 0.19390 |  |  |  |
| 50 | 0.17288 | 0.21895 | 0.21765 | 0.21144 | 0.20708 | 0.19902 | 0.19695 | 0.19695 | 0.19706 | 0.19739 | 0.19793 | 0.19913 | 0.19935 | 0.19902 |  |  |  |

PCC\_Recall

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | PCC | HybirdSimilarity3  (len=2) | HybirdSimilarity3  (len=3) | HybirdSimilarity3  (len=4) | HybirdSimilarity3  (len=5) | HybirdSimilarity3  (len=6) | HybirdSimilarity3  (len=7) | HybirdSimilarity3  (len=8) | HybirdSimilarity3  (len=9) | HybirdSimilarity3  (len=10) | HybirdSimilarity3  (len=11) | HybirdSimilarity3  (len=12) | HybirdSimilarity3  (len=13) | HybirdSimilarity3  (len=14) | HybirdSimilarity3  (len=16) | HybirdSimilarity3  (len=18) | HybirdSimilarity3  (len=20) |
| 5 | 0.06697 | 0.14805 | 0.14710 | 0.14485 | 0.14348 | 0.14027 | 0.13871 | 0.13998 | 0.13985 | 0.14079 | 0.14116 | 0.14227 | 0.14211 | 0.14133 |  |  |  |
| 10 | 0.06461 | 0.14248 | 0.14127 | 0.14118 | 0.13948 | 0.13527 | 0.13422 | 0.13419 | 0.13581 | 0.13500 | 0.13694 | 0.13758 | 0.13755 | 0.13667 |  |  |  |
| 15 | 0.07776 | 0.13972 | 0.13897 | 0.13657 | 0.13554 | 0.12886 | 0.12786 | 0.12783 | 0.12826 | 0.12875 | 0.12887 | 0.12931 | 0.12930 | 0.12859 |  |  |  |
| 20 | 0.09324 | 0.13626 | 0.13609 | 0.13466 | 0.13114 | 0.12660 | 0.12512 | 0.12370 | 0.12335 | 0.12431 | 0.12429 | 0.12481 | 0.12462 | 0.12438 |  |  |  |
| 25 | 0.10058 | 0.13679 | 0.13595 | 0.13209 | 0.13066 | 0.12496 | 0.12349 | 0.12343 | 0.12278 | 0.12271 | 0.12319 | 0.12356 | 0.12357 | 0.12338 |  |  |  |
| 30 | 0.10354 | 0.13850 | 0.13816 | 0.13535 | 0.13241 | 0.12734 | 0.12551 | 0.12398 | 0.12280 | 0.12360 | 0.12363 | 0.12429 | 0.12395 | 0.12450 |  |  |  |
| 35 | 0.10934 | 0.14250 | 0.14225 | 0.13757 | 0.13403 | 0.12895 | 0.12750 | 0.12711 | 0.12664 | 0.12686 | 0.12741 | 0.12848 | 0.12856 | 0.12685 |  |  |  |
| 40 | 0.11286 | 0.14544 | 0.14407 | 0.13979 | 0.13804 | 0.13209 | 0.13007 | 0.12981 | 0.13017 | 0.13058 | 0.13083 | 0.13187 | 0.13137 | 0.13117 |  |  |  |
| 45 | 0.11536 | 0.14699 | 0.14646 | 0.14196 | 0.13885 | 0.13515 | 0.13287 | 0.13317 | 0.13317 | 0.13335 | 0.13378 | 0.13438 | 0.13418 | 0.13402 |  |  |  |
| 50 | 0.11807 | 0.14844 | 0.14717 | 0.14432 | 0.14153 | 0.13730 | 0.13597 | 0.13586 | 0.13586 | 0.13534 | 0.13581 | 0.13632 | 0.13650 | 0.13638 |  |  |  |

## COS

COS\_MAE

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | COS | HybirdSimilarity3  (len=2) | HybirdSimilarity3  (len=3) | HybirdSimilarity3  (len=4) | HybirdSimilarity3  (len=5) | HybirdSimilarity3  (len=6) |
| 5 | 0.822578097787343 | 0.7859098716053611 |  |  |  |  |
| 10 | 0.7914796368834089 | 0.7652607677596734 |  |  |  |  |
| 15 | 0.7809082243053683 | 0.7605953461582428 |  |  |  |  |
| 20 | 0.7753185856163409 | 0.7571792546653316 |  |  |  |  |
| 25 | 0.7727597919718023 | 0.756952796206895 |  |  |  |  |
| 30 | 0.7709456922527254 | 0.7562897406028419 |  |  |  |  |
| 35 | 0.769222281014186 | 0.7565885851569171 |  |  |  |  |
| 40 | 0.7677735385457826 | 0.7563396693934045 |  |  |  |  |
| 45 | 0.7671228195611981 | 0.7572898750459491 |  |  |  |  |
| 50 | 0.7664013663970592 | 0.7578630803748165 |  |  |  |  |

COS\_Precision

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | COS | HybirdSimilarity3  (len=2) | HybirdSimilarity3  (len=3) | HybirdSimilarity3  (len=4) | HybirdSimilarity3  (len=5) | HybirdSimilarity3  (len=6) |
| 5 | 0.09030501089324605 | 0.27483660130718923 |  |  |  |  |
| 10 | 0.08660130718954233 | 0.30424836601307187 |  |  |  |  |
| 15 | 0.09193899782135065 | 0.3100217864923745 |  |  |  |  |
| 20 | 0.09684095860566438 | 0.31100217864923746 |  |  |  |  |
| 25 | 0.10021786492374725 | 0.31078431372549037 |  |  |  |  |
| 30 | 0.1044662309368191 | 0.31002178649237505 |  |  |  |  |
| 35 | 0.10882352941176462 | 0.30969498910675397 |  |  |  |  |
| 40 | 0.11307189542483646 | 0.3099128540305012 |  |  |  |  |
| 45 | 0.11786492374727649 | 0.30947712418300655 |  |  |  |  |
| 50 | 0.12244008714596934 | 0.3094771241830066 |  |  |  |  |

COS\_Recall

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | COS | HybirdSimilarity3  (len=2) | HybirdSimilarity3  (len=3) | HybirdSimilarity3  (len=4) | HybirdSimilarity3  (len=5) | HybirdSimilarity3  (len=6) |
| 5 | 0.04290111846913282 | 0.18758611581293955 |  |  |  |  |
| 10 | 0.04435201727072167 | 0.20013790660634392 |  |  |  |  |
| 15 | 0.04881727126004013 | 0.20000160208693454 |  |  |  |  |
| 20 | 0.05254053868744182 | 0.1990539838372783 |  |  |  |  |
| 25 | 0.05412680477352272 | 0.19784834862512957 |  |  |  |  |
| 30 | 0.05675611433946265 | 0.1963403629000007 |  |  |  |  |
| 35 | 0.05950164055504106 | 0.19500642883000077 |  |  |  |  |
| 40 | 0.06374853668606338 | 0.1934515112987191 |  |  |  |  |
| 45 | 0.06798111442775442 | 0.191621836639551 |  |  |  |  |
| 50 | 0.07078707400315129 | 0.19244651533745832 |  |  |  |  |

|  |  |
| --- | --- |
|  |  |

|  |  |  |
| --- | --- | --- |
|  | UPS | HybirdSimilarity3 |
| 5\_MAE: | 0.7914168244244324 | 跑出来的结果比较差 |
| 10\_MAE: | 0.7680727267154525 |  |
| 15\_MAE: | 0.7605201182316809 |  |
| 20\_MAE: | 0.7572859835176422 |  |
| 25\_MAE: | 0.7560149750520199 |  |
| 30\_MAE: | 0.7556066519728971 |  |
| 35\_MAE: | 0.7552536901754912 |  |
| 40\_MAE: | 0.7552809983384136 |  |
| 45\_MAE: | 0.7556179315255037 |  |
| 50\_MAE: | 0.7562448438690192 |  |

# SML1M

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | |  | PCC | HybirdSimilarity3 | | 5 | 0.9083458780742875 | 0.8860322616364001 | | 10 | 0.8956338868161214 | 0.8738196959879256 | | 15 | 0.8924783085780502 | 0.8711425149407233 | | 20 | 0.8915739874336202 | 0.8707618534129383 | | 25 | 0.8911585604026324 | 0.8710695557326458 | | 30 | 0.891129738114327 | 0.871163745117742 | | 35 | 0.8911651505124332 | 0.8714288841487998 | | 40 | 0.8912257654835664 | 0.8715865398394275 | | 45 | 0.8912588659220019 | 0.8716353482754432 | | 50 | 0.8912731910574779 | 0.8716827236425531 | | |  |  |  | | --- | --- | --- | |  | COS | HybirdSimilarity3 | | 5 | 0.8623801937443291 | 0.8368845396769586 | | 10 | 0.8420901495989048 | 0.8281777158323819 | | 15 | 0.8361174999774003 | 0.8281704540914351 | | 20 | 0.8334485347225545 | 0.8283391326373157 | | 25 | 0.8326393099265716 | 0.828688355023539 | | 30 | 0.8318503283816719 | 0.8290675616029369 | | 35 | 0.8314941515759892 | 0.8295213369477293 | | 40 | 0.8315017416004525 | 0.8302057590597317 | | 45 | 0.8315334883538867 | 0.8305480477443068 | | 50 | 0.8316388994851748 | 0.8309077700192746 | |

## PCC

PCC\_MAE

|  |  |  |
| --- | --- | --- |
|  | PCC | HybirdSimilarity3 |
| 5 | 0.9083458780742875 | 0.8860322616364001 |
| 10 | 0.8956338868161214 | 0.8738196959879256 |
| 15 | 0.8924783085780502 | 0.8711425149407233 |
| 20 | 0.8915739874336202 | 0.8707618534129383 |
| 25 | 0.8911585604026324 | 0.8710695557326458 |
| 30 | 0.891129738114327 | 0.871163745117742 |
| 35 | 0.8911651505124332 | 0.8714288841487998 |
| 40 | 0.8912257654835664 | 0.8715865398394275 |
| 45 | 0.8912588659220019 | 0.8716353482754432 |
| 50 | 0.8912731910574779 | 0.8716827236425531 |

PCC\_Precision

|  |  |  |
| --- | --- | --- |
|  | PCC | HybirdSimilarity3 |
| 5 | 0.003061224489795914 | 0.004363124560168888 |
| 10 | 0.0024982406755805742 | 0.0033779028852920436 |
| 15 | 0.0024630541871921157 | 0.003026038001407457 |
| 20 | 0.002498240675580575 | 0.003131597466572833 |
| 25 | 0.0024982406755805747 | 0.0032019704433497503 |
| 30 | 0.0025334271639690337 | 0.003131597466572833 |
| 35 | 0.0025686136523574927 | 0.003166783954961292 |
| 40 | 0.002709359605911328 | 0.003131597466572833 |
| 45 | 0.0028149190710767048 | 0.003201970443349751 |
| 50 | 0.002744546094299787 | 0.0032371569317382097 |

PCC\_Recall

|  |  |  |
| --- | --- | --- |
|  | PCC | HybirdSimilarity3 |
| 5 | 0.0028914817088274594 | 0.004338312005006906 |
| 10 | 0.0022082031752689657 | 0.0032781445423297182 |
| 15 | 0.002249109514642604 | 0.002701492986895579 |
| 20 | 0.0025233854841925714 | 0.0029414323118510436 |
| 25 | 0.002547428578504144 | 0.0030237835762019767 |
| 30 | 0.002424335775859993 | 0.0028416275121716864 |
| 35 | 0.0025310663469215794 | 0.0028629653808851393 |
| 40 | 0.002616519247293552 | 0.002732483617750315 |
| 45 | 0.0027925148433739855 | 0.002774083406978049 |
| 50 | 0.0027025938174923686 | 0.002911434173055805 |

## COS

COS\_MAE

|  |  |  |
| --- | --- | --- |
|  | COS | HybirdSimilarity3 |
| 5 | 0.8623801937443291 | 0.8368845396769586 |
| 10 | 0.8420901495989048 | 0.8281777158323819 |
| 15 | 0.8361174999774003 | 0.8281704540914351 |
| 20 | 0.8334485347225545 | 0.8283391326373157 |
| 25 | 0.8326393099265716 | 0.828688355023539 |
| 30 | 0.8318503283816719 | 0.8290675616029369 |
| 35 | 0.8314941515759892 | 0.8295213369477293 |
| 40 | 0.8315017416004525 | 0.8302057590597317 |
| 45 | 0.8315334883538867 | 0.8305480477443068 |
| 50 | 0.8316388994851748 | 0.8309077700192746 |

COS\_Precision

|  |  |  |
| --- | --- | --- |
|  | COS | HybirdSimilarity3 |
| 5 | 0.0036945812807881724 | 0.0051020408163265215 |
| 10 | 0.002779732582688245 | 0.0049612948627726855 |
| 15 | 0.0032723434201266674 | 0.004609429978888099 |
| 20 | 0.002955665024630539 | 0.004152005629838136 |
| 25 | 0.002990851513018998 | 0.004081632653061219 |
| 30 | 0.0029556650246305395 | 0.003940886699507384 |
| 35 | 0.00295566502463054 | 0.0038705137227304665 |
| 40 | 0.0031315974665728345 | 0.003659394792399714 |
| 45 | 0.003201970443349751 | 0.0036593947923997143 |
| 50 | 0.0033779028852920444 | 0.0036593947923997143 |

COS\_Recall

|  |  |  |
| --- | --- | --- |
|  | COS | HybirdSimilarity3 |
| 5 | 0.003057947248760149 | 0.005780936703021927 |
| 10 | 0.002627878099159103 | 0.005668951063404942 |
| 15 | 0.0037818888980245915 | 0.004704558100653739 |
| 20 | 0.0025232552743450405 | 0.003935271898556389 |
| 25 | 0.0024745300155245985 | 0.0037517332802684393 |
| 30 | 0.0026673844948573917 | 0.0035236498902352625 |
| 35 | 0.0022949471595410044 | 0.0034443754259281104 |
| 40 | 0.0025145103581191296 | 0.0031626543908045914 |
| 45 | 0.0025622362462735947 | 0.003121863037140539 |
| 50 | 0.002676610762652245 | 0.0030358076202214265 |