|  |  |
| --- | --- |
| 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 |
| 5\_MAE: | 0.81401287162 | 0.80307572036 |
| 10\_MAE: | 0.781908998651648 | 0.7737999620396511 |
| 15\_MAE: | 0.7705317431888051 | 0.7636746968831715 |
| 20\_MAE: | 0.764635042840259 | 0.7587380873975175 |
| 25\_MAE: | 0.7618043097316483 | 0.7563674767991945 |
| 30\_MAE: | 0.7604566653248619 | 0.7555392745159109 |
| 35\_MAE: | 0.7592130913004181 | 0.7549631798597356 |
| 40\_MAE: | 0.758479551933542 | 0.7544509114032824 |
| 45\_MAE: | 0.7579847101695433 | 0.7547301263858996 |
| 50\_MAE: | 0.7577494561177173 | 0.7547744931454026 |

PCC\_Precision

|  |  |  |
| --- | --- | --- |
|  | PCC | HybirdSimilarity3 |
| 5 | 0.10806100217864908 | 0.1775599128540302 |
| 10 | 0.11209150326797371 | 0.1741830065359476 |
| 15 | 0.12483660130718933 | 0.17886710239651413 |
| 20 | 0.13779956427015225 | 0.18322440087145947 |
| 25 | 0.1459694989106752 | 0.18986928104575135 |
| 30 | 0.15174291938997794 | 0.19564270152505414 |
| 35 | 0.1577342047930281 | 0.20370370370370341 |
| 40 | 0.16296296296296273 | 0.20947712418300618 |
| 45 | 0.16753812636165555 | 0.2150326797385618 |
| 50 | 0.17287581699346383 | 0.2189542483660128 |

PCC\_Recall

|  |  |  |
| --- | --- | --- |
|  | PCC | HybirdSimilarity3 |
| 5 | 0.06696765208842986 | 0.14805110943666805 |
| 10 | 0.06460879234859078 | 0.14247774593881346 |
| 15 | 0.07776116693923926 | 0.13972027645566135 |
| 20 | 0.09323600592599557 | 0.13626312915677236 |
| 25 | 0.10058402439878614 | 0.13678747158454765 |
| 30 | 0.10353566284560899 | 0.13849973269110286 |
| 35 | 0.10933768468093143 | 0.14249883009370742 |
| 40 | 0.11286318292199839 | 0.14543510902570572 |
| 45 | 0.11535804042640736 | 0.14699078705996207 |
| 50 | 0.11806725333068481 | 0.14843716503280077 |

## COS

COS\_MAE

|  |  |  |
| --- | --- | --- |
|  | COS | HybirdSimilarity3 |
| 5\_MAE: | 0.822578097787343 | 0.7859098716053611 |
| 10\_MAE: | 0.7914796368834089 | 0.7652607677596734 |
| 15\_MAE: | 0.7809082243053683 | 0.7605953461582428 |
| 20\_MAE: | 0.7753185856163409 | 0.7571792546653316 |
| 25\_MAE: | 0.7727597919718023 | 0.756952796206895 |
| 30\_MAE: | 0.7709456922527254 | 0.7562897406028419 |
| 35\_MAE: | 0.769222281014186 | 0.7565885851569171 |
| 40\_MAE: | 0.7677735385457826 | 0.7563396693934045 |
| 45\_MAE: | 0.7671228195611981 | 0.7572898750459491 |
| 50\_MAE: | 0.7664013663970592 | 0.7578630803748165 |

COS\_Precision

|  |  |  |
| --- | --- | --- |
|  | COS | HybirdSimilarity3 |
| 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 |
| 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 |