# PRÍLOHA E – Detailné vyhodnotenie úspešnosti odporúčania pre dataset ALEF

V tejto časti uvádzame detailné opisy vyhodnotenia úspešnosti odporúčania pre dataset ALEF.

## Vyhodnotenie úspešnosti odporúčania pre skupinu parametrov hľadania frekventovaných množín

V tejto časti sa budeme venovať vyhodnoteniu úspešnosti odporúčania s rôznymi konfiguráciami skupiny parametrov pre algoritmus hľadania uzavretých frekventovaných množín v prúde dát *IncMine*. Ide o parametre (bližšie sú opísané v časti 7.1.4) a ich hodnoty:

* **Minimálna podpora** *(skr. ms)*:{0.005, 0.01, 0.02, 0.03, 0.04, 0.05, 0.1}
* **Miera uvoľnenia** *(skr. rr)*:{0.1, 0.5, 0.9}
* **Dĺžka segmentu** *(skr. sl)*:{25, 50, 100, 150, 200, 500}
* **Maximálna dĺžka množiny** *(skr. mil)*:{10}
* **Veľkosť okna** *(skr. ws)*: {5,10,15}

Spolu je to 270 rôznych konfigurácií. Čo sa týka hodnôt minimalnej podpory tá vyjadruje minimálnu podporu v rámci jedného segmentu. Teda ak dĺžka segmentu je napr. 50 a hodnota minimálnej podpory je 0.02 a menšia tak je každá množina považovaná za frekventovanú. Kombinácie parametrov kde bude teda hodnota budeme ignorovať.

Na grafe 1 možno vidieť hodnoty metriky *presnosť* pre rôzne počty odporúčaných položiek. Vybrali sme tie konfigurácie, ktoré boli medzi dvoma najlepšími výsledkami v odporúčaní pre jednotlivé počty odporúčaných položiek. V tabuľke 4 sú uvedené hodnoty skúmaných parametrov v týchto konfiguráciach aj s dočasným označením konfigurácií. Hodnoty metrík pre menšie počty odporúčaných položiek sú podobné, ale zvlášt konfigurácia *1* dosahuje najlepší výsledok pri počte odporúčaných položiek *1.* Pri väčšom počte odporúčaných položiek dosahuje najlepšie výsledky konfigurácia *2*. Keď sa ďalej pozrieme na výsledky metriky *NDCG* pre tieto konfigurácie uvedené na grafe 2 opäť dosahuje najlepšie výsledky konfigurácia 2.

Tabuľka 1 Hodnoty skúmaných parametrov 5 najlepších konfigurácii

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Konfigurácia (dočasné ozn.)** | **ms** | **rr** | **sl** | **mil** | **ws** |
| 1 | 0.05 | 0.1 | 25 | 10 | 15 |
| 2 | 0.04 | 0.1 | 50 | 10 | 15 |
| 3 | 0.04 | 0.1 | 50 | 10 | 10 |
| 4 | 0.03 | 0.5 | 50 | 10 | 15 |
| 5 | 0.05 | 0.1 | 50 | 10 | 15 |

Graf 1 Hodnoty metriky presnosť pre 5 najlepších konfigurácií zo skupiny parametrov hľadania frekventovaných množín.

Graf 2 Hodnoty metriky NDCG pre 5 najlepších konfigurácií zo skupiny parametrov hľadania frekventovaných množín.

Ďalej sme sledovali vplyv hodnôt jednotlivých parametrov a dvojíc parametrov na výsledky odporúčania. Zistili sme, že konfigurácie s kratšou dĺžkou segmentov dosahovali priemerne lepšie výsledky (*presnosť)* ako dlhšie segmenty. Najlepšie výsledky dávali konfigurácie s dĺžkou segmentu 50. Tiež, že konfigurácie s príliš nízkou hodnotou podpory a príliš vysokou hodnotou podpory dosahovali priemerne horšie výsledky. Príliš nízka hodnota *ms* spôsobuje generovanie veľkého množstva vzorov čo výrazne spomaľuje výpočet až na spodnú hranicu obmedzenia rýchlosti (15 transakcií za sekundu), čo má tiež za následok zhoršovanie výsledkov (pozri spôsob obmedzenia rýchlosti v časti 5.1.7). Príliš vysoká hodnota *ms* zas spôsobí, že veľa vzorov, ktoré by prispeli k dobrému výsledku sa nenájde. Takisto sme pozorovali, že priemerne horšie výsledky dávali tiež konfigurácie s vyššou hodnotou *rr* a vyššou hodnotou *ws*. Tu si možno všimnúť, že hoci sme sledovali takýto trend, tak medzi najlepšími konfiguráciami sú práve konfigurácie s vysokou hodnotou *ws*. Jednotlivé parametre sa navzájom ovplyvňujú. Napríklad zvýšením hodnoty dĺžky segmentu *sl* sa mení účinok minimálnej podpory *ms,* ktorá sa počíta práve vzhľadom na dĺžku segmentu. Sledovali sme teda aj vplyv rôznych dvojíc parametrov na výsledky.

V Tabuľka 2 môžeme sledovať priemerné výsledky metriky *presnosť* pre rôzne kombinácie hodnôt parametrov *ms* a *sl*. Možno pozorovať, že najlepšie výsledky sú pri nízkych hodnotách minimálnej podpory a zároveň nízkych hodnotách dĺžky segmentov.

Tabuľka 2 Priemerné výsledky metriky presnosť pre rôzne kombinácie hodnôt parametrov ms a sl.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ***ms* vs. *sl*** | **0.005** | **0.01** | **0.02** | **0.03** | **0.04** | **0.05** | **0.1** |
| **25** |  |  |  |  |  | **44.43%** | **41.44%** |
| **50** |  |  |  | **46.60%** | **46.38%** | **44.96%** | **41.07%** |
| **100** |  |  | **45.28%** | **44.00%** | **42.11%** | **40.66%** | **34.96%** |
| **150** |  | **43.71%** | **42.89%** | **40.18%** | **38.84%** | **37.93%** | **30.63%** |
| **200** |  | **44.01%** | **40.87%** | **40.76%** | **38.31%** | **35.75%** | **28.86%** |
| **500** | **37.69%** | **38.81%** | **36.92%** | **35.35%** | **32.74%** | **30.59%** | **28.40%** |

V ďalšej Tabuľka 3 si môžeme zas všimnúť, že najlepšie výsledky sú dosahované pre kombináciu krátkych segmentov a väčšej dĺžky okna. To znamená, že často dochádza k aktualizácií vzorov nad menšími časťami dát a počet segmentov v rámci okna je vyšší, čo znamená, že podpora vzorov sa počíta k dlhšej histórií dát.

Tabuľka 3 Priemerné výsledky metriky *presnosť* pre rôzne kombinácie hodnôt parametrov *sl* a *ws*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ***sl* vs *ws*** | **25** | **50** | **100** | **150** | **200** | **500** |
| **5** | **0.4018** | **0.4355** | **0.4222** | **0.4092** | **0.3978** | **0.3432** |
| **10** | **0.4354** | **0.454** | **0.4202** | **0.3882** | **0.3821** | **0.3533** |
| **15** | **0.451** | **0.453** | **0.3996** | **0.3736** | **0.363** | **0.3343** |

Na Graf 3 vidíme priemernú rýchlosť spracovania transakcií jednotlivými konfiguráciami. Všetky konfigurácie okrem konfigurácie 1 spadli až na spodnú hranicu minimálnej požadovanej rýchlosti. Konfigurácia 1 má oproti konfigurácií 2 o polovicu kratšiu dĺžku segmentu a o *0.01* vyššiu hranicu minimálnej podpory. Je zaujímavé, že takáto relatívne malá zmena spôsobí významnú zmenu čo sa týka rýchlosti spracovania aj presnosti odporúčania. Ak teda je dôležitá rýchlosť je dobrou voľbou konfigurácia 1 ak nie je rýchlosť až tak kritická a dôležitá je presnosť tak je lepšou voľbou konfigurácia 2. Do konečného vyhodnocovania teda vyberáme hodnoty z konfigurácií 1 a 2.

Graf 3 Priemerná rýchlosť spracovania transakcií.

## Vyhodnotenie úspešnosti odporúčania pre skupinu parametrov zhlukovania

V tejto časti sa budeme venovať vyhodnoteniu úspešnosti odporúčania s rôznymi konfiguráciami skupiny parametrov pre algoritmus zhlukovania Clustream. Ide o parametre (bližšie sú opísané v časti 5.1.6) a ich hodnoty:

* **Počet zhlukov** *(skr. gc)*: {2, 4, 6, 8}
* **Hraničný počet zmien v modeli používateľa** *(skr. tuc)*: {5,10,15}
* **Hraničný počet zmien v mikrozhlukoch***(skr. tcm)*: {50, 100, 200, 400, 800}
* **Maximálny počet mikrozhlukov***(skr. mmc****)*:** {100,1000}

Spolu je to 120 rôznych konfigurácií. Nastavenie ostatných parametrov sme vybrali z jednej z najlepších konfigurácií z predchádzajúcej časti (konfigurácia 1 v tabuľke 8), ktorá nebola najlepšia čo sa týka celkovej presnosti, ale dosahovala lepšie výsledky z hľadiska rýchlosti spracovania (zvolili sme teda istý kompromis medzi rýchlosťou a presnosťou).

V Tabuľka 4 sú uvedené hodnoty skúmaných parametrov v týchto konfiguráciach aj s dočasným označením konfigurácií. Všetky najlepšie konfigurácie majú hodnotu parametra *tcu=5* a hodnotu parametra *tcm=100.* Čo sa týka počtu skupín sú hodnoty rôzne ale neobsahujú hodnotu 2. Podobne pre parameter *tcm* neobsahujú hodnoty nižšie ako 200.

Na Graf 4 a Graf 5 možno vidieť hodnoty metriky *presnosť,* resp. *ndcg*  pre rôzne počty odporúčaných položiek. Nie je vidno nejaký jasný vzor medzi vybranými konfiguráciami. Rozdiely sú pomerne malé v metrike *presnosť* aj *ndcg.* Na prvých dvoch miestach v oboch metrikách sa striedajú konfigurácie 1 a 2.

Tabuľka 4 Najlepšie konfigurácie vybrané zo skupiny parametrov zhlukovania.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Konfigurácia (dočasné ozn.)** | **gc** | **tuc** | **tcm** | **mmc** |
| 1 | 6 | 5 | 400 | 100 |
| 2 | 8 | 5 | 400 | 100 |
| 3 | 8 | 5 | 800 | 100 |
| 4 | 4 | 5 | 400 | 100 |
| 5 | 4 | 5 | 200 | 100 |

Graf 4 Metrika presnosť pre 5 najlepších konfigurácií zo skupiny parametrov zhlukovania.

Graf 5 Metrika NDCG pre 5 najlepších konfigurácií zo skupiny parametrov zhlukovania.

Ďalej sme sledovali vplyv hodnôt jednotlivých parametrov a dvojíc parametrov na výsledky odporúčania. Zistili sme, že konfigurácie s počtom skupín 2 sú mierne horšie ako konfigurácie s vyšším počtom skupín, kde už väčšie rozdiely nebadáme. Je to spôsobené zrejme tým, že počet skupín 2 je príliš malý aby sa priblížil skutočnému rozdeleniu používateľov do skupín a pri väčšom počte skupín sú identifikované aj niektoré ďalšie dôležité skupiny.Takisto sme zistili, že najlepšie výsledky sú v konfiguráciach s najmenšou hodnotou parametra *tuc.* To znamená, že častejšie dochádza k aktualizácií mikrozhlukov a tým pádom aj k častejšiemu makrozhlukovaniu a tiež, že používateľ je posudzovaný (teda zaradzovaný do skupín) podľa jeho čo najaktuálnejšieho správania (posledných 5 sedení).

Čo sa týka parametra *tcm* tak príliš nízke hodnoty (50 a 100) a príliš vysoká hodnota (800) boli horšie ako hodnoty 200 a 400 (pozri tabuľku Tabuľka 6). Pre parameter *mmc* je situácia celkom jasná, zistili sme, že vyšší počet mikrozhlukov neviedol k lepším výsledkom. Jednotlivé parametre sa navzájom ovplyvňujú. V tabuľke Tabuľka 5 vidíme, že najlepšia kombinácia hodnôt parametrov *gc* (počet skupín) a *tcm* (minimálny počet aktualizácií mikrozhlukov) je pri počte skupín 4 a *tcm = 200.*

Tabuľka 5 Priemerné výsledky metriky presnosť pre rôzne kombinácie hodnôt parametrov gc a tcm.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **gc vs. tcm** | **2** | **4** | **6** | **8** |
| **50** | **46.57%** | **46.66%** | **46.62%** | **46.62%** |
| **100** | **46.65%** | **46.86%** | **46.84%** | **46.82%** |
| **200** | **46.77%** | **47.02%** | **46.91%** | **46.91%** |
| **400** | **46.66%** | **46.77%** | **46.85%** | **46.84%** |
| **800** | **46.26%** | **46.42%** | **46.40%** | **46.41%** |

Tabuľka 6 Priemerné výsledky metriky presnosť pre rôzne kombinácie hodnôt parametrov tcm a tuc.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***tcm vs. tuc*** | **50** | **100** | **200** | **400** | **800** |
| **5** | **47.01%** | **47.12%** | **47.40%** | **47.41%** | **47.07%** |
| **10** | **46.57%** | **46.80%** | **46.81%** | **46.78%** | **46.29%** |
| **15** | **46.26%** | **46.46%** | **46.50%** | **46.16%** | **45.76%** |

Na Graf 6 vidíme priemernú rýchlosť spracovania transakcií jednotlivými konfiguráciami. Čo sa týka rozdielov v rýchlostiach, tie už nie sú také značné ako pri rôznych konfiguráciach zo skupiny parametrov pre hľadanie frekventovaných množín, ale najlepšiu rýchlosť (58.41 transakcií za sekundu) dosiahla konfigurácia 3. Do záverečného vyhodnocovania sme teda vybrali hodnoty parametrov práve z konfigurácií 1 a 2 a kvôli vyššej rýchlosti aj konfigurácie 3.

Graf 6 Priemerná rýchlosť spracovania transakcií pre najlepšie konfigurácie zo skupiny parametrov zhlukovania.

## Vyhodnotenie úspešnosti odporúčania pre skupinu parametrov odporúčania

V tejto časti sa budeme venovať vyhodnoteniu odporúčania v spojitosti s rôznymi konfiguráciami skupiny parametrov pre odporúčanie. Táto skupina je zvláštna, keďže pozorujeme len jeden parameter hoci tu patrí aj parameter *rc* teda počet odporúčaných položiek, ktorého rôzne hodnoty ale sledujeme implicitne pri každom behu aj v ostatných skupinách parametrov. Ide teda o jediný parameter:

* **Veľkosť okna vyhodnocovania** *(skr. ews)*: {1,2,3,4,5,6,7, 8, 9, 10}

Nastavenie hodnôt ostatných parametrov sme zobrali z najlepšej konfigurácie čo sa týka presnosti vybranej z predchádzajúcej skupiny parametrov zhlukovania (konfigurácia 1 v tabuľke Tabuľka 6).

Zmena hodnoty parametra *ews* mení aj množinu sedení, ktoré sú vyhodnocované. Čím dlhšie je vyhodnocované okno tým menej sedení má dostatočnú dĺžku aby mohli byť vyhodnotené. To je závažný fakt, ktorý má vplyv na relevantnosť porovnávania jednotlivých konfigurácií.

Na Graf 7 a Graf 8 a možno vidieť hodnoty metriky *presnosť,* resp. *ndcg*  pre rôzne počty odporúčaných položiek a rôzne veľkosti vyhodnocovacieho okna. Môžeme si zvlášť všimnúť, že hodnoty parametra *ews < 2 a ews > 7* majú horšie výsledky najmä pri malom počte odporúčaných položiek. Malé okno dáva priestor možnosti výberu veľkého množstva vzorov správania (aj kratších aj dlhších), ktoré majú prienik s týmto oknom a môžu byť zapojené do odporúčania. To je však zrejme aj čiastočne obmedzením, pretože okno je príliš malé aby rozpoznalo charakter sedenia a zámer používateľa. Pri väčších oknách sa *presnosť* opäť znižuje, čo je zrejme spôsobené aj tým, že neexistuje dostatok dlhších vzorov, ktoré by vedeli dané správanie používateľa správne charakterizovať.

Najlepšie výsledky *presnosti* sú pri hodnotách *ews > 1 a zároveň ews < 7* pri počte odporúčaných položiek < 10*.* Pri väčšom počte odporúčaných položiek sa javia najlepšie veľkosti *ews > 2* a zároveň *ews < 8*.

Graf 7 Výsledky metriky *presnosť* pre rôzne veľkosti vyhodnocovacieho okna *ews* pre rôzne počty odporúčaných položiek.

Graf 8 Výsledky metriky NDCG pre rôzne veľkosti vyhodnocovacieho okna ews pre rôzne počty odporúčaných položiek.

Tento parameter je teda špecifický a do výsledného hľadania najlepšej konfigurácie (časť 7.2.5) sme vybrali hodnoty *ews* z množiny {2, 3, 4, 5}, pri ktorých nebola množina vyhodnocovaných sedení priveľmi zmenšená a zároveň boli dosiahnuté najlepšie výsledky presnosti.

## Vyhodnotenie úspešnosti odporúčania pre ostatné parametre metódy

V tejto časti sa budeme venovať vyhodnoteniu odporúčania v spojitosti s rôznymi konfiguráciami skupiny osatných parametrov metódy. Ide o parametre (bližšie sú opísané v časti 5.1.6) a ich hodnoty:

* **Hraničná hodnota rozdielu identifikátorov zhlukovania** *(skr. tcdiff)*: {1,2,3,4, 5,6,7,8}
* **Minimálna rýchlosť** *(skr. mts)*: {15}

V tejto skupine sme sa rozhodli testovať rôzne hodnoty len prvého parametra *tcdiff*. Rôzne hodnoty parametra pre obmedzenie minimálnej rýchlosti budeme testovať až vo vyhodnotení rýchlosti spracovania (7.3.2), kde sledujeme k akému zhoršeniu metriky *presnosť* bude dochádzať pri zrýchľovaní spracovania transakcií na už vybranej celkovo najlepšej konfigurácií. Nastavenie hodnôt ostatných parametrov sme zobrali z  konfigurácie vybranej z predchádzajúcej skupiny parametrov odporúčania, kde sme pozorovali len jeden parameter veľkosti vyhodnocovaného okna. Vybrali sme nastavenie vyhodnocovacieho okna s veľkosťou 2, ktorá príliš nezmenšuje množinu vyhodnocovaných sedení a zároveň dosahuje jedny z najlepších výsledkov metriky *presnosť*. Celkovo k makrozhlukovaniu dochádza 9 krát preto sme testovali rôzne hodnoty *tcdiff* po túto hodnotu.

Sledovali sme, že skúmané hodnoty tohto parametra nijak zvlášť neovplyvňovali celkové výsledky (pozri Graf 9). To je dobré znamenie z toho pohľadu, že odstraňovanie neaktívnych používateľov z pamäte v skúmanej doméne už po krátkej dobe nečinnosti zdá sa nejak zásadne neovplyvňuje celkové výsledky.

Do záverečného vyhodnotenia sme vybrali hodnotu parametra *tcdiff* = 5, pri ktorej boli výsledky o niečo málo lepšie než pri ostatných.

Graf 9 Výsledky metriky presnosť pre rôzne hodnoty parametra tcdiff a rôzne počty odporúčaných položiek.

## Celkové vyhodnotenie úspešnosti odporúčania pre najlepšie konfigurácie

Na základe vybraných najlepších konfigurácií zo všetkých skupín parametrov ako sme ich uvádzali v predchádzajúcich častiach sme skonštruovali priestor prehľadávania v ktorom budeme hľadať celkovo najlepšiu konfiguráciu. Ide o parametre a ich hodnoty:

* **Minimálna podpora** *(skr. ms)*:{0.04, 0.05}
* **Miera uvoľnenia** *(skr. rr)*:{0.1, 0.5}
* **Dĺžka segmentu** *(skr. sl)*:{25, 50}
* **Maximálna dĺžka množiny** *(skr. mil)*:{10}
* **Veľkosť okna** *(ws)*: {10,15}
* **Počet zhlukov** *(skr. gc)*: {6, 8}
* **Hraničný počet zmien v modeli používateľa** *(skr. tuc)*: {5}
* **Hraničný počet zmien v mikrozhlukoch***(skr. tcm)*: {400, 800}
* **Maximálny počet mikrozhlukov***(skr. mmc****)*:** {100}
* **Veľkosť okna vyhodnocovania** *(skr. ews)*: {2,3,4,5}
* **Hraničná hodnota rozdielu identifikátorov zhlukovania** *(skr. tcdiff)*: {5}
* **Minimálna rýchlosť** *(skr. mts)*: {15}

V tabuľke 7 uvádzame 7 konfigurácií, ktoré boli najlepšie pre niektorú z rôznych kombinácií hodnôt parametrov *ews* (veľkosť vyhodnocovacieho okna)a *rc* (počet odporúčaných položiek)*.* Jediná z uvedených hodnôt ktorá sa medzi najlepšími konfiguráciami nenachádza je hodnota miery uvoľnenia *rr = 0.5*.

Tabuľka 7 Tabuľka zobrazuje hodnoty parametrov pre 7 najlepších konfigurácií celkovo pre dataset ALEF.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **ID konfigurácie** | **ms** | **rr** | **sl** | **ws** | **gc** | **tuc** | **tcm** | **mmc** | **tcdiff** | **mts** |
| 1 | 0.05 | 0.1 | 25 | 15 | 6 | 5 | 400 | 100 | 5 | 15 |
| 2 | 0.04 | 0.1 | 50 | 15 | 8 | 5 | 800 | 100 | 5 | 15 |
| 3 | 0.05 | 0.1 | 50 | 15 | 8 | 5 | 400 | 100 | 5 | 15 |
| 4 | 0.04 | 0.1 | 50 | 10 | 6 | 5 | 400 | 100 | 5 | 15 |
| 5 | 0.05 | 0.1 | 50 | 15 | 6 | 5 | 800 | 100 | 5 | 15 |
| 6 | 0.04 | 0.1 | 50 | 10 | 8 | 5 | 800 | 100 | 5 | 15 |
| 7 | 0.05 | 0.1 | 50 | 15 | 6 | 5 | 400 | 100 | 5 | 15 |

Výsledky sme sledovali samostatne pre rôzne hodnoty parametra *ews,* ktorý, ako sme už spomínali, mení veľkosť priestoru vyhodnocovania. dNa grafoch Graf 10 a Graf 11 uvádzame výsledky metriky *presnosť* a *ndcg* pre rôzne hodnoty parametra *ews*. Na grafe Graf 12 vidíme porovnanie priemerných rýchlostí jednotlivých konfigurácií. Môžeme si všimnúť konfiguráciu 1, ktorá vyniká naprieč rôznymi veľkosťami vyhodnocovacieho okna pri menšom počte odporúčaných položiek (<= 3). Pri väčšom počte odporúčaných položiek sa najlepšie konfigurácie striedajú. Konfigurácia 1 je výnimočná nielen kvôli svojím výsledkom čo sa týka metrík *presnosť* a *ndcg* ale aj čo sa týka rýchlosti. Všetky ostatné konfigurácie padli až na spodnú hranicu minimálnej požadovanej rýchlosti okrem konfigurácie 1. Z tohto dôvodu sme sa rozhodli konfiguráciu 1 zobrať ako celkovo najlepšiu konfiguráciu hodnôt parametrov pre dataset ALEF a pokračujeme s ňou v ďalších experimentoch.

Na záver sme ešte vyhodnotili priemerný rozdiel medzi výsledkami metódy využívajúcej len globálne vzory a metódy využívajúcej kombináciu skupinových a globálnych vzorov. Ten sa pohybuje okolo 1.7% pre rôzne počty odporúčaných položiek a výnimkou nie je ani vybraná konfigurácia 1 kde sa tento rozdiel pohybuje okolo 1.6% (pozri tabuľku tabuľku 8 a tabuľku tabuľku 9). Tieto výsledky potvrdené jednoduchým štatistickým t-testom súhlasia s hypotézou, ktorú sme uviedli na začiatku kapitoly. Vykonali sme jednoduchý t-test medzi populáciami hodnôt výsledkov metriky *presnosť* všetkých vyhodnocovaných konfigurácií metódy využívajúcej kombináciu skupinových a globálnych vzorov a metódy využívajúcej len globálne vzory. Dosiahnutý výsledok je štandardne vyhodnotený ako vysoko štatisticky signifikantný (p < 0.0001, t=9.8366, df=7740 a rozdiel stredných hodnôt je 0.0165).

Tabuľka 8 Tabuľka zachytáva rozdiel medzi priemernými hodnotami metriky presnosť pre dve metódy odporúčania získané zo všetkých vyhodnocovaných konfigurácií a pre rôzne počty odporúčaných položiek v stĺpcoch.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | [**P@1**](mailto:P@1) | [**P@2**](mailto:P@2) | [**P@3**](mailto:P@3) | [**P@4**](mailto:P@4) | [**P@5**](mailto:P@5) | [**P@10**](mailto:P@10) | [**P@15**](mailto:P@15) |
| **Kombinácia skupinových a globálnych vzorov** | **44.15%** | **44.66%** | **44.71%** | **44.63%** | **44.53%** | **43.79%** | **42.59%** |
| **Iba globálne vzory** | **42.47%** | **43.03%** | **43.09%** | **43.02%** | **42.91%** | **42.13%** | **40.87%** |
| **Rozdiel** | **1.68%** | **1.63%** | **1.62%** | **1.61%** | **1.61%** | **1.66%** | **1.72%** |

Tabuľka 9 Tabuľka zachytáva rozdiel medzi priemernými hodnotami metriky presnosť pre dve metódy odporúčania získané z najlepšej konfigurácie a pre rôzne počty odporúčaných položiek v stĺpcoch. Hodnoty v bunkách sú spriemerované pre rôzne hodnoty parametra ews z {2,3,4};

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | [**P@1**](mailto:P@1) | [**P@2**](mailto:P@2) | [**P@3**](mailto:P@3) | [**P@4**](mailto:P@4) | [**P@5**](mailto:P@5) | [**P@10**](mailto:P@10) | [**P@15**](mailto:P@15) |
| **Kombinácia skupinových a globálnych vzorov** | **50.73%** | **50.98%** | **51.09%** | **51.12%** | **51.34%** | **51.56%** | **50.93%** |
| **Iba globálne vzory** | **49.07%** | **49.51%** | **49.46%** | **49.52%** | **49.73%** | **49.73%** | **49.00%** |
| **Rozdiel** | **1.66%** | **1.47%** | **1.64%** | **1.59%** | **1.61%** | **1.83%** | **1.92%** |

Graf 10 Grafy znázorňujúce výsledky najlepších 7 konfigurácií v metrike presnosť a pri rôznych hodnotách parametra ews.

Graf 11 Grafy znázorňujúce výsledky najlepších 7 konfigurácií v metrike NDCG a pri rôznych hodnotách parametra ews.

Graf 12 Graf znázorňuje priemerné rýchlosti jednotlivých konfigurácií v jednotkách počet transakcií za sekundu.