



Posudek oponenta závěrečné práce

Oponent práce: doc. Ing. Pavel Kordík, Ph.D.
Student: Bc. Kryštof Zindulka
Název práce: Doporučovací algoritmy pro hierarchická data
Obor / specializace: Znalostní inženýrství
Vytvořeno dne: 5. června 2023

Hodnotící kritéria

1. Splnění zadání

- ▶ [1] zadání splněno
- [2] zadání splněno s menšími výhradami
- [3] zadání splněno s většími výhradami
- [4] zadání nesplněno

Práce se věnuje poměrně obtížnému tématu. Cílem je navrhnout doporučení item segmentů, což vyžaduje výzkum a vývoj nových přístupů k doporučení. Navíc se pracuje s reálnými, poměrně velkými daty.

2. Písemná část práce

98/100 (A)

Práce je nadstandardní, obsahuje všechny potřebné komponenty a navíc vidím i zdařilou snahu o formalizaci navržených algoritmů. Některé části experimentů by bylo lepší přestěhovat do přílohy, ale i tak je práce dobře čitelná. U některých vztahů (například výraz 2.5) by bylo dobré diskutovat i zobecnitelnost na jiné datové sady. Například pokud by seriál obsahoval mnoho dílů a většina byla cold-start, pak by zřejmě podělení počtem dílů bylo příliš a hodnoty spíše než v intervalu (0,1) se budou pohybovat někde kolem 0. To zřejmě řeší přidání váhových koeficientů hned v dalším výrazu, přičemž by bylo dobré mít nějakou heuristiku pro nastavení těchto koeficientů na základě vlastností datové sady. Líbí se mi rozsah a kvalita experimentů a to jak v oblasti offline evaluace recall-coverage, tak i těch dalších.

3. Nepísemná část, přílohy

99/100 (A)

Přiložený kód je poměrně rozsáhlý, je zřejmé, že bylo třeba pracovat s reálnými daty. V python skriptech není moc komentářů, na druhou stranu kód je poměrně dobře čitelný.

4. Hodnocení výsledků, jejich využitelnost

99 /100 (A)

Výsledky práce jsou dobře využitelné v oblasti doporučovacíh systémů například k doporučení autorů, žánrů, playlistů a dalších segmentů. Bylo by zajímavé vidět výsledky online AB testů, což je samozřejmě nad rámec této práce. Offline evaluace, i když je v této práci udělána velmi poctivě, je přeci jen špatně využitelné zejména v případech, kdy mícháme doporučení obyčejných item a item segmentů do jednoho scénáře.

Celkové hodnocení

98 /100 (A)

Excelentní práce. Obsahuje i prvky novosti a vidím publikační potenciál, bohužel k tomu by bylo třeba dodělat experimenty nad veřejně přístupnými daty.

Otázky k obhajobě

V části 4.4.1 vizualizujete podobnosti segmentů na základě interakčních dat. Zamýšlel jste se nad nějakým konceptem automatické tvorby a pojmenovávání item segmentů? Co user segmenty, jsou vaše metody zobecnitelné i na ně?

Instrukce

Splnění zadání

Posudte, zda předložená ZP dostatečně a v souladu se zadáním obsahově vymezuje cíle, správně je formuluje a v dostatečné kvalitě naplňuje. V komentáři uveďte body zadání, které nebyly splněny, posudte závažnost, dopady a případně i příčiny jednotlivých nedostatků. Pokud zadání svou náročností vybočuje ze standardů pro daný typ práce nebo student případně vypracoval ZP nad rámec zadání, popište, jak se to projevilo na požadované kvalitě splnění zadání a jakým způsobem toto ovlivnilo výsledné hodnocení.

Písemná část práce

Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále posudte, zda předložená ZP je po věcné stránce v pořádku, případně vyskytují-li se v práci věcné chyby nebo nepřesnosti.

Zhodnoťte dále logickou strukturu ZP, návaznosti jednotlivých kapitol a pochopitelnost textu pro čtenáře. Posudte správnost používání formálních zápisů obsažených v práci. Posudte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 52/2021, článek 3.

Posudte, zda student využil a správně citoval relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami. Zhodnoťte, zda převzatý software a jiná autorská díla, byly v ZP použity v souladu s licenčními podmínkami.

Nepísemná část, přílohy

Dle charakteru práce se případně vyjádřete k nepísemné části ZP. Například: SW dílo – kvalita vytvořeného programu a vhodnost a přiměřenost technologií, které byly využité od vývoje až po nasazení. HW – funkční vzorek – použité technologie a nástroje, Výzkumná a experimentální práce – opakovatelnost experimentů.

Hodnocení výsledků, jejich využitelnost

Dle charakteru práce zhodnoťte možnosti nasazení výsledků práce v praxi nebo uveďte, zda výsledky ZP rozšiřují již publikované známé výsledky nebo přinášející zcela nové poznatky.

Celkové hodnocení

Shrňte stránky ZP, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení nemusí být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích. Obecně platí, že bezvadně splněné zadání je hodnoceno klasifikačním stupněm A.