



Posudek oponenta závěrečné práce

Oponent práce:	Ing. Miroslav Čepek, Ph.D.
Student:	Bc. Peter Kolárovec
Název práce:	Rekonstrukce pseudosociálních sítí z bankovních transakčních dat
Obor / specializace:	Znalostní inženýrství
Vytvořeno dne:	23. srpna 2022

Hodnotící kritéria

1. Splnění zadání

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

Zadání splněno.

2. Písemná část práce

85 /100 (B)

Práce je dobře psaná. Jsou místa, kde by pomohla ilustrace nebo více detailů.

- * Textu by pomohla ilustrace transakčního grafu se všemi typy uzlů a hran. (Rozšířená verze obrázku 4.5b.)
- * V praktické části práce bych ocenil diagram ilustrující návaznost jednotlivých kroků a použitých algoritmů.
- * Ocenil bych více informací o použití algoritmu Node2Vec na transakční graf.
- * Rozhodně bych ocenil více experimentů - s architekturami a hyperparametry.
- * Kapitola 4.6 Deployment říká, že nasazení do produkčního prostředí je náročné, a já s tím souhlasím, ale ocenil bych alespoň návrh a kroky, které k nasazení vedou.

3. Nepísemná část, přílohy

85 /100 (B)

Implementační část diplomové práce používá jazyk Python a využívá knihoven PyTorch, torch-geometric a NetworkX, které reprezentují standardní nástroje pro oblast strojového učení a zpracování grafů. Student jako vývojovou platformu používá Jupyter. Kód je přiměřeně čitelný a srozumitelný. Přesto bych doporučoval rozdělit největší funkce na několik menších - budou srozumitelnější a budou se (potenciálně) lépe testovat. Také bych doporučil intenzivnější dokumentaci.

4. Hodnocení výsledků, jejich využitelnost

100 /100 (A)

Práce vznikla ve spolupráci s Českou Spořitelnou a zdá se jako povedený proof-of-concept projekt, který, podle prezentovaných výsledků a s rozšířením na další typy vztahů může pomoci s různými procesy v bance, ať jde o marketing, KYC, AML nebo compliance.

Celkové hodnocení

85 /100 (B)

I přes výtky k textu práce a prostor pro zlepšení se mi líbí její praktické zaměření a použitelnost.

Otázky k obhajobě

- 1) V kapitole 4.4 a tabulce 4.6, ilustrující dosažené výsledky, používáte model jedné velikosti a jedné architektury. Zkoušel jste i další architektury nebo hyperparametry? Získal jste horší výsledky?
- 2) Zkoušel jste i jiné způsoby kódování informací spojených s hranami? Případně techniky, které hranové informace zpracovávají nativně?
- 3) Jak moc by bylo náročné zopakovat experiment pro další typy vztahů - například tenant-renter zmíněný v zadání?
- 4) Jak byste si představoval nasazení vašeho modelu do produkce? Jak byste model monitoroval?

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.