

Posudek oponenta bakalářské práce

Student: Skyva Petr
Téma: Analýza dat ze sociálních sítí (id 18664)
Oponent: Otrusina Lubomír, Ing., UPGM FIT VUT

1. **Náročnost zadání** průměrně obtížné zadání
2. **Splnění požadavků zadání** zadání splněno
Cílem práce bylo vytvořit systém pro analýzu dat ze sociálních sítí. Zadání práce bylo splněno, i když vytvořený systém je velmi jednoduchý.
3. **Rozsah technické zprávy** je v obvyklém rozmezí
Rozsah technické zprávy je v obvyklém rozmezí. Některé části, ale nejsou pro práci nezbytné. Student např. popisuje soc. síť Instagram a LinkedIn, které nejsou v práci použity. Rovněž je zde popsáno několik českých soc. sítí, ale student se zaměřuje výhradně na angličtinu. Popis bezpečnostních rizik jednotlivých soc. sítí také není příliš relevantní. Kapitola 4 s popisem implementace je až příliš podrobná a popisuje implementaci na příliš nízké úrovni.
4. **Prezentační úroveň předložené práce** 50 b. (E)
Logická struktura technické zprávy je dobrá. Jednotlivé kapitoly práce na sebe dobře navazují. Připomínky k rozsahu některých kapitol jsou uvedeny výše. Práce obsahuje mnohá nepravdivá a nepodložená tvrzení. Vyskytují se zde také některé hůře pochopitelné věty. Velká část práce připomíná spíše vypravování než technický styl. Některé pasáže (např. kapitola 4.4) jsou velmi nudné. Vyhodnocení celého systému je velmi subjektivní a vágní. Práce obsahuje mnoho vět, které nedávají smysl.
5. **Formální úprava technické zprávy** 50 b. (E)
Po formální stránce práce obsahuje mnoho prohřešků. Student v práci nedělá mezery před otevírací závorkou, což působí velmi rušivě. Práce je plná gramatických chyb a překlepů. Některé termíny student nepřekládá (např. topic) a jiné překládá špatně (např. coordinates jako koordinace). Často chybí odkazy na zmiňované soc. síť a knihovny. V práci je spousta ukázek JSON kódu, které nejsou nijak označeny.
6. **Práce s literaturou** 50 b. (E)
V seznamu studijních pramenů je pouze jedna seriózní publikace zabývající se analýzou dat ze soc. sítí. Toto téma je přitom v literatuře dobře pokryto. Ostatní položky jsou různé webové stránky s danou problematikou, případně weby použitých technologií či Wikipedie.
7. **Realizační výstup** 50 b. (E)
Student vytvořil systém pro analýzu dat ze sociálních sítí, omezil se přitom na soc. síť Twitter a Facebook. Použité metody jsou však velmi jednoduché a neumožňují získat žádná zajímavá data. Zcela chybí vizualizace výsledků, což by u takového systému člověk očekával.
8. **Využitelnost výsledků**
Jedná se o práci kompilačního charakteru. Systém sám o sobě pravděpodobně dále využitelný není. Je však možné využít data, která student v práci shromáždil.
9. **Otázky k obhajobě**
 1. Jak vytváříte soubory s aspekty a lokacemi? Nikde v práci není popsáno, jak tyto informace získáváte.
 2. Proč jste se v práci omezil pouze na analýzu sentimentu?
 3. V kapitole 5.6 popisujete analýzu toho, kam uživatelé cestují. Jak tuto informaci získáváte? To, že uživatel zmíní v tweetu nějakou lokaci, přece neznamená, že do dané lokace cestoval.
10. **Souhrnné hodnocení** 50 b. dostatečně (E)
Student v práci vytvořil systém pro analýzu dat ze sociálních sítí. Při analýze však použil pouze velmi jednoduché metody, které aplikoval na malou datovou sadu. Vyhodnocení systému je velmi vágní a nic o systému nevypovídá. Technická zpráva má mnoho nedostatků. Z těchto důvodů navrhuji hodnocení stupněm E (50b).

Prohlášení: Uděluji VUT v Brně souhlas ke zveřejnění tohoto posudku v listinné i elektronické formě.

.....

podpis