



Hodnocení vedoucího závěrečné práce

Student: Martin Koucký
Vedoucí práce: Ing. Daniel Vašata, Ph.D.
Název práce: Analysis of discussion comments and their authors in social media
Obor: Znalostní inženýrství

Datum vytvoření: 15. 6. 2020

Hodnotící kritérium:	Způsob hodnocení – následující škálou 1 až 4:
1. Splnění zadání	1=zadání splněno, 2=zadání splněno s menšími výhradami, 3=zadání splněno s většími výhradami, 4=zadání nesplněno
Popis kritéria: Posuďte, 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, posuďte 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.	
Komentář: Zadání bylo splněno. Student v práci s využitím nástrojů pro strojové zpracování přirozeného jazyka zkoumal možnosti detekce uživatelů, kteří jsou z pohledu příspěvků v sociálních médiích anomální.	
Hodnotící kritérium:	Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):
2. Písemná část práce	92 (A)
Popis kritéria: Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále posuďte, 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. Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 26/2017, článek 3. Posuďte, 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.	
Komentář: Práce je logicky dobře strukturovaná a po typografické a jazykové stránce na výborné úrovni. Zdroje jsou relevantní a správně citované. Jednou z mála drobných výtek je neoptimální formulace a reference v části 2.2.2, kde uvedené tvrzení není formulováno jako věta, byť je tak uvedeno, a také není uveden přímý odkaz na odpovídající větu v uvedeném článku.	
Hodnotící kritérium:	Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):
3. Nepísemná část, přílohy	95 (A)
Popis kritéria: 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	
Komentář: Nepísemnou částí práce byla experimentální analýza možností detekce anomálních uživatelů na datech z médií Reddit a Twitter. Textová data byla převedena do vektorové reprezentace pomocí modelu BERT a následně byla provedena analýza pomocí shlukování a metody Local Outlier Factor. V příloze práce jsou uvedeny zdrojové kódy pro Google Colab, které umožňují případnému zájemci všechny provedené experimenty replikovat.	
Hodnotící kritérium:	Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):
4. Hodnocení výsledků, jejich využitelnost	89 (B)
Popis kritéria: 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.	

Komentář:

V teoretické části práce jsou po rešerši existujících přístupů uvedeny model BERT a způsob detekce anomálií Local Outlier Factor. V experimentální části jsou detailně popsány všechny provedené experimenty a jejich výsledky jsou přehledně prezentovány a diskutovány. Vzhledem k tomu, že se nepodařilo dostatečně přesvědčivě detekovat anomální uživatele, nejsou výsledky publikovatelné. Mohou však sloužit jako základ další práce.

Hodnotící kritérium:

Způsob hodnocení – následující škálou 1 až 5:

5. Aktivita a samostatnost studenta

5a:

1=výborná aktivita,
2=velmi dobrá aktivita,
3=průměrná aktivita,
4=slabší, ale ještě dostatečná aktivita,
5=nedostatečná aktivita

5b:

1=výborná samostatnost,
2=velmi dobrá samostatnost,
3=průměrná samostatnost,
4=slabší, ale ještě dostatečná samostatnost,
5=nedostatečná samostatnost

Popis kritéria:

V souvislosti s průběhem a výsledkem práce posudte, zda byl student během řešení aktivní, zda dodržoval dohodnuté termíny, jestli své řešení průběžně konzultoval a zda byl na konzultace dostatečně připraven (5a). Posudte schopnost studenta samostatně tvůrčí práce (5b).

Komentář:

Student byl aktivní a samostatně pracoval. I v době nouzového stavu jsme měli pravidelné on-line konzultace.

Hodnotící kritérium:

Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):

6. Celkové hodnocení

91 (A)

Popis kritéria:

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.

Text hodnocení:

Práce je celkově na vynikající úrovni a tedy i přes výše uvedené nepřesvědčivé výsledky navrhuji hodnocení stupněm A.

Podpis vedoucího práce: