



Posudek oponenta závěrečné práce

Oponent práce: doc. RNDr. Ing. Marcel Jiřina, Ph.D.
Student: Pavlína Pokorová
Název práce: Rozpoznávání extremistických textů
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

Práce systematicky naplňuje jednotlivé kroky zadání. Dílčí cíle jsou dobře obsahově vymezeny a zpracovány.

2. Písemná část práce

90/100 (A)

Rozsah práce odpovídá rozsahu zadání a požadavkům kladeným na bakalářskou práci. Text je systematicky vystaven a dobře strukturován. Kapitoly na sebe dobře navazují, práce je obsahově vyvážená a mimořádně dobře čtivá. Oceňuji věcně, korektně a výstižně popsané skutečnosti, které svědčí o odborných znalostech studentky. Práce uvádí 65 zdrojů, což je poměrně hodně a rozhodně je to dobře. Zdroje jsou v práci vhodně citovány. Určitou výhradu mám k používání symbolů. Studentka velmi volně mění v práci význam jednotlivých proměnných a to i v rámci dílčích kapitol. Projevuje se to např. u proměnných n , k , D . V rámci jednoho díla by mělo být značení jednotné, ale současně to nevnímám jako zásadní nedostatek, protože to nijak významně nenarušuje pochopení textu. Stylistická i typografická úroveň práce je výborná. Po jazykové stránce je práce také dobře napsána, našel jsem jen minimum chyb, nejvíce jich vnímám u používání čárek ve větě. Zajímavé je také počestvování a využívání anglických výrazů jako např. split, default nebo seed. Neberu to jako nedostatek, ale jen zajímavé pozorování vývoje češtiny.

3. Nepísemná část, přílohy

100/100 (A)

V rámci práce byly vytvořeny dva na sebe navazující programy a to tzv. "trénovací program" a "klasifikační program". Oba tyto programy naplňují požadavky zadání, jsou dobře koncepčně navrženy, implementovány a využity.

4. Hodnocení výsledků, jejich využitelnost

100 /100 (A)

Navržený postup analýzy textových dokumentů i provedené experimenty jsou správně zvoleny a provedeny. Výsledky jsou dobře vyhodnoceny z více pohledů a mají dobrou vypovídací hodnotu. Závěry práce jsou tak použitelné v praxi, zejména v rámci v současnosti řešeného projektu na detekci extremistických textů.

Celkové hodnocení

95 /100 (A)

Práce je dobře zpracována, rozsah práce odpovídá bakalářské práci a požadavkům zadání, navržené postupy jsou promyšlené, kvalitně implementované a dobře vyhodnocené. Výsledky jsou použitelné v praxi. Moje výše uvedené výhrady jsou spíše jen formálního charakteru a nesnižují kvalitu předložené práce.

Otázky k obhajobě

- 1) Uvádíte, že n-gram je n-tice po sobě jdoucích slov (kap. 2.3.1.2). Chci se jen ubezpečit, že máte na mysli opravdu slova a ne znaky.
- 2) U metody Random forest (kap. 2.3.4.5) uvádíte, že se dá využít pro klasifikační i regresní úlohy. V navazující větě pak uvádíte, že při predikci se o výsledku hlasuje, což je nepřesné. Jak se predikuje výsledek u regresní úlohy?

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.