



Hodnocení vedoucího závěrečné práce

Vedoucí práce: Ing. Jaroslav Kuchař, Ph.D.
Student: Bc. Tomáš Stanovčák
Název práce: Extraktor informací o firmách z webových zdrojů
Obor / specializace: Znalostní inženýrství
Vytvořeno dne: 30. května 2022

Hodnotící kritéria

1. Splnění zadání

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

Cílem ZP bylo seznámit se s problematikou extrakce informací z webu a dále připravit nástroj pro získání dat - konkrétně informací o firmách. Z hlediska splnění zadání nevidím nedostatky. V rámci DP je pokryta více teorie extrakce z webu a následně aplikovány vybrané přístupy.

2. Písemná část práce

90/100 (A)

Práce se zabývá zejména teorií a následně aplikací vybraných metod. Text je vhodně strukturován, dobře čtivý a bez zásadních faktických nedostatků. Všechny kapitoly jsou informačně bohaté bez zbytečných částí. Obecně se v rámci práce na DP narazilo na "běžnou" problematiku v této oblasti: nedostatek kvalitně anotovaných dat. To ovlivnilo volbu přístupů extrakce a celkové zaměření DP. Dílčí postupy a rozhodnutí jsou dobře dokumentovány a zdůvodněny. Po jazykové a typografické ani citační stránce nevidím zásadní nedostatky.

3. Nepísemná část, přílohy

90/100 (A)

V rámci práce vzniklo řešení, které je dostupné pro zájemce o tuto problematiku. Vznikla sada nástrojů pro získání obsahu i jejich následné zpracování. Po implementační stránce neshledávám žádné zásadní nedostatky.

4. Hodnocení výsledků, jejich využitelnost

90 /100 (A)

Vzhledem k problematice nedostupnosti kvalitních dat je výstupem práce i základní sada pro vyhodnocení zvolených přístupů. Získání těchto dat je velmi časově náročné. Tato sada je jedním z výstupů, který je dále využitelný pro návazné práce. Výstupem je i implementace, která obsahuje první sadu metod pro extrakci a je možné ji dále rozšiřovat. Kladně hodnotím zvolené přístupy. Kdy mimo klasického hledání vzorů a regulárních výrazů bylo využito hledání vybraných informací pomocí klasifikace a detekce pojmenovaných entit.

5. Aktivita studenta

- ▶ [1] výborná aktivita
- [2] velmi dobrá aktivita
- [3] průměrná aktivita
- [4] slabší, ale ještě dostatečná aktivita
- [5] nedostatečná aktivita

Velmi kladně oceňuji pravidelné konzultace, velmi dobrou připravenost a spolupráci.

6. Samostatnost studenta

- ▶ [1] výborná samostatnost
- [2] velmi dobrá samostatnost
- [3] průměrná samostatnost
- [4] slabší, ale ještě dostatečná samostatnost
- [5] nedostatečná samostatnost

Velmi kladně oceňuji pravidelné konzultace, velmi dobrou připravenost a spolupráci.

Celkové hodnocení

90 /100 (A)

Student se s problematikou velmi dobře seznámil, navrhl a dodal řešení splňující zadání. Velkou část vlastní práce věnoval student získání a přípravě dat aby bylo možné různé přístupy vyhodnotit. Student odzkoušel i zajímavé přístupy pro extrakci s pomocí klasifikace a detekce pojmenovaných entit v textu. Práci doporučuji k obhajobě.

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.

Aktivita studenta

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.

Samostatnost studenta

V souvislosti s průběhem a výsledkem práce posudte schopnost studenta samostatně tvůrčí práce.

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.