



Posudek oponenta závěrečné práce

Student: Bc. Jakub Trhlík
Oponent práce: Ing. Petr Špaček, Ph.D.
Název práce: Fact extraction from Wikipedia article texts
Obor: Webové a softwarové inženýrství

Datum vytvoření: 15. 1. 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: Práce splňuje všechny body zadání bez výhrad.	
Hodnotící kritérium:	Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):
2. Písemná část práce	70 (C)
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: Rozsah práce činí 59 stran bez příloh. Dominantní objem práce má rešeršní charakter, minoritní část je pak popis experimentů, toto je v souladu s povahou práce i zadáním. Textu by prospělo přidání argumentace k výběru technik a technologií, namísto prostých konstatování typu: použijeme toto. Užití externích zdrojů i reference na ně jsou správně dokumentovány a v souladu s licencemi. Práce je psána technickou angličtinou s občasnými chybami, např. jsou zde zdvojená slova nebo chyby ve spellingu.	
Hodnotící kritérium:	Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):
3. Nepísemná část, přílohy	90 (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: I přesto, že je práce experimentálně-rešeršního charakteru, považuji zdrojový kód práce za velmi dobrý, student používá vybrané technologie v souladu s jejich dokumentací, což vede na dobrou znovu-využitelnost.	
Hodnotící kritérium:	Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):
4. Hodnocení výsledků, jejich využitelnost	90 (A)
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:

Za hlavní výstup práce považuji návrh nové metody na extrakci relací, která je založena na bootstrappingu v kombinaci s metodami vzdáleného dohledu (distant supervision) a díky níž se autorovi povedlo obohatit množinu relací DBpedie v oblasti medicíny o zcela nové prvky.

Hodnotící kritérium:

Způsob hodnocení – nehodnotí se

5. Otázky k obhajobě

Popis kritéria:

Uvedte případné dotazy, které by měl student zodpovědět při obhajobě ZP před komisí (body oddělte odřázkami).

Otázky:

Jaké vlastnosti měli specializované datasety užitě pro testování (tabulka 0.9)?

Jakým způsobem budete postupovat při zvyšování míry automatizace (vs. semi-auto) výsledné pipeline?

Hodnotící kritérium:

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

6. Celkové hodnocení

80 (B)

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:

Své finální hodnocení zakládám obtížnosti práce, tj. hlavně na velkém počtu technik a technologií, které student musel nastudovat a použít. Menší vliv pak měla kvalita finálního textu práce.

Podpis oponenta práce: