

Posudek oponenta závěrečné práce

České vysoké učení technické v Praze

Fakulta informačních technologií

Student: Bc. Tomáš Benák
Oponent práce: Ing. Jaroslav Kuchař
Název práce: Učení domén pojmenovaných entit
Obor: Webové a softwarové inženýrství

Datum vytvoření: 30. 5. 2016

Hodnotící kritérium:	Způsob hodnocení - následující škálou 1 až 5:
1. Náročnost a další komentář k zadání	1=mimořádně náročné zadání, 2=náročnější zadání, 3=průměrně náročné zadání, 4=lehčí, ale ještě dostatečně náročné zadání, 5=nedostatečně náročné zadání
Popis kritéria: Podrobněji charakterizujte diplomovou (bakalářskou) práci a její případné návaznosti na předchozí nebo běžící projekty. Dále posuďte, čím je zadání této ZP náročné. (U obtížnější ZP lze dále tolerovat některé nedostatky, které by u ZP standardní obtížnosti tolerovány nebyly; a naopak u jednoduché ZP mohou být zjištěné nedostatky hodnoceny přísněji.)	
Komentář: Cílem ZP je seznámit se s problematikou vytváření/učení modelů pro pojmenované entity, které by umožňovaly jejich klasifikaci do odpovídajících domén zájmů. Práce staví na předchozích výzkumech a rozšiřuje možná řešení o další přístupy, které ověřuje experimenty.	
Hodnotící kritérium:	Způsob hodnocení - následující škálou 1 až 4:
2. 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 splňuje zadání. V komentáři uveďte body zadání, které nebyly zcela splněny, případně rozšíření ZP oproti původnímu zadání. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.	
Komentář: Zadání bylo splněno v plném rozsahu. Student se seznámil s problematikou, navrhl metody a ověřil experimenty.	
Hodnotící kritérium:	Způsob hodnocení - následující škálou 1 až 4:
3. Rozsah písemné zprávy	1=splňuje požadavky, 2=splňuje požadavky s menšími výhradami, 3=splňuje požadavky s většími výhradami, 4=nesplňuje požadavky
Popis kritéria: Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části.	
Komentář: ZP splňuje požadavky na rozsah. Práce obsahuje všechny důležité části, které jsou informačně bohaté. Pouze části zabývající se analýzou a návrhem jsou spojené do jedné kapitoly a v některých situacích je to nepřehledné.	
Hodnotící kritérium:	Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):
4. Věcná a logická úroveň práce	75 (C)
Popis kritéria: 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.	
Komentář: Struktura práce je až na drobnosti v pořádku. Pouze některé pojmy jsou nejprve používány a až později definovány (např. 2.1.2 Pojmenované entity nebo 2.4 Wikipedia a DBpedia). Po věcné stránce je práce v pořádku. Některá tvrzení a rozhodnutí by bylo dobré lépe prezentovat (např. volba dat, metod, nebo „nejlepe dopadl algoritmus kNN, jehož vyřazení? c?as tre?nova?ni? je te?me?r? nulovy?“). Pojem záloha bych v tomto kontextu považoval za standardní mechanismus pro cache mezivýsledků.	
Hodnotící kritérium:	Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):
5. Formální úroveň práce	95 (A)
Popis kritéria: Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 12/2014, článek 3.	
Komentář: Z pohledu typografické a jazykové stránky je práce v pořádku.	
Hodnotící kritérium:	Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):

6. Práce se zdroji

95 (A)

Popis kritéria:

Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení ZP. Charakterizujte výběr studijních pramenů. Posuďte, zda student využil všechny relevantní zdroje nebo zda se pokoušel řešit již vyřešené problémy. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.

Komentář:

V práci je čerpáno především z odborné literatury. Práce se zdroji je v pořádku, není porušena citační etika a bibliografické citace jsou v souladu s citačními zvyklostmi.

Hodnotící kritérium:

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

7. Hodnocení výsledků, publikační výstupy a ocenění

95 (A)

Popis kritéria:

Vyjádřete se k úrovni dosažených hlavních výsledků ZP, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, apod. Případně také zhodnoťte, zda software nebo zdrojové texty, které nevytvořil sám student, byly v ZP použity v souladu s licenčními podmínkami a autorským právem. Popište případnou publikační činnost a získaná ocenění související s řešením této ZP.

Komentář:

V rámci ZP vzniklo řešení, které staví na současném výzkumu. Aplikace je použitelná v reálném nasazení a je možné ji integrovat do existujících nástrojů pro detekci pojmenovaných entit v textu.

Hodnotící kritérium:

Způsob hodnocení - nehodnotí se

8. Komentář o využitelnosti výsledků

Popis kritéria:

Uveďte, zda hlavní výsledky ZP rozšiřují již publikované známé výsledky a/nebo přinášející zcela nové poznatky. Uveďte možnosti využití výsledků ZP v praxi.

Komentář:

Výstupy ZP jsou použitelné v současném výzkumu i praktickém použití. Výsledky experimentů jsou publikovatelné pro komunitu v oblasti sémantického webu a detekce entit.

Hodnotící kritérium:

Způsob hodnocení - nehodnotí se

9. Otázky k obhajobě

Popis kritéria:

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

Otázky:

Jakým způsobem jsou přenášeny modely mezi dvěma implementovanými aplikacemi?

Pokuste se zdůvodnit proč mají výsledky při použití Manhattan i Euclidean vzdálenosti stejné výsledky vzhledem k charakteru metody a dat.

Popište možné výhody/nevýhody využití metody GET v RESTful API vzhledem k časovým nárokům na některé modely.

Hodnotící kritérium:

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

10. Celkové hodnocení

89 (B)

Popis kritéria:

Shrňte stránky ZP studenta, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení **nemusí** být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích 1 až 9.

Text hodnocení:

Student se seznámil s problematikou učení domén entit. Navrhl, implementoval a provedl experiment s metodami, které rozšiřují současné metody. Výsledky jsou použitelné v praktických nasazeních pro nástroje pro detekci entit ale i vhodné k publikování pro odbornou komunitu. Doporučuji k obhajobě.

Podpis oponenta práce: