

Posudek oponenta závěrečné práce

České vysoké učení technické v Praze

Fakulta informačních technologií

Student: Bc. Michal Kopp

Oponent práce: Ing. Petr Pulc

Název práce: System for structure discovery in multimedial metadata in the project Open-Narra

Obor: Webové a softwarové inženýrství

Datum vytvoření: 7. 11. 2015

Hodnotící kritérium:	Způsob hodnocení - následující škálou 1 až 5:
1. Náročnost a další komentář k zadání	<u>1=mimořádně náročné zadání,</u> 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: Práce zahrnuje pokročilé metody strojového učení, které student musel nastudovat samostatně. Zároveň se musel potýkat s poměrně netriviální interpretací výstupů z použitých metod. K větší složitosti zadání přispěly i nedostatečnosti datové sady dodané FAMU a nutnost seznámit se s výrazně rozsáhlejšími vstupy.	
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: V práci zcela chybí diskuse nad metadaty a existujícími ontologiemi, což mi přijde škoda. Je sice pravdou, že reálně použitá data tyto ontologie (z důvodu různých úrovní pohledu na data) příliš nedodržují, ale systém Narra, kterého se tato práce dotýká, dodržuje jistý systém rozšíření nad Dublin Core. Obecná diskuse nad metadaty použitelnými pro shlukování pomocí samoorganizujících se map je také poměrně krátká. V praktické části je pak jen krátce popsána použitá metoda a použitá data, opět bez rozsáhlejší diskuse. Osobně by mne ale například zajímaly výsledky při použití jen vybraných modalit nebo při použití různých vzdáleností (vektory získané z konceptualizace textových dokumentů jsou nejčastěji porovnávány kosinovou vzdáleností).	
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: Práce je vzhledem k obecné nesamozřejmosti některých probíraných témat velmi stručná. Na druhou stranu ale žádné zásadní informace z provedených experimentů nebyly zamlčeny a přesnou implementaci v programu Matlab je možné dohledat na přiloženém médiu. Je škoda, že některé části kódu nebyly podrobněji vysvětleny v textu. S přihlédnutím k obsahu CD a množství komentářů v kódu ale považuji požadavky za splněné.	
Hodnotící kritérium:	Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):
4. Věcná a logická úroveň práce	85 (B)
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: Text je členěn logicky a postup řešení shledávám vcelku bezchybným. Použité metody jsou vždy alespoň zevrubně popsány. Text tedy klade poměrně vysoké nároky na znalosti čtenáře, což je ale u magisterské závěrečné práce možné. Naopak oceňuji zapojení hierarchického shlukování, které je snadno pochopitelné.	

<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):</i>
5. Formální úroveň práce	60 (D)
<i>Popis kritéria:</i> Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 12/2014, článek 3.	
<i>Komentář:</i> Student se nedopustil žádných zásadních prohřešků proti typografii, ale po jazykové stránce je práce poměrně slabá. Množství překlepů a českých gramatických postupů v anglickém textu práce poměrně znesnadňovalo čtení. Příště bych tedy doporučil zaměřit se výrazně více na textovou část práce.	
<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):</i>
6. Práce se zdroji	80 (B)
<i>Popis kritéria:</i> 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.	
<i>Komentář:</i> Všechny části textu byly poměrně dobře pokryty kvalitní literaturou. Nicméně bylo by vhodné čtenáře odkázat na další literaturu i u témat, která student považuje za samozřejmá. Bohužel mi ale v textu obecně chybí některé pasáže, chybí mi tedy i některé zdroje.	
<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):</i>
7. Hodnocení výsledků, publikační výstupy a ocenění	100 (A)
<i>Popis kritéria:</i> 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.	
<i>Komentář:</i> Celý výsledek bádání studenta považuji za výborný a po textových úpravách publikovatelný pro vědeckou komunitu. Část dokonce publikovaná již byla na konferenci ITAT 2015: Search for Structure in Audiovisual Recordings of Lectures and Conferences. In: ITAT 2015 conference proceedings. Aachen: CEUR Workshop Proceedings, 2015, pp. 150-158. ISSN 1613-0073. ISBN 1515120651.	
<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení - nehodnotí se</i>
8. Komentář o využitelnosti výsledků	
<i>Popis kritéria:</i> Uvedte, zda hlavní výsledky ZP rozšiřují již publikované známé výsledky a/nebo přinášející zcela nové poznatky. Uvedte možnosti využití výsledků ZP v praxi.	
<i>Komentář:</i> Výsledek práce je po drobných úpravách použitelný podle původního zadání v systému Narra vyvíjeného na FAMU v Praze. Po teoretické stránce je výsledek významnou aplikací neuronových sítí, konkrétně samoorganizující se mapy, na multimediální obsah se všemi problémy s tím spojenými.	
<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení - nehodnotí se</i>
9. Otázky k obhajobě	
<i>Popis kritéria:</i> Uvedte případné dotazy, které by měl student zodpovědět při obhajobě ZP před komisí (body oddělte odrážkami).	
<i>Otázky:</i> - Plánuje student pokračovat v tématu? Jaké by byly další kroky? - Bylo testováno i použití jiných vzdáleností, například kosinové? Pokud ano, k jakým vedlo výsledkům? - Bylo shlukování prováděno i jen nad jednotlivými modalitami multimédia, nebo byly použity všechny dostupné najednou?	
<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):</i>
10. Celkové hodnocení	80 (B)
<i>Popis kritéria:</i> Shrňte stránky ZP studenta, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení nesmí být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích 1 až 9.	
<i>Text hodnocení:</i> Práce přináší další neotřelý pohled na práci s multimédií v oblasti strojového učení. Celkově tedy práci hodnotím jako velmi inspirativní a náročnou. Krátká textová část a formální úroveň ale bohužel práci činí náročnou i pro případné čtenáře. A spíše než že by dávala odpovědi, otevírá prostor pro další otázky...	

Podpis oponenta práce: