

Posudek oponenta závěrečné práce

České vysoké učení technické v Praze

Fakulta informačních technologií

Student: Bc. Šimon Hlaváč
Oponent práce: Ing. Pavel Kordík, Ph.D.
Název práce: Automatic text summarization
Obor: Znalostní inženýrství (magisterský)

Datum vytvoření: 5. 6. 2015

<p><i>Hodnotící kritérium:</i></p> <p>1. Náročnost a další komentář k zadání</p> <p><i>Popis kritéria:</i> 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.)</p> <p><i>Komentář:</i> Sumarizace textu je poměrně rozsáhlá oblast, která souvisí s NLP. Samozřejmě záleží do jaké hloubky se jde.</p>	<p><i>Způsob hodnocení - následující škálou 1 až 5:</i></p> <p>1=mimořádně náročné zadání, 2=náročnější zadání, 3=průměrně náročné zadání, 4=lehčí, ale ještě dostatečně náročné zadání, 5=nedostatečně náročné zadání</p>
<p><i>Hodnotící kritérium:</i></p> <p>2. Splnění zadání</p> <p><i>Popis kritéria:</i> 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.</p> <p><i>Komentář:</i> Zadání bylo splněno.</p>	<p><i>Způsob hodnocení - následující škálou 1 až 4:</i></p> <p>1=zadání splněno, 2=zadání splněno s menšími výhradami, 3=zadání splněno s většími výhradami, 4=zadání nesplněno</p>
<p><i>Hodnotící kritérium:</i></p> <p>3. Rozsah písemné zprávy</p> <p><i>Popis kritéria:</i> Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části.</p> <p><i>Komentář:</i> Práce obsahuje všechny důležité části, jen analytická část je trochu podměrečná.</p>	<p><i>Způsob hodnocení - následující škálou 1 až 4:</i></p> <p>1=splňuje požadavky, 2=splňuje požadavky s menšími výhradami, 3=splňuje požadavky s většími výhradami, 4=nesplňuje požadavky</p>
<p><i>Hodnotící kritérium:</i></p> <p>4. Věcná a logická úroveň práce</p> <p><i>Popis kritéria:</i> 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.</p> <p><i>Komentář:</i> Největší výhrady mám hned k teoretickému úvodu - proč je takový prostor věnován genetickým algoritmům, které a jsou vysvětlovány i základní koncepty? Našel bych jiné algoritmy, které by si zasloužily v teoretickém úvodu více prostoru. Autor by také mohl lépe generalizovat a provést pečlivější taxonomii metod, protože část kolem 1.2.1.3 Swarm-based approaches budí podezření, že z oblasti přečetl jen jeden článek.</p> <p>V práci není analýza a hned se jde na návrh. Bylo by dobré vědět, jaké algoritmy-nástroje již existují, jak se dají použít, proč ne, apod.</p>	<p><i>Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):</i></p> <p>72 (C)</p>
<p><i>Hodnotící kritérium:</i></p> <p>5. Formální úroveň práce</p> <p><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.</p> <p><i>Komentář:</i> Vadí mi úprava některých grafů, jinak pěkné.</p>	<p><i>Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):</i></p> <p>92 (A)</p>
<p><i>Hodnotící kritérium:</i></p>	<p><i>Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):</i></p>

6. Práce se zdroji

94 (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:

Podstatné zdroje uvedeny.

Hodnotící kritérium:

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

7. Hodnocení výsledků, publikační výstupy a ocenění

93 (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:

Druhá část práce je poměrně zdařilá. Ladění parametrů algoritmů je metodicky správné a výsledky jsou přesvědčivé.

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:

Metody jsou zaimplementovány do systému, ve kterém se dají snadno použít. Nenašel jsem zmínku o tom, zda budou součástí další verze, k tomu se zřejmě vyjádří vedoucí práce.

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 odrážkami).

Otázky:

Existují algoritmy, které jsou použitelné na sumarizaci textů na úrovni slov pro český jazyk? Jaká jsou jejich omezení?

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 **nesmí** být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích 1 až 9.

Text hodnocení:

Vzhledem k nižší kvalitě teoretické části a absenci analýzy hodnotím práci jako velmi dobrou přesto, že bylo vykonáno mnoho experimentů a výsledky se zdají použitelné.

Podpis oponenta práce: