

# Posudek oponenta závěrečné práce

České vysoké učení technické v Praze

Fakulta informačních technologií

**Student:** Jan Švejda  
**Oponent práce:** Ing. Karel Klouda, Ph.D.  
**Název práce:** Web service for advanced text analysis  
**Obor:** Softwarové inženýrství

**Datum vytvoření:** 13. 6. 2017

|  |  |
|--|--|
| <b>Hodnotící kritérium:</b>  | <b>Způsob hodnocení - následující škálou 1 až 5:</b>   |
| <b>1. Náročnost a další komentář k zadání</b>  | <b>1=mimořádně náročné zadání,<br/>2=náročnější zadání,<br/>3=průměrně náročné zadání,<br/>4=lehčí, ale ještě dostatečně náročné zadání,<br/>5=nedostatečně náročné zadání</b> |
| <b>Popis kritéria:</b><br>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.)  |  |
| <b>Komentář:</b><br>Zadání spojuje dvě části: jedna je softwarové inženýrská a požaduje vytvoření webové služby a RESTového rozhraní, druhá spadá do oblasti znalostního inženýrství a požaduje nastudování metod zpracování přirozeného jazyka a i jejich (byť základní) aplikaci na reálná data. Obě tyto části obsahují v dané oblasti obvyklé a nekomplikované úkoly, jejich spojením ale podle mého vzniklo zadání spíše náročnější.  |  |
| <b>Hodnotící kritérium:</b>  | <b>Způsob hodnocení - následující škálou 1 až 4:</b>   |
| <b>2. Splnění zadání</b>   | <b>1=zadání splněno,<br/>2=zadání splněno s menšími výhradami,<br/>3=zadání splněno s většími výhradami,<br/>4=zadání nesplněno</b>  |
| <b>Popis kritéria:</b><br>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.  |  |
| <b>Komentář:</b><br>Zadání bylo bezezbytku splněno. Ve výsledné aplikaci není zahrnuta "prediction of popularity", sám autor ale v části 3.6.1 vysvětluje (za použití rozumných a velmi uvěřitelných argumentů), proč byla tato funkcionální vynechána.  |  |
| <b>Hodnotící kritérium:</b>  | <b>Způsob hodnocení - následující škálou 1 až 4:</b>   |
| <b>3. Rozsah písemné zprávy</b>  | <b>1=splňuje požadavky,<br/>2=splňuje požadavky s menšími výhradami,<br/>3=splňuje požadavky s většími výhradami,<br/>4=nesplňuje požadavky</b>                                |
| <b>Popis kritéria:</b><br>Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části.   |  |
| <b>Komentář:</b><br>Textová část práce je lehce nadstandardně dlouhá a není to tím, že by se v některé části autor zbytečně rozepisoval. Naopak, díky širokému tématickému záběru bylo nutné dělat kompromisy a dobře zvážit, kdy zacházet do detailů a kdy se spolehnout na to, že čtenář do dané problematiky alespoň trochu vidí. Myslím, že tyto kompromisy byly udělané s rozvahou a rozumně: např. tam, kde byly popisovány běžné softwarové inženýrské postupy, byl popis stručný a detaily bylo možné vyčíst z příloženého UML diagramu. Z popisů algoritmů strojového učení apod. je také cítit, že jim autor dobře porozuměl (nebo to umí alespoň velmi dobře imitovat). |  |
| <b>Hodnotící kritérium:</b>  | <b>Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):</b>  |
| <b>4. Věcná a logická úroveň práce</b>   | <b>100 (A)</b>   |
| <b>Popis kritéria:</b><br>Posuďte, zda předložená ZP je po věcné stránce v pořádku, případně vyskytly-li se v práci věcné chyby nebo nepřesnosti. Zhodnoťte dále logickou strukturu ZP, návaznosti jednotlivých kapitol a pochopitelnost textu pro čtenáře.  |  |
| <b>Komentář:</b><br>Jedná se o jednu z nejlépe napsaných prací, kterou jsem kdy četl. Toto hodnocení je o to významnější, že je práce psána anglicky.  |  |
| <b>Hodnotící kritérium:</b>  | <b>Způsob hodnocení - bodové hodnocení 0 až 100 bodů (známka A až F):</b>  |
| <b>5. Formální úroveň práce</b>  | <b>95 (A)</b>  |
| <b>Popis kritéria:</b><br>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 14/2015, článek 3.  |  |

**Komentář:**

V práci se vyskytují jen drobné formální nesrovnalosti. Např. za rovnicemi někdy chybí tečky, při odkazování na obrázky je ve slově "figure" použito malé f namísto velkého a občas jsem měl pochybnosti o správnosti použité interpunkce.

*Hodnotící kritérium:*

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

**6. Práce se zdroji**

100 (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:**

K práci se zdroji nemám připomínek a považuji ji za naprosto adekvátní.

*Hodnotící kritérium:*

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

**7. Hodnocení výsledků, publikační výstupy a ocenění**

100 (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:**

Výsledná aplikace dělá přesně to, co bylo požadováno v zadání. Výsledky samotných modelů nejsou nikterak oslňující, ale to se, myslím, ani čekat nedalo a cílem práce nebylo ladit tyto modely.

*Hodnotící kritérium:*

*Způsob hodnocení - nehodnotí se*

**8. Komentář o využitelnosti výsledků**

*Popis kritéria:*

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

**Komentář:**

Výsledná webová služba a rozhraní mohou velmi usnadnit používání různých knihoven a frameworků pro zpracování (textových) dat pro uživatele ne-experta, resp. zpřístupní výsledky a funkce těchto knihoven a frameworků běžným uživatelům či, pomocí RESTového rozhraní, jiným aplikacím. V appendixu C je navíc popsáno, jak lze do výsledného softwaru přidávat další modely a funkcionality.

*Hodnotící kritérium:*

*Způsob hodnocení - nehodnotí se*

**9. Otázky k obhajobě**

*Popis kritéria:*

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

**Otázky:**

1. Je v současné době zajištěno, aby se k API nedostal nikdo nežádoucí? Pokud ano, tak jak?
2. Myslíte, že by výsledky některých modelů zlepšila lemmatizace vstupních dat v českém jazyce? Jak by se příp. lemmatizér do aplikace dal přidat?

*Hodnotící kritérium:*

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

**10. Celkové hodnocení**

98 (A)

*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 výše uvedenému navrhuji práci hodnotit jako výbornou.

Podpis oponenta práce: