



Posudek oponenta závěrečné práce

Student: Vojtěch Paukner
Oponent práce: Ing. Magda Friedjungová
Název práce: Word sense representation for the Czech language
Obor: Znalostní inženýrství

Datum vytvoření: 9. 6. 2019

<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – následující škálou 1 až 4:</i>
1. Splnění zadání	<u>1=zadání splněno,</u> 2=zadání splněno s menšími výhradami, 3=zadání splněno s většími výhradami, 4=zadání nesplněno
<i>Popis kritéria:</i> 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.	
<i>Komentář:</i> Student splnil zadání v plném rozsahu. Téma považuji za náročnější vzhledem k doméně, které na fakultě není při výuce věnován příliš velký prostor, a vzhledem k úskalím českého jazyka.	
<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):</i>
2. Písemná část práce	95 (A)
<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. 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.	
<i>Komentář:</i> Práce je logicky členěna a je dobře čitelná. Teoretická část práce je velmi pěkně zpracovaná. Student si vyhrál s obrázky a příklady, pomocí kterých velmi dobře demonstruje teoretickou stránku věci. Po praktické stránce student čelil několika výzvám a problémům. Největším úskalím je kvalita dat (především "word sense disambiguation" dataset) a výpočetní náročnosti. Context2vec neobsahuje veřejně dostupný předtrénovaný model pro češtinu. Implementace ELMo modelu zase uvažuje pouze ASCII kódování, které je nepřípustné pro český jazyk. Při přípravě experimentů na studenta podle mě číhala past vedle pasti. Samotné předzpracování dat považuji za náročnější část práce. Několik poznámek: U Context2vec mi chybí popis modelu pro němčinu. Chybí mi zdůvodnění (motivace) zvolených binárních klasifikátorů. Stejně tak metoda určení výsledné velikosti k u k-NN klasifikátoru (přestože u k-means se student určením k zabývá). U tabulky 5.6 mi uniká, proč je přesnost prezentována na trénovací a validační množině a ne na testovací (nicméně testovací se vyskytuje v závěru). Student uvádí, že nedosáhl dobrých výsledků při použití alg. AdaBoost, nicméně v tabulce 5.6 algoritmus dosahuje srovnatelných výsledků jako logistická regrese či k-NN. Což si vysvětluji tak, že tyto tři metody jsou všechny stejně špatné, neboť dosahují přesnosti okolo 50%. V sekci 5.6.2 bych uvítala tabulku obsahující dosažené přesnosti na učících sadách jako tomu bylo pro češtinu v tabulce 5.6. Přestože se zadání práce soustředí na český jazyk, je v práci poskytnuto srovnání s německým jazykem - v závěru bych tak očekávala nějaké větší porovnání výsledků. Nicméně to jsou za mě pouze drobnosti.	
<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):</i>
3. Nepísemná část, přílohy	92 (A)
<i>Popis kritéria:</i> 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, Významná a experimentální práce – opakovatelnost experimentů	

Komentář:

Jedná se o práci experimentálního typu. Technologie i příslušné algoritmy jsou dobře zvolené, k implementaci nemám výhrady. Popis experimentů i kódu je dosatčující. U českého jazyka je práce replikovatelná, u němčiny bohužel z důvodu nemožnosti publikovat použitý dataset.

Hodnotící kritérium:

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

4. Hodnocení výsledků, jejich využitelnost

95 (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:

Rozpoznávání různých významů slov není vůbec jednoduchým problémem. Stejně tak NLP aplikovaný nad češtinou. Práce je velmi zajímavá, přestože dosažené přesnosti klasifikátorů nejsou zářné. To však není chybou studenta, ale dostupností potřebných dat a obtížností jejich přípravy. Kladně hodnotím srovnání supervizovaných a nesupervizovaných metod, stejně tak porovnání výsledků pro český a německý jazyk. Na práci lze navázat dalšími experimenty - využití jiného korpusu, rozšíření dat, zohlednění jiných klasifikátorů či shlukovacích metod, srovnání mezi slovanskými jazyky apod.

Hodnotící kritérium:

Způsob hodnocení – nehodnotí se

5. Otázky k obhajobě

Popis kritéria:

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

Otázky:

- 1) Jaká byla motivace pro srovnání češtiny a němčiny (slovanský vs. germánský jazyk)? Např. lexikální podobnost, vývoj?
- 2) Dokážete odhadnout, při jak velkých trénovacích datech byste dosáhl vyšší přesnosti u Context2vec modelu? Ostatně, jaké stávající přesnosti jste dosáhl?
- 3) U hlubokých neuroných sítích pro češtinu uvádíte jako aktivační funkci ReLU, zkoušel jste i jiné? Proč jste zvolil tuto? U němčiny jste postupoval stejně?
- 4) Stručně prosím shrňte dělení dat na trénovací, validační a testovací sadu a jejich použití při experimentech.

Hodnotící kritérium:

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

6. Celkové hodnocení

95 (A)

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:

Student se zabýval zpracováním jazyka, konkrétně state-of-the-art metodami pro detekci rozdílných významů slov. Své poznatky demonstroval na českém a německém jazyce. Experimenty byly provedeny jak se supervizovanými metodami, tak nesupervizovanými. Práci doporučuji hodnotit stupněm A, neboť splňuje veškeré náležitosti.

Podpis oponenta práce: