FAKULTA INFORMAČNÍCH TECHNOLOGIÍ ČVUT V PRAZE

Posudek oponenta závěrečné práce

Student: Bc. Jiří Kožusznik
Oponent práce: Ing. Martin Flusser
Název práce: Using deep neural networks for sentiment analysis from utterances
Obor: Znalostní inženýrství

Datum vytvoření: 28. 1. 2019

Hodnotící kritérium:	*Způsob hodnocení – následující škálou 1 až 4:*
1. 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 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.

Komentář:
Student splnil převážnou část zadání. Vytknout jen mohu pár věcí například v bodech:
1) Mohlo být lépe popsáno, jaké nástroje Matlab nabízí pro použití neuronových sítí. Možnosti použití GPU a virtual cloud Metacentra práce nepopisuje.
3) Je diskutabilní, jestli se použité sítě dají označit jako hluboké.
Zde je možná souvislost s doporučenou změnou datasetu, jak je popsáno v posudku školitele.

Hodnotící kritérium:	*Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):*
2. Písemná část práce	*25 (F)*

Popis kritéria:
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.

Komentář:
První polovina práce má rešeršní charakter a nese samozřejmě výrazně více textu, než druhá polovina, která popisuje provedený experiment a výsledky. Již na prví pohled je vidět jistá (nejen jazyková) nevyrovnanost mezi první a druhou půlkou práce, což by mohlo indikovat, že v první půlce práce došlo k přejímání cizích textů a tedy, že se student mohl dopustit hrubého porušení citačních standardů, vzhledem k tomu že v práci není nikde naznačena přímá citace (citát). V následujícím textu shrnu sekce u kterých se mi povedlo dohledat originální text a tedy potvrdit zmíněné porušení konvencí.

1.1 (Audio-visual content) celá sekce shodná s https://www.toprankblog.com/2013/03/what-is-content/
1.2 (Bark scale) celá sekce shodná s wikipedií, včetně uváděného zdroje na původní literaturu. V té však tento blok textu nefiguruje. https://en.wikipedia.org/wiki/Bark_scale
1.3 (Audio Descriptors) po blocích zkopírováno z uvedeného zdroje, ale není správně citováno.
http://recherche.ircam.fr/anasyn/peeters/ARTICLES/Herrera_1999_ICMC_MPEG7.pdf
1.3.1 (MPEG-7 Audio Descriptors) První odstavec zkopírován z wikipeide. https://en.wikipedia.org/wiki/MPEG-7
1.4 a 1.5: Celé sekce dohromady čítající přes 5 stránek zkopírovány z https://courses.fit.cvut.cz/MI-ADM/media/lectures/02/kniha1-v1.pdf. Zde je zdroj uveden vždy jednou na konci sekce. Nejedná se o korektní citaci.
Fig. 1.2 (SVM) Celý popis obrázku zkopírován a nesprávně citován
1.7.1 (Cross validation) Celá sekce zkopírována bez citace či reference. https://www.openml.org/a/estimation-procedures/1
1.7.2 (Friedman test) celá sekce zkopírována bez citace či reference.
https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0184834
http://www.lsi.us.es/redmidas/CEDI07/%5B9%5D.pdf
1.7.3 (Holm correction) Úvodní odstavec (bez první věty) zkopírován bez citace či reference. Zde pravděpodobně v důsledku kopírování vnikla první věta bez slovesa: "Bonferroni-Holm correction [7] for multiple comparisons."
https://www.frontiersin.org/articles/10.3389/fpsyg.2013.00853/full

Opakuji, že problém výše uvedený se vyskytuje jen v první půlce práce.
V druhé části práce bych určitě doporučil trochu zapracovat na přehlednosti a jazykové kvalitě. Např. sekce "Testing" je sestavená z více různých experimentů a zasloužila by lepší strukturu a vysvětlení. Dále sekce 3.3 má 20 podsekcí, kdy téměř každá začíná slovy "This script loads", případně "This script creates". Zde by autor mohl alespoň ukázat snahu o práci s jazykem. Nutno ale říct, že jazykové ztvárnění, i přes nedokonalosti, nebrání v pochopení a porozumění textu.

Hodnotící kritérium:	*Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):*
3. Nepísemná část, přílohy	*85 (B)*

Popis kritéria:
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, Výzkumná a experimentální práce – opakovatelnost experimentů

Komentář:
Stěžejní část přílohy tvoří Matlab soubory. Student vytvořil zhruba dvacet skriptů, které mají, dle mého názoru, na diplomovou úroveň dostatečnou kvalitu. Mohu jen vytknout, absenci nebo nedostatek komentářů, které jsou klíčové pro následné použití.

Hodnotící kritérium:	*Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):*
4. Hodnocení výsledků, jejich využitelnost	*90 (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:
Co se týče experimentu, student odvedl solidní práci. Osvojil si práci se zvukovými deskriptory a také správně využil několika různých metod strojového učení. Pravděpodobně i díky dodržení správné metodologie (volba datasetu, cross validace, statistické testy, adt..) měl možnost prezentovat svoje výsledky na worshopu ITAT 18th. Práce měla být zaměřená na hluboké neuronové sítě a v těch, jak jsem pochopil, se nepovedlo přinést průlomové výsledky. Zde student nese vinu částečně tím, že na radu školitele nepoužil nově doporučená data. Přesto hodnotím práci a její výsledky jako převážně úspěšné.

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 odrážkami).

Otázky:

Student by určitě měl dostat prostor vyjádřit se k originalitě rešeršní části. Tam by mě zajímalo, které sekce student skutečně psal na klávesnici, ať už z vlastní hlavy, nebo parafrázovaně ze zdrojů.

Student pro LSTM použil jednu a dvě skryté vrstvy, kde počty neuronů volil z předem definovaných hodnot. Zajímalo by mě, podle čeho tyto hodnoty volil. Také by mě zajímalo, jestli student zkoušel i více vrstev a případně o kolik byl výsledek horší (lepší)?

Použitý dataset, obsahující sedm emocí vyjádřených deseti herci, čítal celkem 535 mluvených záznamů. Jak si vysvětlit, že 535 není dělitelné čísly 7 a 10?

Hodnotící kritérium:	*Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):*
6. Celkové hodnocení	*70 (C)*

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:

Důležitý bod pro hodnocení této práce je posouzení toho, do jak velké míry je závažné ono zmiňované porušení citačních standardů v první kapitole. Nemám porovnání s ostatními DP na FIT, ani se zde uplatňovanými citačními zvyklostmi. V mých očích je to důvodem k hodnocení nejlépe za C, ale tento názor může ještě ovlivnit diskuze na obhajobě. Nutno zdůraznit, že praktická část je originálním dílem studenta, stejně jako popis algoritmů a experimentů v druhé části práce. Proto z mého pohledu práce přináší originální poznatky a není ve své hlavní části plagiátem, ale samozřejmě si zaslouží snížené hodnocení za nesprávné použití citací. Jako nesporný důkaz kvality této práce vnímám i to, že student měl na základě této práce příspěvek na konferenci s recenzním řízením.

Podpis oponenta práce: