

Posudek oponenta diplomové práce

Student: Obrtlík Petr, Bc.
Téma: Počítač jako inteligentní spoluhráč ve slovně-asociační hře Krycí jména (id 20086)
Oponent: Hradiš Michal, Ing., Ph.D., UPGM FIT VUT

- 1. Náročnost zadání** **obtížnější zadání**
Téma vyžaduje využití pokročilejších metod zpracování přirozeného jazyka. Zároveň se jedná o pionýrské téma, kde nejsou známa vhodná řešení.
- 2. Splnění požadavků zadání** **zadání splněno**
- 3. Rozsah technické zprávy** **je v obvyklém rozmezí**
- 4. Prezentací úroveň předložené práce** **65 b. (D)**
Struktura, jednoznačnost a pochopitelnost práce pro čtenáře jsou horší hlavně v částech popisujících samotné řešení.
Kapitola 2 je vhodná a smysluplná. Naopak Kapitola 3 se příliš nevztahuje k samotné práci a uvedené informace dále nejsou využity. Další kapitola o zpracování textu je relevantní, ale nepůsobí příliš přesvědčivě, protože obsahuje nepřesnosti a nevysvětluje metody důsledně. Je ale potřeba přihlídnout k vyšší náročnosti tématu. Kapitola 5 je v některých pasážích obtížně pochopitelná, protože zachází do implementačních aspektů místo toho, aby uceleně popsala řešení na úrovni metod a algoritmů na vhodné úrovni abstrakce. Velmi nevhodné je uvádění jmen použitých scriptů při popisu metod a jejich parametrů. Experimenty by měly být definovány důsledněji a jejich výsledky by měly být prezentovány takovou formou, aby bylo možné přímo vyčíst informace relevantní ke konkrétním poznatkům - rozhodně není potřeba prezentovat detailně všechny výsledky, když nejsou využity v nějaké interpretaci.
- 5. Formální úprava technické zprávy** **70 b. (C)**
Po jazykové stránce nemám k práci výraznější výhrady kromě určité zkratkovitosti a heslovitosti textu. V typografické úpravě bych ocenil méně odrážkových seznamů, správnou interpunkci za rovnicemi a lepší sazbu některých rovnic (4.1, 4.2, 5.1, 5.4, 5.5), lepší umístění rovnic a obrázků, konzistentnější podobu tabulek bez šedých částí a vektorové obrázky (např. u obrázku 5.1 je nízká rastrová podoba a nízká kvalita nečekaná).
- 6. Práce s literaturou** **70 b. (C)**
Práce se odkazuje celkem na 25 zdrojů, které jsou pro dané téma dostatečné. Text ale působí, že se student v metodách a celkově v tématu úplně neorientuje.
- 7. Realizační výstup** **80 b. (B)**
Student vytvořil funkční řešení náročnější úlohy ke které neexistují standardní a uznávaná řešení. Zároveň si vytvořil nástroje pro sběr dat a připravil datovou sadu, na které své řešení vyhodnotil. Bohužel jeho hloubka znalostí nebyla dostatečná, aby řešení bylo doopravdy kvalitní a experimenty poskytují jen základní informace.
- 8. Využitelnost výsledků**
Jedná se velmi zajímavou práci s velkým potenciálem i s ohledem na možný přínos ke stavu poznání jazykového modelování. Bylo by vhodné práci dále rozvíjet směrem k webové aplikaci, která by umožňovala sběr informací ze hry ve velkém rozsahu. Samotné řešení umělého hráče je spíše základní a nesystematické, ale může sloužit jako referenční řešení při dalším výzkumu.
- 9. Otázky k obhajobě**
 - Proč nevyužíváte skutečnosti, že vzdálenosti slov reprezentují pravděpodobnosti kookurence?
 - Informaci z předchozích kol hry nevyužíváte příliš efektivně. Máte se současnými zkušenostmi nějaké návrhy, jak by je bylo možné využít lépe?
 - Přinesla vám vaše práce nějaké poznatky ohledně vlastností embeddingů?
- 10. Souhrnné hodnocení** **75 b. dobře (C)**
Student zpracoval velmi zajímavé téma, pro které neexistují známá řešení. Vytvořil řešení, které vhodně využívá metody embeddingu slov a vytvořené řešení vhodně vyhodnotil. Bohužel ale jeho úroveň pochopení problému a metod byla dostatečná pouze pro základní řešení a neexperimentoval s různými variantami. Hodnocení snižují nedostatky textové části práce.

V Brně dne: 11. června 2018

.....
podpis