

Posudek bakalářské práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce Barbora Štěpánková
Název práce Generation of Czech Lyrics to Cover Songs
Rok odevzdání 2024
Studijní program Informatika
Studijní obor IPP

Autor posudku Mgr. Rudolf Rosa, Ph.D. Vedoucí
Pracoviště Ústav formální a aplikované lingvistiky

K celé práci

lepší OK horší nevyhovuje

	lepší	OK	horší	nevyhovuje
Obtížnost zadání	X			
Splnění zadání	X			
Rozsah práce <i>... textová i implementační část, zohlednění náročnosti</i>	X			

Práce splňuje zadání, které je přitom spíše náročnější, neboť automatické generování překladů písňových textů je úloha, která není dosud uspokojivě vyřešena; pro český jazyk jde pak pravděpodobně o první práci, která se vůbec pokouší tento problém řešit.

Práce je primárně výzkumná, vyžadovala od studentky seznámení se s mnoha pokročilými koncepty a postupy, a to jak z počítačové lingvistiky a strojového učení, tak z versologie a analýzy zpěvnosti (singability). Jádrem práce je metoda, která vychází z nejnovější existující odborné literatury věnující se tomuto problému a souvisejícím problémům, avšak nejedná se o pouhou reimplementaci. Vzhledem ke specifikům úlohy studentka kriticky zhodnotila existující literaturu, ze které vhodně vychází, ale existující postupy dále upravuje, vylepšuje, adaptuje a rozšiřuje, tak aby odpovídaly potřebám úlohy, navrhuje přitom řadu vlastních zcela nových metod a postupů. Studentka se také naučila pracovat s řadou nástrojů, některé z nich dále upravila a vylepšila.

Podstatnou součástí práce je i sběr a zpracování vhodných dat (pro úlohu neexistoval již připravený dataset), diskuze možných postupů a zkoumání různých možností řešení.

Zvláště ocenit bych chtěl kvalitní a detailní vyhodnocení provedených experimentů, zahrnující jak automatické vyhodnocení pomocí řady existujících, adaptovaných i nově navržených automatických metrik, tak na úrovni bakalářské práce nadprůměrně extenzivní manuální vyhodnocení za pomoci 25 dobrovolníků.

Generované výstupy pochopitelně nedosahují kvality lidských překladů (což ani se současným stavem poznání a technologií nelze očekávat), avšak vzhledem k tomu, že jde o první práci řešící tuto úlohu pro češtinu, lze výsledky považovat za velmi dobré. Ukazuje se tedy, že zvolená metoda řešení je vhodná, byť práce správně identifikuje i další neprozkoumané ale slibné možné postupy (zejména možnost postavit řešení na technologii strojového překladu namísto jazykového modelování).

Celkově se jedná o nadprůměrně rozsáhlou a kvalitní bakalářskou práci, v některých aspektech se již blíží spíše práci diplomové.

Textová část práce

lepší OK horší nevyhovuje

Formální úprava <i>... jazyková úroveň, typografická úroveň, citace</i>	X			
--	---	--	--	--

Struktura textu ... <i>kontext, cíle, analýza, návrh, vyhodnocení, úroveň detailu</i>	X			
Analýza	X			
Vývojová dokumentace		X		
Uživatelská dokumentace		X		
Text je kvalitní, dobře strukturovaný, dostatečně obsáhlý, dobře popisující všechny aspekty problému a jeho řešení. Práce vychází z desítek citovaných odborných publikací, což je u bakalářské práce značně nadprůměrné. Vývojová a uživatelská dokumentace je poměrně minimalistická, což je u výzkumného typu práce v pořádku.				

Implementační část práce

lepší OK horší nevyhovuje

Kvalita návrhu ... <i>architektura, struktury a algoritmy, použité technologie</i>		X		
Kvalita zpracování ... <i>jmenné konvence, formátování, komentáře, testování</i>	X			
Stabilita implementace		X		
Práce je primárně výzkumná, kvalita implementace jednotlivých experimentů odpovídá standardu běžnému pro tento typ prací. Oceňuji, že je přiložen i spustitelný skript, který lze použít pro spuštění hlavní části řešení, tedy generování českého textu pro vstupní anglický text. Zároveň je přiložena i grafická aplikace, která taktéž umožňuje spuštění programu.				

Celkové hodnocení Výborně
Práci navrhuji na zvláštní ocenění Ano

Datum

Podpis