Posudek bakalářské práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce	Tomáš Sourada
Název práce	Automatic inflection in Czech language
Rok odevzdání	2023
Studijní program	Informatika
Studijní obor	Umělá inteligence
Autor posudku	Mgr. Rudolf Rosa, Ph.D. Vedoucí
Pracoviště	Ústav formální a aplikované lingvistiky

K celé práci

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

Student implementoval několik metod automatického skloňování podstatných jmen (retrográdní přístup, rekurentní neuronová síť, neuronová síť architektury Transformer); zároveň zprovoznil a v některých případech i nově natrénoval řadu dalších nástrojů a modelů jako baselines. Kromě konverze a zpracování existujícího datasetu (MorfFlex) vytvořil i nový malý testovací dataset vyskloňovaných neologismů.

Velmi oceňuji velké množství kvalitně provedených a vyhodnocených experimentů. Autor zkouší mnohé varianty jednotlivých metod a mnohé hodnoty hyperparametrů, tak aby nalezl co nejlepší nastavení. Následně existenzivně a důsledně vyhodnocuje svůj postup na několika test setech a porovnává s mnohými dalšími systémy včetně současného state-of-the-art. Zcela nad rámec zadání a nad rámec toho, na co byl postup určen a vyladěn, trénuje a vyhodnocuje svůj systém i na datech pro jiné jazyky, tak aby se mohl skutečně zodpovědně porovnat s jinými existujícími systémy.

Dosažené výsledky předčí očekávání, neboť student v mnoha případech dosahuje srovnatelných výsledků či dokonce překonává dosavadní nejlepší systémy a výsledky. Výsledkem práce je tedy systém, který lze označit za pravděpodobně nejlepší aktuálně existující systém řešící danou úlohu: student v rámci své práce dosahuje state-of-the-art.

Přestože práce je primárně výzkumná, součástí odevzdané práce je i snadno použitelná knihovna pro automatické skloňování českých podstatných jmen.

Celkově tedy student zadání velmi dobře splnil, ba v mnoha aspektech dokonce překonal. Práce jako taková se kvalitou a rozsahem spíše blíží již práci magisterské.

Textová část práce

	lepší	OK	horší	nevyhovuje
Formální úprava *... jazyková úroveň, typografická úroveň, citace*	X			
Struktura textu *... kontext, cíle, analýza, návrh, vyhodnocení, úroveň detailu*	X			
Analýza	X			
Vývojová dokumentace		X		
Uživatelská dokumentace		X		

Text práce je velmi podrobný (asi 90 stran), student detailně diskutuje všechny důležité aspekty problému a jeho řešení. Přesto je práce dobře strukturovaná a text je kvalitní. Student vhodně vychází z existujícícho výzkumu, vhodně využívá existující nástroje a postupy (bibliografie odkazuje asi 40 prací), které však kriticky hodnotí a dle potřeby upravuje a rozvíjí, tak aby dosáhl dobrých výsledků na dané úloze.

Implementační část práce

	lepší	OK	horší	nevyhovuje
Kvalita návrhu *... architektura, struktury a algoritmy, použité technologie*		X		
Kvalita zpracování *... jmenné konvence, formátování, komentáře, testování*		X		
Stabilita implementace	X	X		

Práce je primárně výzkumná, jsou však přiloženy veškeré důležité zdrojové kódy a datasety. Kvalita zpracování je dobrá, odpovídá výzkumnému typu práce. Součástí práce je i snadno spustitelný a dobře a rychle fungující nástroj, což pro daný typ práce lze považovat za mírný nadstandard.

Celkové hodnocení Výborně

Práci navrhuji na zvláštní ocenění Ano

Datum Podpis