

Posudek oponenta bakalářské práce

Student: Prexta Dávid

Téma: Porovnávání anotačních nástrojů (id 20061)

Oponent: Otrusina Lubomír, Ing., UPGM FIT VUT

1. **Náročnost zadání** průměrně obtížné zadání
2. **Splnění požadavků zadání** zadání splněno

Zadání práce bylo splněno. Student se seznámil s nástroji pro automatické sémantické anotování textu a s existujícími řešeními pro srovnávání anotačních nástrojů a vyhodnocení jejich úspěšnosti. Na základě těchto znalostí navrhl a implementoval nástroj pro porovnávání anotačních nástrojů.
3. **Rozsah technické zprávy** je v obvyklém rozmezí

Rozsah technické zprávy je v obvyklém rozmezí. Všechny části technické zprávy jsou informačně bohaté a pro práci skutečně nezbytné.
4. **Prezentační úroveň předložené práce** 80 b. (B)

Prezentační úroveň technické zprávy je dobrá. Jednotlivé kapitoly na sebe dobře navazují a práce je dobře pochopitelná pro čtenáře. I přesto mám k práci několik připomínek. V úvodu práce jsou pouze informace o tom, co se nachází v jaké kapitole, zcela chybí jakýkoliv úvod do dané problematiky. Ten je popsán až v kapitole 2. Student občas v práci používá některé pojmy ještě před tím, než je vysvětlí (např. D2KB, Etyping). V práci se občas vyskytují některé nepřesné nebo neúplné informace. V sekci "Chybne rozpoznané entity" by bylo dobré uvést nějaké příklady.
5. **Formální úprava technické zprávy** 85 b. (B)

Formální úprava technické zprávy je na velmi dobré úrovni. Vytkl bych snad jen chybějící odkazy na použité technologie či nástroje (např. AIDA, ClueWeb, Gerbil Nif transfer) a fakt, že kapitola občas začíná hned podkapitolou bez jakéhokoliv úvodu. Výpis 7.1 na straně 34 je velmi nepřehledný.
6. **Práce s literaturou** 85 b. (B)

Výběr studijních pramenů je přiměřený danému tématu. Přesto bych zde ale očekával více vědeckých publikací věnujících se danému tématu. Student se v průběhu celé práce na příslušné studijní prameny vhodně odkazuje.
7. **Realizační výstup** 90 b. (A)

Realizační výstup práce je na velmi dobré úrovni. Student navrhl a implementoval nástroj, který dokáže porovnávat anotační nástroje na různých datových sadách. Oproti konkurenčním řešením má nástroj několik výhod.
8. **Využitelnost výsledků**

Jedná se o práci kompilačního charakteru. Výsledky práce jsou zcela jistě použitelné v praxi.
9. **Otázky k obhajobě**
 - V práci uvádíte, že po dokončení testů bylo zjištěno, že průměrně 80 % času identifikace chyb zabírá komunikace s MediaWiki API. Nebylo by tedy lepší ke zjišťování přesměrování používat stažený dump Wikipedie? Proč k tomu používáte MediaWiki API?
 - Jak Váš nástroj pozná, jestli budete číst uspořádaný nebo neuspořádaný soubor ve formátu NIF? Je potřeba mu to explicitně uvést a nebo to umí detekovat sám?
10. **Souhrnné hodnocení** 85 b. velmi dobře (B)

Prezentační úroveň i formální úprava technické zprávy jsou na velmi dobré úrovni. Realizační výstup práce je velmi zdařilý a má oproti konkurenčním nástrojům mnohé výhody. Výsledky práce jsou zcela jistě použitelné v praxi. Z těchto důvodů navrhuji hodnocení stupněm B.

Prohlášení: Uděluji VUT v Brně souhlas ke zveřejnění tohoto posudku v listinné i elektronické formě.

V Brně dne: 1. června 2017

.....
podpis

