

Posudek bakalářské práce

předložené na Matematicko-fyzikální fakultě
Univerzity Karlovy v Praze

posudek oponenta

Autor/ka: Tomáš Kypta

Název práce: Rozšířený HMM tagger a jeho aplikace na morfologické značkování
češtiny

Studijní program a obor: Informatika, programování

Rok odevzdání: 2007

Jméno a tituly vedoucího/opponenta: Mgr. Barbora Vidová Hladká, PhD.

Pracoviště: ÚFAL MFF UK

	excelentní	odpovídající	slabší	nevyhovující
Náročnost zadaného tématu	X			
Míra splnění zadání	X			
Struktura textové části práce		X		
Jazyková a typografická úroveň		X		
Analýza	X			
Vývojová dokumentace		X		
Uživatelská dokumentace		X		
Kvalita zpracování softwarové části	X			
Stabilita aplikace	X			

Nejvýznamnější klady:

V předložené práci autor prokazuje výborné zvládnutí tématu experimentálního charakteru. Písemné zpracování dokládá správné porozumění zadání a jeho důkladné zpracování včetně hodnotné evaluace výsledků.

Nejzávažnější nedostatky:

Hlavní nedostatky shledávám v uživatelské dokumentaci (kapitola 3) – jejich výčet následuje:

- nikde není uvedeno jméno programu
- uživatel, který neumí pracovat s programy v Java, nenajde návod, jak program spustit
- Část 3.1.1 Spouštění programu
 - o u přepínače '-di' není zřejmé, které soubory se z daného adresáře použijí
- část 3.1.3 Běh EM algoritmu
 - o není zřejmé, jestli EM algoritmus musí být vždy spuštěn při trénování či tagování. Pokud ne, s jakými hodnotami koeficientů se pracuje?
- část 3.1.10 Příklady vstupních parametrů
 - o příkazové řádky by bylo vhodné pro lepší porozumění doplnit komentáři

Pár poznámek k textu práce:

- V textu práce by bylo vhodné pro plynulejší čtení zaměnit pořadí kapitol 3 a 4.
- Časová složitost trénování a tagování by měla být začleněna přímo do části s vyhodnocením výsledků (a ne do přílohy, jak je tomu v textu).
- Aby mohly být prezentované výsledky rigorózně porovnány s jinými experimenty, je dobré u výčtu dat neuvádět pouze kvantitativní charakteristiky, ale i zdroje dat - v případě PZK míněno uvést adresáře, z kterých data pocházejí.
- Na několika místech chybí čárky ve větách (např. 2.1.4, bod 1 za „Otázkou“).
- Autor systematicky píše „viz.“ namísto „viz“ (rozkazovacího způsobu slovesa „vidět“).

Další poznámky:

Uvedené nedostatky se týkají výhradně dokumentace. Nijak nesnižují kvalitu odvedené práce.

	výborně	velmi dobře	dobře	neprospěš/a
Návrh známky	X			

Datum: 3.září 2007

Podpis: