

Posudek bakalářské práce

předložené na Matematicko-fyzikální fakultě
Univerzity Karlovy v Praze

X posudek vedoucího □ posudek oponenta

Autor/ka: Martin Kirschner

Název práce: Konstrukce sémantického slovníku z neanotovaných dat

Studijní program a obor: Informatika, Ústav formální a aplikované lingvistiky

Rok odevzdání: 2008

Jméno a tituly vedoucího/oponenta: RNDr. Jiří Semcecký, PhD.

Pracoviště: ÚFAL MFF UK

	e x c e l e n t n í	o d p o v í d a j í c í	s l a b š í	n e v y h o v u j í c í
Náročnost zadaného tématu	X			
Míra splnění zadání			X	
Rozsah práce			X	
Struktura textové části práce		X		
Analýza		X	X	
Vývojová dokumentace		X		
Uživatelská dokumentace		X		
Jazyková a typografická úroveň		X		X
Návrh a design implementace		X		
Kvalita zpracování softwarové části		X		X
Stabilita aplikace		X		

Nejvýznamnější klady:

Použití zajímavých metod strojového učení.
Návrh a implementace obecného datového formátu.
Kategorizace sémantických vztahů.
Dobře napsaná webová aplikace pro zobrazování sémantických vztahů.

Nejzávažnější nedostatky:

Práce obsahuje nepřesné formulace a velké množství překlepů a často se objevují termíny bez předchozí definice. Např.:

v kapitole 3.1 “význam každého slova je tak zastoupen vektorem v sémantickém prostoru” - zde termín sémantický prostor není nijak vysvětlen, ani v daném kontextu nedává smysl, autor zřejmě myslel prostor dokumentů.

3.1: “šum v datech tady odpovídá výkladům významů slov, ...” je podivné

3.2.1: “výber kontextu... ne tak rozsáhlý, aby ... nezabíral výpočetní kapacitu na úkor výstižnějších slov” - nejsem si jistý, jak to autor myslel, pravděpodobně to není skutečný důvod

3.2.1 “... slova, jichž slou zkoumaná slova součástí” - těžko se představuje, jak může být slovo součástí slova.

...

Je také škoda, že autor uvádí v práci velmi málo detailů ohledně metody, kterou použil (např. jsem ani nenalezl informaci, jakou implementaci LSA použil) a velmi málo konkrétních výsledků (téměř žádné konkrétní příklady z dat).

Další poznámky:

Domnívám se, že se jedná o zajímavé téma a přístup, na práci je však bohužel vidět, že byla odevzdána “narychlo” a je v ní řada nedodělaných a slabých míst.

	v ý b o r n ě	v e l m i d o b ř e	d o b ř e	n e p r o s p ě l / a
Návrh známky		X		

Datum: 17. 6. 2008

Podpis:

