

Posudek bakalářské práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce Jiří Kolář
Název práce Utilization of latent semantic analysis in virtual screening
Rok odevzdání 2017
Studijní program Informatika **Studijní obor** Obecná informatika

Autor posudku Petr Škoda **Role** Oponent
Pracoviště KSI

Prosím vyplňte hodnocení křížkem u každého kritéria. Hodnocení *OK* označuje práci, která kritérium vhodným způsobem splňuje. Hodnocení *lepší* a *horší* označují splnění nad a pod rámec obvyklý pro bakalářskou práci, hodnocení *nevyhovuje* označuje práci, která by neměla být obhájena. Hodnocení v případě potřeby doplňte komentářem. Komentář prosím doplňte všude, kde je hodnocení jiné než *OK*.

K celé práci	lepší	OK	horší	nevyhovuje
Obtížnost zadání	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Splnění zadání	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Rozsah práce ... <i>textová i implementační část, zohlednění náročnosti</i>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Komentář Click here to enter text.				

Textová část práce	lepší	OK	horší	nevyhovuje
Formální úprava ... <i>jazyková úroveň, typografická úroveň, citace</i>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Struktura textu ... <i>kontext, cíle, analýza, návrh, vyhodnocení, úroveň detailu</i>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Analýza	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vývojová dokumentace	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Uživatelská dokumentace	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Komentář Významná část textu je věnována popisu deskriptorů používaných pro metody virtuálního screeningu. Tato část textu ukazuje autorův přehled v dané doméně. Celkově tato část působí spíše přehledově, neb některé popsané metody se dále v práci již nepoužívají. Opačným dojmem bohužel působí popis navržené metody a implementace. Z popisu metody není zřejmé, jak metoda funguje – zejména pak pro druhou variantu s použitím vektorů vlastností pro fragmenty. Zde by mohl být popis trochu více detailní případně s příkladem. Ačkoliv se nejedná o práci implementační popis vzniklého programu by mohl být poněkud rozsáhlejší a úplnější. Programátorská dokumentace je stručná a nesouvislá (popis použitých knihoven je v kapitole "3 Model - general information"). Přiložený diagram sice poskytuje základní přehled o modulech aplikace, nicméně je na různých úrovních abstrakce a detailu. Dalším problémem je v zásadě absence uživatelské dokumentace. Její základ je možné najít v README.md souboru (mimo textovou část práce). Uživatel se zde dočte informaci o použitelných parametrech, ale již se třeba nedozví, v jakém formátu mají být vstupní data. Zde bych ocenil lepší popis ideálně odkaz na nějaké příklady konfigurací a vstupů.

Implementační část práce

	lepší	OK	horší	nevyhovuje
Kvalita návrhu ... <i>architektura, struktury a algoritmy, použité technologie</i>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Kvalita zpracování ... <i>jmenné konvence, formátování, komentáře, testování</i>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Stabilita implementace	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Komentář Některé zdrojové soubory obsahují zakomentovaný kód. Extrémním případem je soubor "configGenerator.py" jenž obsahuje v zásadě pouze komentář. Výsledný zdrojový kód tak působí poněkud zanedbaným/uspěchaným dojmem. Dalším příkladem je použití různých konvencí pro pojmenování proměnných a to jak pomocí lower-CamelCase tak i snake_case i v rámci jednoho souboru. Za pozitivní by šlo naopak označit na první pohled rozumné pojmenování metod. Pozitivním aspektem je umístění výsledné implementace na github, což otevírá potenciál pro další využití a modifikace. Bohužel horší prezentace a nedostatečná dokumentace bude dle mého názoru výrazně odrazovat případné zájemce o tento projekt.

Celkové hodnocení Velmi dobře

Práci navrhuji na zvláštní ocenění Ne

Datum 27. srpna 2017

Podpis