

Hodnocení vedoucího bakalářské práce

Autor/autorka práce: Michel Soběhart

Název práce: Agregátor novinových zpráv

Cíl práce

Cílem práce bylo vytvoření programu pro monitorování novinových zpráv pomocí RSS, jejich stahování a shlukování.

Spolupráce

Student pravidelně docházel na konzultace a vždy byl připraven. Jedinou výtku mám k občasné pasivitě studenta při návrhu řešení. Celkově byla spolupráce na velmi dobré úrovni.

Text práce

Po vizuální stránce působí práce dobře. Výhrady mám k obrázkům v kapitolách 3 a 4 (např. různé měřítko u vývojových diagramů). Zvláštní je také prázdné místo na str. 9. Stylisticky je práce průměrná - některé obraty jsou zvláštní, ale pochopitelné.

Text dostatečně popisuje jednotlivé technologie používané v novinových agregátorech i navržený systém. V teoretické části mi chybí ucelený popis některého existujícího agregátoru. Některé části jsou popsány povrchně, např. shlukování. Text obsahuje některé nepřesnosti, např. Čebyševova vzdálenost (Čebyševova), záměna slova objekt a příznak na str. 8, definice lemmatizace a stemmingu na str. 10-11 atd. Tyto nepřesnosti zásadně nesnižují pochopitelnost práce.

Řešení

Oba programy (stahování zpráv a agregace) fungují stabilně. Překlad i spuštění jsou bezproblémové pomocí přiložených skriptů. Úspěšnost agregace je na pohled velmi vysoká a pravděpodobně odpovídá úspěšnosti udané v práci (96%).

Kód je strukturovaný a komentovaný. Návrh tříd a jejich metod je logický, i když ne zcela optimální.

V práci mi celkově chybí větší zaměření na experimenty, např. s různými příznaky, metrikami podobnosti a parametry agregace.

Otázky k obhajobě

1. Obrázek 4.5 naznačuje, že při zpracování dat pokaždé „trénujete příznakový model“, přitom na str. 36 máte uložen natrénovaný model. Jak to tedy je?

Závěr

Práce splňuje všechny body zadání.

Navrhuji hodnocení známkou **velmi dobře** a práci doporučuji k obhajobě.

V Plzni 28.5.2014


Ing. Michal Konkol