

Posudek oponenta bakalářské práce

Student: Kohout Tomáš
Téma: Programovací jazyk Scala a jeho využití pro analýzu dat (id 21604)
Oponent: Bartík Vladimír, Ing., Ph.D., UIFS FIT VUT

- Náročnost zadání** průměrně obtížné zadání
Cílem této bakalářské práce bylo důkladně seznámení s programovacím jazykem Scala a zhodnocení jeho použitelnosti s ohledem na analýzu rozsáhlých dat a s využitím těchto znalostí provést případovou studii a odpovídající experimenty. Zadání hodnotím jako vhodné pro bakalářskou práci.
- Splnění požadavků zadání** zadání splněno
Všechny body zadání byly splněny. Student se kromě samotné analýzy dat zaměřil také na problém paralelního zpracování, což lze vnímat jako rozšíření oproti zadání.
- Rozsah technické zprávy** je v obvyklém rozmezí
Rozsah technické zprávy je okolo 60 normostran, je tedy v obvyklém rozmezí.
- Prezentační úroveň předložené práce** 90 b. (A)
Logická struktura práce je bez větších nedostatků, snad jen kapitola 3 popisující případovou studii bych rozdělil na teoretickou a praktickou část. Po obsahové stránce je uvedeno vše potřebné a bez faktických chyb. Text je celkově velmi dobře srozumitelný.
- Formální úprava technické zprávy** 76 b. (C)
Typografická úroveň je na velmi dobré úrovni, vytknul bych pouze lehce rozmazané obrázky a ojedinělé výskyty textu přetékajícího přes okraj. Jazyková úroveň textu je spíše průměrná, text obsahuje především v první polovině řadu překlepů a jazykových chyb. Jejich množství ale zásadním způsobem neovlivňuje srozumitelnost textu.
- Práce s literaturou** 88 b. (B)
K použití literatury nemám výhrady, snad až ne jeden odkaz na wikipedii v seznamu literatury. Převzaté výsledky jsou jasně odlišeny od vlastních, k porušení citační etiky tak určitě nedošlo.
- Realizační výstup** 90 b. (A)
Realizační výstup představuje klasifikátor e-mailů, který rozpoznává spam od běžných e-mailů. Zahrnuje celkem 4 různé klasifikátory, jejichž úspěšnost je podrobně vyhodnocena v experimentální části. V implementaci bylo také využito souběžné zpracování, jehož podpora je v jazyce Scala lepší než v ostatních zmiňovaných jazycích.
- Využitelnost výsledků**
Práce je spíše kompilačního charakteru, přínos vidím v podrobném srovnání jazyků využívaných pro analýzu dat. Práce může také čtenáři sloužit pro základní seznámení s jazykem Scala.
- Otázky k obhajobě**
 - Uvažoval jste o využití váhovacích metod (nepř. TF-IDF) ve vaší práci?
 - Popište stručně, jak pracuje knihovna SymSpell pro opravu pravopisných chyb v textu.
 - Proved'te stručné shrnutí hlavních výhod jazyka Scala oproti Pythonu nebo R.
- Souhrnné hodnocení** 85 b. velmi dobře (B)
Jde o kvalitní bakalářskou práci s mírnými nedostatky v technické zprávě. Realizační výstup je nadprůměrný. Hodnotím proto tuto bakalářskou práci stupněm B (velmi dobře).

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

V Brně dne: 29. května 2019

.....
podpis