

Posudek oponenta diplomové práce

Autor/autorka práce: **Miroslav Vozábal**

Název práce: **Tools and Methods for Big Data Analysis**

Práce je na první pohled nadstandardně dlouhá a čítá přes 100 stran samotného textu plus přílohy. Přesto, až na pár výjimek, z nichž jedna je součástí dotazu, jsou informace v práci relevantní a autorovi myšlenky neodbíhají od tématu. Práce tak jde především do velké hloubky dané problematiky. Teoretická složka práce na počet stran výrazně převyšuje část popisující realizaci samotného řešení. S ohledem však na komplikovanost dané úlohy, která byla postavena především na analýze dostupných možností, je toto pochopitelné.

Výsledky práce považuji za zcela postačující. Na první pohled se zdá, že metody na zpracování Big Data jsou pro účely relativně malých a dobře strukturovaných EEG dat zbytečné až nepoužitelné. Diplomant však přesto k tomuto našel cestu a konečně měření přináší kladné výsledky v otázce rychlosti zpracování EEG dat. Navíc autor předvedl, že je schopen se zorientovat napříč mnoha technologiemi ve složité problematice a implementovat vlastní rozšíření. Rovněž část zhodnocující důvody proč se „běžné“ Big Data modely jako map-reduce nehodí na EEG data je přínosná. Hlavní výtku mám k porovnání dle mého neporovnatelných databází, které působí samoučelně. Mám pocit, že vybraná EEGBase databáze byla vybrána již před samotnou analýzou. K tomuto se vracím v dotazech k práci.

Po formální stránce lze práci máloco vytknout. Obsahuje naprosté minimum překlepů, úroveň anglického jazyka je vysoká, obrázky jsou přehledné a dobře čitelné a všechny náležitosti jako seznamy zkratk, obrázků a tabulek jsou k dispozici. Kde může autor doprovodit text obrázkem či tabulkou, činí tak a text se tím stává ještě srozumitelnějším. Občas jsem sice měl při čtení pocit, že právě čtená část by měla být zmíněna o pár stránek dříve a naopak, ale vzhledem k rozsahu je toto přijatelné a čtenář nenarazí přímo na místo, kde by mu unikal kontext.

Práce s literaturou je rovněž výborná a 22 zmíněných zdrojů obsahuje velké množství knih a odborných časopiseckých publikací.

Autor splnil zadání v plném rozsahu. Samotným výsledkem práce je spíše paralelizace existujících procesů zpracovávajících EEG data nežli aplikace „běžných“ Big Data metod/modelů na EEG data. Na druhou stranu toto odráží fakt, že autor své řešení cílil předně na praktičnost. Tomu odpovídají i možnosti budoucího rozšíření.

Dotazy k práci

V práci srovnáváte na první pohled nesrovnatelné databáze. Proč padnul výběr právě na tyto reprezentanty namísto srovnání databází v EEG doméně?

V práci jdete občas do takové hloubky, že zmíněná fakta poté působí nadbytečně. Příkladem budiž tabulka cen Matlab licencí (Table 11, str. 80). Jak je konkrétně tato pro práci přínosná?

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

V Plzni 24. 8. 2016

Ing. Václav Papež

SOUHLASÍ
3 ORIGINALÉM

Západočeská univerzita v Plzni
Fakulta aplikovaných věd
Katedra informatiky a výpočetní techniky