

Posudek práce

předložené na Přírodovědecké fakultě JU

☒ posudek vedoucího
☒ bakalářské práce

☐ posudek oponenta
☐ diplomové práce

Autor/ka: Jiří Burda
Název práce: Zpracování dat v prostředí Hadoop
Studijní program a obor: Aplikovaná informatika
Rok odevzdání: 2018

Jméno a tituly vedoucího/opponenta: doc. Ing. Ladislav Beránek, CSc.
Pracoviště: Ústav aplikované informatiky
Kontaktní e-mail: beranek@ef.jcu.cz

Odborná úroveň práce:

☐ vynikající ☐ velmi dobrá ☐ průměrná ☒ podprůměrná ☐ nevyhovující

Věcné chyby:

☐ téměř žádné ☒ vzhledem k rozsahu přiměřený počet ☐ méně podstatné četné ☐ závažné

Výsledky:

☐ originální ☒ původní i převzaté ☐ netriviální kompilace ☐ citované z literatury ☐ opsané

Rozsah práce:

☐ veliký ☐ standardní ☒ dostatečný ☐ nedostatečný

Grafická, jazyková a formální úroveň:

☐ vynikající ☐ velmi dobrá ☒ průměrná ☐ podprůměrná ☐ nevyhovující

Tiskové chyby:

☐ téměř žádné ☒ vzhledem k rozsahu a tématu přiměřený počet ☐ četné

Celková úroveň práce:

☐ vynikající ☐ velmi dobrá ☐ průměrná ☒ podprůměrná ☐ nevyhovující

Slovní vyjádření, komentáře a připomínky vedoucího/oponenta:

Práce se zabývá frameworkem Hadoop a možnostmi zpracování dat pomocí tohoto frameworku. Přitom musel využít i další návazné aplikace jako je Apache Flume nebo Apache Hive a další. Formulace některých vět je někde poměrně neobratná nebo komplikovaná, naopak některé příklady skriptů popisované autorem jsou zase poměrně jednoduché, například popisu založení účtu na Twitteru si mohl autor odpustit.

V zásadě se však domnívám, že autor splnil zadání, i když jsem měl původně představu, že aplikace bude rozsáhlejší. Nicméně musím konstatovat, že se autor odvedl poměrně velký kus práce a strávil nad prací hodně času. Vypořádání se s těmito technologiemi není jednoduché, vyžaduje prostudování mnoha zdrojů.

Doporučuji práci k obhajobě a hodnotím známkou dobře.

Případné otázky při obhajobě a náměty do diskuze:

Pro vaše aplikace byste mohl využít i komponentu Hbase. Mohl byste stručně popsat rozdíl mezi komponentami pro zpracování velkých objemů dat Hadoop Distributed File System (HDFS) a Hbase (Hadoop database)?

Práci

☒ doporučuji

☐ nedoporučuji

uznat jako diplomovou/bakalářskou.

Navrhuji hodnocení stupněm:

☐ výborně ☐ velmi dobře ☒ dobře ☐ neprospěl/a

Místo, datum a podpis vedoucího/oponenta:

V Českých Budějovicích 14.5.2018

