

POSUDEK OPONENTA DIPLOMOVÉ PRÁCE

Jméno studenta: Jaromír Hnik
Název práce: Hadoop a jeho využití pro Data Science
Autor posudku: Pavel Kříž
Cíl práce: Analýza problematiky a návrh řešení implementace pro analýzu dat za použití strojového učení a vybraných knihoven.

Povinná kritéria hodnocení práce	Stupeň hodnocení					
	A	B	C	D	E	F
Práce svým zaměřením odpovídá studovanému oboru	X					
Vymezení cíle a jeho naplnění	X					
Zpracování teoretických aspektů tématu			X			
Zpracování praktických aspektů tématu		X				
Adekvátnost použitých metod	X					
Hloubka a správnost provedené analýzy		X				
Práce s literaturou	X					
Logická stavba a členění práce		X				
Jazyková a terminologická úroveň			X			
Formální úprava a náležitosti práce	X					
Vlastní přínos studenta		X				
Využitelnost výsledků práce v teorii (praxi)				X		

Vyjádření k výsledku anti-plagiátorské kontroly:

Práce v rámci automatické kontroly anti-plagiátorským systémem nevykazuje žádné nedostatky.

Díličí připomínky a náměty:

- Terminologické nepřesnosti, např.
 - „Primární klíč je entita...“, kdy pojem entita má v modelování databázových systémů svůj specifický význam,
 - „pracuje s k -nejbližším prvkem“ – jedná se o k nejbližších sousedů,
- absence elektronické přílohy.

Celkové posouzení práce a zdůvodnění výsledné známky:

Student Jaromír Hnik předkládá diplomovou práci v rozsahu 63 stran textu, ve které se věnoval problematice zpracování tzv. *Big Data* a to konkrétně pomocí nástrojů využívajících *Apache Hadoop*. Teoretická část je věnována historii databázových systémů, pojmu *Big Data*, popisu *Apache Hadoop* a vědnímu oboru *Data Science*. Praktická část se zaměřuje na použití *Hadoopu* na vybrané úlohy.

Na práci kladně hodnotím výběr tématu, které je velmi aktuální. Zpracování je na dobré úrovni. Vzhledem k šíři zadání mohlo být některým dílčím kapitolám věnováno méně prostoru (historie DBS) a jiným naopak více (aplikace v kap. 5). Absence elektronické přílohy u tohoto typu práce výrazně snižuje její využitelnost. Čtenáři pak nezbyváá, než „opisovat“ skripty a vymýšlet si vlastní ukázková data.

Cíl práce považuji za splněný.

Otázky k obhajobě:

- Doporučil byste případnému zájemci o tuto problematiku nějaké online kurzy, tutoriály nebo online „laboratorní“ prostředí?

Práci doporučuji k obhajobě.

Navržená výsledná známka: B

V Hradci Králové, dne 3. září 2019

podpis