



POSUDEK VEDOUCÍHO DIPLOMOVÉ PRÁCE

Jméno studenta: Jaromír Hnik

Název práce: Hadoop a jeho využití pro Data Science

Autor posudku: Ing. Karel Mls, Ph.D.

Cíl práce: Analýza problematiky a návrh řešení implementace pro analýzu dat za použití strojového učení a vybraných knihoven.

Povinná kritéria hodnocení práce	Stupeň hodnocení (známka)					
	A	B	C	D	E	F
Práce svým zaměřením odpovídá studovanému oboru	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vymezení cíle a jeho naplnění	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zpracování teoretických aspektů tématu	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zpracování praktických aspektů tématu	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Adekvátnost použitých metod, způsob jejich použití	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Hloubka a správnost provedené analýzy	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Práce s literaturou	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Logická stavba a členění práce	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Jazyková a terminologická úroveň	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Formální úprava a náležitosti práce	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vlastní přínos studenta	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Využitelnost výsledků práce v teorii (v praxi)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Vyjádření k výsledku anti-plagiátorské kontroly

Míra shody práce s jinými evidovanými texty dle Odevzdej.cz je 0%, jedna identifikovaná shoda – definice, která byla citována přímo, včetně uvedení původního zdroje. Práce je originální.

Dílejší připomínky a náměty:

V textu přes značné úsilí přetrvalo menší množství pravopisných nedostatků (paměťové čipy... nahrazovali, občasná problémy s interpunkcí ve větách, používání zkratk a číslovek – např., tzn., 2 logické stavy). Po formální stránce nepůsobí dobře ani osamocené podkapitoly (2.2.1, 2.4.1, 3.1.1, ...).

Celkové posouzení práce a zdůvodnění výsledné známky:

Práce se zabývá problematikou efektivního zpracování velkých objemů dat z pohledu strojového učení a Data Science.

Autor v práci zúročil dlouhodobý praktický zájem o problematiku a zkušenosti s frameworkem Hadoop.

V teoretické části autor postupně představuje pojmy jako databáze, OLAP, OLTP, Big Data, dílčí podsystémy frameworku Hadoop, Data Science, strojové učení.

V praktické části navazuje na teoretické okruhy a popisuje implementaci frameworku Hadoop, jazyka Python a příslušných knihoven na příkladu typických úloh při zpracování velkých dat. Konkrétní příklad představuje úlohu doporučování obsahu na základě podobnosti.

Autor pracoval samostatně, pravidelně konzultoval a doporučení pečlivě implementoval do práce.

Práce splňuje po formální i obsahové stránce požadavky kladené na diplomové práce, je přínosná a přímo využitelná při řešení praktických projektů zaměřených na zpracování Big Data, včetně projektů řešených na FIM UHK.

Na základě uvedeného práci doporučuji k obhajobě a navrhuji hodnocení stupněm A.

Otázky k obhajobě:

nemám

Práci doporučuji k obhajobě.

Navržená výsledná známka: A

V Hradci Králové, dne 4. září 2019

podpis