

POSUDEK OPONENTA DIPLOMOVÉ PRÁCE

Student: Hanzlík Roman

Oponent: prof. Ing. Jan Platoš, Ph.D.

Studijní program: **Inženýrská informatika**

Studijní obor: **Informační technologie**

Akademický rok: **2020/2021**

Téma diplomové práce: **Data science v prostředí Apache Spark**

Hodnocení práce:

Hodnocení může přesahovat na další strany.

Práce pojednává o problematice Data Science a jeho aplikace v prostředí Apache Spark. Práce je rozdělena do dvou částí. První část – teoretická, je velmi důkladně sepsaná a obsahuje historii zpracování znalostí a dat od pravěku do současnosti. V rámci popisu se sice vyskytují drobné opakované části, ale jinak se mi zdá práce vyčerpávající.

Druhá část, praktická, obsahuje kromě popisu prostředí Apache Spark i reálné příklady s výpisem zdrojového kódu. Tato část je na nižší úrovni, než je část teoretická.

Student se velmi dobře chopil zpracování obecné rešerše, která je dobře vystavěná a všechny potřebné náležitosti jsou velmi kvalitní. PO formální stránce je práce na velmi dobré úrovni v rámci textu jsem nezaznamenal gramatické nebo jiné problémy. Struktura kapitol i jednotlivých částí je také velmi dobrá.

Praktická stránka není nijak rozsáhlá, výpisy zdrojového kódu zabírají většinu místa, ale v rámci praktických ukázek jsou i drobná srovnání a výsledné výpisy. Nejsem si jist, zda tato část odpovídá čtvrtému bodu zadání, ale v minimální míře se s tímto lze smířit.

Diplomant vytvořil velmi dobrou kompilaci a souhrn problematiky Data Science, která je velmi dobře sepsaná a mohla by klidně sloužit jako základ pro studium této problematiky. Sice se nejedná o nové informace, ale jsou velmi dobře kompilovány. Praktická část je po teoretické stránce také velmi dobrá, ale praktická část je velmi malá a je spíše ukázkou použití.

Celkově práci hodnotím jako velmi dobrou. Praktická část mohla být na vyšší úrovni, ale po teoretické stránce je práce opravdu velmi dobrá.

Celkové hodnocení práce:

Známku uvede oponent dle svého uvážení dle klasifikační stupnice ECTS:

A – výborně, B – velmi dobře, C – dobře, D – uspokojivě, E – dostatečně, F – nedostatečně.

Stupeň F znamená též „nedoporučuji práci k obhajobě“.

**Předloženou diplomovou práci doporučuji k obhajobě a navrhuji hodnocení
B - velmi dobře.**

**V případě hodnocení stupněm „F – nedostatečně“ uveďte do připomínek a slovního vyjádření
hlavní nedostatky práce a důvody tohoto hodnocení.**

Datum 29. 8. 2021

Podpis oponenta diplomové práce