

Posudek oponenta bakalářské práce

Student: Stoika Anastasiia
Téma: Případová studie na dolování z dat v jazyce Python (id 22015)
Oponent: Burgetová Ivana, Ing., Ph.D., UIFS FIT VUT

1. **Náročnost zadání** průměrně obtížné zadání
Jedná se o zadání s náročnější studijní částí, která vyžaduje studium problematiky dolování z dat a možných prostředků jazyka Python pro tuto oblast, a méně náročnou implementační částí (implementace skriptů, které provedou analýzu konkrétní datové sady).
Celkově se tedy jedná o průměrně obtížné zadání.
2. **Splnění požadavků zadání** zadání splněno
3. **Rozsah technické zprávy** je v obvyklém rozmezí
4. **Prezentační úroveň předložené práce** 60 b. (D)
Logická struktura předložené technické zprávy není ideální. Zpráva obsahuje dvě teoretické kapitoly, které obě popisují různé aspekty procesu získávání znalostí z dat. Tím pádem jsou některé pojmy vysvětlovány opakovaně. Horší orientace v technické zprávě je také způsobena nekonzistentním využíváním nadpisů 3 úrovně. Studentce se také nepodařilo odstranit všechny poznámky, které do finální verze technické zprávy nepatří (str. 5 a 14).
5. **Formální úprava technické zprávy** 58 b. (E)
Jazyková stránka předložené technické zprávy je špatná. Zpráva obsahuje mnoho chyb a nelogických slovních spojení. Z typografického hlediska se jedná o kvalitní práci.
6. **Práce s literaturou** 70 b. (C)
Výběr studijních pramenů odpovídá tématu práce. Jako kvalitu však poněkud kazí odkazy na wikipedii a dvě nekompletní citace (29, 34). Převzaté prvky jsou odlišeny od vlastních výsledků a úvah.
7. **Realizační výstup** 70 b. (C)
Realizačním výstupem této bakalářské práce je kód v jazyce Python, který provádí počáteční analýzu vybrané datové sady a následně vyhledání odlehklých hodnot v této datové sadě s využitím metod shlukové analýzy a klasifikačních metod. Výsledky dosažené s využitím shlukovacích metod jsou srovnatelné s výsledky publikovanými v případové studii, která byla vybrána pro tuto bakalářskou práci. Všechny prostředky, které studentka využila pro řešení této diplomové práce, byly použity v souladu s licenčními podmínkami.
8. **Využitelnost výsledků**
Tato bakalářská práce nepřináší žádné nové poznatky, pouze ukazuje možnosti analýzy dat prostředky jazyka Python.
9. **Otázky k obhajobě**
 - Vysvětlíte, jakým způsobem byly kategorické atributy *ID* a *Prod* transformovány pro klasifikační model *Gaussian Naive Bayes*. Neovlivnila tato transformace výsledný klasifikační model (např. vytvořením určitých nových vztahů mezi jednotlivými hodnotami atributů)?
10. **Souhrnné hodnocení** 68 b. uspokojivě (D)
Realizační výstup této diplomové práce odpovídá požadavkům zadání a splňuje všechny body zadání. Bohužel technická zpráva vykazuje mnoho nedostatků, proto navrhuji hodnocení stupněm D.

Prohlášení: Uděluji VUT v Brně souhlas ke zveřejnění tohoto posudku v listinné i elektronické formě.

V Brně dne: 29. května 2019

.....
podpis