

Posudek oponenta diplomové práce

Student: Bartoš Stanislav, Bc.
Téma: Využití dolování dat pro identifikaci plateb (id 22880)
Oponent: Burget Radek, Ing., Ph.D., UIFS FIT VUT

1. **Náročnost zadání** průměrně obtížné zadání
2. **Splnění požadavků zadání** zadání splněno
Všechny body zadání byly splněny a jsou zdokumentovány v technické zprávě.
3. **Rozsah technické zprávy** je v obvyklém rozmezí
Technická zpráva se svým rozsahem pohybuje na dolní hranici rozmezí obvyklého pro diplomové práce.
4. **Prezentační úroveň předložené práce** 76 b. (C)
Technická zpráva pokrývá jak teoretickou část zaměřenou zejména obecně na metody získávání znalostí z dat a na související dostupné prostředky v jazyce Python, tak i praktickou část řešící konkrétní aplikaci vybraných metod pro párování plateb. Struktura technické zprávy je srozumitelná a jednotlivé části jsou dostatečně podrobně zpracované. Drobné výhrady mám ke kapitole 4 - Návrh, která je naopak velmi stručná a očekával bych zde podrobnější popis vstupních dat a zejména návrh konkrétního způsobu využití klasifikačních metod pro danou úlohu včetně volby rysů pro klasifikaci. Tyto informace jsou však postupně doplněny v popisu experimentálního vyhodnocení.
5. **Formální úprava technické zprávy** 75 b. (C)
Po typografické i jazykové stránce je práce standardní pouze s drobnými nedostatky, které zahrnují např. různou velikost písma v obrázcích a tabulkách (např. tab. 2.1) nebo poněkud neúčelně využitou 'heat map' pro matice záměn, kde se vzhledem k velkým rozdílům hodnot vyskytují pouze dvě rozpoznatelné barvy. Z jazykového pohledu bych očekával poněkud strukturovanější popis navržených metod a postupů, text práce často spíše připomíná volné vyprávění (např. kap. 5.7).
6. **Práce s literaturou** 72 b. (C)
Seznam použité literatury je přiměřeně rozsáhlý o obsahuje zdroje relevantní pro danou oblast. Zdroje nejsou v textu citovány příliš často, autor se ve většině případů spokojí s výčtem zdrojů na začátku kapitoly a např. celá kapitola 2.2 už v textu na zdroje nedokazuje.
7. **Realizační výstup** 80 b. (B)
Realizačním výstupem je zejména sada vytvořených nástrojů umožňující experimentální vyhodnocení různých klasifikátorů dostupných v jazyce Python na zdrojových datech společnost Roger. Dále pak student integroval vybrané metody přímo do platebního systému firmy Roger tak, aby se zjednodušilo ruční přiřazování plateb k fakturám. Všechny části řešení mi student demonstroval, jsou funkční a dávají výsledky popsané v technické zprávě.
8. **Využitelnost výsledků**
Vytvořené nástroje byly nasazeny do platebního systému Roger a budou zřejmě využívány v praxi.
9. **Otázky k obhajobě**
 - Mohl byste stručně shrnout, jaké vlastnosti přichozích plateb ve výsledném řešení využíváte pro jejich identifikaci?
10. **Souhrnné hodnocení** 76 b. dobře (C)
Pan Bartoš poměrně detailně prostudoval relevantní metody získávání znalostí z dat a navrhl způsob jejich využití pro identifikaci plateb. Vytvořil experimentální nástroje pro ověření jednotlivých metod na základě výsledků experimentů zvolil řešení, které bude zřejmě využíváno v praxi. Přes některé výhrady k technické zprávě považuji celkově tuto práci za standardní a navrhuji hodnocení stupněm C.

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

V Brně dne: 29. června 2020

Burget Radek, Ing., Ph.D.
oponent