

Posudek oponenta diplomové práce

Student: Fojtík David, Bc.
Téma: Analýza firemních dat o vytiženosti služeb s využitím dolování dat (id 22153)
Oponent: Burgetová Ivana, Ing., Ph.D., UIFS FIT VUT

1. **Náročnost zadání** průměrně obtížné zadání
2. **Splnění požadavků zadání** zadání splněno s drobnými výhradami

Moje výhrady se týkají bodu 1, kdy úkolem studenta bylo seznámit se s problematikou získávání znalostí z databází. Z technické zprávy jasně vyplývá, že toto seznámení se s problematikou bylo dosti povrchní a pro zadání tohoto typu podle mě nedostatečné. Předpokládala bych zde podrobnější nastudování klasifikačních metod, způsobu jejich vyhodnocení a problematiky nerovnoměrného rozdělení tříd objektů.
3. **Rozsah technické zprávy** je v obvyklém rozmezí
4. **Prezentační úroveň předložené práce** 62 b. (D)

Logická struktura předložené technické zprávy je poměrně dobrá, ale některé pasáže nejsou dostatečně podrobné. Jedná se především o popis úlohy 2 - predikce zákazníka. Chybí zde popis trénovací množiny (počet objektů, počet objektů z jednotlivých tříd) a lepší zhodnocení dosažených výsledků (především z pohledu nevyváženosti jednotlivých tříd). Dále bych v kapitole 2 očekávala lepší popis klasifikátorů, které byly využity pro experimenty v nástroji RapidMiner.

Student také nepoužívá obvyklou terminologii pro popis testování klasifikačních modelů (zejména pak pro rozdělení dat na trénovací a testovací množinu). Dále zpráva obsahuje nezanedbatelné množství matoucích informací (prohození supervised a unsupervised filtrů na str 41/42, chyba ve vzorci 2.3).
5. **Formální úprava technické zprávy** 68 b. (D)

Z jazykového hlediska se jedná o průměrnou práci, která obsahuje obvyklé množství chyb. Z typografického hlediska mám připomínky ke kvalitě obrázků prezentujících výsledky provedených úloh v nástroji RapidMiner (obrázky 3.3 - 3.13) a vytvořenou aplikaci (obrázky 5.1 - 5.6) a ke kvalitě sekvenčních diagramů v příloze, které nejsou vůbec čitelné.
6. **Práce s literaturou** 60 b. (D)

Seznam použité literatury není špatný, ale myslím si, že student uvedené zdroje dobře nevyužil, především při zpracování kapitoly 2. Zde student evidentně vycházel především ze studijní opory do předmětu ZZN, což podle mě pro zadání tohoto typu není dostatečné. Kapitola 2 také obsahuje mnoho vět převzatých ze studijní opory, v nichž jsou v lepším případě změněna 1-2 slova, přičemž studijní opora není citována u jednotlivých vět nebo odstavců, ale pouze na začátku celé kapitoly je uvedeno, že teorie vychází mj. ze studijní opory. Seznam literatury také obsahuje jednu nekompletní citaci.
7. **Realizační výstup** 80 b. (B)

V rámci diplomové práce byla vytvořena aplikace, která umožňuje dolování asociačních pravidel a predikci zájmu zákazníka o vybraný produkt. Student sám implementoval především uživatelské rozhraní pro tuto aplikaci, implementace samotných dolovacích algoritmů byly převzaty z nástroje Weka. Vytvořená aplikace je funkční, výsledky získané pomocí této aplikace byly porovnány s výsledky získanými pomocí nástroje RapidMiner na zvolených firemních datech. V obou nástrojích bylo dosaženo srovnatelných výsledků. Jako drobné negativum vnímám to, že aplikace byla vytvořena pro konkrétní firemní data.
8. **Využitelnost výsledků**

Vytvořená aplikace umožňuje provádět analýzu dat konkrétní firmy. Může tedy být použitelná pro potřeby dané firmy. Není ale jasné, zda vytvořenou aplikaci lze přímo použít pro libovolnou firmu s obdobným zaměřením.
9. **Otázky k obhajobě**
 - Je vytvořená aplikace použitelná i pro data jiných firem podobného zaměření?
 - Řešil jste nějakým způsobem problém nerovnoměrného zastoupení objektů v jednotlivých třídách během predikce toho, zda si zákazník koupí nějaký produkt?
10. **Souhrnné hodnocení** 62 b. uspokojivě (D)

Celkové hodnocení diplomové práce je ovlivněno především nekvalitní technickou zprávou, která svědčí o tom, že orientace studenta v dané problematice je spíše povrchní. I přesto ale student vytvořil funkční aplikaci, a proto navrhuji hodnocení stupněm D.

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

V Brně dne: 6. června 2019

Burgetová Ivana, Ing., Ph.D.
oponent