

OPONENTSKÝ POSUDEK DIPLOMOVÉ PRÁCE

Autor/ka: Bc. Ondřej Pelinka
Název práce: Analýza výpůjček Univerzitní knihovny Pardubice

Vedoucí práce: Ing. Jan Panuš, Ph.D.
Oponent: Ing. Pavel Jirava, Ph.D.

Hodnocení práce:

(hodnocení: 1 je výborný, ... , 4 je nevyhovující)

	1	1-	2	2-	3	4
1. Náročnost tématu	x					
2. Volba vhodné metodiky zpracování		x				
3. Splnění cílů práce	x					
4. Odborný přínos autora/rky		x				
5. Logický postup řešení	x					
6. Využití teoretických znalostí		x				
7. Návrh modelu řešení	x					
8. Sběr dat, materiálu pro řešení	x					
9. Výstižný souhrn práce a vhodná klíč. sl.	x					
10. Průběžná citace použité literatury	x					
11. Formální úprava textu		x				
12. Formální úprava grafická (obrázky, ...)		x				
13. Celkové hodnocení práce	x					

Další připomínky, vyjádření, doporučení nebo nedoporučení k obhajobě:

Cíl práce je formulován v zadání – autor měl analyzovat výpůjčky v knihovně UPCE. Měl zvolit vhodné metody, realizovat zpracování, přípravu a analýzu dat. Vstupním datovým souborem byla data ve formátu *.xls o 383 tisících záznamů. Ty jsou následně upravovány, formátovány a poměrně podrobně analyzovány. Následně se autor zaměřuje na vlastní přípravu dat pro modelování ve zvoleném SW nástroji. Výstupem je datový soubor o 100 592 záznamech, se kterým autor dále pracuje. V analytické části využívá jednu ze známých metod dataminingu- rozhodovací stromy. Pokusil se klasifikovat studenty podle MDT (mezinárodní desetinné třídění) jejich výpůjček. Ačkoli výstupy a výsledky testování vytvořených klasifikačních modelů nejsou příliš přesvědčivé (nejlepší dosažená hodnota je okolo 90% správně klasifikovaných objektů, ostatní výsledky se pohybují níže) metodicky postupoval autor správně a představené modely jsou korektní. Ukázalo se, že klasifikace studentů celé UPCE podle MDT je s použitím rozhodovacích stromů velmi nepřesná a to zřejmě proto, že je slabá závislost mezi zkoumanými vstupy. Problém autor následně zúžil pouze na studenty FES. Zde právě bylo pro 1. kategorii MDT dosaženo nejlepších výsledků. Pokud by však byly zařazeny i další kategorie, rozhodovací stromy by již

nebyly příliš úspěšné pro dostatečně přesnou klasifikaci – opět z důvodu slabé závislosti mezi vstupy.

Cíle práce formulované v zadání byly dosaženy. Předložený text dle oponenta splňuje požadavky na diplomovou práci. Formálně i obsahově je text až na drobnosti v pořádku.

Práci doporučuji k obhajobě

Hodnocení práce známkou: **Výborně**

Otázky k obhajobě:

- 1) Vysvětlete, jak jste z původní vstupní datové matice (na str. 21 uvádíte 383 t řádků) vytvořil vstup pro modelování (na str. 22 a 27 uvádíte 283 t záznamů respektive 239t , následně na straně 37 - 100 592 záznamů). Především mi není jasný skok z 383t na 283t.
- 2) Na str. 65 uvádíte, že velkou roli ve slabé závislosti mezi vstupy hrají vyučované předměty napříč celou Univerzitou. S tímto tvrzením by se dalo souhlasit a zní logicky - spíše by mne zajímalo, o co jej opíráte. Nebylo by přesnější říci, že „zvolená metoda pro klasifikaci (rozhodovací stromy) neprokázala silnou, jasnou závislost mezi vstupy. Významným faktorem může být (další výzkum by to mohl prokázat) vyučování předmětů napříč celou Univerzitou“.
To, že jedna zvolená metoda pro klasifikaci dává špatné výsledky, snad ještě neznamená, že neexistují skryté závislosti odhalitelné jinou metodou.

V Pardubicích, dne 29.5.2014

Ing. Pavel Jirava, Ph.D.