

OPONENTNÍ POSUDEK ZÁVĚREČNÉ KVALIFIKAČNÍ PRÁCE

Autor závěrečné práce: Petr Franz

Název práce: Klasifikační algoritmy v dataminingových úlohách

Oponent práce Ing. Sylvie Sýkorová

Pracoviště opONENTA vedoucí prodejce

- | | |
|---|-----------------|
| A. Kvalita abstraktu, klíčová slova odpovídají náplni práce | Výborně (1) |
| B. Rozsah a zpracování rešerše | Výborně (1) |
| C. Řešení práce po teoretické stránce | Výborně (1) |
| D. Vhodnost, přiměřenost použité metodiky | Výborně (1) |
| E. Úroveň zpracování výsledků a diskuse | Velmi dobře (2) |
| F. Vlastní přínos k řešení problematice | Výborně (1) |
| G. Formulace závěru práce | Výborně (1) |
| H. Splnění zadání (cílů) práce | Splněno |
| I. Skladba, správnost a úplnost citací literárních údajů | Výborně (1) |
| J. Typografická a jazyková úroveň (vč. pravopisu) | Výborně (1) |
| K. Formální náležitosti práce
(struktura textu, řazení kapitol, přehlednost ilustrací) | Výborně (1) |

Komentáře či připomínky:

Autor ve své práci uvádí základní principy dataminingu a popisuje klasifikační algoritmy. Svoje postupy a praktické řešení rozdělil do dvou částí. V první části prezentuje datamining a nástroj IBM SPSS Modeler, kde se zaměřil na zavedení pojmů a proces dolování dat. Zde autor poukazuje na rozhodovací stromy a algoritmy, které jsou autorem zřetelně popsány a vyjádřeny ve vzorcích. Ve své druhé části autor popisuje tvorbu programu pro vizualizaci vybraných klasifikačních algoritmů se zobrazením výsledného stromu s jednotlivými kroky a výpočty. Desktopovou aplikaci vytvořil v objektovém jazyce C++ s použitým prostředím Visual Studio 2012. Pro vstup do aplikace použil tabulková data z externího souboru. Načtení dat bylo přes vytvořený Streamreader, který pomůže přečíst řádek a rozdělit ho podle daného oddělovače. V hlavní části programu je třída kategorie, která počítá četnosti, jejichž výsledné hodnoty jsou použity ve vybraných algoritmech Entropie a Informačního zisku. Příklady výpočtů jsou autorem prezentovány v tabulkách. Pro dosažení výsledných hodnot je použito čítání četností v závislosti na predikovaném atributu. Celkový design aplikace prezentuje propojení všech popsanych částí, které autor přehledně odprezentoval ve vývojových diagramech. Odborně působí i prověření správnosti algoritmů, kde autor upozorňuje na kritický bod, a to jsou reálná data, která jsou často neúplná. Výsledné hodnoty a samotná diskuse jsou z mého pohledu úzce popsány. Pro lepší vizualizaci práce by byla vhodná úvaha o vlastním přínosu a také doplnění o diskusi nad zkušeností z e-learningového kurzu.

...pokračuje na straně 2

Celkové zhodnocení:

Závěrem uvádím, že písemné zpracování bakalářské práce je provedeno přehlednou formou, text a grafické znázornění je výstižné. Jsou tu prezentovány všechny body, které jsou uvedeny v zadání bakalářské práce. Autor tu prezentuje práci, která je na velmi dobré odborné úrovni.

Otázky k obhajobě:

Otázky:

1. V jakém případě by jste vhodně použil algoritmus GINI index?
2. O jaké další algoritmy by autor rozšířil stávající program?
3. Uveďte příklady kam je aplikaci možné přenést?

Celková klasifikace:

Práce splňuje požadavky na udělení akademického titulu, a proto ji doporučuji k obhajobě
Navrhuji tuto práci klasifikovat stupněm Výborně (1)

V Jablonci nad Nisou

dne 22.5.2015

Podpisem současně potvrzuji, že nejsem v žádném osobním vztahu k autorovi práce


.....
podpis oponenta