



Posudek diplomové práce studenta Hlaveše

Cílem diplomové práce studenta Vojtěcha Hlaveše byl teoretický návrh a praktické ověření vhodných modifikací jedné či více existujících měr kvality souborů pravidel získávaných z dat. Uvedený cíl je obecně důležitý vzhledem ke skutečnosti, že metody získávání pravidel z dat představují nejrozšířenější typ metod dobývání znalostí z dat, a že bez vhodných měr různých aspektů jejich kvality nelze mezi sebou srovnávat soubory pravidel získané z týchž dat pomocí různých metod. Zvláštní důležitost však má v kontextu toho, že míry kvality souborů pravidel byly dosud rozpracovány pouze pro klasifikační pravidla. Pro obecnější typy pravidel existuje pouze velký počet měr kvality jednotlivých pravidel, zatímco soubory pravidel se dosud uvažovaly pouze v případech kontrastních množin pravidel (které lze na jednotlivá pravidla převést) nebo jako kontext pro měření kvality jednotlivých pravidel. Nejpřínosnějším naplněním cíle práce by proto byl pokus o zobecnění některých měr kvality ze souborů klasifikačních pravidel na obecnější soubory pravidel. Protože tento specifický cíl je teoreticky náročný a vyžaduje znalosti statistiky a fuzzy logiky výrazně překračující úroveň absolventů oboru informatika, bylo zadání práce formulováno obecněji, aby umožnilo splnění vytčeného cíle i méně ambiciózním způsobem.

Vojtěch Hlaveš projevil po prostudování doporučené literatury ochotu pustit se i do výše zmíněného ambiciózního zobecnění některých měr kvality souborů klasifikačních pravidel na soubory pravidel obecnějšího typu. Navrhl celkem tři typy takového zobecnění a v průběhu následujícího roku je rozpracoval pro hlavní zástupce dvou nejběžnějších typů měr kvality klasifikačních pravidel a celkem čtyři metody získávání pravidel z dat, většinu z nich v několika modifikacích (např. v případě metody Guha rozpracoval tato zobecnění pro implikační i ekvivalenční kvantifikátory). Zároveň studoval i závislost získaných souborů pravidel na hodnotách nastavitelných parametrů metody, podobně jako se studuje v případě klasifikačních pravidel pomocí ROC křivek. Jeho rozhodnutí přistoupit k zadanému úkolu tímto náročným způsobem bylo zúročeno již po prvním roce práce na tématu, kdy koncepce jeho zobecnění byla přijata jako příspěvek na konferenci ITAT 2006. Po dvou letech je nyní zúročeno mimořádně kvalitní diplomovou prací.

K dosažení vytčeného cíle se diplomant musel důkladně seznámit jak s principy zobecňovaných měr, tak i s principy použitých metod získávání pravidel z dat. To vyžadovalo netriviální znalosti ze značně širokého spektra disciplin: obzervační logiky, fuzzy logiky, teorie umělých neuronových sítí i některých tradičních oblastí strojového učení, a v neposlední řadě z matematické statistiky. Předložená práce svědčí o tom, že všechny tyto znalosti si dokázal v průběhu uplynulých dvou let v potřebné míře osvojit. Testování navrženého zobecnění od něj naopak vyžadovalo seznámit se s alespoň jednou existující implementací každé z použitých metod. Přitom se s nimi musel seznámit natolik důkladně, aby mohl naprogramovat rozhraní mezi jejich výstupy a jeho vlastní implementací navržených

zobecnění měř kvality a aby mohl zautomatizovat získávání a vyhodnocování tisíců souborů pravidel z dat pomocí těchto metod. Provedené testování zároveň svědčí i o jeho velké svědomitosti a disciplinovanosti při experimentálním ověřování.

Vzhledem k uvedeným skutečnostem se domnívám, že diplomová práce Vojtěcha Hlaveše si jednoznačně zaslouží klasifikaci známkou *výborně*.

V Praze, 4. května 2007



Martin
