

**Posudek na diplomovou práci**  
**ARTIFICIAL NEURAL NETWORKS FOR CLUSTERING AND**  
**RULE EXTRACTION**

(Umělé neuronové sítě pro klastrování a extrakci pravidel)

**Autor posudku:** Mgr. Pavel Jiroutek

**Autor práce:** Jiří Iša

Tato práce zkoumá velmi zajímavou oblast umělé inteligence – využití speciálních architektur neuronových sítí inspirovaných teorií fuzzy množin pro automatickou extrakci pravidel z neznámých dat.

V první části práce autor shrnuje základní principy klastrování a představuje v této oblasti nejčastěji používané metody. Zejména se jedná o Kohonenovy samoorganizující mapy a jejich modifikace, které stručně srovnává na jednoduché klastrovací úloze. Další pozornost je věnována možnosti využití fuzzy teorie ve zkoumané oblasti a její kombinace s neuronovými sítěmi. Jsou představeny dvě existující architektury fuzzy neuronových sítí (ANFIS, FINN) pro extrakci pravidel. Na základě diskuse výhod a nevýhod všech výše zmíněných metod autor navrhuje vlastní architekturu – rostoucí fuzzy inferenční síť (GFINN) a novou metodu výběru signifikantních vstupních příznaků. V poslední části práce autor ukazuje vlastnosti navržených metod na dvou sadách dat - na umělé množině dat ukazuje schopnost nové metody zpětně extrahovat pravidla, podle kterých tato množina vznikla. Na reálných datech ("Boston Housing" z databáze UCI) je demonstrována schopnost navržených metod pracovat s hodnotami, které jsou zatíženy reálným šumem i procedurálními chybami při jejich sběru.

Autor ve své práci jasně ukazuje, že má velmi dobrý přehled ve všech výše uvedených oblastech. V souladu se zadáním obsahuje práce rekapitulaci a porovnání základních metod pro extrakci znalostí. Autor rovněž splnil další požadavek - navržení a experimentální ověření vlastní metody pro extrakci znalostí a detekci relevantních vstupních příznaků. Disk CD, přiložený k práci, obsahuje autorem vytvořené programy pro prostředí Weka, které realizují popisované experimenty, a animovanou demonstraci průběhu adaptace některých metod.

K práci nemám závažnějších připomínek, pouze bych upozornil na formální nedostatky v jinak výborném textu:

Jedná se zejména o překlepy, které se v práci vyskytují.

Porovnávané tabulky 5.1 a 5.2, které by měly obsahovat podle textu stejná pravidla, se v několika pravidlech liší.

Tabulka 5.3 by měla být alespoň pro některou z testovaných konfigurací znázorněna ve stejném formátu, jako obrázek 5.3, aby bylo porovnání GFINN a RBF na první pohled zřejmé.

Autor v experimentální části práce dospěl k některým překvapivým výsledkům. Jejich obecnou platnost by bylo vhodné ověřit v dalších experimentech. Podobně, jak autor sám navrhuje v závěru, by bylo vhodné provést porovnání s některou z architektur rozhodovacích stromů, která by mohla na úloze vykazovat porovnatelné výsledky. Tím by ovšem tento text pravděpodobně přerostl formát diplomové práce.

Co je však třeba velmi vysoko vyzdvihnout je vysoký standard prezentace popisovaných metod i vlastních výsledků, přijatelný pro kvalitní vědeckou publikaci. Práce je psána výbornou angličtinou a má přehlednou logickou strukturu. Tyto atributy, kromě zajímavých výsledků, dále rozšiřují její publikační potenciál.

Doporučuji, aby práce byla přijata jako diplomová a byla připuštěna k obhajobě.

V Praze dne 17.5.2007

A rectangular area of the document is redacted with a solid black fill, obscuring the signature of the reviewer.

Pavel Jiroutek