

Posudok bakalárskej práce

Martin Švantner: „Klasifikační a regresní stromy“

V práci sa autor venuje populárnej klasifikačnej technike, klasifikačným stromom, umožňujúcej spracovanie dát s náročnou štruktúrou a v praxi často obľúbenej najmä kvôli názornej a prehľadnej interpretácii výsledkov. Rozšírenie tejto techniky umožňuje navyše riešiť aj úlohy regresné. Práca je koncipovaná tak, aby čitateľa stručne oboznámila s podstatou klasifikačných a regresných stromov a na jednom zvolenom algoritme demonštrovala s akými rôznymi otázkami sa stretávame pri konštrukcii takéhoto klasifikátora a ako ich môžeme riešiť, teda akým spôsobom môžeme stromy „pestovať“.

Autor sa pri písaní práce oboznámil s problematikou klasifikácie, ktorou svoju prácu aj začal. Po tomto úvode nasleduje kapitola venovaná klasifikačným stromom, ktorá je zameraná na popis konštrukcie stromu „zhora nadol“. Tento typ je základom pre obľúbenú konštrukciu „rast-prerezávanie“, na ktorej je založený algoritmus popísaný v štvrtej kapitole. V tretej kapitole autor vysvetľuje spojitosť klasifikačnej úlohy s úlohou regresnou a ako možno použiť stromy k riešeniu regresnej úlohy. Je ďalej predstavená modifikácia konštrukcie klas. stromov z kapitoly dva pre prípad regresných stromov. V závere kapitoly nájdeme aj prehľad niektorých známych algoritmov používaných na tvorbu stromov. Štvrtá kapitola sa potom zameriava na asi najznámejší algoritmus na tvorbu binárnych stromov, algoritmus CART. Okrem podrobného popisu sa dozvieme aj ako sa tento algoritmus vie vysporiadať s chýbajúcimi pozorovaniami, čo je jeho veľkou výhodou. V poslednej kapitole autor na reálnych dátach zrozumiteľne prezentuje použitie algoritmu CART na vyriešenie dvoch klasifikačných a jednej regresnej úlohy.

Charakter práce je kompilačný, ale autor sa dobre oboznámil s danou problematikou a študovaný algoritmus sa naučil aj prakticky aplikovať. Práca má dobrú jazykovú a grafickú úroveň a mohla by poslúžiť čitateľovi ako stručný úvod do problematiky rozhodovacích stromov. Čo možno autorovi vytknúť je, že teoretickú časť práce mohol spracovať viac po svojom, miestami sú totiž niektoré pasáže veľmi blízke použitým referenciám.

Súdim ale, že práca splňuje nároky kladené na bakalársku prácu a doporučujem ju prijať k obhajobe.

V Prahe dňa 19.6.2008

Mgr. Jaroslav Ševčík
KPMS MFF UK

