

Posudek oponenta diplomové práce

1. Identifikační údaje

Název práce: **Predikce časových řad návštěvnosti webové domény pomocí SVM**
Autor práce: **Bc. Michal Říha**

2. Cíl práce

Cílem diplomové práce je návrh modelu pro predikci návštěvnosti webu Univerzity Pardubice pomocí metody Support Vector Machines.

3. Obsahové zpracování a přístup k řešení

Vzhledem k současnému rostoucímu počtu publikací v oblasti web miningu lze téma práce označit za aktuální. Autor si jako téma své diplomové práce zvolil predikci návštěvnosti webových stránek. Ta je užitečná zejména pro efektivní správu webových serverů a aplikací. Cíle práce lze označit jako středně obtížné.

Autor nejprve definuje web mining a jeho podoblasti. V tomto přehledu bych uvítal častější citace současné světové literatury v oblasti web miningu. Celkově považuji počet položek v seznamu literatury za minimální možný. V první kapitole jsou dále definovány časové řady a klouzavé průměry.

Ve druhé kapitole jsou definovány Support Vector Machines. Tato kapitola zahrnuje dostatečně podrobný popis uvedené metody a veškeré klíčové vztahy. Ve třetí kapitole je realizován návrh modelu pro predikci návštěvnosti webu upce.cz pomocí Support Vector Machines. Model zahrnuje rozdělení dat na trénovací a testovací množinu, volbu parametrů učení a verifikaci modelu. Kromě toho jsou data předzpracována pomocí standardizace. Jednotlivé kroky jsou jasně a srozumitelně popsány. K přehlednosti práce přispívají také dílčí závěry u jednotlivých kapitol. Autor realizuje celou řadu experimentů, ve výsledcích bych však uvítal také uvedení počtu nalezených jádrových funkcí.

4. Formální náležitosti a úprava

Práce je psaná jasně a srozumitelně. Práci lze vytknout několik překlepů a chybějící interpunkce. Obrázky i tabulky jsou zpracovány v odpovídající kvalitě. Seznam literatury je řazen podle průběžného citování, je zde hned několik překlepů a chyb v citování. Přílohy obsahují grafy s predikovanými hodnotami.

5. Hodnocení a otázky k obhajobě

Diplomant splnil cíle diplomové práce v plném rozsahu. Navrhl model pro predikci návštěvnosti webových stránek, tento model verifikoval v programovém prostředí Statistica. V rámci obhajoby by bylo vhodné, aby se diplomant vyjádřil k tomu, jaké proměnné použil jako vstupy modelu. Dále uveďte, jaký časový posun reprezentují výstupní hodnoty datového souboru. Dále vysvětlete funkci parametru γ u zvolených jádrových funkcí.

Práci doporučuji k obhajobě

Navržené hodnocení: **velmi dobře**
Pardubice, 17. května 2011

Ing. Petr Hájek, Ph.D.