

Posudek diplomové práce

předložené na Matematicko-fyzikální fakultě
Univerzity Karlovy v Praze

Autor/ka: Václav Koudelka

Název práce: Sémantická personalizace

Rok odevzdání: 2009

Práce se zabývá obecnou personalizací webových stránek, která by šla využít na libovolné doméně, nebyla by tedy svázaná s konkrétní webovou aplikací.

Nejprve práce rozebírá vstupní data od uživatele, která budou použita pro vlastní návrh adaptace. Poté popisuje tři různé algoritmy pro zpracování takových dat a pro návrh adaptace. Dále je podrobně rozebrána implementace jednotlivých částí aplikace i databáze pod ní. V následující kapitole je představena ukázková aplikace využívající personalizace.

Práce má dvě stěžejní části – analýzu problému a implementaci vlastního frameworku. Analýza je provedena dobře, pouze mi zde chybí odkazy na použití regresních metod (neuronové sítě, SVM,...), které by šly dobře použít spolu s vlastnostmi navštívených stránek. Implementace je také dobrá - obecně oceňuji fakt, že aplikace je schopna pracovat nezávisle na konkrétním webu a také přístup k předávání doporučených adaptací.

V práci mi chybí srovnání jednotlivých metod klasifikace uživatelů mezi sebou. V ukázkové aplikaci se pracuje s daty od uživatelů sebranými za týden, obsahujícími stovky návštěv. Na těchto datech by se krásně dala otestovat schopnost predikce jednotlivých metod – rozdělit jednotlivé návštěvy na trénovací část, a na testovací části nechat metodu predikovat další uživatelskou akci a porovnat ji se skutečností. Na základě toho by se dalo kvantitativně určit, která z metod je nejvhodnější pro danou aplikaci.

Požadavek, aby stránky měly ve svém kódu obsaženy své vlastnosti v tagu meta, je pochopitelný – metody pro klasifikaci uživatelů tak mají informace navíc. Ovšem omezení těchto vlastností na nominální data s malou doménou mi přijde neprůhledné a zbytečně omezující – nemůže si diskretizaci dělat metoda sama? Nějaká jiná metoda by např. informaci o přesné ceně (12 499Kč místo 4) využila. Nerozumím také tomu, proč vlastnosti stránek musí být pro všechny stránky společné – je jasné, že Notebooky a Procesory budou mít vlastnosti různé, a srovnávat jejich ceny mezi sebou je přinejmenším zvláštní.

Celkově se mi práce líbila a doporučuji ji k obhajobě.

Datum: 28.8.2009

Podpis:

