

Posudek diplomové práce

předložené na Matematicko-fyzikální fakultě
Univerzity Karlovy v Praze

Autor/ka: Bc. Martin Major

Název práce: Použití metod předpovídání budoucích uživatelských hodnocení pro doporučování filmů

Rok odevzdání: 2012

Práce se zabývá predikcí hodnocení uživatelů, zaměřuje se na doménu doporučování filmů. V práci jsou stručně představeny dva nejhlavnější směry doporučování, kolaborativní filtrování a filtrování založené na obsahu. Dále je rozebrána situace na CSFD.cz, z tohoto serveru byla využita data pro experimenty. Autor popisuje dvě naivní metody a poté metodu FilmDNA, kterou navrhl. Nakonec jsou popsány provedené experimenty a možná rozšíření práce.

Práce popisuje rozhodně zajímavé a užitečné téma, které navíc má šanci být využito v reálném provozu na CSFD. Také bych vyzdvihl použití reálných dat pro experimenty. Kód je psaný jako procedury v databázi, je stručný a poměrně přehledný.

Bohužel práce v několika důležitých oblastech není dostačující, za prvé je to současný stav. Představení dvou typů doporučování se odbude na jedné stránce, práce by si zasloužila delší pojednání a např. teoretické srovnání autorova přístupu s existujícími. Navíc by autorovi mohla literatura poskytnout zajímavé možnosti na zlepšení, např. nebrat hodnocení jako absolutní, ale pro každého uživatele odečíst jeho průměrné hodnocení a tím zjistit, zda se daný film uživateli líbil víc než ostatní filmy, co viděl.

Druhá část, které by chtěla vylepšit, jsou experimenty. Autorova metoda se sice umístila nejlépe, ovšem mezi FilmDNA a průměrem byl velmi malý rozdíl. Není uvedena žádná statistická významnost nebo rozptyl. Práce by měla použít pro srovnání existující volně dostupné implementace predikce preferencí (<http://lenskit.grouplens.org/>, <http://mahout.apache.org/>, <http://www.cs.waikato.ac.nz/~ml/weka/>), srovnání s průměrem je vhodné pro představu, jak dopadne naivní metoda, ale není možné jako jediná informace. Provedení experimentu 3x rozhodně nedostačuje pro vyloučení závislosti výsledku na rozložení trénovacích dat. Doporučuji alespoň 20x – 100x. Zvláštní také je, že varianta FilmDNA s váženým průměrem se konzistentně umísťovala jako horší než varianta s obyčejným průměrem a zřejmě i než průměrné hodnocení, což ukazuje, že podobnost uživatelů není moc přesná.

Nakonec několik obecných poznámek.

- Experimenty byly provedeny s uživateli, kteří mají alespoň 500 hodnocení, těch je 15% a dodali 62% všech hodnocení. Tento fakt sice pro metodu znamená, že využívá více než polovinu dostupných dat, ovšem také to znamená, že doporučování bude pouze pro 15% uživatelů, což se zdá jako velmi málo.
- Data se nedělí na „testovací“ a „validační“, ale na „trénovací“ a „testovací“. Validací data jsou částí trénovacích dat a používá je sám algoritmus.
- Poslední poznámka se týká výběru atributů žánr a země původu filmů. Tyto atributy by měly být vybrány na základě nějakého rozhodnutí, ovšem v práci není výběr nijak zdůvodněn. V sekci Možná zlepšení je navrženo zjišťování vhodnosti atributů, což zní jako dobrý nápad.

Celkově tuto práci hodnotím jako problematickou. Zabývá se zajímavým tématem, pracuje s reálnými daty, experimenty jsou hezky popsány. Práci ovšem táhne ke dnu malé studium literatury, malý výběr metod pro srovnání a obecně zvláštní a nepřesvědčivé výsledky experimentů. Přesto práci s velkými výhradami doporučuji k obhajobě.

Datum: 17.12.2012

Podpis: