

Evaluační algoritmy lokálně senzitivního hashování (LSH) v doporučovacíh systémech

Ladislav Martínek

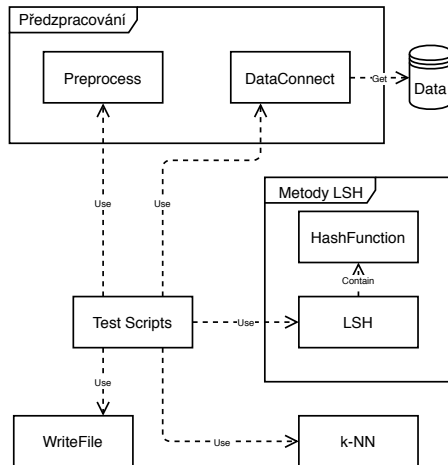
6. května 2018

- Analýza metod LSH
- Návrh aplikace metod do algoritmu k-NN
- Návrh a implementace testovací knihovny
- Testování různých parametrizací LSH
- Diskuze výsledků

- Osobní zájem v oblasti doporučovacích systémů
- Narůstajícího množství dat
- Osobní přínos

- Kolaborativní filtrování
- User-based k-NN
- Kosinova podobnost
- Lokálně senzitivní hashování

- Předzpracování
- Modul k-NN
- Modul LSH
- Testování přesnosti k-NN a úspěšnosti doporučování



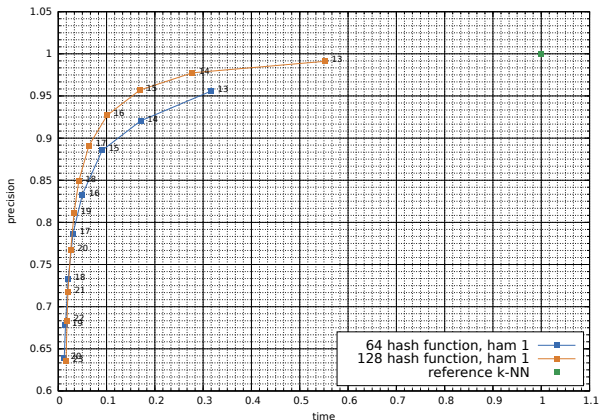
- Projekční matice
- Počet hashovacích funkcí
- Multi-probe
- Popularita položek

- Python
- Modul pro tvorbu grafů
- Argumenty
 - d** Dimenze hashovací funkce
 - c** Nastavení metody multi-probe
 - m** Hashovací funkce
 - M** Standartní hashovací funkce






- Testování na přesnost 10-NN
- Úspěšnost doporučování na vybraných modelech
- Databáze A
 - záznamu uživatel-položka:** 1 280 189
 - uživatelů:** 290 625
 - položek:** 11 101
- Databáze B
 - záznamu uživatel-položka:** 54 234 951
 - uživatelů:** 1 028 287
 - položek:** 52 689

Výsledky testování

- Přesnost 10-NN
- Úspěšnost doporučování (recall a catalog coverage)
- Závislost přesnosti 10-NN a hodnoty recall



- Dosažené výsledky
- Nastavení modelů
- Budoucí práce

-  RICCI, F.; ROKACH, L.; SHAPIRA, B.; aj.: *Recommender systems handbook*. Springer New York Dordrecht Heidelberg London, 2011, ISBN 978-0-387-85819-7.
-  XUE, G.-R.; LIN, C.; YANG, Q.; aj.: Scalable collaborative filtering using cluster-based smoothing. [online], August 15 - 19, 2005, [cit. 2018-03-08]. Dostupné z: <https://dl.acm.org/citation.cfm?id=1076056>
-  WANG, J.; SHEN, H. T.; SONG, J.; aj.: Hashing for Similarity Search: A Survey. [online], Submitted on 13 Aug 2014, [cit. 2018-03-11], 1408.2927. Dostupné z: <http://arxiv.org/abs/1408.2927>
-  LV, Q.; JOSEPHSON, W.; WANG, Z.; aj.: Multi-probe LSH: efficient indexing for high-dimensional similarity search. [online], September 23 - 27, 2007, [cit. 2018-03-17]. Dostupné z: <http://dl.acm.org/citation.cfm?id=1325851.1325958>
-  STECK, H.: Item Popularity and Recommendation Accuracy. [online], October 23 - 27, 2011, [cit. 2018-03-19]. Dostupné z: [doi:10.1145/2043932.2043957](https://doi.org/10.1145/2043932.2043957)

Děkuji za pozornost!

Dotazy?