

Рекомендательная система Авито

Василий Рубцов

24 июня 2017

Модель LightFM [1].

$$\hat{r}_{ui} = f(b_i + b_u + \langle q_i, p_u \rangle)$$

$$q_i = \sum_{j \in f_i} e_j, \quad p_u = \sum_{j \in f_u} e_j,$$

$$b_i = \sum_{j \in f_i} b_j, \quad b_u = \sum_{j \in f_u} b_j,$$

$$f(x) = \frac{1}{1 + \exp(-x)},$$

где $f_i \subset F^I$, $f_u \subset F^U$, $e_j \in \mathbb{R}^f$, $b_j \in \mathbb{R}$,
 F^I — множество признаков товаров,
 F^U — множество признаков пользователей.

LightFM

$$\hat{r}_{ui} = f\left(\sum_{j \in f_i} b_j + \sum_{j \in f_u} b_j + \left\langle \sum_{j \in f_i} e_j, \sum_{j \in f_u} e_j \right\rangle\right)$$

$$\hat{r}_{ui} = f\left(\sum_{j \in f_i} b_j + \sum_{j \in f_u} b_j + \sum_{j \in f_i} \sum_{k \in f_u} \langle e_j, e_k \rangle\right)$$

Факторизационные машины

$$\hat{r}(x) = w_0 + \sum_{j=1}^d w_j x_j + \sum_{j_1=1}^d \sum_{j_2=j_1+1}^d \langle v_{j_1}, v_{j_2} \rangle x_{j_1} x_{j_2}$$

Функции потерь:

- 1) BRP (bayesian personalized ranking).
Оптимизирует AUC.
- 2) WARP (weighted approximately ranked pairwise) [2].
Оптимизирует precision@k.

Функции потерь:

- 1) BRP (bayesian personalized ranking).
Оптимизирует AUC.
- 2) WARP (weighted approximately ranked pairwise) [2].
Оптимизирует precision@k.

Оптимизация:

- adagrad
- adadelat

Рекомендательная система

Признаки для товаров:

- id товара,
- id микрокатегории
- топ-1000 слов из описания

Рекомендательная система

Признаки для товаров:

- id товара,
- id микрокатегории
- топ-1000 слов из описания

Веса взаимодействий пользователя с товаром:

$$w_i = \left(\frac{t_i - t_{min}}{t_{max} - t_{min}} \right)^8,$$

где t_{max} и t_{min} — максимальное и минимальное время соответственно.

Рекомендательная система

Признаки для товаров:

- id товара,
- id микрокатегории
- топ-1000 слов из описания

Веса взаимодействий пользователя с товаром:

$$w_i = \left(\frac{t_i - t_{min}}{t_{max} - t_{min}} \right)^8,$$

где t_{max} и t_{min} — максимальное и минимальное время соответственно.

Стратегия	Публичный рейтинг
LightFM	5270
+ микрокатегория + веса	14256
+ топ-1000 слов из описание	15467

Таблица: Качество модели



Maciej Kula.

Metadata embeddings for user and item cold-start recommendations.

In Toine Bogers and Marijn Koolen, editors, *Proceedings of the 2nd Workshop on New Trends on Content-Based Recommender Systems co-located with 9th ACM Conference on Recommender Systems (RecSys 2015), Vienna, Austria, September 16-20, 2015.*, volume 1448 of *CEUR Workshop Proceedings*, pages 14–21. CEUR-WS.org, 2015.



Jason Weston, Hector Yee, and Ron Weiss.

Learning to rank recommendations with the k-order statistic loss.

In *ACM International Conference on Recommender Systems (RecSys)*, 2013.