Session-based Recommendation Systems

Контаев Илья, 16ПМИ

Goal

A recommendation system presents **items** to **users** in a relevant way to improve audience retention and engagement by modeling user **preferences**.

Long-term user preferences prediction are classical matrix completion based methods.

Session-based approach works with **short-term** preferences by predicting the best next item from sequential user interaction logs.

The core of recommendation systems is the relevance-based ranking of items.

Data

(user_id, item_id, timestamp, *optionally*: user_features, item_features) $\mapsto$ (value)

Explicit Feedback (*clean, expensive, low coverage)*

rating $\{1, 2, \ldots, n\}$, the user rated some movie 4 out of 5

binary $\{-1, 1\}$, the user likes or dislikes some post

Implicit Feedback (*noisy, cheap, high coverage*)

interaction the user listened to a song for 5 seconds

mixed the user clicked on a product and added it to wishlist

Usually, we don't have large quantities of explicit ratings available in sequence-based RecSys.

Models

- Simple Association Rules
- Markov Chains, Hidden Markov Models
- **Nearest Neighbors**: Item-based kNN and Session-based kNN
- **Matrix Factorization**
- **RNN(**GRU4Rec and improvements)

Evaluation of Session-based Recommendation Algorithms, https://arxiv.org/abs/1803.09587

k-NN

The next item is recommended by the similarity of an item from previous session and other items.

Variety of modifications exist which use heuristics such as sampling, caching and sequence awareness.

Usually a strong baseline.

Unifying Nearest Neighbors Collaborative Filtering. RecSys '14.

When Recurrent Neural Networks meet the Neighborhood for Session-Based Recommendation. RecSys '17

Matrix Factorization

Combines classic matrix factorization for long-term preference and factorized Markov chains to incorporate sequential information into the model.

- Factorized Personalized Markov Chains (FPMC)
- Factored Item Similarity Models(FISM)
- Factorized Sequential Prediction with Item Similarity Models(FOSSIL)
- Session-based Matrix Factorization (SMF)

Factorizing Personalized Markov Chains for Next-basket Recommendation, 2010

FISM: Factored Item Similarity Models for top-N Recommender Systems, 2013

Fusing Similarity Models with Markov Chains for Sparse Sequential Recommendation, 2016

RNN

RNNs for session-based, or more generally, sequential prediction problems is a natural choice.

GRU4Rec models user sessions in order to predict the probability of the subsequent events given a session beginning.

In other articles authors propose improvements of loss functions, negative sampling strategy and adding of user/item embeddings.

Session-based Recommendations with Recurrent Neural Networks,
https://arxiv.org/abs/1511.06939

Recurrent Neural Networks with Top-k Gains for Session-based Recommendations ,
https://arxiv.org/abs/1706.03847

Personalizing Session-based Recommendations with Hierarchical Recurrent Neural Networks,
https://arxiv.org/abs/1706.04148

Pairwise ranking losses for GRU4Rec and SMF

Assumption: *(u, i, j)* - user *u* prefers item *i* to item *j*

BPR: Bayesian Personalized Ranking

$$L = -\frac{1}{N_s} \sum_{j=1}^{N_s} log(\sigma(r_i - r_j))$$

Compares the score of a positive and a negative sampled item.

Where N_S is the sample size, r_k is the score on item *k* at the given point of the session, *i* is the desired item (next item in the session) and *j* are the negative samples.

TOP1 (ranking negative scores with regularization)

$$L = -\frac{1}{N_s} \sum_{j=1}^{N_s} log(\sigma(r_i - r_j)) + \sigma(r_j^2)$$

Thank you! Any questions?