



Sequential Recommendation with User Memory Networks

11/13/2018

Xiaosong Jia



Motivation

- Compress all of a user's previous records into a fixed hidden representation?
- Weaken the signal of highly correlated items for sequential recommendation
- Overlooking such signal makes it difficult for us to understand and explain the sequential recommendations.



General Framework

- N users, M item
- p_u -user u 's vector, q_i -item i 's vector
- M^u the personalized memory matrix
- $p_u^m = READ(M^u, q_i)$
- $p_u = MERGE(p_u^*, p_u^m)$
- $MERGE(x, y) = x + ay$
- Prediction: $PREDICT(p_u, q_i) = \delta(p_u^T q_i)$
- $M^u \leftarrow -WRITE(M^u, q_i)$

Matrix Factorization

$$\begin{array}{c}
 \begin{array}{ccccc}
 & s_1 & s_2 & s_3 & s_4 & s_5 \\
 u_1 & 1.4 & ? & 1.1 & 0.7 & ? \\
 u_2 & ? & 0.3 & ? & 0.7 & 0.5 \\
 u_3 & 0.4 & 0.3 & ? & ? & 0.3 \\
 u_4 & 1.4 & ? & 1.2 & ? & 0.8
 \end{array}
 \end{array}
 \Rightarrow
 \begin{array}{c}
 \begin{array}{cc}
 & U^T \\
 \begin{array}{c} 0.8 \\ 0.9 \\ 0.1 \\ 0.9 \end{array} & \begin{array}{c} 0.6 \\ 0.1 \\ 0.3 \\ 0.5 \end{array}
 \end{array}
 \times
 \begin{array}{c}
 \begin{array}{ccccc}
 & S \\
 \begin{array}{ccccc} 1.0 & 0.2 & 1.0 & 0.8 & 0.4 \\ 1.0 & 1.0 & 0.5 & 0.1 & 0.9 \end{array}
 \end{array}
 =
 \begin{array}{c}
 \begin{array}{ccccc}
 & s_1 & s_2 & s_3 & s_4 & s_5 \\
 u_1 & 1.4 & 0.8 & 1.1 & 0.7 & 0.9 \\
 u_2 & 1.0 & 0.3 & 1.0 & 0.7 & 0.5 \\
 u_3 & 0.4 & 0.3 & 0.3 & 0.1 & 0.3 \\
 u_4 & 1.4 & 0.7 & 1.2 & 0.8 & 0.8
 \end{array}
 \end{array}$$



Item-level method

- Read:

$$w_{ik} = (\mathbf{q}_{v_i^u})^T \cdot \mathbf{m}_k^u, \quad z_{ik} = \frac{\exp(\beta w_{ik})}{\sum_{i'} \exp(\beta w_{i'k})}, \quad \forall k = 1, 2, \dots, K \quad (7)$$

$$\mathbf{p}_u^m = \sum_{k=1}^K z_{ik} \cdot \mathbf{m}_k^u$$

- Write: Recent K items.

Feature-level method

- Global latent feature matrix $F = \{f_1, f_2, \dots, f_k\}$
- User preference matrix $M^u = \{m_1^u, \dots, m_k^u\}$
- Read: $w_{ik} = q_i^T \cdot f_k, z_{ik} = \frac{\exp(\beta w_{ik})}{\sum_j \exp(\beta w_{ij})}, \forall k = 1, 2, \dots, K$
 $p_u^m = \sum_{k=1}^K z_{ik} \cdot m_k^u$
- Write: $erase_i = \sigma(E^T q_i + b_e) \quad m_k^u \leftarrow m_k^u \odot (1 - z_{ik} \cdot erase_i)$
 $add_i = \tanh(A^T q_i + b_a), m_k^u \leftarrow m_k^u + z_{ik} \cdot add_i$