

Рекомендательные системы

Постановка задачи

i — user

a — item

$D = \{i, a, r_{i,a}\}$

$\hat{r}_{i,a} = ?$

$$i \begin{matrix} & & & a \\ \left(\begin{array}{ccccc} r_{11} & r_{12} & \cdot & \cdot & r_{1a} \\ r_{21} & \cdot & \cdot & r_{24} & r_{2a} \\ r_{31} & \cdot & r_{33} & r_{34} & r_{3a} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ r_{i1} & r_{i2} & \cdot & \cdot & r_{ia} \end{array} \right) \end{matrix}$$

User-based collaborative filtering

Найти других пользователей с похожими вкусами.

$$\hat{r}_{ia} = \frac{\sum_j r_{ja} * w_{ij}}{\sum_j |w_{ij}|},$$

где

w_{ij} — величина схожести

r_{11}	\cdot	r_{13}	\cdot	r_{1a}
r_{21}	\cdot	\cdot	r_{24}	r_{2a}
r_{31}	\cdot	r_{33}	r_{34}	r_{3a}
\vdots	\vdots	\vdots	\vdots	\vdots
r_{i1}	r_{i2}	\cdot	\cdot	r_{ia}

Item-based collaborative filtering

$$\hat{r}_{ia} = \frac{\sum_b r_{ib} * w_{ab}}{\sum_j |w_{aj}|},$$

где

w_{ab} — величина схожести

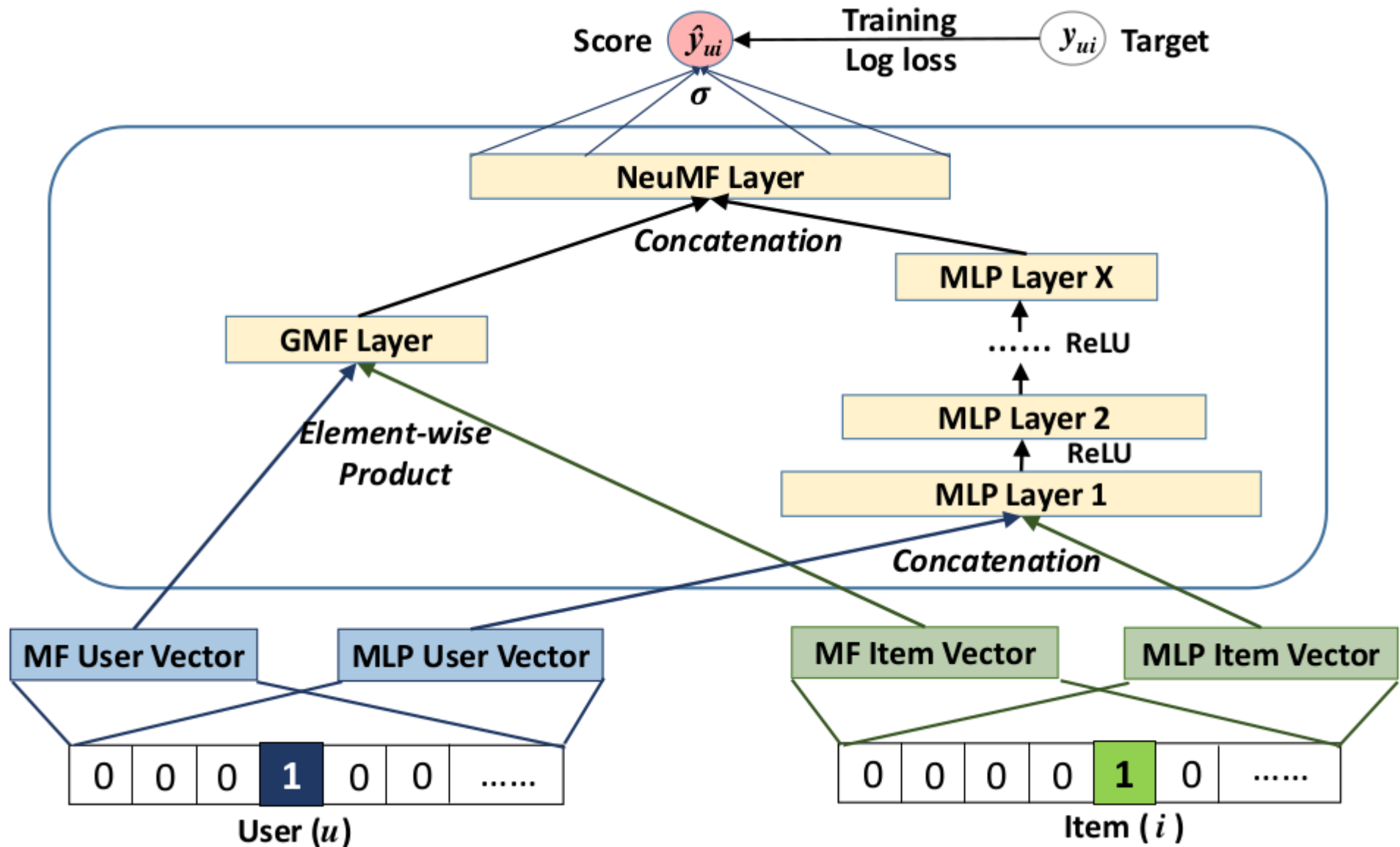
$$\begin{pmatrix} r_{11} & \cdot & r_{13} & \cdot & r_{1a} \\ \color{red}{r_{21}} & \cdot & \cdot & \color{red}{r_{24}} & \color{red}{r_{2a}} \\ r_{31} & \cdot & r_{33} & r_{34} & r_{3a} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ r_{i1} & r_{i2} & \cdot & \cdot & r_{ia} \end{pmatrix}$$

Alternating least squares

$$\begin{array}{c}
 R \\
 m \times n
 \end{array}
 \begin{pmatrix}
 2 & . & . & 4 & 5 & . \\
 5 & . & 4 & . & . & 1 \\
 . & . & 5 & . & 2 & . \\
 . & 1 & . & 5 & . & 4 \\
 . & . & 4 & . & . & 2 \\
 4 & 5 & . & 1 & . & . \\
 3 & 1 & . & 3 & . & .
 \end{pmatrix}
 \approx
 \begin{array}{c}
 P \\
 m \times k
 \end{array}
 \begin{pmatrix}
 . & . & . \\
 . & . & . \\
 . & . & . \\
 . & . & . \\
 . & . & . \\
 . & . & . \\
 . & . & .
 \end{pmatrix}
 \begin{array}{c}
 Q^T \\
 k \times n
 \end{array}
 \begin{pmatrix}
 . & . & . & . & . & . \\
 . & . & . & . & . & . \\
 . & . & . & . & . & .
 \end{pmatrix}$$

$$\min_{q,p} \sum_{(u,i) \in R} (r_{u,i} - q_i^T p_u)^2 + \lambda (\|q_i\|^2 + \|p_u\|^2)$$

Neural matrix factorization model



$$\text{precision @ } K$$
$$p @ K = \frac{\sum_{k=1}^K \text{relev}(k)}{K}$$

$$\text{average precision @ } K$$
$$AP @ K = \frac{1}{K} \sum_{k=1}^K \text{relev}(k) * p @ k$$

mean average precision at K

$$\text{map @ } K = \frac{1}{N} \sum_{j=1}^N ap @ K_j$$

Метрики качества

- *Cumulative Gain at K*
- *Discounted Cumulative Gain at K*
- *Normalized Discounted Cumulative Gain at K*
- *Mean reciprocal rank*
- *Ранговый коэффициент корреляции Кендэлла*
- *Ранговый коэффициент корреляции Спирмена*
- *Expected reciprocal rank*
- *PFound*