$$R = \begin{array}{c} \\ u_1 \\ \vdots \\ u_n \end{array} \begin{array}{c} \begin{array}{ccc} i_1 & \dots & i_m \end{array} \\ \begin{pmatrix} & \vdots & \\ \dots & r_{ui} & \dots \\ & \vdots & \end{pmatrix} \end{array}$$

– матрица рейтингов

$$\tilde{R} = \tilde{r}_{ui} = \begin{cases} r_{ui}, & \exists\, r_{ui} \\ 0, & \nexists\, r_{ui} \end{cases}$$

$\tilde{R}$ – полная матрица

$$\sum_{u \in U} \sum_{i \in I} \left(\tilde{r}_{ui} - \langle p_u, q_i \rangle \right)^2 =$$

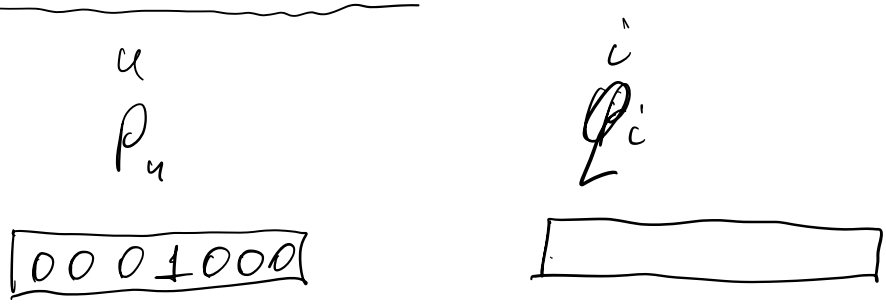
$$\| \tilde{R} - P^T Q \|_F^2 \to \min_{P,Q}$$

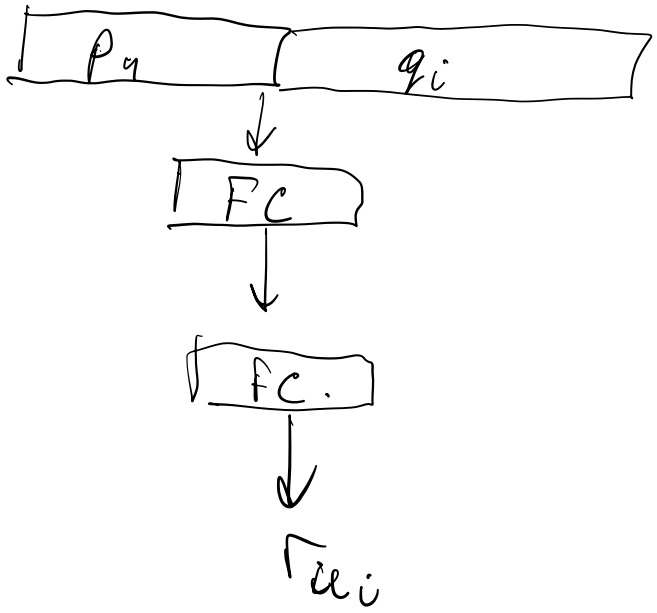
↓

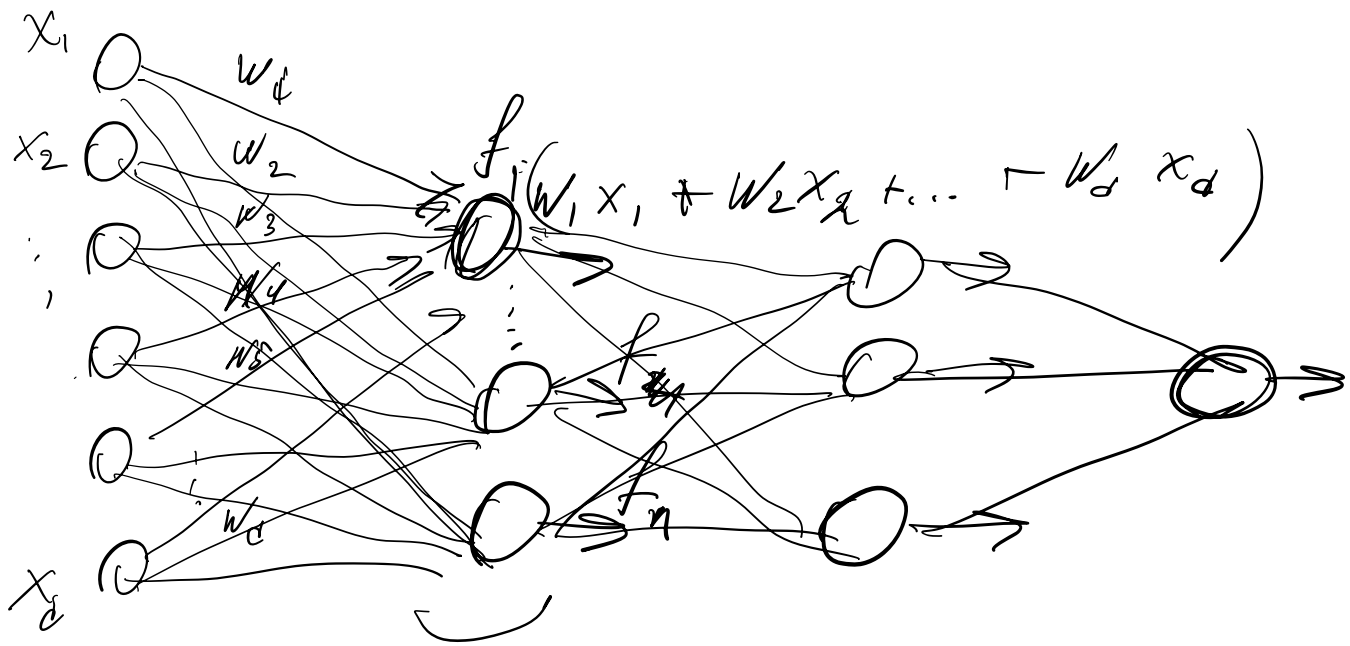
задача низкорангового приближения

Ридж SVD.

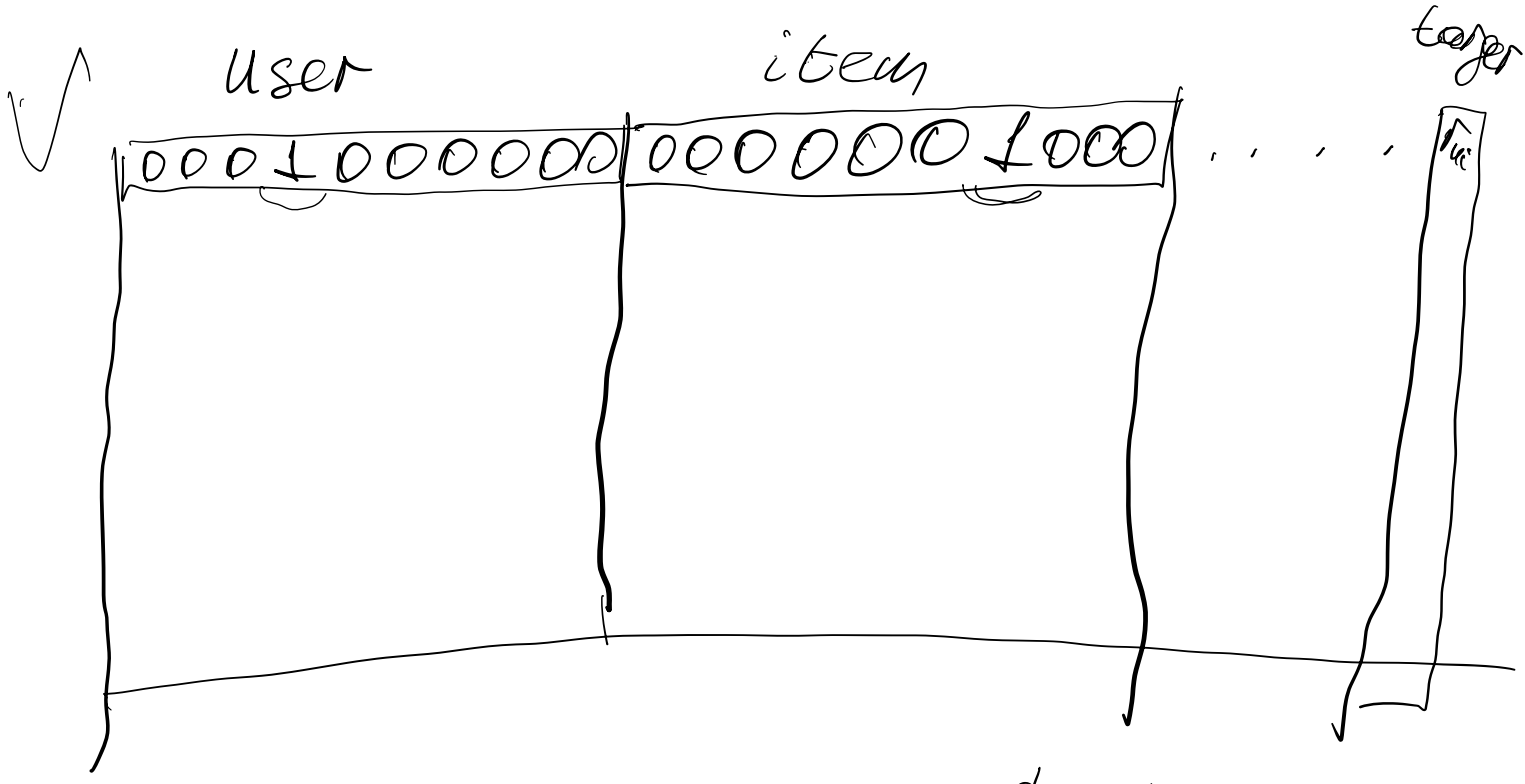
Neural CF


u
p_u
i
q_i
0001000


p_u
q_i
FC
FC.
r_{ui}


x_1
x_2
x_d
w_1
w_2
w_3
w_4
w_5
w_d
$f_1(w_1 x_1 + w_2 x_2 + \dots + w_d x_d)$

Факторизационные машины



$$a(x) = w_0 + \sum_{j=1}^{d} w_j x_j + \sum_{j_1=1}^{d} \sum_{j_2=j_1+1}^{d} w_{j_1 j_2} x_{j_1} x_{j_2}$$

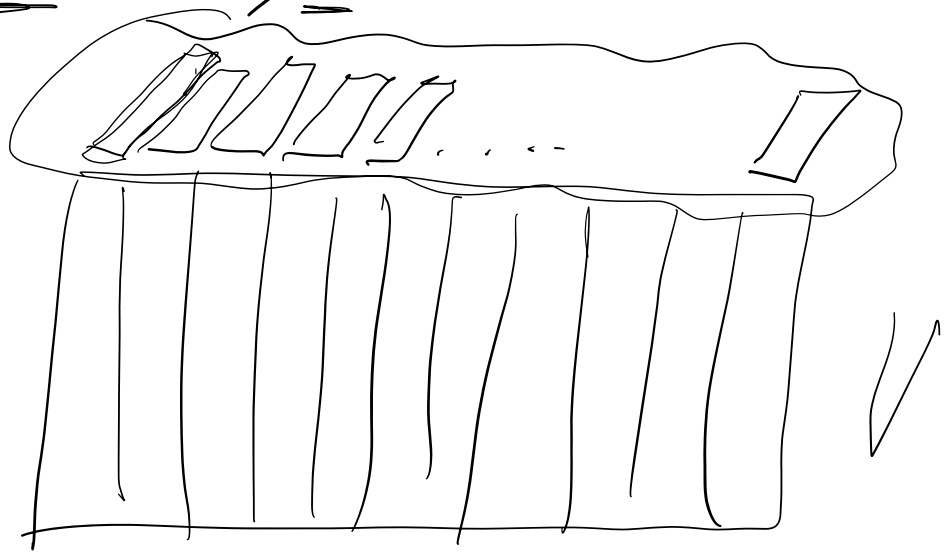
1 $\quad$ d $\quad$ $d(d-1)/2$.

$$w_{j_1 j_2} = \langle v_{j_1}, v_{j_2} \rangle$$

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

вектор данных

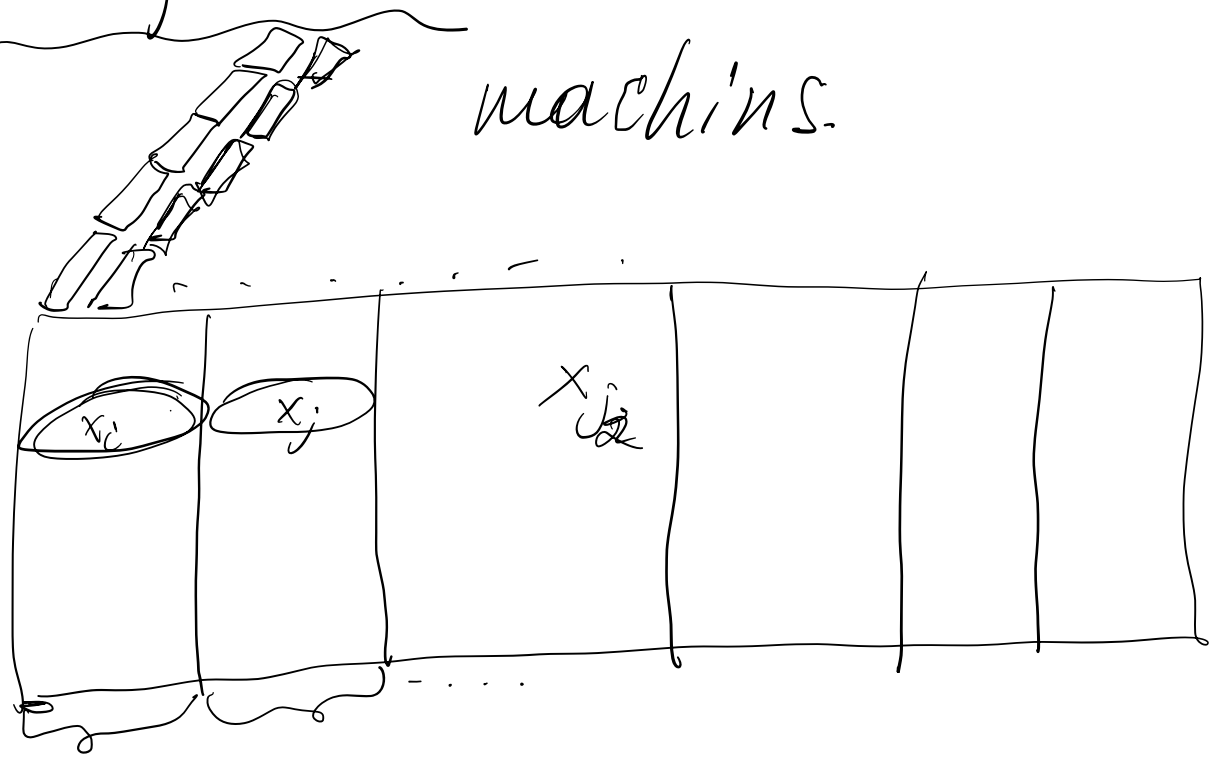
l $\quad$ d $\quad$ $d\,r$



r_{ui}

libFM.

FFM field-aware factorization machins.



items
из одного альбома

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

f_{j_1}, f_2 – индексы групп признаков x_{j_1} и x_{j_2}.

неявные Модели

iALS.

фидбек → явный (лайк. купил. звездочки).

→ неявный (кликнул. долго смотрел)

$$s_{ui} = \begin{cases} 1, & \exists r_{ui} \\ 0, & \nexists r_{ui} \end{cases}$$

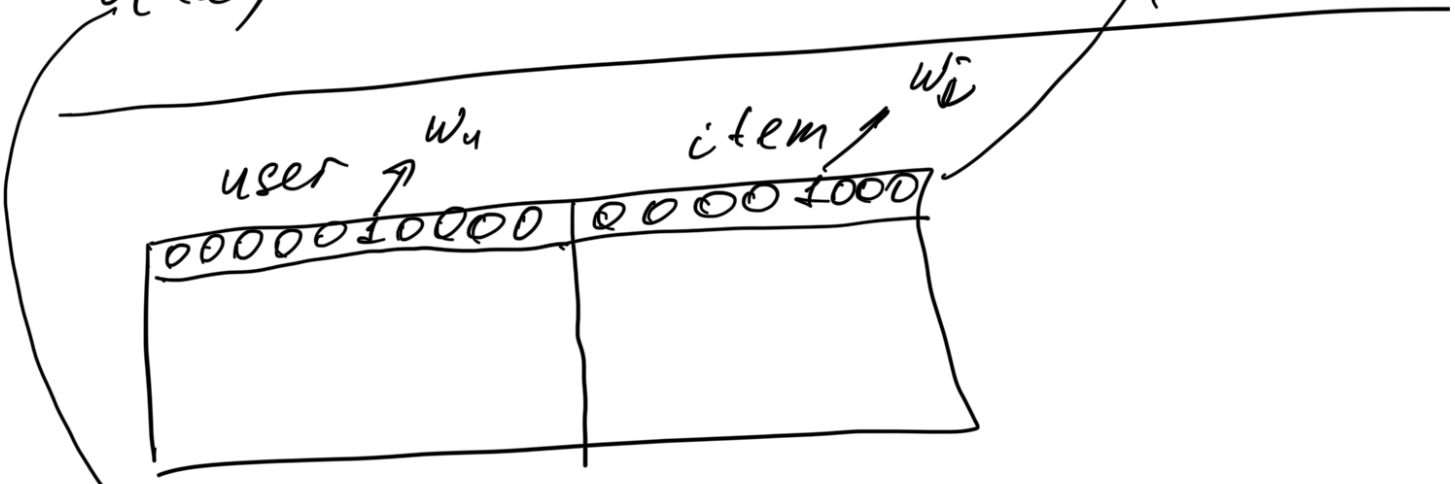
$$c_{ui} = 1 + \alpha s_{ui} \qquad \alpha \approx 100$$

$$\sum_{\substack{u \in U \\ i \in I}} c_{ui} \, L\left(s_{ui}, \langle p_u, q_i \rangle\right) \Rightarrow \min_{P, Q}$$

i - индекс item

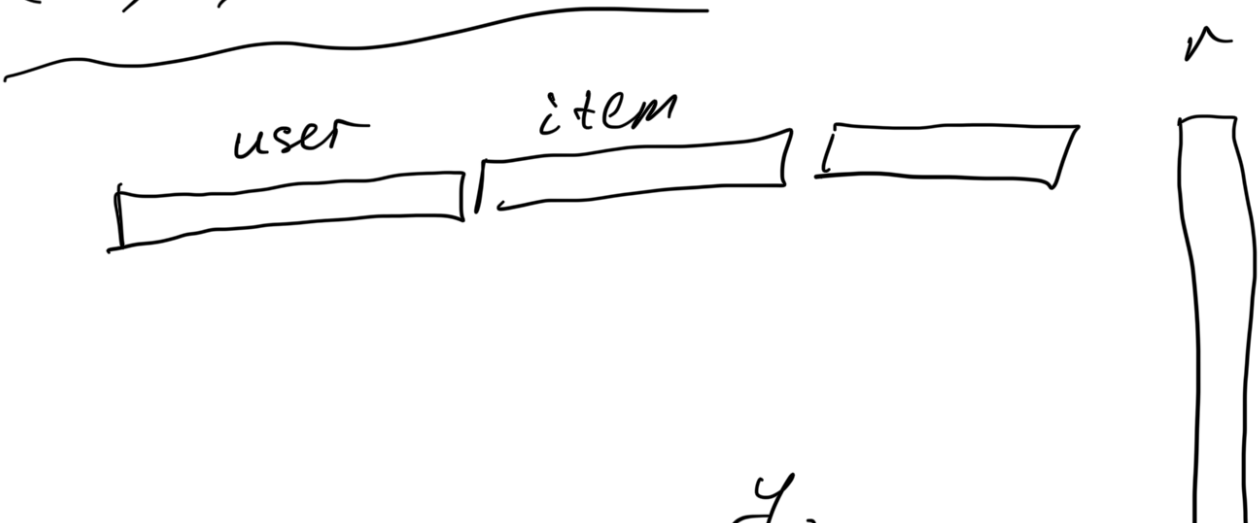
u - индекс пользователя.

$$a(x) = w_0 + w_u + w_i + \langle v_u, v_i \rangle$$



$$\sum_{(u,i) \in R} (a(x) - r_{ui})^2 \to \min$$

(u, i) r_{ui}

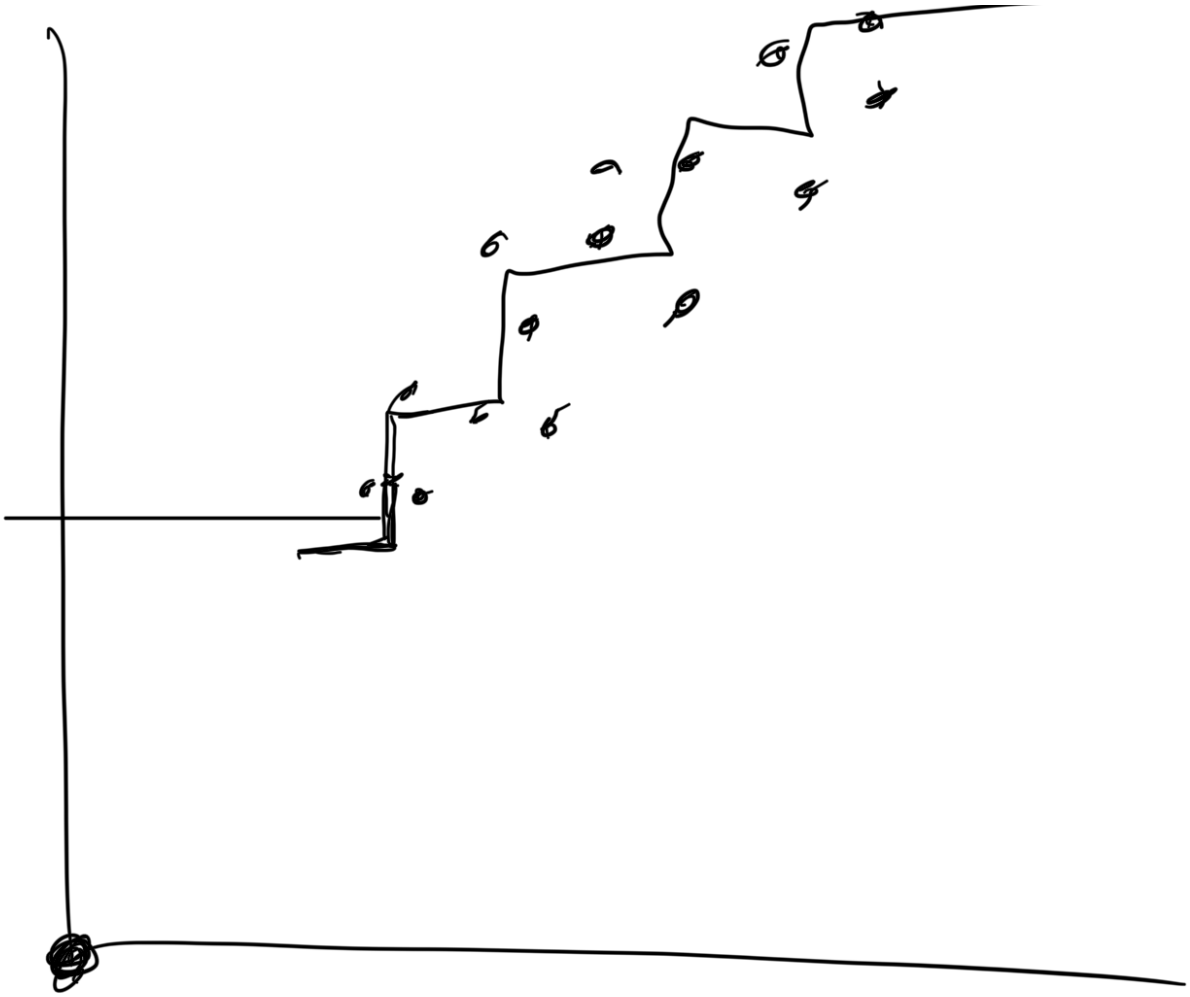


r_{ui}

Метрики.

- MSE, RMSE, MAE
- AUC-RR, F мера, ...

$R_u(K)$ - лучшие и товаров для ч. с точки зрения модели

L_u - товары заинтересовавшие пользователя

- hitrate@k = $[R_u(K) \cap L_u \neq \emptyset]$
- точность precision@k = $\frac{|R_u(K) \cap L_u|}{|R_u(K)|}$
- recall@k = $\frac{|R_u(K) \cap L_u|}{|L_u|}$