

Рекомендательные системы (RecSys)

$$\text{Уточ} = \text{Округл} \left(0.3 \cdot \underset{\substack{Rz \times 2 \\ T \times 15}}{Rz} + 0.25 \cdot \underset{P \times 1}{T} + 0.15 \cdot P + 0.3 \cdot \underset{I \times 1}{I} \right)$$

$$\text{Наклон} = \text{Округл} (0.3 \cdot Rz + 0.25 \cdot T + 0.15 \cdot P) / 0.7$$

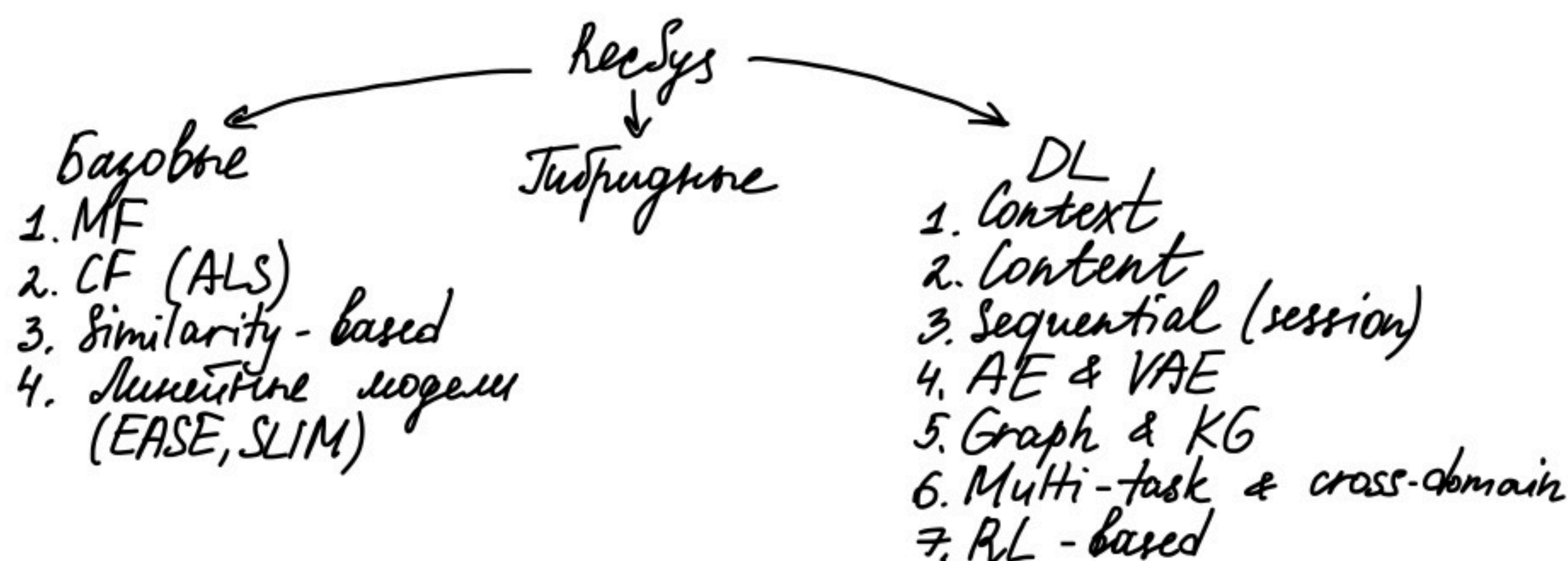
Если $\text{Наклон} > m$

$$T^u > n \quad (m=n=8) - \text{предварительно}$$

16 лекций, 16 семинаров

План:

- Формализация задачи, метрики качества + функции ранжирования
- Занятия 2-12 Подходы в RecSys



• Extra:

- Интерпретируемость
- LLM в RecSys
- A/B тесты online & offline
- API сервис

Ранжирование как ML задача

В общей постановке:

X - мн-во айтелов (= объектов рекомендаций)

$X^L = \{x_1, x_2, \dots, x_L\}$ - обучающая выборка

$i < j$ на парах $(i, j) \in \{1, \dots, L\}^2$

$a: X \rightarrow \mathbb{R}$, что $i < j \Leftrightarrow a(x_i) < a(x_j)$

I - мн-во айтелов, $i \in I$

1) totality (полнота)

$$\forall i, j \in I: i \neq j \Rightarrow i \succsim j \vee j \succsim i$$

2) антисимметричность

$$\forall i, j \in I: i \succsim j \wedge j \succsim i \Rightarrow i = j$$

3) транзитивность

$$\forall i, j \in I: i \succsim j \wedge j \succsim k \Rightarrow i \succsim k$$

свойства
отношения
линейного
порядка
 $\succsim \subset I^2$

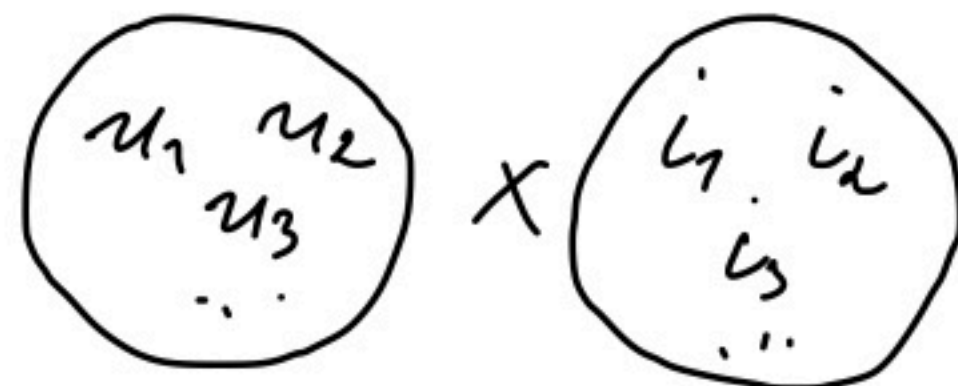
I Поиск (ранжирование выдачи)

query $\rightarrow d_1, d_2, \dots, d_n \xrightarrow{\quad} d_2 \quad 0.98$
 $\mathcal{D} = \{(q, d, y)\}$ y - оценки релевантности $d_1 \quad 0.96$
 $y \in [0, 1]$ (и.б. и $y \in \mathbb{R}$) \vdots
 $d_n \quad 0.1$
 $(q, d) \rightarrow p(y=1)$

\mathcal{D} - мн-во док-в
 \mathcal{Q} - мн-во запросов
 $\mathcal{D}_q \subseteq \mathcal{D}$ - мн-во док-в к запросу q
 $X = \mathcal{Q} \times \mathcal{D}$ - пар (q, d)
 $x \equiv (q, d), q \in \mathcal{Q}, d \in \mathcal{D}_q$
 \mathcal{Y} - упоряд. мн-во рейтингов (оценки rel-ции)
 E.g: CTR/клики, ассессорная разметка
 $y: X \rightarrow \mathcal{Y}$
 $(q, d_1) < (q, d_2) \Leftrightarrow y(q, d_1) < y(q, d_2)$

II Рекомендации пользователям

\mathcal{U} - мн-во пользователей, $u \in \mathcal{U}$
 \mathcal{I} - мн-во айтемов, $i \in \mathcal{I}$
 $X = \mathcal{U} \times \mathcal{I}$ - пар (пользователь, айтем)
 $(u, i_1) < (u, i_2) \Leftrightarrow y(u, i_1) < y(u, i_2)$
 $\forall u \in \mathcal{U}: \{y(u, i_1), y(u, i_2), \dots, y(u, i_k)\}$
 $|\mathcal{U}| = m$ где $k \leq n$
 $|\mathcal{I}| = n$



← какому
 пользователю
 выдам топ-к
 рекомендаций

Данные $\begin{cases} \rightarrow \text{explicit feedback} \\ \rightarrow \text{implicit feedback} \end{cases}$

Матрица взаимодействий users x items

	i_1	i_2	i_3
u_1	3	0	4
u_2	0	1	0
u_3	5	0	0

explicit $\begin{cases} \text{e.g.} \\ \text{(оценки} \\ \text{пользов.)} \end{cases}$ implicit $\begin{cases} \text{e.g.} \\ \text{(число} \\ \text{просмотров)} \end{cases}$

	i_1	i_2	i_3
u_1	1	0	1
u_2	0	1	0
u_3	1	0	0

explicit $\begin{cases} \text{e.g.} \\ \text{лайк/} \\ \text{дизлайк} \end{cases}$ implicit $\begin{cases} \text{e.g.} \\ \text{клик/} \\ \text{не клик} \end{cases}$

$r(u, i)$ - оценка rel-ции / результат взаимодействия юзера u с айтемом i

проблема - разреженность матрицы

тип итогового списка топ- n рек-ий:

- 1) новинки
- 2) только новые
- 3) микс

Функции ранжирования

1. point-wise
2. pair-wise
3. list-wise

Примеры 2, 3: BPR, WARP, YetiRank

Метрики оценки качества ранжирования

r_i - истинная релевантность
 i -го айтема для пользователя, $r_i \in \mathbb{R}$
 y_i - предсказание релевантности, $y_i \in \mathbb{Y}$

Пусть $r_i \in \{0, 1\}$, $y_i \in \{0, 1\}$ (либо м.б. $r_i \in \mathbb{R}$ и/или $y_i \in \mathbb{R}$)

1) Hitrate = $\frac{\#hits}{\#hits + \#misses}$

1.1. Hitrate@k = $\frac{\sum_{i=1}^k [y_i = 1]}{|Y|}$ - для 1 юзера
~ accuracy (доля правильных ответов)

1.2. На юзере hr@k=1, если хотя бы 1 правильный, k-длина списка рек-и
Hitrate@k = $\frac{1}{|u|} \sum_{u \in U} [\sum_{i=1}^k [y_{ui} = 1] > 0]$ - доля юзеров, у кого хотя бы 1 рек-и правильная

2) Precision@k = $\frac{\sum_{i=1}^k [y_i = 1]}{k}$

3) Recall@k = $\frac{\sum_{i=1}^k [y_i = 1]}{|R|}$

	1	2	3	4	5	6	7	
u_1								- нерелевантный
u_2								- релевантный
u_3								

$|I| = 7$

• u_1

$p@1 = 0$

$r@1 = 0$

$p@3 = \frac{1}{3}$

$r@3 = \frac{1}{3}$

$p@5 = \frac{2}{5}$

$r@5 = \frac{2}{3}$

• для выборки

$hr@3 = 1$

$p@3 = \frac{1}{|u|} \sum_{u \in U} p@k_u = \frac{\frac{1}{3} + \frac{2}{3} + \frac{2}{3}}{3} = \frac{5}{9}$

$r@3 = \frac{1}{|u|} \sum_{u \in U} r@k_u = \frac{\frac{1}{3} + \frac{2}{3} + \frac{2}{3}}{3} = \frac{5}{9}$

Плюсы:

- + интерпретируемость
- + простота
- + применимость и к бинар. классификации
- + расчет и на юзере, и на всей выборке

Минусы:

- чувствительность к порогу k
- не учитывают порядок
- бинаризуем оценки релевантности

4) Average Precision (AP@k)

$$AP@k = \frac{\sum_{i=1}^k y_i}{\sum_{j=1}^k r_j} \cdot p@i$$

5) Mean Average Precision (MAP@k)

$$MAP@k = \frac{1}{|u|} \sum_{u \in u} AP@k_u$$

	1	2	3	4	5
p@i	1	1	1/3	2/4	2/5
y_i	1	1	0	0	0

	1	2	3	4	5
p@i	0	0	1/3	2/4	3/5
y_i	0	0	1	1	1

$$AP@5 = 1$$

$$AP@5 = \frac{1 \cdot \frac{1}{3} + 1 \cdot \frac{1}{2} + 0 \cdot \frac{1}{5}}{3} \approx 0.44$$

Плюс: + учитывает порядок

Минус: $|R_u| \leq k \Rightarrow$ метрика занижается при больших k

$$MNAP@k = \frac{1}{|u|} \sum_{u \in u} \frac{1}{\min(n_u, k)} \cdot \sum_{i=1}^k y_i \cdot p@i$$

	1	2	3	4	5	6	7
u ₁							
u ₂							
u ₃							

где n_u - число айتمов, с которыми юзер взаимодействовал

(вопрос с семинара:)

$$MAP@3 = \frac{1}{3} \cdot \left(\frac{1}{2} + \frac{5}{6} + \frac{7}{12} \right) = \frac{23}{36} \approx 0.64$$

$$AP@3 = \frac{1 + \frac{1}{2} \cdot 0 + \frac{1}{3} \cdot 1}{2} = \frac{5}{6}$$

$$MNAP@3 = \frac{1}{3} \cdot \left(\frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{5}{3} + \frac{1}{2} \cdot \frac{7}{6} \right) = \frac{5}{9} \approx 0.56$$

6) Discounted cumulative gain (DCG@k)

$$DCG@k = \sum_{i=1}^k g(y_i) \cdot d(i) = \sum_{i=1}^k \frac{2^{y_i} - 1}{\log(i+1)}$$

\uparrow функция
 гипотеза
 \uparrow релевантность
 \uparrow релевантность
 \uparrow discount

$$NDCG@k = \frac{DCG@k}{IDCG@k}$$

это идеальное ранжирование (перестановка, где все релевантные айتمы в начале)