

# **DL в рекомендательных системах**

Спецкурс "*Введение в эффективные системы ML*"

Алексеев Илья

26 октября 2024

# **План лекции**

- Введение
- Отбор кандидатов
- Ранжирование

# Введение

- основные понятия
- постановка задачи
- почему *система*

## Понятия

- Два типа объектов: **пользователи и предметы.**
- Цель: обеспечить **взаимодействие** пользователя с предметами.

# Данные в рекомендательных системах

- о пользователе:
  - пол, возраст, место жительства, история взаимодействий с предметами
- о предмете:
  - размер, дата выхода, вкус, история взаимодействий пользователей
- о взаимодействии:
  - тип, дата, оценка

## Постановка задачи

Для пользователя  $u$  выдать упорядоченный набор  $i_1, \dots, i_k$ , с предметами из которого он точно провзаимодействует.

## **Рекомендательная система**

- отбор кандидатов
- ранжирование кандидатов и формирование выдачи

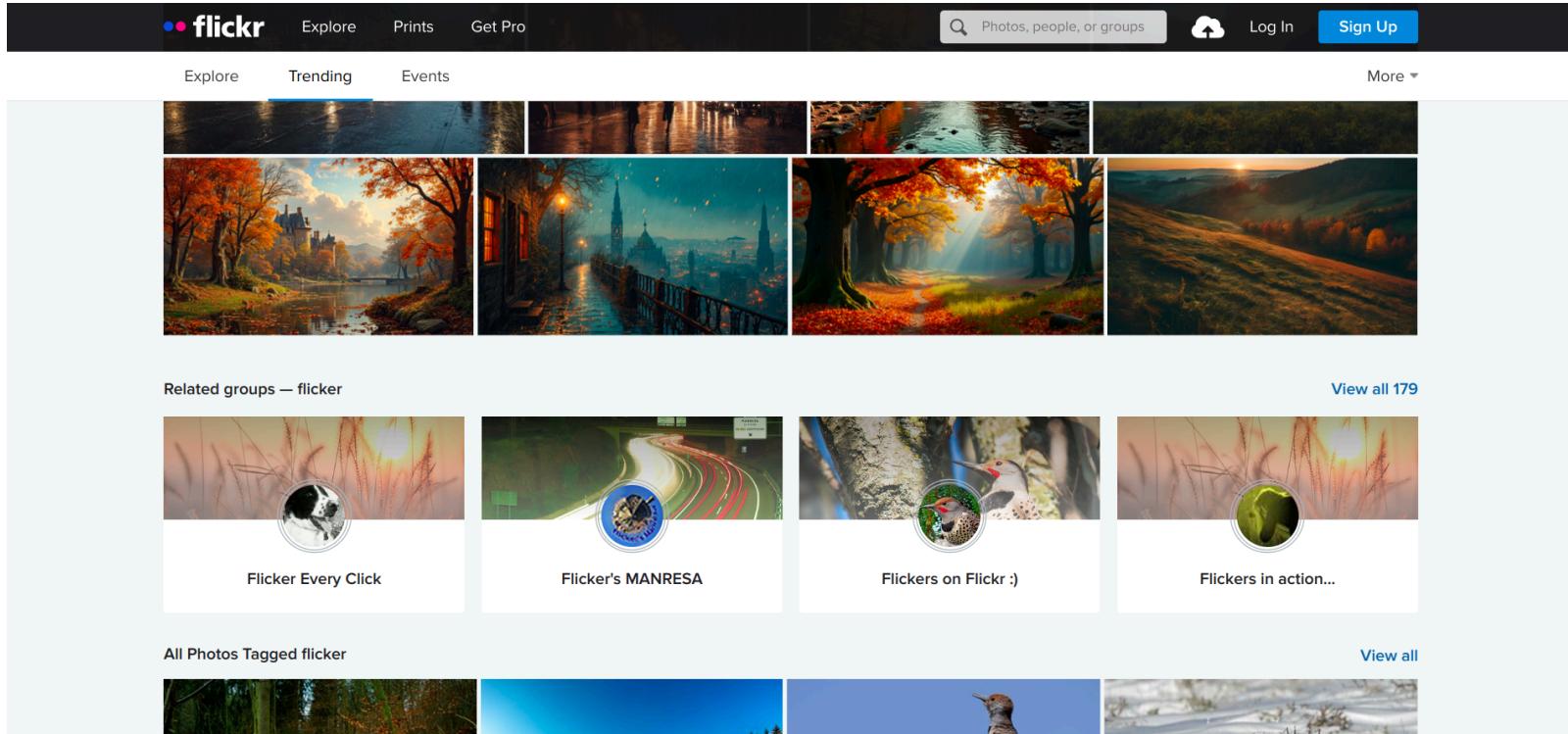
# Вопросы??



# Отбор кандидатов

- Поиск по контенту
- Колaborативная фильтрация
- Контрастивное обучение
- Sequential recommenders

# Пример: обмен фото



The screenshot shows the Flickr homepage with a dark header bar. The header includes the Flickr logo, navigation links for 'Explore', 'Prints', and 'Get Pro', a search bar with placeholder text 'Photos, people, or groups', a cloud icon for 'Log In', and a blue 'Sign Up' button.

Below the header, there are three tabs: 'Explore' (selected), 'Trending' (highlighted in blue), and 'Events'. To the right of these tabs is a 'More' dropdown menu.

The main content area displays a grid of 12 thumbnail images, all depicting autumn scenes such as colorful trees, fallen leaves, and scenic landscapes.

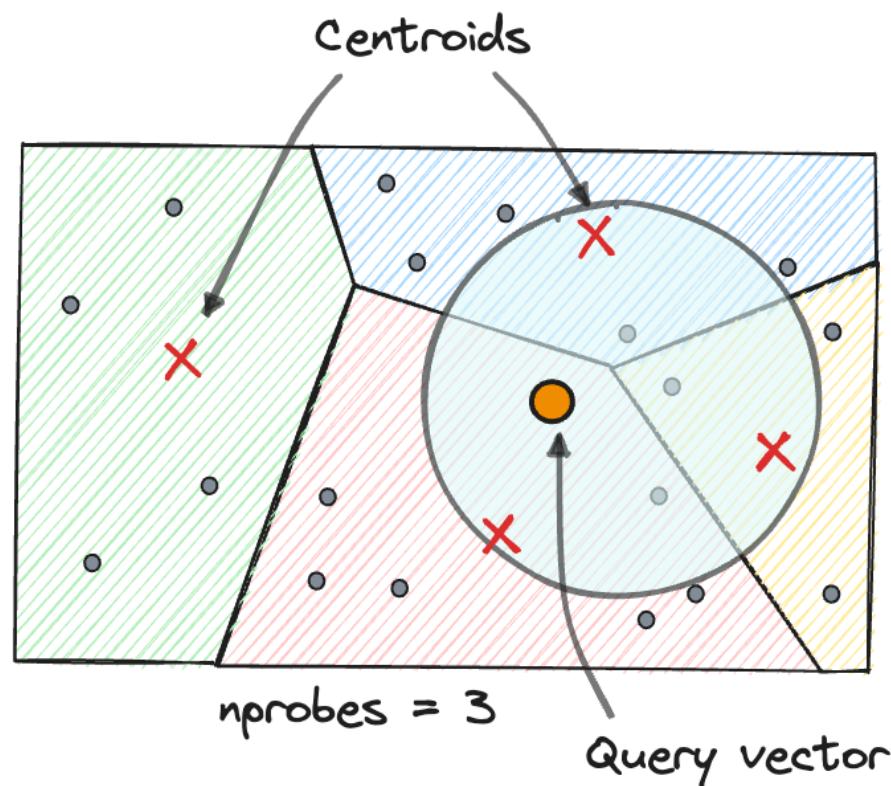
Below the grid, there is a section titled 'Related groups — flicker' with four group cards:

- Flicker Every Click
- Flicker's MANRESA
- Flickers on Flickr :)
- Flickers in action...

Each group card features a thumbnail image, the group name, and a 'View all 179' link.

At the bottom, there is another section titled 'All Photos Tagged flicker' with four thumbnail images, followed by a 'View all' link.

# Поиск по контенту



## **Поиск похожих: плюсы и минусы**

Плюсы:

- требует мало данных
- учитываем предпочтения пользователя

Минусы:

- не предлагаем пользователю ничего нового
- не учитываем сторонние признаки о пользователе и взаимодействии

**Вопросы??**



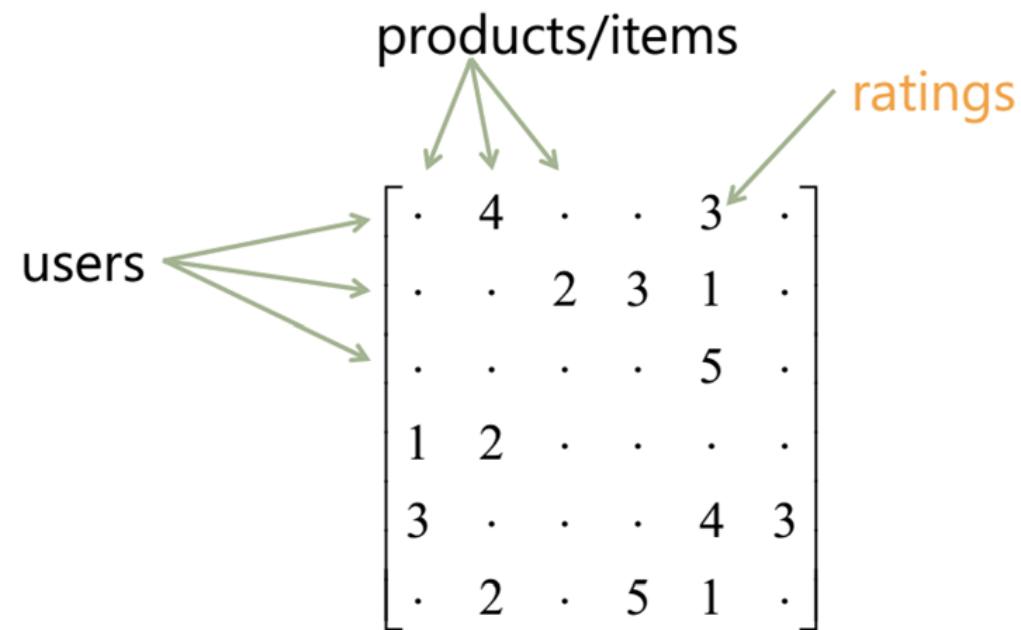
# Колаборативная фильтрация: KNN

Матрица взаимодействий

$$R \in \text{Mat}(n_u \times n_i).$$

Мера близости объекта и пользователя:

$$\hat{R}_{ui} = \frac{1}{|N(u)|} \sum_{\tilde{u} \in N(u)} R_{\tilde{u}i}.$$



## Колаборативная фильтрация: MF

Обучаем эмбединги  $P \in \mathbb{R}^{n_u \times d}$  и  $Q \in \mathbb{R}^{n_i \times d}$  на задачу регрессии:

$$\hat{R}_{ui} = \langle p_u, q_i \rangle + b_u + b_i + \mu.$$

Скаляры  $b_u, b_i, \mu$  выступают в роли bias'a.

# Колаборативная фильтрация: плюсы и минусы

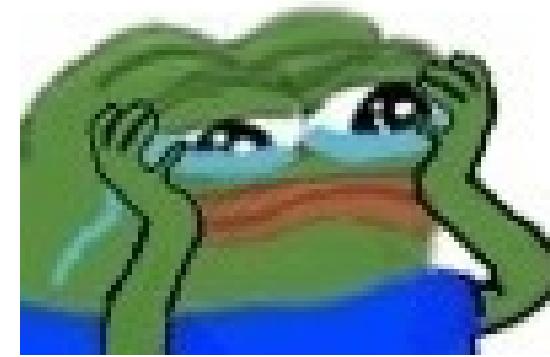
Плюсы:

- рекомендуем новое!
- обучения требуется только матрица взаимодействий
- понятно как деплоить

Минусы:

- проблема холодного старта
- не учитываются сторонние признаки

**Вопросы??**



## Контрастивное обучение: идея

Что если обучать близости

$$\hat{R}_{ui} \sim \langle p_u, q_i \rangle$$

$\implies$

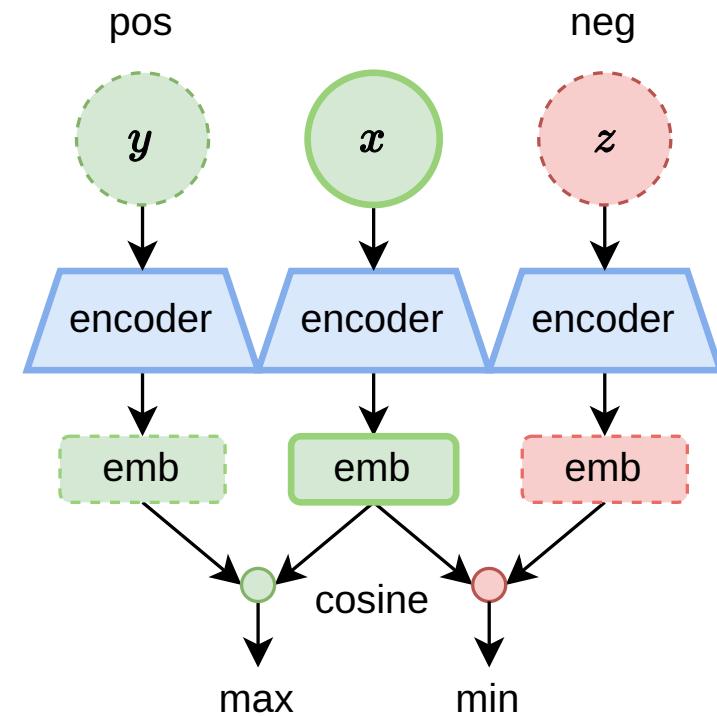
$$\hat{R}_{ui} \sim \langle p_\phi(u), q_\psi(i) \rangle$$

## Контрастивное обучение: лосс

In-batch negative sampling

$$\mathcal{L}_{ui} = -\log \frac{\exp(s(u, i))}{\sum_z \exp(s(u, z))},$$

где  $s(u, i) = \langle p_\phi(u), q_\psi(i) \rangle$



# Контрастивное обучение: плюсы и минусы

Плюсы:

- предсказываем новое
- простота деплоя (опять векторный индекс)
- сторонние признаки
- теплый старт

Минусы:

- сложное обучение
- слабо учитывается контекст и время

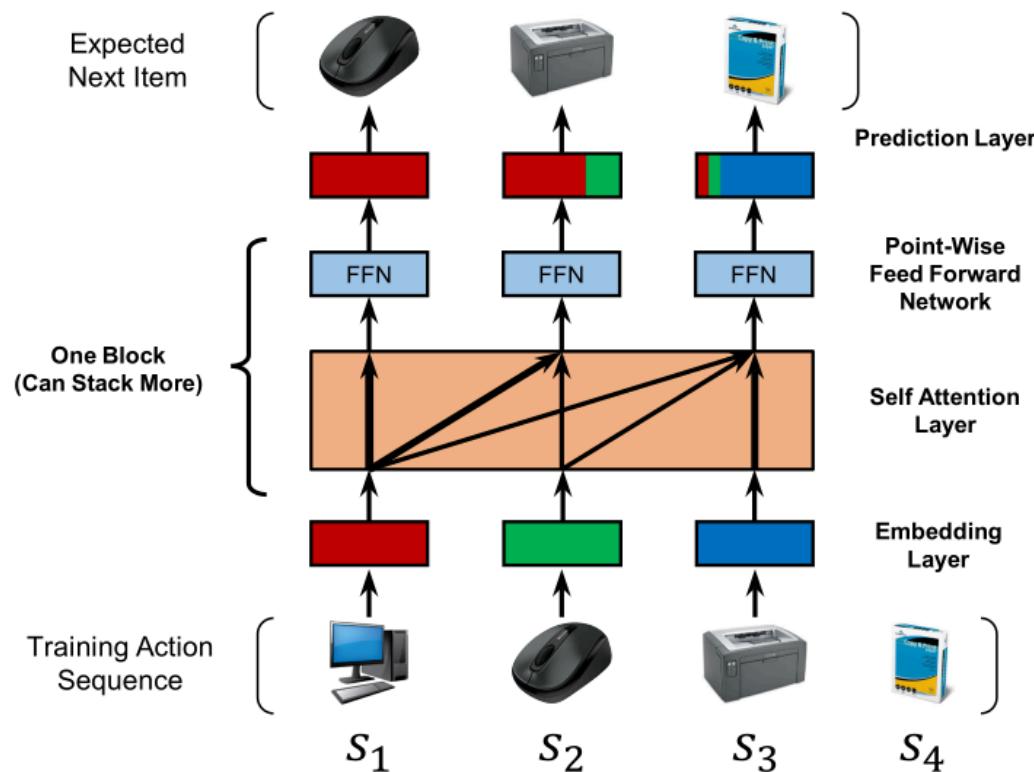
# Sequential Recommenders



# Sequential Recommenders



# SASRec: идея



## **SASRec: связь с колаборативной фильтрацией**

Финальная классификация на число всех предметов:

$$\text{Softmax}(W h_t) \in \mathbb{R}^{n_i},$$

где  $W \in \mathbb{R}^{n_i \times d}$ .

# Вопросы??

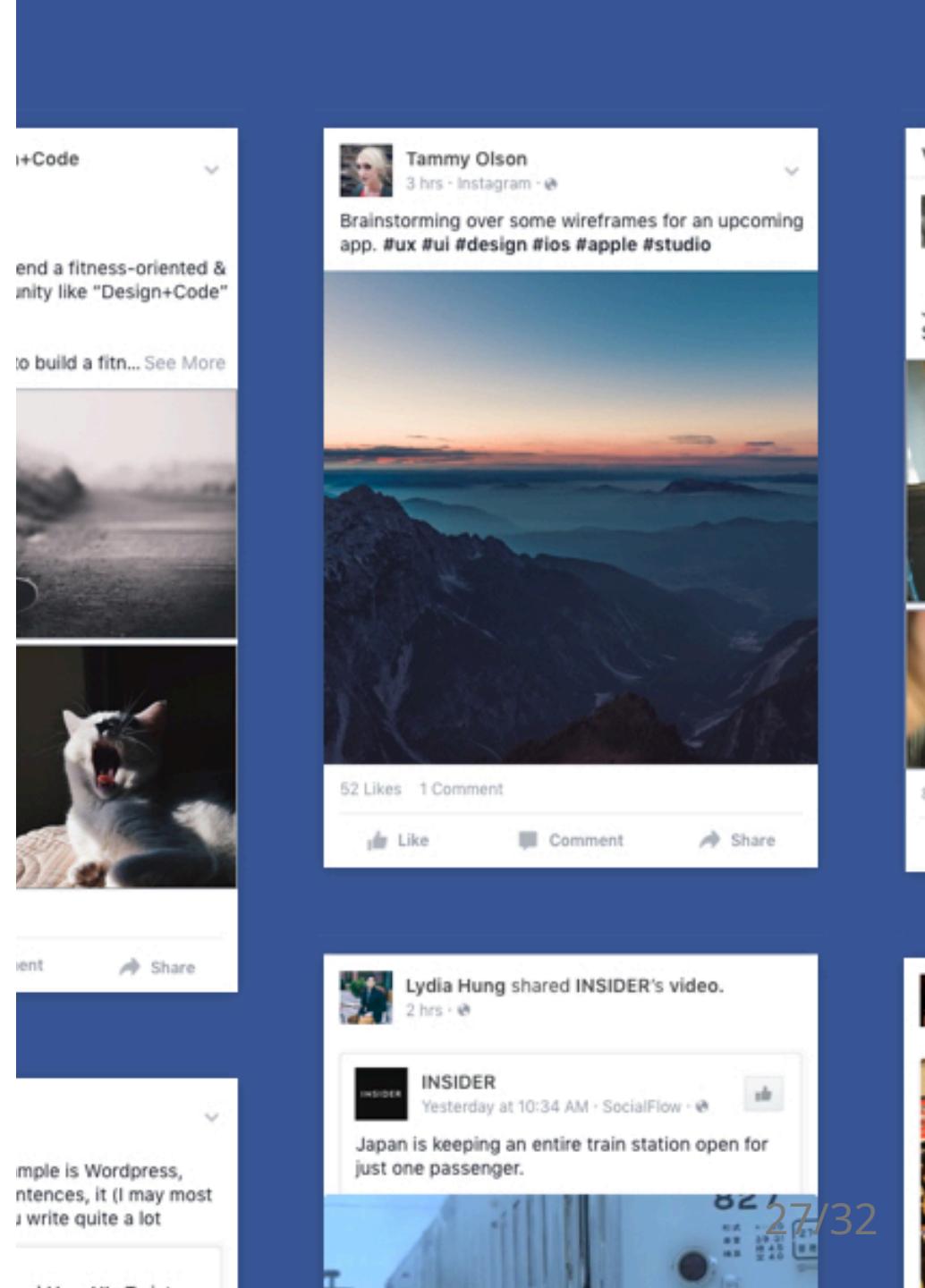


# Ранжирование

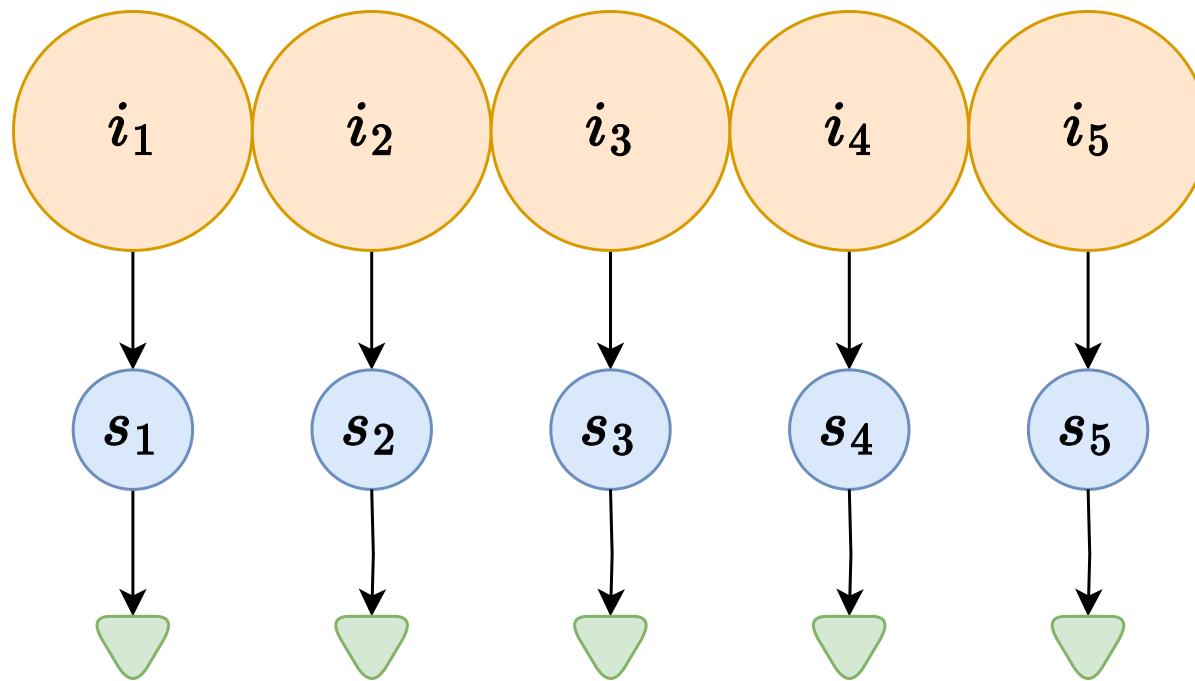
- Методы: pointwise, pairwise, listwise
- LambdaRank

# Ранжирование: постановка задачи

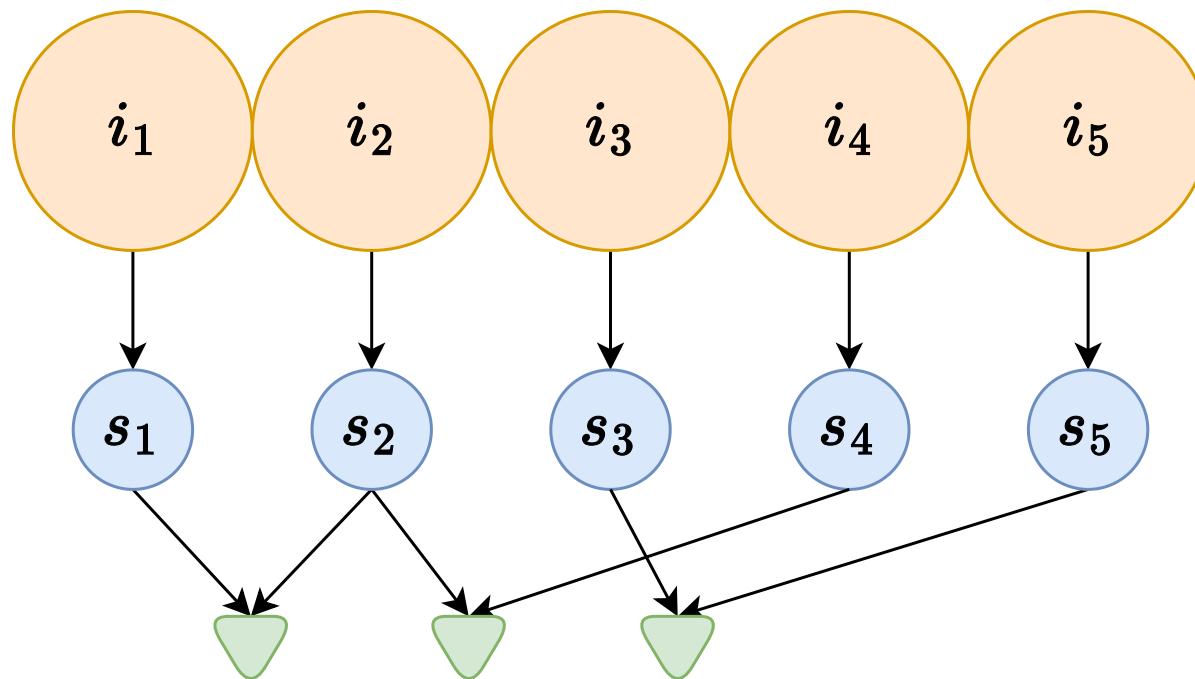
Имеется пользователь  $u$  и набор кандидатов  $\{i_1, \dots, i_k\}$ . Задача: отсортировать (=отранжировать) кандидатов по релевантности.



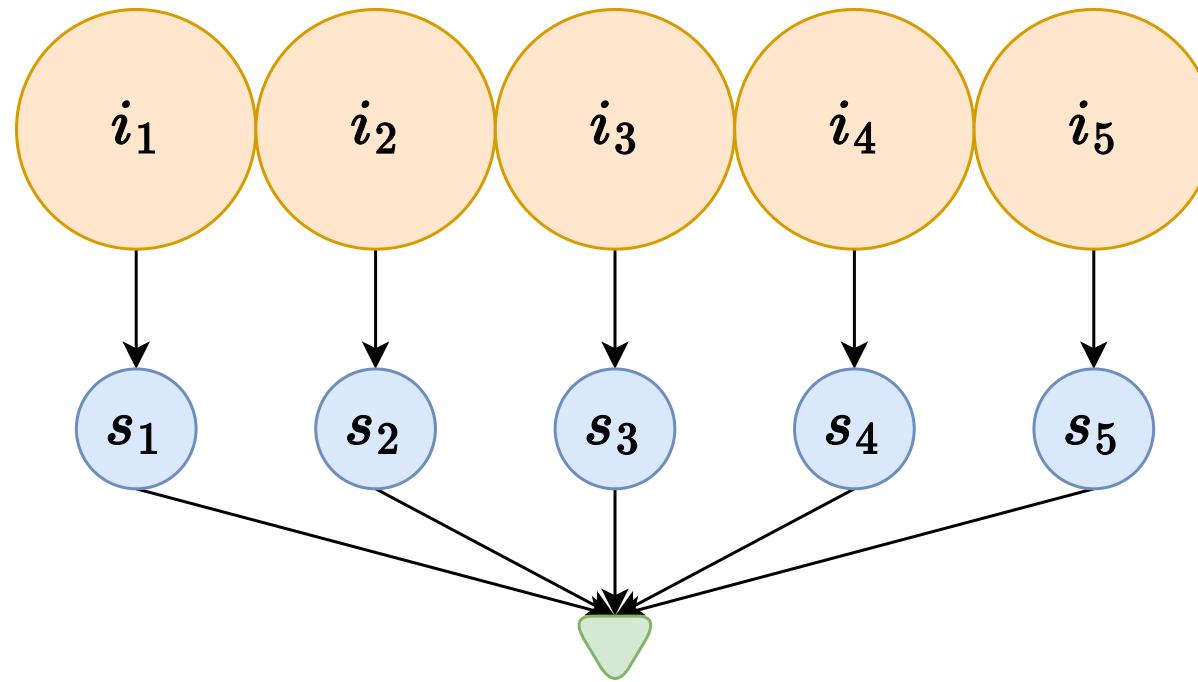
## Pointwise ранжирование



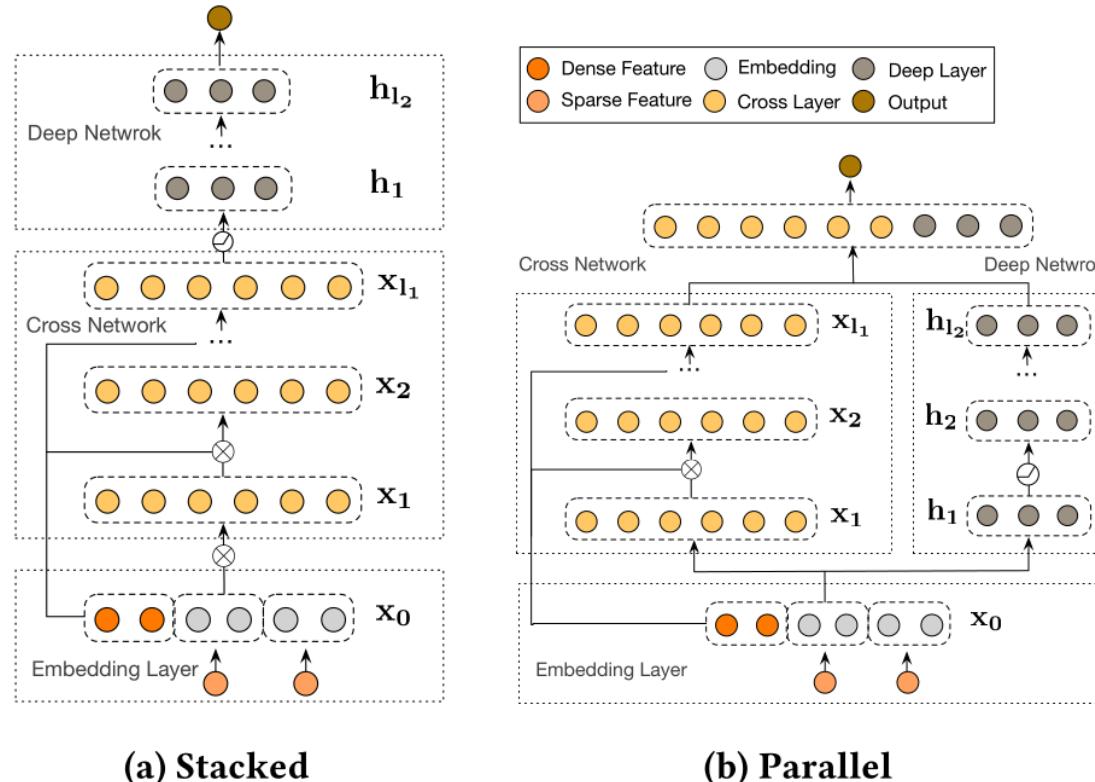
## Pairwise ранжирование



## Listwise ранжирование



# Важнее сеть, а не лосс



**Figure 1: Visualization of DCN-V2.**  $\otimes$  represents the cross operation in Eq. (1), i.e.,  $x_{l+1} = x_0 \odot (W_l x_l + b_l) + x_l$ .

## Вывод LambdaRank

обратимся к доске! если вы еще живы... 💀

