# Введение в рекомендательные системы. часть 2

# Холодный старт

- У нового пользователя нет истории смотрения
- У нового объекта нет истории взаимодействия с пользователями

# Холодный старт. Почему это важно?

- Новинки часто интересны пользователям
- Объекты могут иметь короткий срок жизни на сервисе (Новости)
- Нужно сразу показывать всю привлекательность сервиса для нового пользователя

# Проблема новых объектов

#### Как решить?

- Жанры
- Тэги
- Авторы
- Режиссер, актеры
- Исполнитель
- Страна
- Год производства

и т. д.

# Проблема новых объектов

Обработка сырых данных

- Видео
- Тексты
- Музыка

# Проблема новых пользователей

#### Как решить?

- Использовать информацию о пользователе (пол, возраст, страна и т. д.)
- Онбординг (предложение выбрать интересный контент из списка при первом заходе)

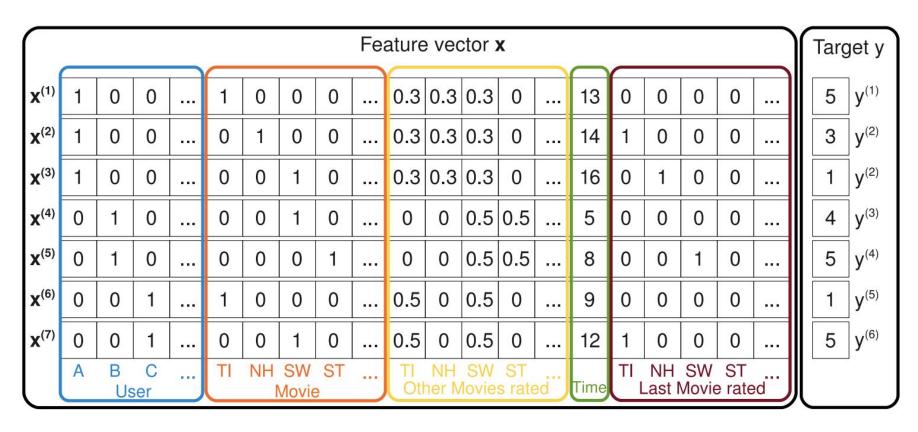
# Content-based рекомендации

Используем схожесть контента по его фичам, а не по совместному смотрению

$$\hat{\mathbf{r}}_{ui} = \bar{r}_i + \frac{\sum_{j \in I_u} sim(i,j)(r_{ui} - \bar{r}_j)}{\sum_{j \in I_u} sim(i,j)}$$
, где  $I_u$  - топ  $\mathbf{N}$ 

айтемов, наиболее похожих на айтем і

#### **Factorization Machines**



### **Factorization Machines**

 $\hat{y}(\mathbf{x}) := w_0 + \sum w_i x_i + \sum \sum \langle \mathbf{v}_i, \mathbf{v}_j \rangle x_i x_j$ 

 $w_0 \in \mathbb{R}, \quad \mathbf{w} \in \mathbb{R}^n, \quad \mathbf{V} \in \mathbb{R}^{n \times k}$ 

i = 1, j = i + 1

i=1

#### **Factorization Machines**

#### Преимущества:

- Гибридный подход. Используем сразу коллаборативную часть и фичи объектов и пользователей
- Не надо думать о проблеме холодного старта для айтемов (не совсем верно)

## Onboarding для новых пользователей

- Самый популярный контент внутри каждого жанра
- Просто выбрать любимые жанры
- Разделить контент на непохожие группы и показать пользователю самый популярный из каждой группы

Onboarding для новых пользователей

Идея:

Разделить контент на непохожие группы и показать пользователю самый популярный из каждой группы

#### Что почитать?

- Recommender Systems Handbook, F. Ricci, 2011
- Recommender Systems: The Textbook, Charu C. Aggarwal, 2016
- Collaborative Filtering for Implicit Feedback Datasets,
  Yifan Hu
- Factorization Machines, Steffen Rendle
- Netflix blog

# Сообщество, конференции, курсы, данные

- RecSys конференция
- Recommender-systems specialization Coursera.org
- Open Data Science канал #recommender\_systems
- Датасет MovieLens
- Million Songs Challenge
- Job Recommendation Challenge

Спасибо за внимание!