

# Введение в рекомендательные системы

Игнатов Дмитрий Игоревич ♦

- ♦ Национальный исследовательский университет Высшая школа экономики  
Факультет компьютерных наук  
Департамент анализа данных и искусственного интеллекта &  
Лаборатория интеллектуальных систем и структурного анализа

2017

# План занятия

На первых двух встречах:

- 1 Введение в рекомендательные системы (РС). Подходы на основе сходства по пользователям и на основе сходства по объектам рекомендации. Оценка качества рекомендательных систем (Точность, полнота, F-мера. Бимодальная кроссвалидация).
- 2 Булева матричная факторизация и сингулярное разложение (SVD). Сравнение результатов на основе метода ближайшего соседа по MAE на примере рекомендации фильмов.
- 3 Гибридные рекомендации на примере сервиса онлайн-радиостанций.

Дополнительные темы.

- 1 Модели SVD, SVD++ и time-SVD. Градиентный спуск.
- 2 Контекстные рекомендации.
- 3 Свободно-распространяемые библиотеки РС.
- 4 Пример РС для краудсорсинга на основе бикластеризации.
- 5 РС на основе бикластеризации и ассоциативных правил.
- 6 Системы совместного пользования ресурсами. РС для фолксономий.

# Оглавление

## 1 Case-study

- Неформальное введение
- User-based и item-based подходы. Оценка качества.
- BMF & SVD
- Рекомендация радиостанций

## 2 Полезные ссылки

- Что почитать?
- Сообщество, конференции, курсы
- Примеры компаний
- Источники данных
- Свободно-распространяемые PC

## 3 Дополнительные темы

- Модель на основе идеи SVD. Градиентный спуск.
- Контекстные рекомендации

# Таксономия рекомендательных систем

- Неперсонифицированные (Non-personalised)
  - ▶ Средняя оценка (пример: <https://www.kinopoisk.ru/>)
- Основанные на фильтрации содержания (Content-based RS)
  - ▶ по сходству описания (пример: <http://www.spellsmell.ru/>)
- Коллаборативная фильтрация (Collaborative Filtering)  
(пример: <https://movielens.org/>)
  - ▶ по сходству пользователей (user-based RS)
  - ▶ по сходству объектов (item-based RS)
  - ▶ матричная факторизация (BMF, NMF, SVD, PLSA, LDA и т.п.)
  - ▶ ассоциативные правила и частые множества (пример объяснения: <https://www.amazon.com/>)
- Гибридные РС

# Рекомендательные системы

<http://Amazon.com>

## Frequently Bought Together



Price For All Three: **\$86.01**

[Add all three to Cart](#)

[Add all three to Wish List](#)

[Show availability and shipping details](#)

- ☒ **This item:** Machine Learning for Hackers by Drew Conway Paperback **\$33.87**
- ☒ Machine Learning in Action by Peter Harrington Paperback **\$25.75**
- ☒ Programming Collective Intelligence: Building Smart Web 2.0 Applications by Toby Segaran Paperback **\$26.39**

## Customers Who Bought This Item Also Bought



Programming Collective Intelligence: Building ...  
> Toby Segaran  
★★★★☆ (84)  
Paperback  
**\$26.39**



Machine Learning in Action  
> Peter Harrington  
★★★★☆ (10)  
Paperback  
**\$25.75**



Mining the Social Web: Analyzing Data from ...  
> Matthew A. Russell  
★★★★☆ (19)  
Paperback  
**\$26.36**



Data Analysis with Open Source Tools  
> Philipp K. Janert  
★★★★☆ (29)  
Paperback  
**\$24.05**



R Cookbook (O'Reilly Cookbooks)  
> Paul Teetor  
★★★★☆ (18)  
Paperback  
**\$32.43**



The Art of R Programming: Tour of Statistical ...  
Norman Matloff  
★★★★☆ (29)  
Paperback  
**\$25.06**

Are any of these items inappropriate for this page? [Let us know](#)

# Рекомендательные системы

<http://Imhonet.ru>

## Оценки фильма Любопытное стечение обстоятельств

Een Bizarre Samenloop Van Omstandigheden, A Curious Conjunction of Coincidences

[Фильмы](#) / [Комедии](#) / [обстоятельств»](#) /



**Да, Вам стоит смотреть фильм «Любопытное стечение обстоятельств»**

Людам, с оценками, похожими на [Ваши](#), этот фильм **нравится**

**А ещё они рекомендуют Вам [31 фильм](#)**

Ваша прогнозируемая оценка фильма после его просмотра  
8.2

Смотрели? Оцените

[Не рекомендовать](#)

[Про фильм](#)

[Онлайн](#)

[Скачать](#)

[Отзывы](#)

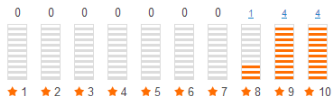
[Персоны](#)

[Кадры](#)

[Оценки](#)

[Похожие](#)

### Распределение оценок



Всего оценок — 9

### Кому больше нравится

Кому нравится:



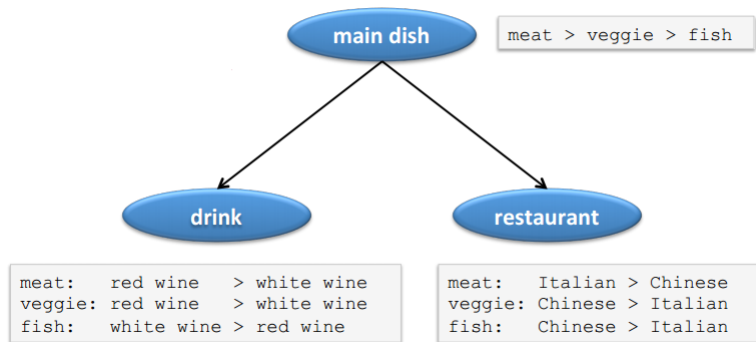
99%

1%



# Обучение предпочтениям

<http://www.preference-learning.org/>



# Таксономия рекомендательных систем

- Неперсонифицированные (Non-personalised)
  - ▶ Средняя оценка
- Основанные на фильтрации содержания (Content-based RS)
  - ▶ по сходству описания
- Коллаборативная фильтрация (Collaborative Filtering)
  - ▶ по сходству пользователей (user-based RS)
  - ▶ по сходству объектов (item-based RS)
  - ▶ матричная факторизация (BMF, NMF, SVD, PLSA, LDA и т.п.)
  - ▶ ассоциативные правила и частые множества
- Гибридные РС



# Рекомендательные системы

<http://Amazon.com>

## Frequently Bought Together



**Price For All Three: \$86.01**

[Add all three to Cart](#)

[Add all three to Wish List](#)

[Show availability and shipping details](#)

- ☒ **This item:** Machine Learning for Hackers by Drew Conway Paperback **\$33.87**
- ☒ Machine Learning in Action by Peter Harrington Paperback **\$25.75**
- ☒ Programming Collective Intelligence: Building Smart Web 2.0 Applications by Toby Segaran Paperback **\$26.39**

## Customers Who Bought This Item Also Bought



Programming Collective Intelligence: Building ...  
> Toby Segaran  
★★★★☆ (84)  
Paperback  
**\$26.39**



Machine Learning in Action  
> Peter Harrington  
★★★★☆ (10)  
Paperback  
**\$25.75**



Mining the Social Web: Analyzing Data from ...  
> Matthew A. Russell  
★★★★☆ (19)  
Paperback  
**\$26.36**



Data Analysis with Open Source Tools  
> Philipp K. Janert  
★★★★☆ (29)  
Paperback  
**\$24.05**



R Cookbook (O'Reilly Cookbooks)  
> Paul Teetor  
★★★★☆ (18)  
Paperback  
**\$32.43**



The Art of R Programming: Tour of Statistical ...  
Norman Matloff  
★★★★☆ (29)  
Paperback  
**\$25.06**

Are any of these items inappropriate for this page? [Let us know](#)

# Рекомендательные системы

<http://Imhonet.ru>

## Оценки фильма Любопытное стечение обстоятельств

Een Bizarre Samenloop Van Omstandigheden, A Curious Conjunction of Coincidences

[Фильмы](#) / [Комедии](#) / [обстоятельств»](#) /



**Да, Вам стоит смотреть фильм «Любопытное стечение обстоятельств»**

Людам, с оценками, похожими на [Ваши](#), этот фильм **нравится**

**А ещё они рекомендуют Вам [31 фильм](#)**

Ваша прогнозируемая оценка фильма после его просмотра  
8.2

Смотрели? Оцените

[Не рекомендовать](#)

[Про фильм](#)

[Онлайн](#)

[Скачать](#)

[Отзывы](#)

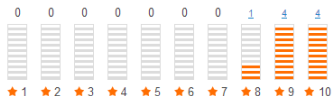
[Персоны](#)

[Кадры](#)

[Оценки](#)

[Похожие](#)

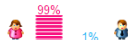
### Распределение оценок



Всего оценок — 9

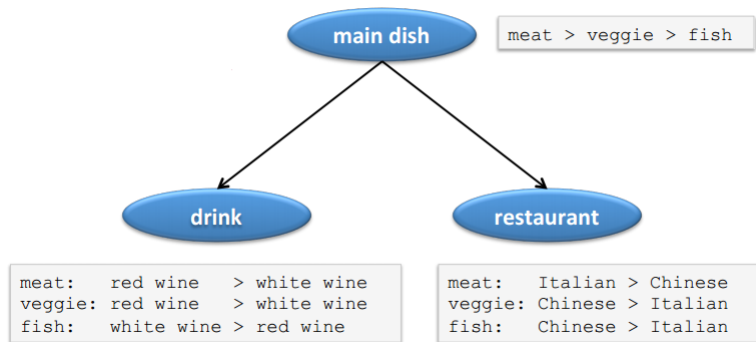
### Кому больше нравится

Кому нравится:



# Обучение предпочтениям

<http://www.preference-learning.org/>



# Case-study 1: Введение

- Введение в рекомендательные системы (РС). Подходы на основе сходства по пользователям и на основе сходства по объектам рекомендации. Оценка качества рекомендательных систем (Точность, полнота, F-мера. Бимодальная кроссвалидация).
- Dmitry I. Ignatov, Jonas Poelmans, Guido Dedene, Stijn Viaene: A New Cross-Validation Technique to Evaluate Quality of Recommender Systems. PerMIn 2012: 195-202.
- Игнатов Д. И., Каминская А. Ю., Магизов Р. А. Метод скользящего контроля для оценки качества рекомендательных интернет-сервисов // КИИ-2010, Т. 1. М. : Физматлит, 2010. С. 175-182.
- Слайды по запросу.

## Case-study 2: BMF & SVD

- Булева матричная факторизация и сингулярное разложение (SVD). Сравнение результатов на основе метода ближайшего соседа по MAE на примере рекомендации фильмов.
- Dmitry I. Ignatov, Elena Nenova, Natalia Konstantinova, Andrey V. Konstantinov: Boolean Matrix Factorisation for Collaborative Filtering: An FCA-Based Approach. AIMS 2014: 47-58
- Elena Nenova, Dmitry I. Ignatov, Andrey V. Konstantinov: An FCA-based Boolean Matrix Factorisation for Collaborative Filtering. FCA Meets IR 2013: P. 57-73
- Слайды

## Case-study 3: Рекомендация радиостанций

- Гибридные рекомендации на примере сервиса онлайн-радиостанций.
- Dmitry I. Ignatov, Sergey I. Nikolenko, Taimuraz Abaev, Natalia Konstantinova: Online Recommender System for Radio Station Hosting: Experimental Results Revisited. WI-IAT (2) 2014: 229-236
- Dmitry I. Ignatov, Andrey V. Konstantinov, Sergey I. Nikolenko, Jonas Poelmans, Vasily Zaharchuk: Online Recommender System for Radio Station Hosting. BIR 2012: 1-12
- Слайды

# Оглавление

## 1 Case-study

- Неформальное введение
- User-based и item-based подходы. Оценка качества.
- BMF & SVD
- Рекомендация радиостанций

## 2 Полезные ссылки

- Что почитать?
- Сообщество, конференции, курсы
- Примеры компаний
- Источники данных
- Свободно-распространяемые РС

## 3 Дополнительные темы

- Модель на основе идеи SVD. Градиентный спуск.
- Контекстные рекомендации

# Что почитать?

- [Recommender Systems Handbook, 2011](#) — “First comprehensive handbook dedicated entirely to the field of recommender systems.”
- [O.Celma. Music Recommendation and Discovery, 2010](#) — “... this book presents the state of the art of all the different techniques used to recommend items, focusing on the music domain as the underlying application.”
- [Тоби Сегаран. Программируем коллективный разум. Глава 2, 2012\(2008\)](#)
- [Greg Linden et al. Amazon.com Recommendations: Item-to-Item Collaborative Filtering. Industry Report. 2003](#)
- [Y. Koren, Matrix Factorization Techniques for Recommender Systems, IEEE Computer, 2009](#)
- [Вводные статьи Сергея Николенко на Habrahabr](#)
- ...



# Сообщество, конференции, курсы

- Серия конференций RecSys
- RecSys Wiki
- International Society for Music Information Retrieval (ISMIR)
- J.A. Konstan, M.D. Ekstrand. Introduction to Recommender Systems. Coursera
- С.Николенко. Рекомендательные системы. Лекции по машинному обучению в КФУ, 2014
- К.В. Воронцов Коллаборативная фильтрация. Курс лекций по машинному обучению
- ...

# Примеры компаний

- IMXONET – рекомендательный сервис
- “Retail Rocket – платформа для мультиканальной персонализации интернет-магазина на основе big data, созданная разработчиками рекомендательных систем Ozon.ru и Wikimart.ru ”
- Surfingbird – рекомендательная система веб-страниц
- “Gravity – Rock Solid Recommendations is a fast-growing company that helps customers with its state-of-the-art recommendation solutions.”
- ...

# Источники данных

- RecSys Challenge 2015
- Group Lens Datasets
- ECML/PKDD Discovery Challenge 2011
- BibSonomy :: dumps for research purposes
- Million Song Dataset Challenge
- Job Recommendation Challenge at Kaggle
- ...

# Свободно-распространяемые РС

- MyMediaLite
- Easyrec
- Python-recsys
- LibRec
- LensKit
- MRec
- SVDFeature
- ...

# Оглавление

## 1 Case-study

- Неформальное введение
- User-based и item-based подходы. Оценка качества.
- BMF & SVD
- Рекомендация радиостанций

## 2 Полезные ссылки

- Что почитать?
- Сообщество, конференции, курсы
- Примеры компаний
- Источники данных
- Свободно-распространяемые РС

## 3 Дополнительные темы

- Модель на основе идеи SVD. Градиентный спуск.
- Контекстные рекомендации

# Модель на основе идеи SVD. Градиентный спуск.

Yehuda Koren, Robert M. Bell: Advances in Collaborative Filtering. Recommender Systems Handbook 2011: 145-186

- Базовая модель:

$$r_{ui} = \mu + b_i + b_u + q_i^T p_u + e_{ui}$$

- Целевая функция:

$$\min_{b, q, p} = \sum_{(u, i) \in R} (r_{ui} - \mu - b_i - b_u - q_i^T p_u)^2 + \lambda(b_i^2 + b_u^2 + \|q_i\|^2 + \|p_u\|^2), \text{ где}$$

$$R = \{(u, i) \mid \text{оценка } r_{ui} \text{ определена}\}$$

- Метод градиентного спуска:

$$e_{ui} = r_{ui} - \hat{r}_{ui}$$

$$b_u \leftarrow b_u + \gamma \cdot (e_{ui} - \lambda \cdot b_u)$$

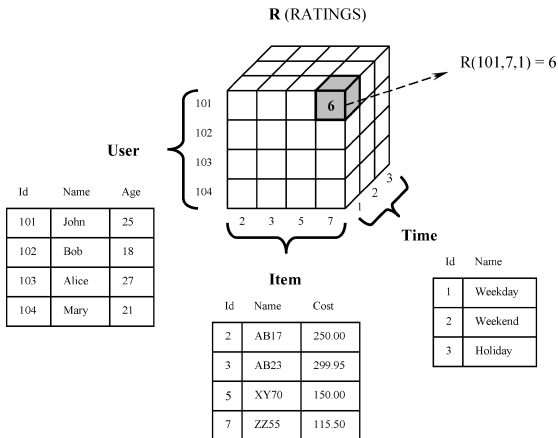
$$b_i \leftarrow b_i + \gamma \cdot (e_{ui} - \lambda \cdot b_i)$$

$$q_i \leftarrow q_i + \gamma \cdot (e_{ui} \cdot p_u - \lambda \cdot q_i)$$

$$p_u \leftarrow p_u + \gamma \cdot (e_{ui} \cdot q_i - \lambda \cdot p_u)$$

# Контекстные рекомендации

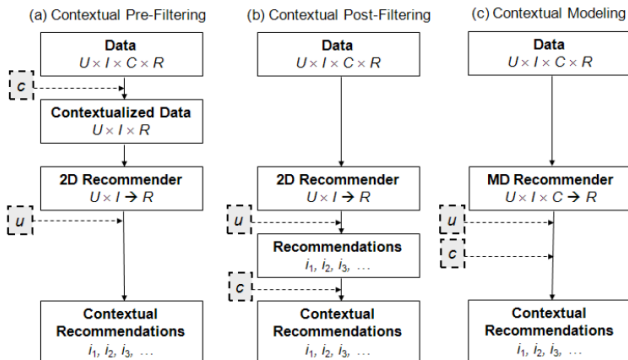
Gediminas Adomavicius, Alexander Tuzhilin: Context-Aware Recommender Systems. Recommender Systems Handbook 2011: 217-253



Многомерная модель  $User \times Item \times Time$  рекомендательного пространства  
(контекстная информация – время)

# Контекстные рекомендации

G. Adomavicius, A. Tuzhilin, 2011(2005)



Способы встраивания контекста в пространство рекомендаций



# Вопросы и контакты

[www.hse.ru/staff/dima](http://www.hse.ru/staff/dima)

Спасибо!

dmitrii.ignatov[at]gmail.com

dignatov[at]hse.ru