

Прикладные задачи анализа данных

Рекомендательные системы

Дьяконов А.Г.

**Московский государственный университет
имени М.В. Ломоносова (Москва, Россия)**



Системы рекомендаций



★★★★★ 5.0 из 5

Матерь Тьма ✓ В наличии

Скидка 20%

~~188 р.~~ 150 р.

Добавить в корзину

Автор: [Воннегут К.](#)

Серия: [Эксклюзивная классика](#)

Жанр: [Классическая проза](#)

Издательство: [Издательство «АСТ»](#)

ISBN: 978-5-17-099474-8

Артикул: p1600862

Возрастное ограничение: 16+

Похожие товары



Системы рекомендаций

**что интересно данному пользователю
в данный момент времени
в данном контексте**

Системы рекомендаций

товары	книги фильмы музыка игры приложения
контент	новости сайты статьи
досуг	рестораны отели театральные представления выставки туры
социальные связи	друзья группы

Виды рекомендаций

по контенту Content-based	Рекомендация похожих по описанию товаров
коллаборативная фильтрация Collaborative Filtering	Рекомендация по статистике покупок Проблема холодного старта: новый товар новый пользователь
гибридная Hybrid	

Информация

Описание пользователя

+ лог пользователя (поиск, ожидания и т.п.)

Описание товара

Взаимодействие (пользователь, товар)

Взаимодействие (пользователь, пользователь)

Взаимодействия (товар, товар)

Что рекомендуют

заменители (alternative)
сопутствующие товары (cross sell)
бандлы
аксессуары (up sell)
популярные товары (best sellers)
персональные / неперсональные

оффлайн / онлайн

Как рекомендуют

max вероятность покупки
max матожидание прибыли
товары из категории (long-tail)

Разные каналы рекомендаций

сайт

почта

смс, приложения

Сбор данных

явный (explicit)

- оценка объекта
- ранжирование группы объектов
 - выбор одного товара из двух
- создание списка любимых объектов

неявный (implicit)

- что искал, смотрел, клал в корзину, купил
 - лог поведения
- анализ содержимого компьютера

Мифы о рекомендательных системах

**Если улучшить ленту на главной странице,
то покупки с неё увеличатся на 30%**

Меньше 10% смотрят на ленту главной страницы

**70% всех покупок совершаются по рекомендациям
2/3 всех фильмов смотрят по рекомендациям**

**Настоящая эффективность рекомендаций (сколько покупок только
благодаря им) меньше 10%**

Соревнование Netflix

2006 год

~ 100.5 миллионов оценок 1,2,...,5

~ 480 000 пользователей

17 770 фильмов

RMSE

Netflix = 0.9514

надо = 0.8563

~ 20 000 участников

RBM = 0.8990

SVD = 0.8914

Для бизнеса > 0.88

Content-based-методы

**Если есть хорошие признаковые описания пользователей и объектов
(и только они), тогда**

$$u \sim f_u$$
$$i \sim f_i$$

Можно решать как обычную задачу обучения с учителем

$$\{([f_u, f_i], r_{ui})\}$$

Цель: $u \rightarrow i_1, \dots, i_k : \hat{r}_{ui_1} \geq \hat{r}_{ui_2} \geq \dots$

Content-based-методы

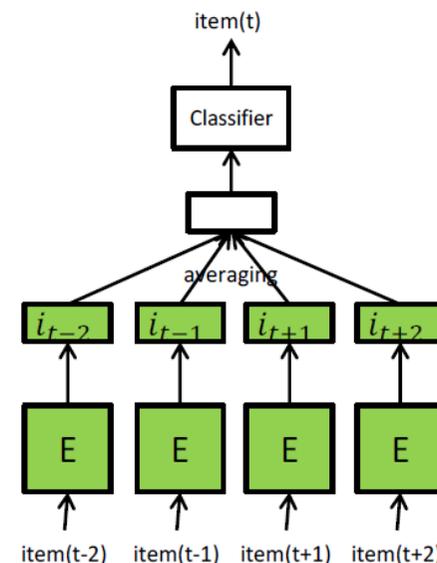
Хорошие признаки

- даны

пол, возраст, рейтинг, число лайков и т.п.

- **word2vec, GloVe**
- **матричные разложения**
- **Deep Walk (графовые)**
- **Автокодировщики**

word2vec → **prod2vec**



Коллаборативная фильтрация

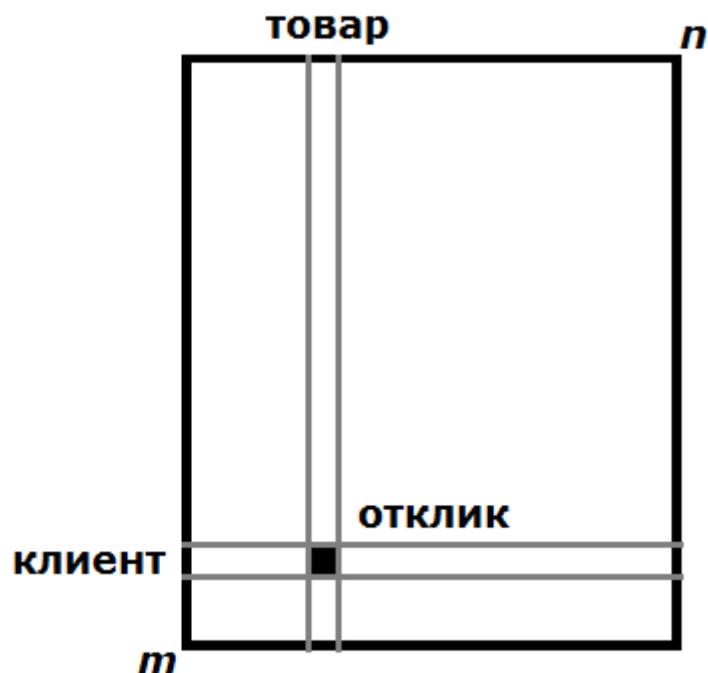
Если известна лишь статистика:

$$\{(u, i, r_{ui})\}$$

нет содержательных признаков!

**Решение на статистике поведения лучше,
чем на описаниях!**

Статистика



	item1	item2	item3	item4
user1	1	2	5	
user2		2		5
user3	3	3	5	
user4		4		5
user5	5		3	

Разреженная матрица

Цель: фактически уметь дозаполнять матрицу...

GroupLens-алгоритм

По пользователям (User-based)

$$\hat{r}_{ui} = \bar{r}_u + \frac{\sum_v \text{sim}(u, v)(r_{vi} - \bar{r}_v)}{\sum_v \text{sim}(u, v)}$$

По товарам (Item-based)

$$\hat{r}_{ui} = \bar{r}_i + \frac{\sum_j \text{sim}(i, j)(r_{uj} - \bar{r}_j)}{\sum_j \text{sim}(i, j)}$$

Проблема холодного старта

Плохие предсказания, если мало статистики

Долгие вычисления (нужен пересчёт)

Похожесть

корреляция

$$\text{sim}(u, v) = \frac{\sum_i (r_{ui} - \bar{r}_u)(r_{vi} - \bar{r}_v)}{\sqrt{\sum_i (r_{ui} - \bar{r}_u)^2} \sqrt{\sum_i (r_{vi} - \bar{r}_v)^2}}$$

но м.б.

похожесть по описанию,

похожесть по кластерам,

...

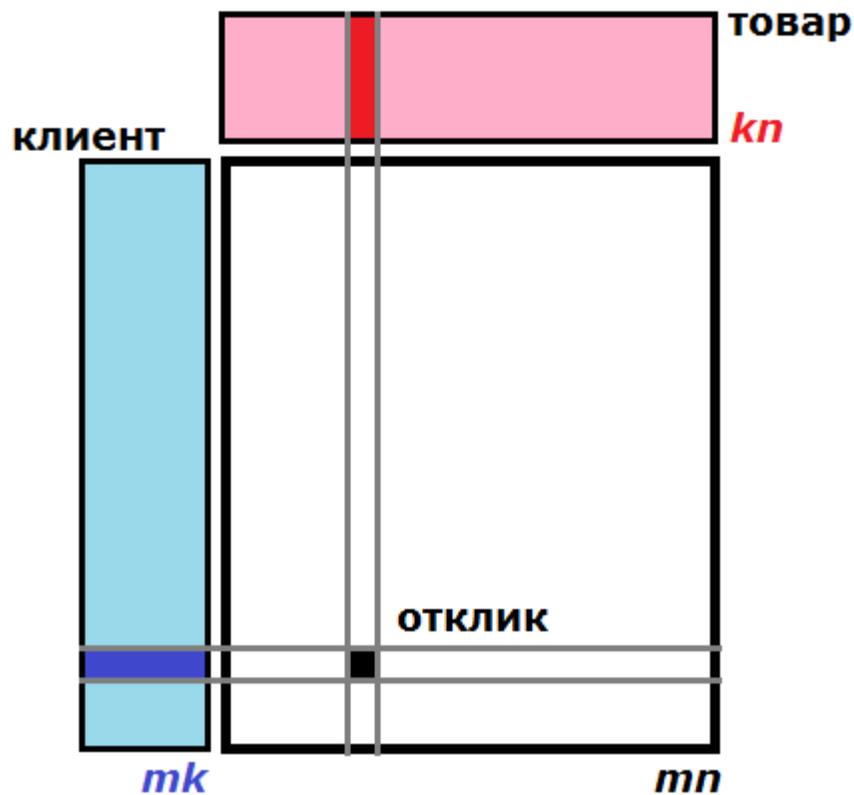
Не обязательно такую близость...

выбор k самых близких

YouTube

$$\text{sim}(i, j) = \frac{\text{view}(\{i, j\})}{\text{view}(\{i\}) \cdot \text{view}(\{j\})}$$

SVD



$$R = U \cdot \Lambda \cdot V^T$$

$$R_{m \times n} \approx U_{m \times k} \cdot \Lambda_{k \times k} \cdot V_{n \times k}^T$$

$$R \approx U' \cdot V'$$

$$\hat{r}_{ui} = \langle p_u, q_i \rangle$$

SVD

$$r_{u,i} \approx \langle p_u, q_i \rangle$$

$$J = \sum_{(u,i)} (\langle p_u, q_i \rangle - r_{u,i})^2 + \lambda_1 \sum_u \|p_u\|^2 + \lambda_2 \sum_i \|q_i\|^2$$

Одновременно получили признаковое описание пользователей и товаров

Минимизация

- градиентный спуск
- **ALS (Alternating Least Squares)**

$$p_u(t+1) = \left(\sum_{i:r_{u,i}>0} (\langle q_i, q_i \rangle + \lambda_1 I) \right)^{-1} \left(\sum_{i:r_{u,i}>0} r_{u,i} q_i \right)$$

Улучшения модели

$$r_{u,i} \approx r + r_u + r_i + \langle p_u, q_i \rangle$$

**Учитываем смещения
«добрый/злой» пользователь
«плохой/хороший» товар**

SVD++

$$r_{u,i} \approx r + r_u + r_i + \left\langle p_u + \frac{1}{\sqrt{|\text{view}(u)|}} \sum_{j \in \text{view}(u)} y_j, q_i \right\rangle$$

+ что просматривал, но не покупал пользователь

**Легко обобщать на разное число факторов:
(пользователь, канал, товар)**

Simon Funk статья в блоге во время конкурса Netflix

timeSVD++

Неизвестные зависят от времени...

Когда нет явного отклика

**Если оценки даны не в шкале,
а перечислены только отклики на услугу...**

$$\{(u, i, 1)\}$$

иногда решение вырождается в константное

выход: пропуски = нули

На практике:

часто знаем, что видел пользователь...

и почему-то не отреагировал

содержание рассылки

баннеры на странице

Факторизационные машины



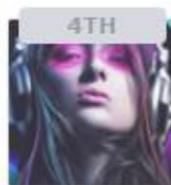
1st/239



4th/657



4th/163



4th/133

Steffen Rendle**libFM: Factorization Machine Library**<http://www.libfm.org/>**Супермодель, иммитирует****SVD, SVD++, FPMC, Pairwise interaction tensor factorization,
SVM с полином. ядром и т.п.**

Ask Peter Norvig

Q5: What, say, 3 recent papers in machine learning do you think will be influential to directing the cutting edge of research these days? (41 Up-votes, 26.08.2014)

I've never been able to pick lasting papers in the past, so don't trust me now, but here are a few:

Rendle's "Factorization Machines"

Wang et al. "Bayesian optimization in high dimensions via random embeddings"

Dean et al. "Fast, Accurate Detection of 100,000 Object Classes on a Single Machine"

Факторизационные машины

	Feature vector x															Target y						
$x^{(1)}$	1	0	0	...	1	0	0	0	...	0.3	0.3	0.3	0	...	13	0	0	0	0	...	5	$y^{(1)}$
$x^{(2)}$	1	0	0	...	0	1	0	0	...	0.3	0.3	0.3	0	...	14	1	0	0	0	...	3	$y^{(2)}$
$x^{(3)}$	1	0	0	...	0	0	1	0	...	0.3	0.3	0.3	0	...	16	0	1	0	0	...	1	$y^{(2)}$
$x^{(4)}$	0	1	0	...	0	0	1	0	...	0	0	0.5	0.5	...	5	0	0	0	0	...	4	$y^{(3)}$
$x^{(5)}$	0	1	0	...	0	0	0	1	...	0	0	0.5	0.5	...	8	0	0	1	0	...	5	$y^{(4)}$
$x^{(6)}$	0	0	1	...	1	0	0	0	...	0.5	0	0.5	0	...	9	0	0	0	0	...	1	$y^{(5)}$
$x^{(7)}$	0	0	1	...	0	0	1	0	...	0.5	0	0.5	0	...	12	1	0	0	0	...	5	$y^{(6)}$
	A	B	C	...	TI	NH	SW	ST	...	TI	NH	SW	ST	...	Time	TI	NH	SW	ST	...		
	User				Movie					Other Movies rated						Last Movie rated						

$$r_{ui} \sim w_0 + w_u + w_i + v_u^T v_i$$

модель второго порядка:

$$w_0 + \sum_{i=1}^n w_i x_i + \sum_{1 \leq i < j \leq n} v_i^T v_j x_i x_j \sim w_0 + w^T x + x^T \underbrace{W}_{\sim \text{rg}=k} x$$

«факторизация» – в предположении, какая у нас матрица весов, иначе была бы просто «модель второго порядка»

Факторизационные машины

Что ещё...

- факторизация отдельных блоков (FFM – field-aware factorization machine)
- эффективное блочное хранение

FFM – field-aware factorization machine



Линейная модель

$$\phi(\mathbf{w}, \mathbf{x}) = \mathbf{w}^T \mathbf{x} = \sum_{j \in C_1} w_j x_j$$

Полиномиальная модель (Poly2)

$$\phi(\mathbf{w}, \mathbf{x}) = \sum_{j_1, j_2 \in C_2} w_{j_1, j_2} x_{j_1} x_{j_2}$$

Факторизационная машина

$$\phi(\mathbf{w}, \mathbf{x}) = \sum_{j_1, j_2 \in C_2} \langle \mathbf{w}_{j_1}, \mathbf{w}_{j_2} \rangle x_{j_1} x_{j_2}$$

Факторизационная машина с полями

$$\phi(\mathbf{w}, \mathbf{x}) = \sum_{j_1, j_2 \in C_2} \langle \mathbf{w}_{j_1, f_2}, \mathbf{w}_{j_2, f_1} \rangle x_{j_1} x_{j_2}$$

Оптимизационная задача

$$\min_{\mathbf{w}} \sum_{i=1}^L \left(\log(1 + \exp(-y_i \phi(\mathbf{w}, \mathbf{x}_i))) + \frac{\lambda}{2} \|\mathbf{w}\|^2 \right),$$

$$\phi(\mathbf{w}, \mathbf{x}) = \sum_{j_1, j_2 \in C_2} \langle \mathbf{w}_{j_1, f_2}, \mathbf{w}_{j_2, f_1} \rangle x_{j_1} x_{j_2},$$

LogLoss + регуляризация

Что такое поля...

Field name		Field index
User	→	field 1
Movie	→	field 2
Genre	→	field 3
Price	→	field 4

Что ещё?

- неотрицательные матричные разложения
 - вероятностные разложения
 - специальные регуляризаторы
 - локальная низкоранговость
 - бикластеризация
- тензоры (тензорное разложение)

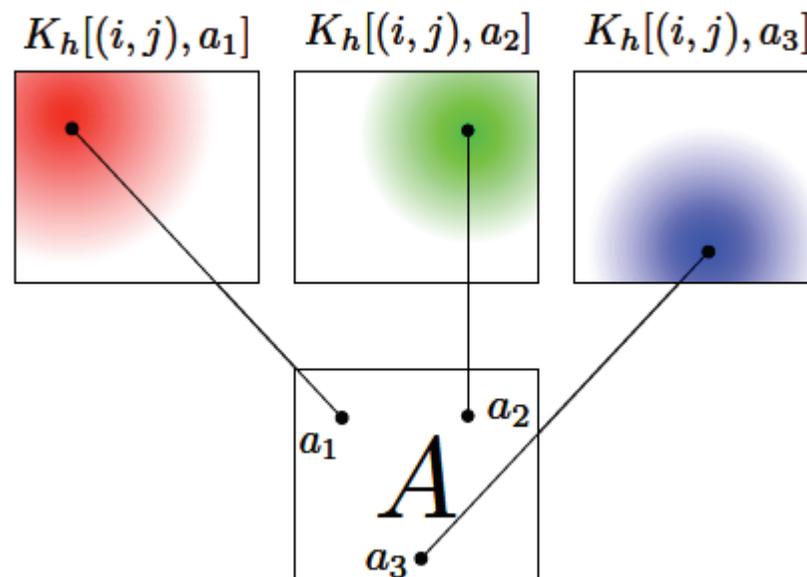
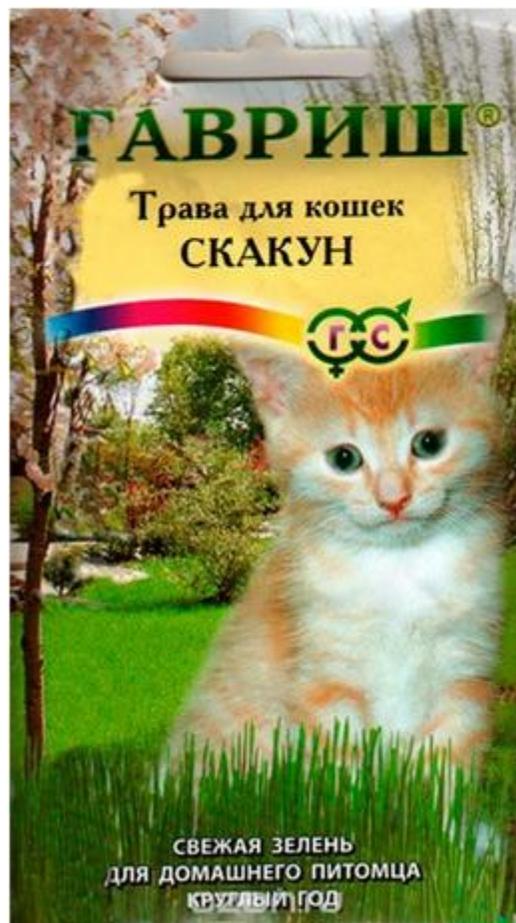


рис. из дипломной работы М.Трофимова

Простые методы



Трава для кошек Скакун, 10 г

Тип	Комнатные растения
Вид	Разнообразные комнатные
Время посадки в грунт	Январь, Февраль, Март, Апрель, Май, Июнь, Июль, Август, Сентябрь, Октябрь, Ноябрь, Декабрь
Время урожая	Январь, Февраль, Март, Апрель, Май, Июнь, Июль, Август, Сентябрь, Октябрь, Ноябрь, Декабрь
Назначение	Для контейнеров

15 ₺

[Добавить в корзину](#)

Вместе с этим товаром покупают



Трава для кошек Скакун,
10 г

+



Фигус Притупленный, 3
шт.

+



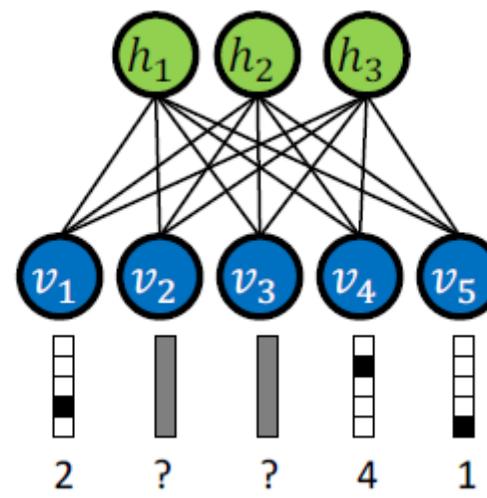
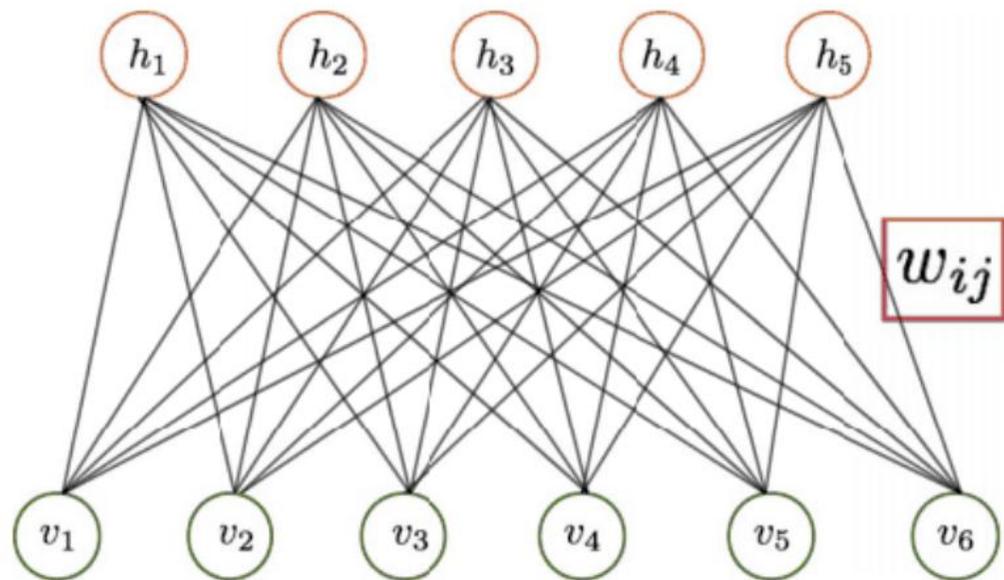
Нолина (бокарнея отогнутая) Бутылочное
дерево, 3 шт.

= 85 ₺

[В корзину](#)

Бандлы ~ по статистике

RBM



Spotify

Рекуррентные нейронные сети для предсказания последовательности треков

Функционалы качества

Уже было...

- **RMSE (Netflix)**
- **Precision, Recall**
- **NDCG**

Желаемые свойства рекомендаций

Разнообразие (diversity) ~ непохожие на другие товары из списка

Плохо: к ноутбуку только ноутбуки того же производителя

Новизна (novelty) ~ для пользователя

Плохо: каждый день одно и то же

Доверие ~ обосновать рекомендацию

«с товаром покупают», «скидка за комплект», ...

Контекст

канал захода / просмотра / покупки
состояние корзины / счёта / предыстория
география
время (года, суток)
погода и т.п.

Рекомендовать надо только то, что без рекомендации не купит...

Неидентифицируемые пользователи + новые товары

многорукие бандиты

Замечание

Тоже есть подготовка данных

- удаление выбросов**

**слишком популярные товары /
активные пользователи (оптовики)**

CASE: LenKor

**Пример решения (рекомендательных) задач
методом ближайшего соседа**

Задача «Predict Grant Applications»

Прогноз результата выполнения гранта

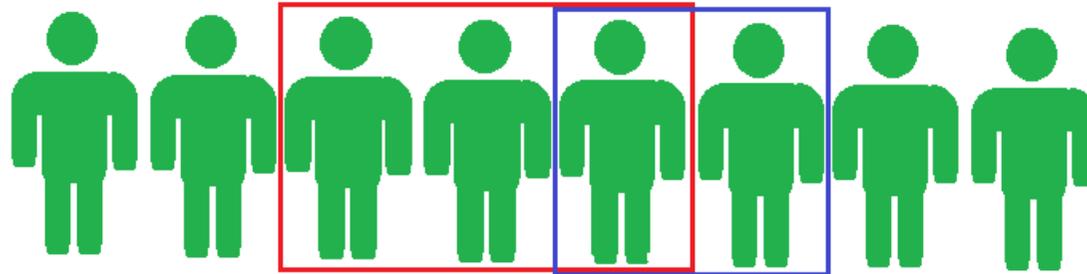
'Grant.Application.ID'	'RFCD.Percentage.4'	'Person.ID.1'
'Grant.Status'	'RFCD.Code.5'	'Role.1'
'Sponsor.Code'	'RFCD.Percentage.5'	'Year.of.Birth.1'
'Grant.Category.Code'	'SEO.Code.1'	'Country.of.Birth.1'
'Contract.Value.Band...see.note.A'	'SEO.Percentage.1'	'Home.Language.1'
'Start.date'	'SEO.Code.2'	'Dept.No..1'
'RFCD.Code.1'	'SEO.Percentage.2'	'Faculty.No..1' 'With.PHD.1'
'RFCD.Percentage.1'	'SEO.Code.3'	'No..of.Years.in.Uni.at.Time.of.Grant.1'
'RFCD.Code.2'	'SEO.Percentage.3'	'Number.of.Successful.Grant.1'
'RFCD.Percentage.2'	'SEO.Code.4'	'Number.of.Unsuccessful.Grant.1'
'RFCD.Code.3'	'SEO.Percentage.4'	'A..1'
'RFCD.Percentage.3'	'SEO.Code.5'	'A.1'
'RFCD.Code.4'	'SEO.Percentage.5'	'B.1'
		'C.1'

По описанию гранта ~ будет ли его выполнение успешным.

	Название	Область	Фин.	Коллектив	Статьи А
11	Топологические инварианты	021 – 100%	300.000	Иванов	10
				Пеший	3
12	Написание рекомендательной системы	217 – 60%	550.000	Печенкин	2
		218 – 49%		Белых	1
				Абашидзе	0

Технология LENKOR - именно для этой задачи и была разработана!

1. Не ясно, как измерить похожесть проектов
2. Но ясно, как измерить похожесть коллективов, спонсоров, областей и т.д.

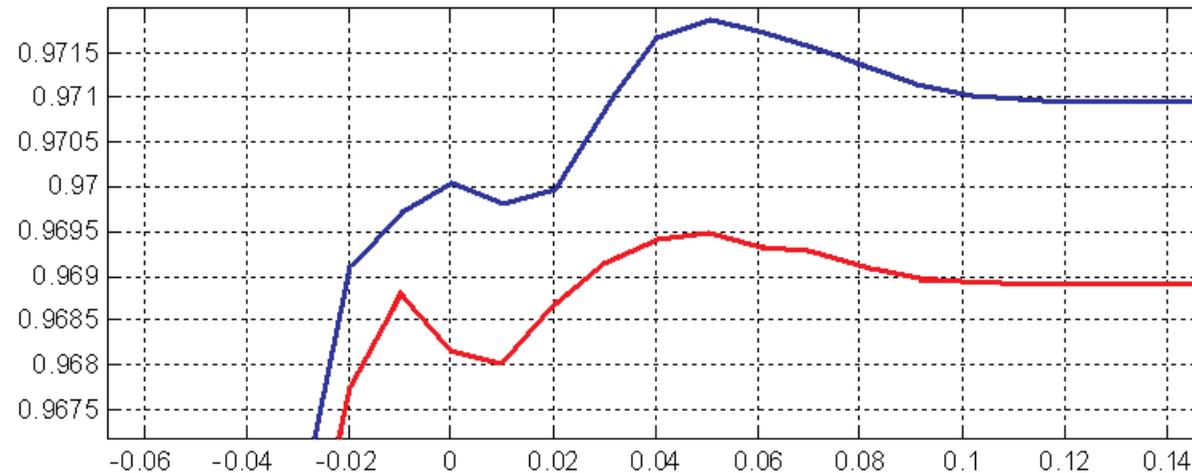


3. Близость (метрика) = сумма близостей (метрик)

$$B(x, x_i) = \sum_{\omega} c_{\omega} B_{\omega}(x, x_i)$$

4. Коэффициенты можно настраивать
5. Потом можно добавить нелинейность

Вариация коэффициента при фиксированных остальных



Добавление признака «язык» в линейную комбинацию

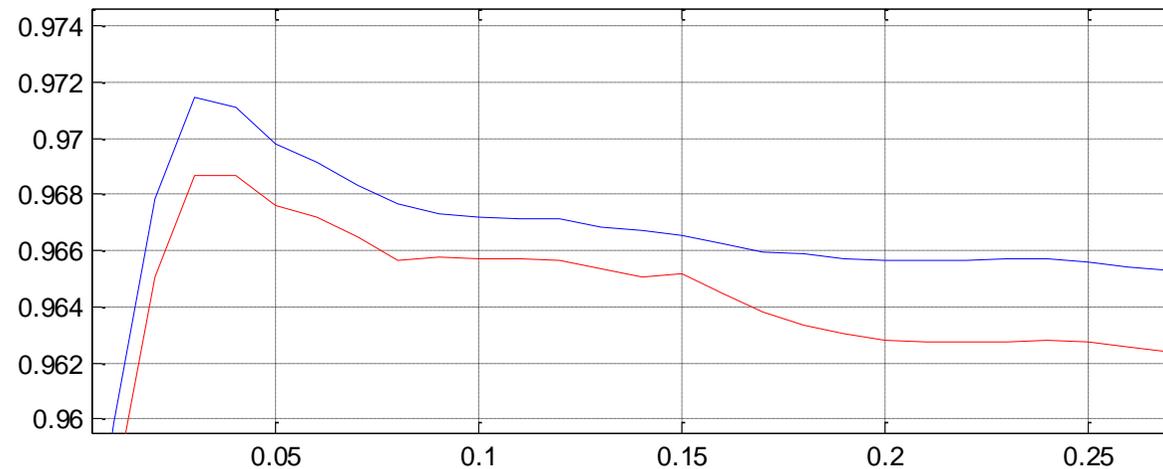
Сразу смотрим значение **нужного функционала качества!**

Настройка напоминает ту, что описана выше...

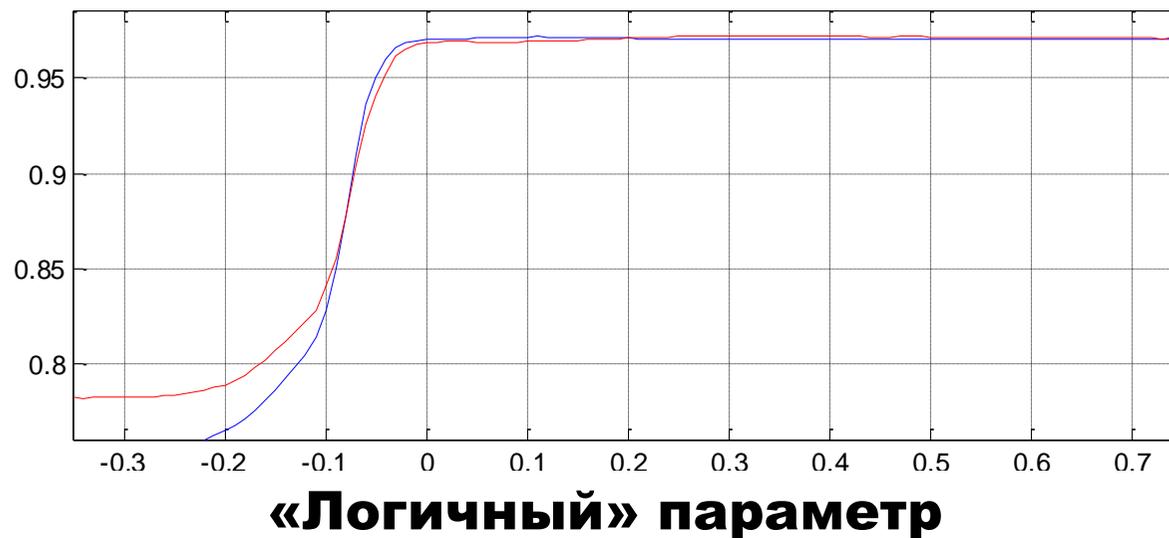
Применима во многих задачах...

Например, прогноз объёмов продаж в зависимости от рекламы.

Фиксация всех параметров, кроме одного (Grant)

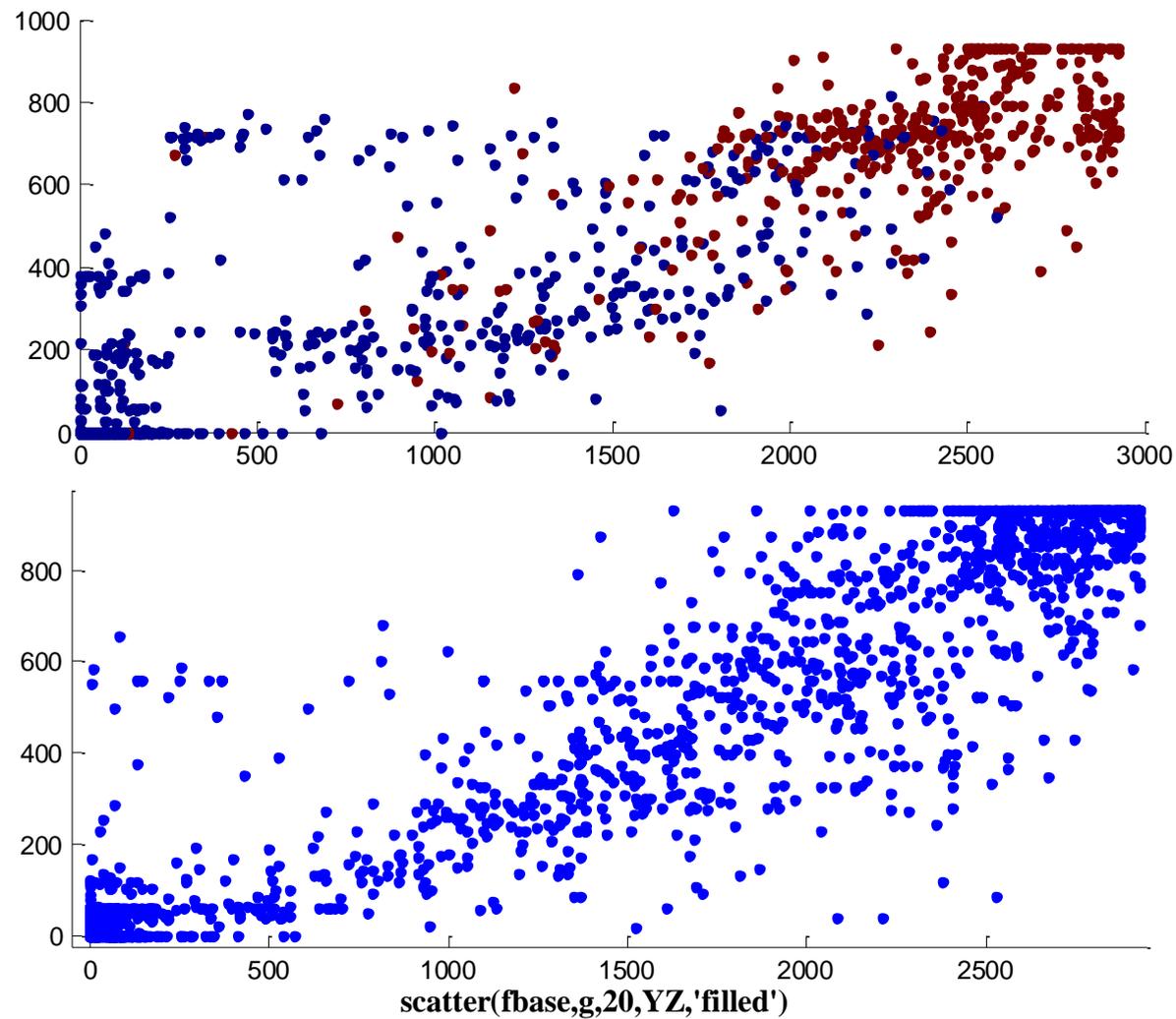


Можно улучшить (здесь: обучение и контроль)

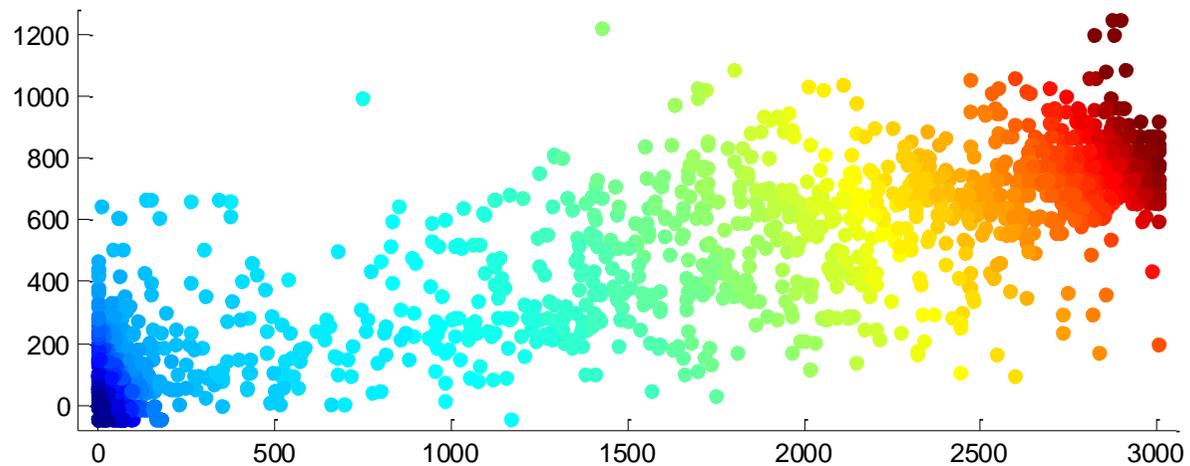
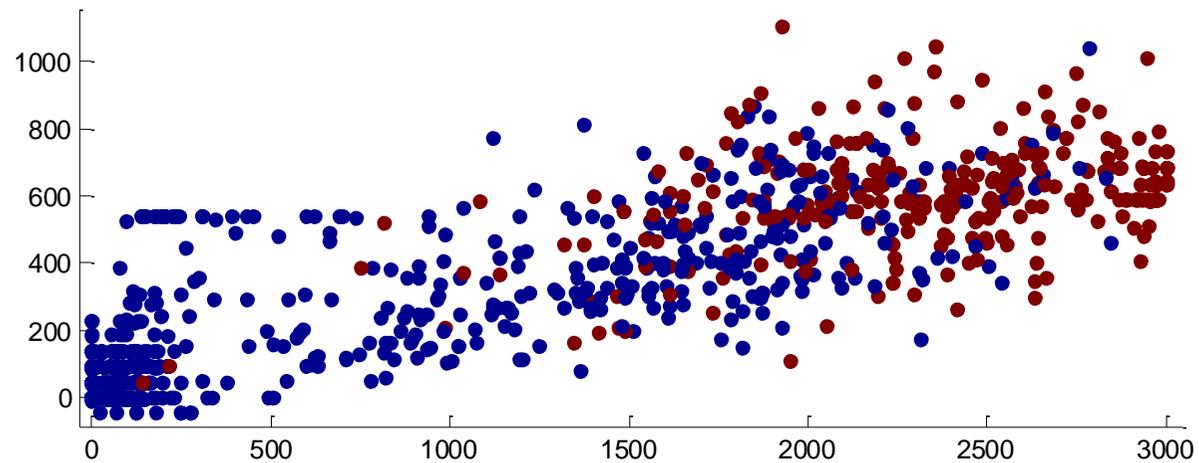


«Логичный» параметр

Кстати, ещё раз «как выглядят ответы»



Что послал (Grant)



**«VideoLectures.Net Recommender System Challenge»
(ECML/PKDD Discovery Challenge 2011)**

Рекомендация лекций для просмотра

Задание соревнования – написать рекомендательную систему для ресурса VideoLectures.net

Первый подконкурс (cold start) – необходимо по **одной просмотренной лекции рекомендовать лекции из множества «**НОВЫХ** лекций», которые были недавно выложены на сайт, и для них ещё нет статистики популярности, только подробное описание.**

N. Antulov-Fantulin, M. Bošnjak, T. Šmuc, M. Jermol, M. Žnidaršič, M. Grčar, P. Keše, N. Lavrač, ECML/PKDD 2011 - Discovery challenge: "VideoLectures.Net Recommender System Challenge", <http://tunedit.org/challenge/VLNetChallenge/>

Описание лекции

101, 'Lecture', 'eng', 'biology', '2008-12-04', '2009-02-12', 'Implementing a common framework on business', 'Professor Rudolf Smith', ...

Функционал качества

$$\frac{1}{|Z|} \sum_{z \in Z} \frac{|\{r_1, \dots, r_{\min(S, R, z)}\} \cap \{s_1, \dots, s_{\min(S, R, z)}\}|}{\min(S, R, z)}$$

r_1, \dots, r_R – **рекомендации**

s_1, \dots, s_S – **правильные ответы**

$Z = \{5, 10, 15, 20, 25, 30\}$

Надо рекомендовать 30 лекций.

Нахождение метрики

$$\rho(\text{Lecture}_1, \text{Lecture}_2) = c_1 \cdot \rho_1(\text{Author}_1, \text{Author}_2) + c_2 \cdot \rho_2(\text{Title}_1, \text{Title}_2) + \dots \\ + c_n \cdot \rho_n(\text{Subject}_1, \text{Subject}_2)$$

Не обязательно метрики по непересекающимся описаниям

Название

Аннотация

Текст

Название + Аннотация

Название + Аннотация + Текст

Дьяконов А.Г. Алгоритмы для рекомендательной системы: технология LENKOR // Бизнес-Информатика, 2012, №1(19), С. 32–39.

A. D'yakonov Two Recommendation Algorithms Based on Deformed Linear Combinations // Proc. of ECML-PKDD 2011 Discovery Challenge Workshop, pp 21-28 (2011).

Что использовано в решении

Близость двух лекций оценивалась используя **только**

1. Близость категорий.
2. Близость авторов.
3. Близость языков.
4. Близость названий.
5. Близость названий, описаний, названий и описаний событий.

+ статистика

m_{ij} – число пользователей: смотрели и i -ю лекцию и j -ю лекцию

Не использовано

Аналогичные данные по событиям (конференции, на которых они прочитаны, школы-семинары, циклы лекций и т.д.)

Описания слайдов

Даты съёмки

Обработка текста

**Использовалось приведение к общей основной форме
(стеммер Портера)**

Нестандартное TF-IDF-преобразование – изменение качества на 5%.

Porter, 1980, An algorithm for suffix stripping, Program, Vol. 14, № 3, pp. 130–137.

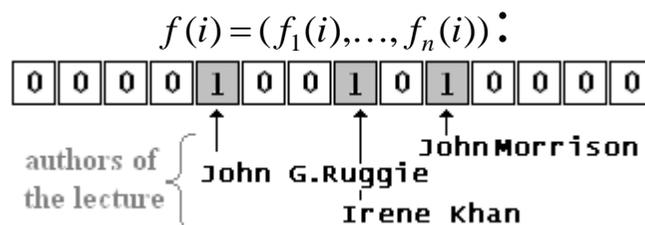
<http://tartarus.org/~martin/PorterStemmer/>

Как строилась метрика

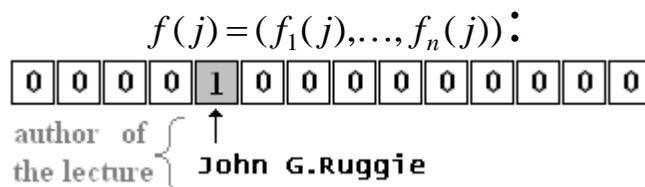
n – число авторов всех лекций,

$f(i) = (f_1(i), \dots, f_n(i))$ – бинарный вектор, описывает авторов i -й лекции:
единицы помечают номера соответствующих авторов.

```
7445, 'debate', 'en', 7436, 25, '2008-12-04', '2009-02-11', 'Questions, Reactions from the audience',
NULL, NULL
```



```
7442, 'lecture', 'en', 7436, 112, '2008-12-04', '2009-02-11', 'Implementing a common framework on
business and human rights', 'Professor John Ruggie, UN Special Representative on Business and
Human Right\nIrene Khan, Secretary General, Amnesty International\nModerated by: John
Morrison, Programme Director, Business Leaders Initiative on Human Rights\n(BLIHR)', NULL
```



Измененная косинусная мера

$$\langle f(i), f(j) \rangle = \frac{f_1(i)f_1(j) + \dots + f_n(i)f_n(j)}{\sqrt{f_1(i)^2 + \dots + f_n(i)^2 + \varepsilon} \sqrt{f_1(j)^2 + \dots + f_n(j)^2 + \varepsilon}}$$

$$\gamma(i, j) = \left\langle \sum_{t \in I} \left(\frac{m_{it}}{\sum_{s=1}^L m_{is}} \cdot \frac{f(t)}{\sqrt{f_1(t)^2 + \dots + f_n(t)^2 + \varepsilon}} \right), f(j) \right\rangle$$

from the co-view statistics

– близость между **новой** j -й лекции и **старой** i -й лекцией (точнее, похожими на неё **старыми** лекциями с точки зрения пользователей)

I – множество индексов «старых лекций»,

m_{ij} – число пользователей, которые просмотрели и i -ю лекцию и j -ю лекцию при $i \neq j$, и m_{ii} – число пользователей, которые просмотрели i -ю, делённое пополам

Почему: пользователь посмотрел раздел «Биология», а новых лекций в нём нет... нечего рекомендовать?!

Окончательное решение

Вычисляем близость по формуле:

$$\gamma = 0.19 \cdot \sqrt{0.6 \cdot \gamma_{\text{cat}} + 5.6 \cdot \gamma_{\text{auth}}} + \sqrt{4.5 \cdot \gamma_{\text{lang}} + 5.8 \cdot \gamma_{\text{dic}} + 3.1 \cdot \gamma_{\text{dic2}}} \cdot$$

Получено перебором различных форм ответа:

$$\gamma = C_1 \cdot \gamma_{\text{cat}} + C_2 \cdot \gamma_{\text{auth}} + C_3 \cdot \gamma_{\text{lang}} + C_4 \cdot \gamma_{\text{dic}} + C_5 \cdot \gamma_{\text{dic2}}$$

$$\gamma = C_1 \cdot \gamma_{\text{cat}} + \sqrt{C_2 \cdot \gamma_{\text{auth}} + C_3 \cdot \gamma_{\text{lang}} + C_4 \cdot \gamma_{\text{dic}} + C_5 \cdot \gamma_{\text{dic2}}}$$

$$\gamma = C_1 \cdot \gamma_{\text{cat}} + C_2 \cdot \gamma_{\text{auth}} + (C_3 \cdot \gamma_{\text{lang}} + C_4 \cdot \gamma_{\text{dic}})^2 + C_5 \cdot \gamma_{\text{dic2}}$$

$$\gamma = (C_1 \cdot \gamma_{\text{cat}} + C_2 \cdot \gamma_{\text{auth}}) + C_3 \cdot \gamma_{\text{lang}} + \sqrt{C_4 \cdot \gamma_{\text{dic}} + C_5 \cdot \gamma_{\text{dic2}}}$$

При решении задачи оптимизации использовался метод
покоординатного спуска

Рекомендуем 20 лекций с наибольшими значениями γ .

Окончательное решение

Решение, выложенное на сайте

$$\left(\gamma_1 \cdot \left(1.07 - 0.07 \frac{t_1 - t_{\min}}{t_{\max} - t_{\min}} \right), \dots, \gamma_N \cdot \left(1.07 - 0.07 \frac{t_N - t_{\min}}{t_{\max} - t_{\min}} \right) \right),$$

Rank	Team	Time of Submission	Preliminary Result	Final Result
1	+ D'yakonov Alexander	Jul 08, 09:27:22	0.37281	0.35857
2	+ lefman	Jul 07, 00:24:15	0.31063	0.30743
3	+ Nitram	Jul 08, 06:47:32	0.30661	0.27684
4	sofos	Jul 06, 23:22:55	0.27433	0.27151
5	+ Inner Peace	Jul 08, 10:52:35	0.27268	0.25773
6	+ DMIR	Jul 08, 11:39:37	0.26813	0.25498
7	+ ddi	Jul 08, 11:07:37	0.26298	0.24920
8	+ Haibin Liu	Jul 08, 08:45:05	0.25172	0.24559
9	+ tao	Jul 08, 09:50:24	0.22465	0.24044
10	+ Team SIG	Jul 08, 10:25:03	0.22465	0.24044

t_j – время выкладывания
на сайт j -й новой
лекции,

t_{\min} – минимальное t_j ,

t_{\max} – максимальное
(вычислялось в днях).

**Увеличение качества
примерно на 5%.**

Литература

**Дьяконов А.Г. Алгоритмы для рекомендательной системы: технология LENKOR // Бизнес-Информатика, 2012, №1(19), С. 32–39.
[https://bijournal.hse.ru/2012--1\(19\)/53535879.html](https://bijournal.hse.ru/2012--1(19)/53535879.html)**

Y. Koren, R.M. Bell, C. Volinsky Matrix Factorization Techniques for Recommender Systems // IEEE Computer 42(8): 30-37 (2009).

S. Funk Netflix Update: Try This at Home // <http://sifter.org/~simon/journal/20061211.html>

libFM: Factorization Machine Library // <http://www.libfm.org/>

FFM – field-aware factorization machine (слайды) // <http://www.csie.ntu.edu.tw/~r01922136/slides/ffm.pdf>