



ТЕХНОСФЕРА

Лекция 7 Recomendation Systems

Владимир Гулин

24 марта 2018 г.

План лекции

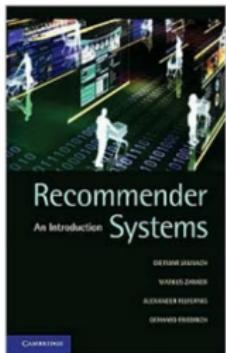
Зачем это все нужно?

Эмпирические алгоритмы

Подключаем науку

Пример 1

Товары



Recommender Systems An Introduction

DETTMAR JANNACH
MARKUS ZANKER
ALEXANDER FELFERNIG
GERHARD FRIEDRICH

Table of Contents

Customers who bought this also bought



Recommender Systems: An Introduction

by Dietmar Jannach, Markus Zanker, Alexander Felfernig, Gerhard Friedrich

AVERAGE CUSTOMER RATING:

★★★★★ (Be the first to review)



Registrieren, um sehen zu können, was deinen Freunden gefällt.

FORMAT:

Hardcover

NOOKbook (eBook) - not available

[Tell the publisher you want this in NOOKbook format](#)

NEW FROM BN.COM

\$65.00 List Price

\$52.00 Online Price

(You Save 20%)

Add to Cart

NEW & USED FROM OUR

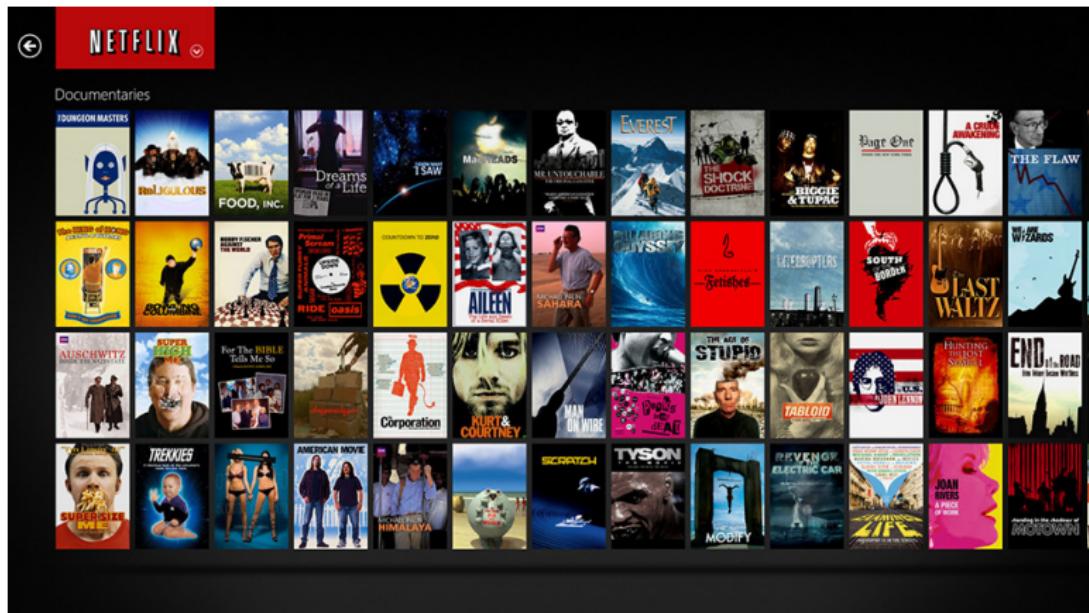
New starting at **\$56.46** (You S

Used starting at **\$51.98** (You t

See All Prices

Пример 2

Фильмы



Пример 3

Друзья

Поиск друзей

Расширенный поиск

Начните вводить любое имя, название или слово



Марина Попова
5 общих друзей



Темирхан Омаров
6 общих друзей



Юлия Гришко
9 общих друзей



Ксения Рагозина
6 общих друзей



Алексей Петров
10 общих друзей



Алина Романова
9 общих друзей



Евгений Миронов
5 общих друзей



Родион Насакин
6 общих друзей



Светлана Данильченко
10 общих друзей

Пример 4

Работа

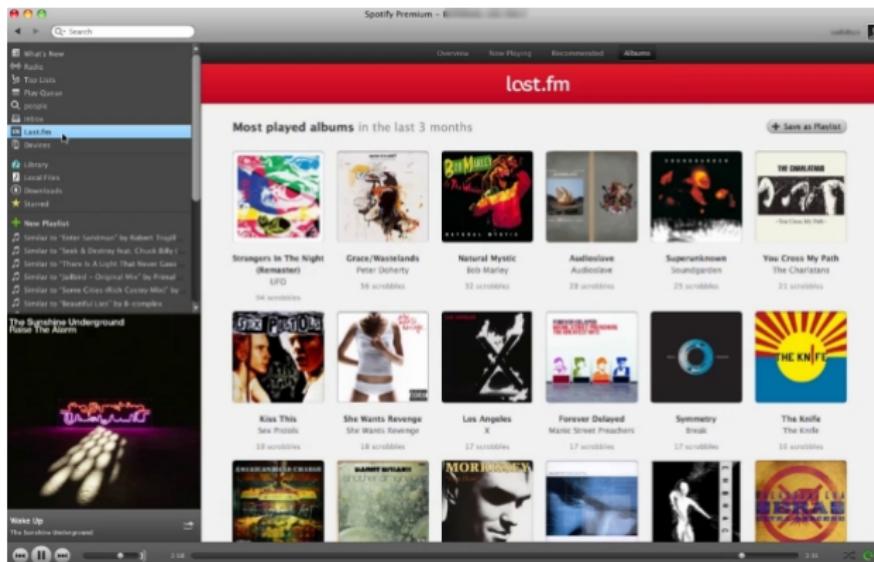
Jobs You May Be Interested In *beta*

	Research Scientist, NLP, Machine... Adchemy - San Francisco Bay Area	
	Data Scientist (PL) LinkedIn - Mountain View, CA	
	Senior Principal Data Scientist Symantec - San Francisco Bay Area	

[Feedback | See more »](#)

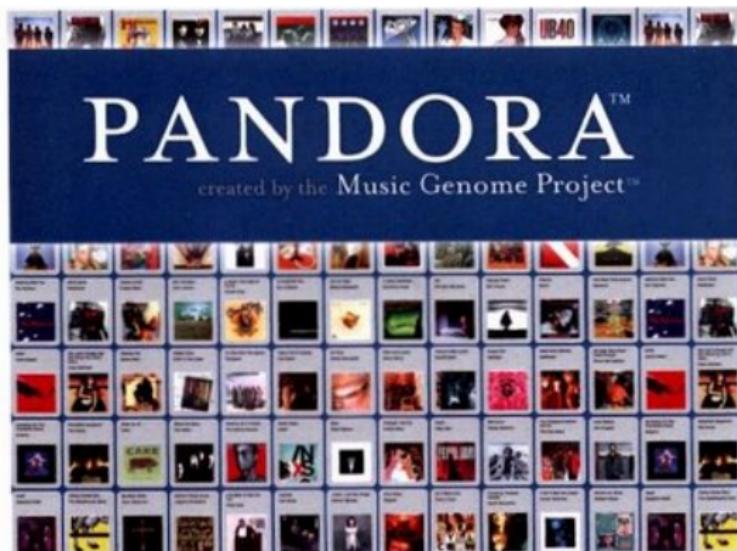
Пример 5

Музыка



Пример 6

Радио



Пример 7

И даже персонализированный поиск о_0

[Windows Phone \(United States\)](#)

www.windowsphone.com/

"I love my Windows Phone. The stuff that's important to me is pinned to my Start screen, like the Live Tile for my All Recipes app which updates regularly with ...

My Phone - Windows Phone 7 - Canada - India

[Windows Phone Apps+Games Store \(United ... - WindowsPhone.com](#)

www.windowsphone.com/en-us/store

Browse Apps+Games for Windows Phone. ... Popular and Famous. This collection features the most famous and innovative apps in the Store. First Row ...

[Windows Phone - Wikipedia, the free encyclopedia](#)

en.wikipedia.org/wiki/Windows_Phone

Windows Phone is a family of mobile operating systems developed by Microsoft, and is the successor to its Windows Mobile platform, although incompatible with ...

[Phones | Windows Phone \(United States\)](#)

www.windowsphone.com/en-us/phones

Features · Phones · Apps+Games · News · How-to. Explore. My Phone. Have a Windows Phone? Sign in · My Phone · Find My Phone · Purchase history ...

[Windows Phone 8 - New Windows 8 Phones, Windows 7 ... - AT&T](#)

www.att.com › Wireless › Devices

AT&T has the newest Windows Phone 8 & Windows 7 Phones, including smartphones and devices from HTC, Samsung, and Nokia. Choose the best Windows 8 ...

[Windows Phone \(windowsphone\) on Twitter](#)

<https://twitter.com/windowsphon>

The Official Windows Phone Twitter Channel - keeping you updated with the latest ... If iOS bores you and Android intrigues you, #WindowsPhone will likely ...

[Windows Phone \(United States\)](#)

www.windowsphone.com/

"I love my Windows Phone. The stuff that's important to me is pinned to my Start screen, like the Live Tile for my All Recipes app which updates regularly with ...

My Phone - Windows Phone 7 - Canada - India

[Windows Phone Apps+Games Store \(United ... - WindowsPhone.com](#)

www.windowsphone.com/en-us/store

Browse Apps+Games for Windows Phone. ... Popular and Famous. This collection features the most famous and innovative apps in the Store. First Row ...

You've visited this page 2 times. Last visit: 10/29/12

[Yes, you can use Google+ through Internet Explorer on Windows ...](#)

<https://plus.google.com/.../posts/bc57m7wedyF>

Danny Sullivan – 46 minutes ago – Yes, you can use Google+ through Internet Explorer on Windows Phone 8, but it's not a pleasant experience, I've found. I can haz app please

[Windows Phone - Wikipedia, the free encyclopedia](#)

en.wikipedia.org/wiki/Windows_Phone

Windows Phone is a family of mobile operating systems developed by Microsoft, and is the successor to its Windows Mobile platform, although incompatible with ...

[Images for windows phone - Report images](#)



Рекомендации в MRG

Одноклассники

Сообщения Обсуждения Оповещения Другие Пости События Музыка Видео Поиск

От меня привет! 0 OK

Напишите заметку

Фото Видео Трансляция Настроение
Объявление

Всё Рекомендации Обсуждаемое Избранное Друзья Группы Игры Поиск

Одноклассники

Украсить фото

Жора КИРПИЧ Хинтон
Лента
Рекомендации
Друзья
Фото
Группы
Игры
Подарки
Товары
Объявления
Фотоконкурс

Юла 12+

Популярно сейчас

Чеховские Медведи - Спартак
36 065 просмотров

Чтобы оставить комментарий, вам нужно войти в систему.

Евгений Журин написать
Александр Шарак написать
Алексей Воропаев написать
Игорь Алферов написать
Дмитрий Бугайченко написать

Твиты 50 классов

3 короля e-commerce

16:30 MTS RUS

Смартфоны на Android
Покупай по безопасной сделке

2,9 км

4 900 ₽

Куртка зимняя му...

Юла Прод

3,2 км

900 ₽

Книга Гарри потер...

16:28 MTS RUS

PANDAO

Поиск товаров в Китае

Лучшее Товары за баллы Красота и зд.

60% 68%

62 205 заказов ★★★★★

Наушники с экранированным пр...

98-р 49 ₽

30 146 заказов ★★★★★

Подставкой для мобильных телефон...

133-р 16 ₽

Кооперативная улица, Зк5

Сортировка и фильтры

IL Патио

★★★★★ Заказ от 300 ₽ • Итальянская

Delivery Express

45 мин

IL Патио

★★★★★ Заказ от 300 ₽ • Итальянская

Delivery Express

45 мин

BB&Burgers

★★★★★

DC Hot

45 мин

10 / 59

С чего все началось?

Netflix Prize

- ▶ 480189 users
- ▶ 17770 movies
- ▶ 100480507 оценок {1, 2, 3, 4, 5}
- ▶ 2.10.2006 - 21.09.2009
- ▶ Призовой фонд 1 000 000\$
- ▶ Задача улучшить RMSE на 10% (с 0.9514 до 0.8563)

Специфика задачи построения рекомендательных систем

- ▶ Отсутствует традиционное признаковое описание
- ▶ Имеется в наличии очень малое количество данных
- ▶ Данные смещены в сторону положительных рекомендаций
- ▶ Практически невозможно задействовать ассесоров

Типы рекомендательных систем

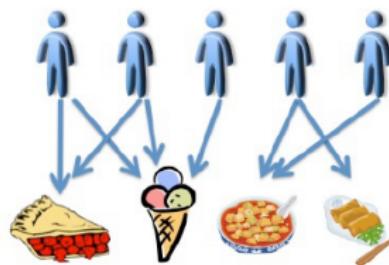
Collaborative Filtering

Content-based systems

Knowledge-based systems

Collaborative Filtering

- ▶ Имеются данные о предпочтениях пользователей (пользователь-объект)



Collaborative Filtering

Восстанавливаем пропущенные значения матрицы

Collaborative Filtering

Алгоритмы:

- ▶ Neighborhood Information
 - ▶ User-based
 - ▶ Item-based
- ▶ Model-based
 - ▶ SVD
 - ▶ Latent factor models
 - ▶ Factorization machines
 - ▶ Compositions

Постановка задачи

Дано:

- ▶ Users, $u \in U$
- ▶ Items, $i \in I$
- ▶ Marks, $r_{ui} \in R$
- ▶ Events, $(u, i, r_{ui}, \dots) \in \mathcal{K}$

Задача:

- ▶ Выделить похожие объекты

$$i \rightarrow \{i_1, \dots, i_{l_s}\}$$

- ▶ Найти похожих пользователей

$$u \rightarrow \{u_1, \dots, u_{U_s}\}$$

- ▶ Предсказать

$$\hat{r}_{ui} = h(u, i, \dots) \approx r_{ui}$$

Как оценить рекомендации?

Метрики:

- ▶ MAE
- ▶ RMSE
- ▶ Precision@k
- ▶ Recall@k
- ▶ Map@k
- ▶ DCG@k



Метрики

MAE

$$MAE = \frac{1}{k} \sum_{i=1}^k |r_i - y_i|$$

RMSE

$$RMSE = \sqrt{\frac{1}{k} \sum_{i=1}^k (r_i - y_i)^2}$$

Метрики

DCG

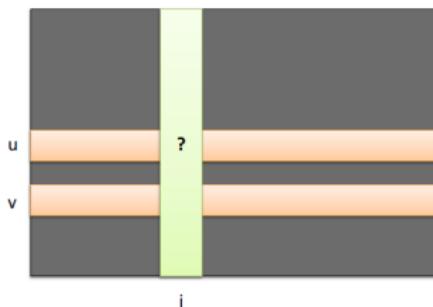
$$DCG@k = \sum_{i=1}^k \frac{2^{rel_i} - 1}{log_2(i + 1)}$$

Есть вещи, которые не понятно как мерить...

- ▶ Разнообразие
- ▶ Неожиданность
- ▶ Новизна
- ▶ Доверие
- ▶ Способность быстро адаптироваться
- ▶ Масштабируемость
- ▶ Удобство в использовании
- ▶ ...

User-based

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

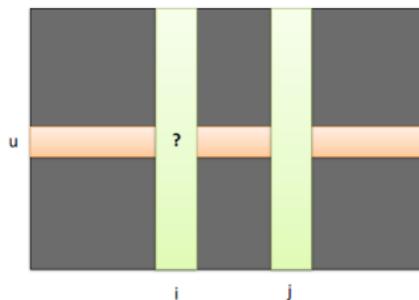


Меры похожести

- ▶ Косинусная мера
- ▶ Коэффициент корреляции Пирсона
- ▶ Что-нибудь специфичное для задачи

Item-based

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



Вопрос:

- ▶ Какие проблемы у этих методов?

Пример. Amazon.com (2003)

Идея:

Посчитаем статистику, какие товары часто покупают вместе и будем рекомендовать по ней

Frequently Bought Together



Price for all three: \$221.04

Add all three to Cart

Add all three to Wish List

Show availability and shipping details

- This Item: Machine Learning: A Probabilistic Perspective (Adaptive Computation and Machine Learning series) by Kevin P. Murphy Hardcover \$76.50
- Pattern Recognition and Machine Learning (Information Science and Statistics) by Christopher Bishop Hardcover \$67.68
- The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition (Springer Series in Statistics) by Trevor Hastie Hardcover \$76.86

Customers Who Bought This Item Also Bought



Learning From Data
Yaser S. Abu-Mostafa
★★★★★ 78
#1 Best Seller in
Computer Neural Networks
Hardcover



Pattern Recognition and
Machine Learning...
Christopher Bishop
★★★★★ 111
Hardcover
\$67.68



The Elements of Statistical
Learning...
Trevor Hastie
★★★★★ 45
#1 Best Seller in
Bioinformatics
Hardcover
\$76.86



Probabilistic Graphical
Models: Principles and...
Daphne Koller
★★★★★ 27
#1 Best Seller in Natural
Language Processing
Hardcover
\$94.81



An Introduction to
Statistical Learning...
Gareth James
★★★★★ 56
#1 Best Seller in
Mathematical & Statistical...
Hardcover
\$73.58



Bayesian Reasoning and
Machine Learning
David Barber
★★★★★ 12
Hardcover
\$85.49

Пример. Amazon.com (2003)

Алгоритм:

1. Для каждого товара в каталоге i_1
2. Для каждого пользователя u , который заказал i_1
3. Для каждого товара i_2 , который также заказывал пользователь u
4. Запомнить, что пользователь заказывал вместе i_1 и i_2
5. Для каждого товара i_2
6. Расчитать похожесть i_1 и i_2

Вопрос:

- ▶ Как распараллелить?

Пример. Amazon.com (2003)

Анализ:

- ▶ Сложность алгоритма $O(|I|^2|U|)$
- ▶ На практике все близко к $O(|I||U|)$
- ▶ Требование по памяти $O(|I|^2)$ (на практике сильно меньше)
- ▶ Тривиально масштабируется в случае большого объема данных
- ▶ Очень быстро работает в продакшене

Пример. Youtube (2010)

Item-to-item

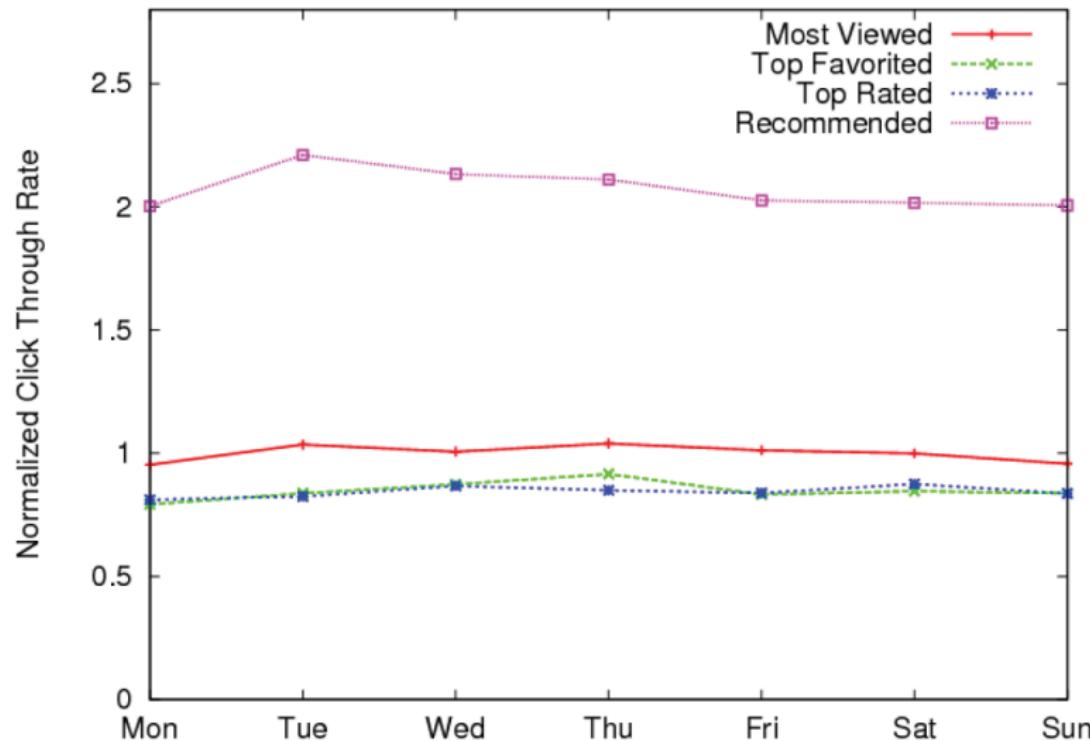
- ▶ Снова оцениваем похожесть роликов друг на друга
- ▶ Рассматриваем логи за последние сутки
- ▶ c_{ij} - количество раз, когда лайкали/смотрели/долго смотрели ролик i вместе с роликом j

$$r(i,j) = \frac{c_{ij}}{f(i,j)}, \quad f(i,j) = c_i c_j$$



Пример. Youtube (2010)

Результаты:



Низкоранговое представление матрицы

Пусть $A \in R^{n \times m}$

- ▶ Если $\text{rank}(A) = k$, то найдутся матрицы $X \in R^{n \times k}$, $Y \in R^{m \times k}$, что

$$A = XY^T$$

- ▶ Если $\text{rank}(A) > k$, то низкоранговое приближение

$$A \approx XY^T$$

- ▶ SVD разложение

$$\|A - XY^T\|_2 \rightarrow \min_{X, Y}$$

Model-based CF

Идея:

Будем искать решение в виде

$$\hat{r}_{ui} = \mathbf{p}_u^T \mathbf{q}_i$$

		Item			
		W	X	Y	Z
User	A		4.5	2.0	
	B	4.0		3.5	
	C		5.0		2.0
	D		3.5	4.0	1.0

=

		W	X	Y	Z
User	A	1.2	0.8		
	B	1.4	0.9		
	C	1.5	1.0		
	D	1.2	0.8		

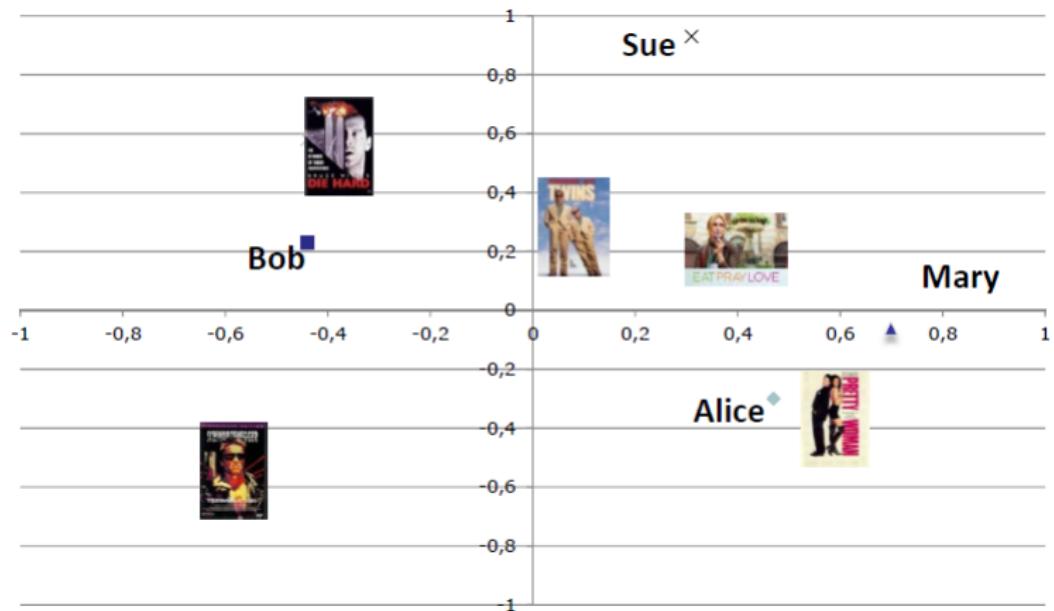
\times

		W	X	Y	Z
		1.5	1.2	1.0	0.8
		1.7	0.6	1.1	0.4

Rating MatrixUser MatrixItem Matrix

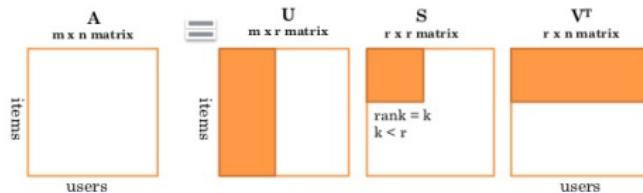
$$\sum_{(u,i) \in \mathcal{K}} (r_{ui} - \mathbf{p}_u^T \mathbf{q}_i)^2 \rightarrow \min_{p,q}$$

Latent Factors



Модель SVD

$$A = U \times S \times V^T$$



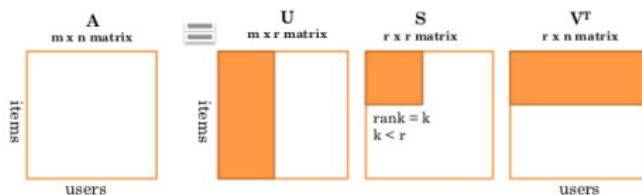
$$A_k = U_k \times S_k \times V_k^T$$

Замечание:

Однако на практике такую модель не используют. Почему?

Модель SVD

$$A = U \times S \times V^T$$



$$A_k = U_k \times S_k \times V_k^T$$

Замечание:

- ▶ Однако на практике такую модель не используют. Почему?
- ▶ Стандартная модель SVD предполагает наличие нулей на всех пропущенных значениях матрицы, что приводит к низкому качеству предсказания. Поэтому делают не так...

Модель SVD

Решают оптимизационную задачу

$$J(\mathbf{P}, \mathbf{Q}) = \sum_{(u,i) \in \mathcal{K}} (r_{ui} - \mathbf{p}_u^T \mathbf{q}_i)^2 \rightarrow \min_{p,q}$$

Параметры подбирают при помощи градиентных методов оптимизации. Например с помощью SGD:

Обозначим

$$e_{ui} = r_{ui} - \mathbf{p}_u^T \mathbf{q}_i$$

Тогда

$$\mathbf{p}_u \leftarrow \mathbf{p}_u + \eta e_{ui} \mathbf{q}_i$$

$$\mathbf{q}_i \leftarrow \mathbf{q}_i + \eta e_{ui} \mathbf{p}_u$$

Модель SVD с регуляризацией

Чтобы побороться с проблемой переобучения добавляют регуляризующее слагаемое

$$J(\mathbf{P}, \mathbf{Q}) = \sum_{(u,i) \in \mathcal{K}} (r_{ui} - \mathbf{p}_u^T \mathbf{q}_i)^2 + \lambda (\|\mathbf{p}_u\|^2 + \|\mathbf{q}_i\|^2) \rightarrow \min_{p,q}$$

Алгоритм оптимизации соответственно модифицируется

$$e_{ui} = r_{ui} - \mathbf{p}_u^T \mathbf{q}_i$$

$$\mathbf{p}_u \leftarrow \mathbf{p}_u + \eta(e_{ui}\mathbf{q}_i - \lambda\mathbf{p}_u)$$

$$\mathbf{q}_i \leftarrow \mathbf{q}_i + \eta(e_{ui}\mathbf{p}_u - \lambda\mathbf{q}_i)$$

РАБОТАЕТ ДОЛГО!!!

ALS

Alternationg Least Squares

$$J(\mathbf{P}, \mathbf{Q}) = \sum_{(u,i) \in \mathcal{K}} (r_{ui} - \mathbf{p}_u^T \mathbf{q}_i)^2 + \lambda (\|\mathbf{p}_u\|^2 + \|\mathbf{q}_i\|^2) \rightarrow \min_{p,q}$$

Идея:

Зафиксируем \mathbf{q}_i , тогда оптимизация распадется на независимые по пользователям задачи линейной регрессии:

$$\sum_u \left\{ \sum_{i \in K(u)} (r_{ui} - \mathbf{p}_u^T \mathbf{q}_i)^2 + \lambda \|\mathbf{p}_u\|^2 \right\} \rightarrow \min_{\mathbf{p}_u}$$

- ▶ Последовательно настраиваем параметры \mathbf{p}_u и \mathbf{q}_i

ALS

$$\mathbf{p}_u^* = (\mathbf{Q}_u^T \mathbf{Q}_u + \lambda \mathbf{I})^{-1} \mathbf{Q}_u^T \mathbf{r}_u$$

$$\mathbf{q}_i^* = (\mathbf{P}_i^T \mathbf{P}_i + \lambda \mathbf{I})^{-1} \mathbf{P}_i^T \mathbf{r}_i$$

Итеративный алгоритм ALS:

$$\mathbf{p}_u^{2t} = \mathbf{p}_u^*|_{2t-1}, \quad \forall u \in U$$

$$\mathbf{q}_i^{2t+1} = \mathbf{q}_i^*|_{2t}, \quad \forall i \in I$$

Обозначим

$$\mathbf{W}_u = (\mathbf{Q}_u^T \mathbf{Q}_u + \lambda \mathbf{I})^{-1}, \quad \mathbf{d}_u = \mathbf{Q}_u^T \mathbf{r}_u$$

ALS

Тогда

$$\mathbf{p}_u^* = (\mathbf{Q}_u^T \mathbf{Q}_u + \lambda \mathbf{I})^{-1} \mathbf{Q}_u^T \mathbf{r}_u = \mathbf{W}_u \mathbf{d}_u = \sum_{j \in K(u)} \mathbf{W}_u \mathbf{q}_j r_{uj}$$

\mathbf{Q}_u - матрица из \mathbf{q}_i , таких что $i \in K(u)$, \mathbf{r}_u - вектор оценок пользователя u

- ▶ Способ быстро обновлять оптимальный профиль пользователя
- ▶ Способ объяснить рекомендацию

Объяснение предсказаний

$$\hat{r}_{ui} = \mathbf{q}_i^T \mathbf{p}_u^* = \sum_{j \in K(u)} \mathbf{q}_i^T \mathbf{W}_u \mathbf{q}_j r_{uj}$$

Матрица \mathbf{W}_u положительна, допускает разложение Холецкого
 $\mathbf{W}_u = \mathbf{V}_u^T \mathbf{V}_u$

$$\hat{r}_{ui} = \sum_{j \in K(u)} r_{uj} (\mathbf{V}_u \mathbf{q}_i)^T \mathbf{V}_u \mathbf{q}_j$$

- ▶ $sim_u(i, j) = (\mathbf{V}_u \mathbf{q}_i)^T \mathbf{V}_u \mathbf{q}_j$ - персонализированная мера схожести
- ▶ \mathbf{V}_u - оператор предпочтений пользователя

Усложнение модели SVD

Добавляем смещение:

- ▶ Данные смещены
- ▶ Юзеры оптимисты
- ▶ Популярные фильмы

$$\hat{r}_{ui} = \mu + b_i + b_u + \mathbf{p}_u^T \mathbf{q}_i$$

$$J(\mathbf{P}, \mathbf{Q}) = \sum_{(u,i) \in \mathcal{K}} (r_{ui} - \mu - b_i - b_u - \mathbf{p}_u^T \mathbf{q}_i)^2 + \lambda (\|\mathbf{p}_u\|^2 + \|\mathbf{q}_i\|^2 + b_u^2 + b_i^2) \rightarrow \min_{p,q}$$

Implicit Feedback

- ▶ Имеются только тройки $(u, i, r = 1)$
- ▶ Например: лайки

Вопрос:

- ▶ Какую модель обучит SVD на таких данных?

Implicit Feedback

- ▶ Имеются только тройки $(u, i, r = 1)$

$$\sum_{(u,i) \in \mathcal{K}} c_{ui} (r_{ui} - \mathbf{p}_u^T \mathbf{q}_i)^2 + \lambda (\|\mathbf{p}_u\|^2 + \|\mathbf{q}_i\|^2) \rightarrow \min_{p,q}$$

- ▶ Элементы, которые не наблюдаем берем с малым весом (типа частичное обучение)
- ▶ Есть алгоритм I-ALS, который использует дополнительные предподсчеты и работает со сложностью обычного ALS

Усложнение модели SVD

- ▶ В реальной жизни всегда имеет место проблема холодного старта
- ▶ Также есть большое количество пользователей, для которых имеется совсем немного информации о предпочтениях
- ▶ Хочется добавить дополнительную информацию в модель, чтобы побороться с этими проблемами
- ▶ Добавим информацию о неявных предпочтениях пользователей

$$|N(u)|^{-0.5} \sum_{i \in N(u)} x_i$$

- ▶ Добавим информацию о самом юзере

$$\sum_{a \in A(u)} y(a)$$

$$\hat{r}_{ui} = \mu + b_i + b_u + \mathbf{q}_i^T [\mathbf{p}_u + |N(u)|^{-0.5} \sum_{i \in N(u)} x_i + \sum_{a \in A(u)} y(a)]$$

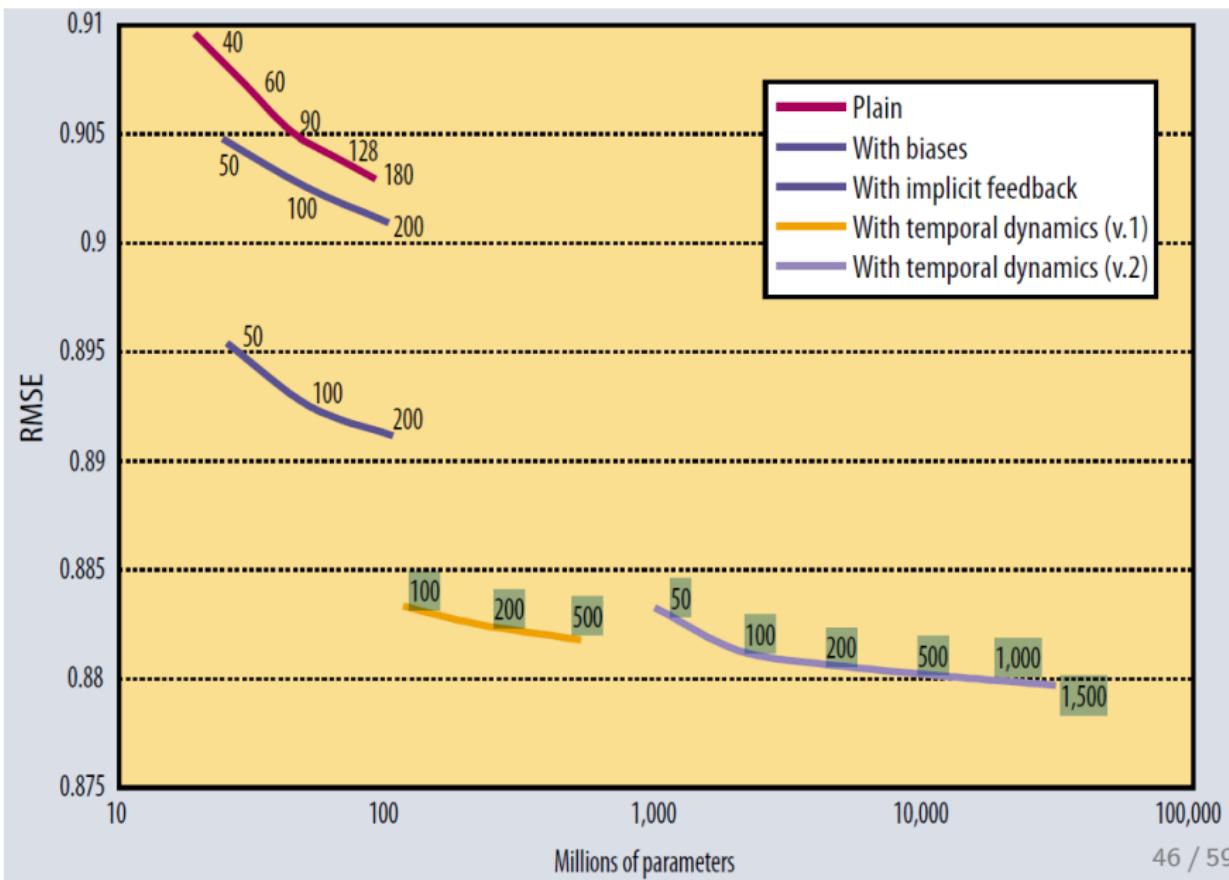
Усложнение модели SVD

Temporal Dynamics

- ▶ Добавим изменения предпочтений пользователей во времени
- ▶ Товары могут терять свою популярность

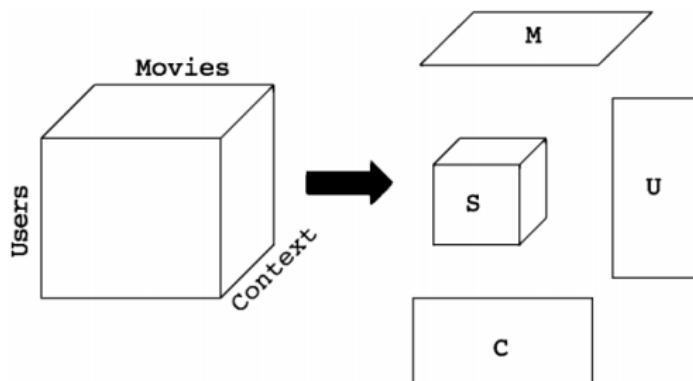
$$\hat{r}_{ui} = \mu + b_i(t) + b_u(t) + \mathbf{p}_u^T(t)\mathbf{q}_i$$

Сравнительный анализ на данных Netflix Prize



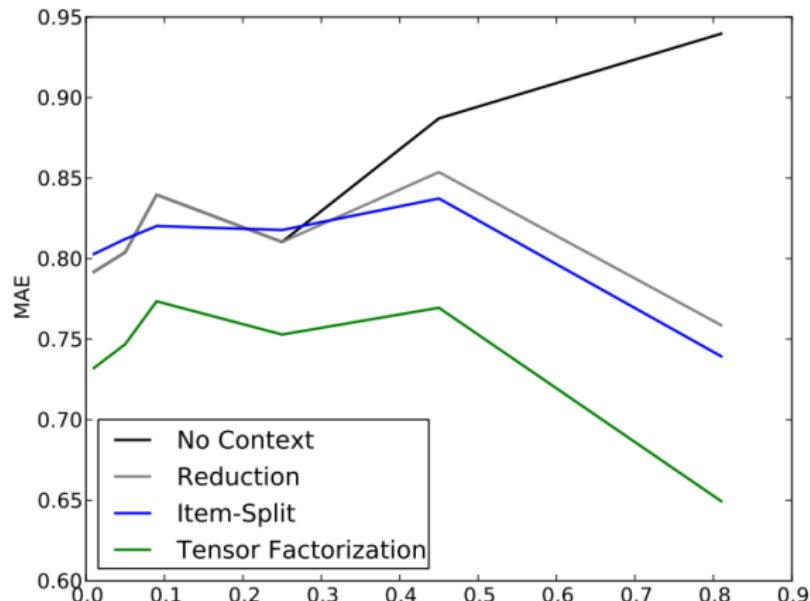
Tensor Factorization

High Order Singular Value Decomposition



Tensor Factorization

Результаты



Linear model

Сделаем признаки \mathbf{x} для предсказания r_{ui}

- ▶ Номер пользователя u (one-hot кодирование)
- ▶ Номер объекта i (one-hot кодирование)
- ▶ Любая дополнительная информация про пользователя или объект

Линейная модель

$$h(\mathbf{x}) = w_0 + \sum_{j=1}^p w_j x_j$$

Feature vector \mathbf{x}								Target y	
$x^{(1)}$	1	0	0	...	1	0	0	0	...
$x^{(2)}$	1	0	0	...	0	1	0	0	...
$x^{(3)}$	1	0	0	...	0	0	1	0	...
$x^{(4)}$	0	1	0	...	0	0	1	0	...
$x^{(5)}$	0	1	0	...	0	0	0	1	...
$x^{(6)}$	0	0	1	...	1	0	0	0	...
$x^{(7)}$	0	0	1	...	0	0	1	0	...
A	B	C	...	T1	NH	SW	ST	...	
User				Movie					

Factorization Machines

Факторизационные машины представлены как универсальная модель коллаборативной фильтрации, обобщающая многие из известных моделей

$$h(\mathbf{x}) = w_0 + \sum_{j=1}^p w_j x_j + \sum_{j=1}^p \sum_{j'=j+1}^p x_j x_{j'} \mathbf{v}_j^T \mathbf{v}_{j'}$$

- ▶ $\mathbf{x} \in R^p$ - вектор признак объекта
- ▶ $h(\mathbf{x})$ - предсказание
- ▶ $w_0, \mathbf{w}, \mathbf{v}$ - параметры
- ▶ Модель “квадратичной” регрессии

Factorization Machines

$$\hat{y}(x) = w_0 + \sum_{j=1}^p w_j x_j + \sum_{j=1}^p \sum_{j'=j+1}^p x_j x_{j'} \mathbf{v}_j^T \mathbf{v}_{j'}$$

	Feature vector \mathbf{x}													Target y						
\mathbf{x}_1	1	0	0	...	1	0	0	0	...	0.3	0.3	0.3	0	...	13	0	0	0	0	...
\mathbf{x}_2	1	0	0	...	0	1	0	0	...	0.3	0.3	0.3	0	...	14	1	0	0	0	...
\mathbf{x}_3	1	0	0	...	0	0	1	0	...	0.3	0.3	0.3	0	...	16	0	1	0	0	...
\mathbf{x}_4	0	1	0	...	0	0	1	0	...	0	0	0.5	0.5	...	5	0	0	0	0	...
\mathbf{x}_5	0	1	0	...	0	0	0	1	...	0	0	0.5	0.5	...	8	0	0	1	0	...
\mathbf{x}_6	0	0	1	...	1	0	0	0	...	0.5	0	0.5	0	...	9	0	0	0	0	...
\mathbf{x}_7	0	0	1	...	0	0	1	0	...	0.5	0	0.5	0	...	12	1	0	0	0	...
	A	B	C	...	TI	NH	SW	ST	...	TI	NH	SW	ST	...	Time	TI	NH	SW	ST	...
	User				Movie					Other Movies rated					Last Movie rated					

Factorization Machines

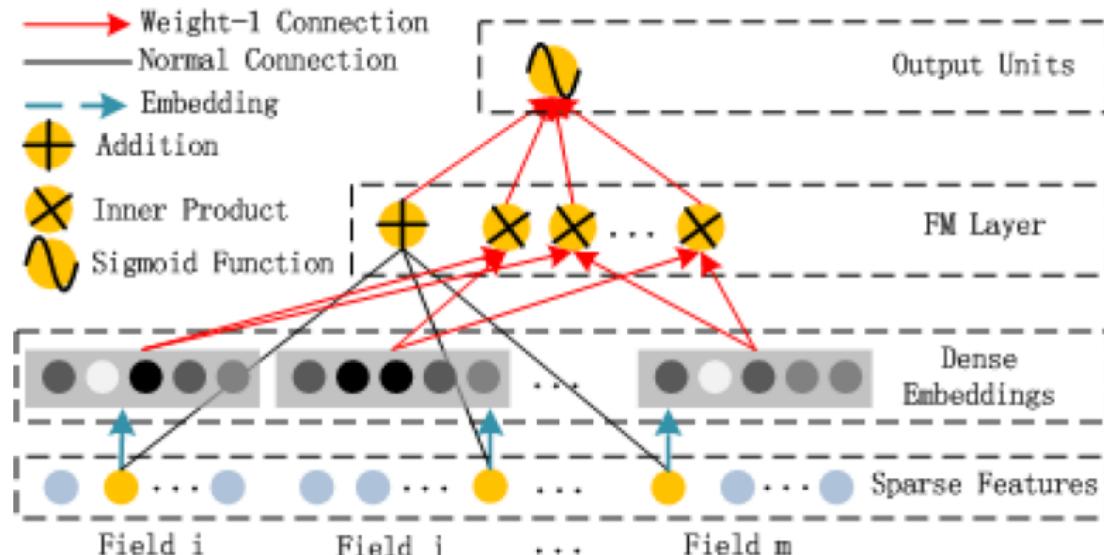
- ▶ Позволяет моделировать популярные модели коллаборативной фильтрации
- ▶ Позволяет добавить в модель коллаборативной фильтрации (обычные признаки, контекст и т.д)
- ▶ Имеется эффективный и масштабируемый метод обучения
- ▶ Показывает хорошие результаты на Kaggle
- ▶ Имеется opensource библиотека libfm

Factorization Machines

Вопрос:

Какова связь между рассмотренными сегодня моделями коллаборативной фильтрации и нейросетями?

Factorization Machines As Neural Network



$$\hat{y}(x) = w_0 + \sum_{j=1}^p w_j x_j + \sum_{j=1}^p \sum_{j'=j+1}^p x_j x_{j'} \mathbf{v}_j^T \mathbf{v}_{j'}$$

Итоги

- ▶ Neighborhood methods
 - ▶ Проблема холодного старта
 - ▶ Матрица оценок используется непосредственно для получения предсказаний
 - ▶ У больших компаний с большим количеством данных чаще всего реализуется эвристический item-to-item подход
- ▶ Model-based methods
 - ▶ Основаны на оффлайновых расчетах
 - ▶ В продакшене используется уже обученные модели
 - ▶ Модели необходимо часто обновлять и перестраивать
 - ▶ Есть множество методов, которые показывают хорошие результаты
- ▶ Можно применять комбинированный подход
- ▶ Все это можно использовать в композициях

Открытые вопросы

- ▶ Как улучшать качество рекомендаций а не RMSE?
- ▶ Как обосновывать рекомендации?
- ▶ Как решать проблему холодного старта?

Открытые вопросы

- ▶ Как учитывать неявные предпочтения?
- ▶ Как учитывать контекст в рекомендациях?
- ▶ Как дополнительно учитывать признаки объектов?
- ▶ Как учитывать связи между пользователями?
- ▶ Как учитывать информацию из других источников (например из поиска)?

Технические вопросы

- ▶ Как быстро обновлять рекомендации?
- ▶ Как масштабировать расчет рекомендаций?
- ▶ Как быстро выбирать топ документов для предсказаний?
- ▶ Как измерять качества контекстных рекомендаций?

Вопросы

