

Персонализация важна в различных сервисах



+\$2.93 млрд дополнительно к бюджету компании после внедрения рекомендательной системы во все сервисы



80% того, что смотрят пользователи является результатом рекомендаций; итог: экономия \$1 млрд



Spotify достиг 96 миллионов премиальных подписок за 2018; средняя выручка с подписки — €4.89

Что такое рекомендательная система?



Примеры:

- Amazon
- Netflix
- Spotify
- Ozon

Множество областей применения: e-commerce, новости, туризм, развлечения, образование...

Цель: предсказать пользовательские предпочтения на основе его предыдущих действий

В более общем смысле

Рекомендательные системы ставят целью
восстановить **частично наблюдаемые взаимоотношения**
между **двумя или более типами объектов**.

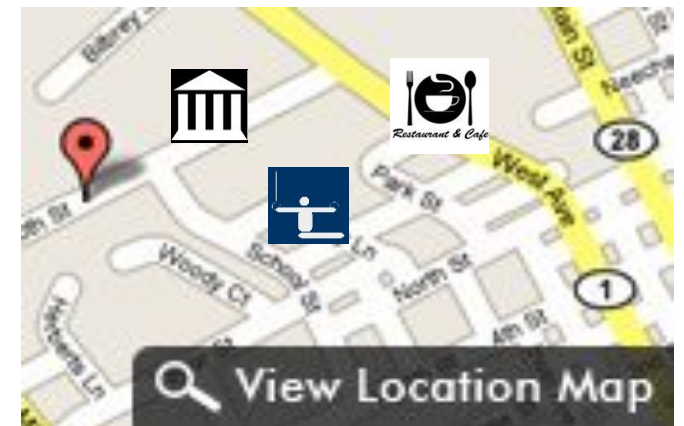
Социальные сети:
пользователь ↔ пользователь



Данные на основе
последовательностей:
предмет → предмет
(порядок имеет значение)

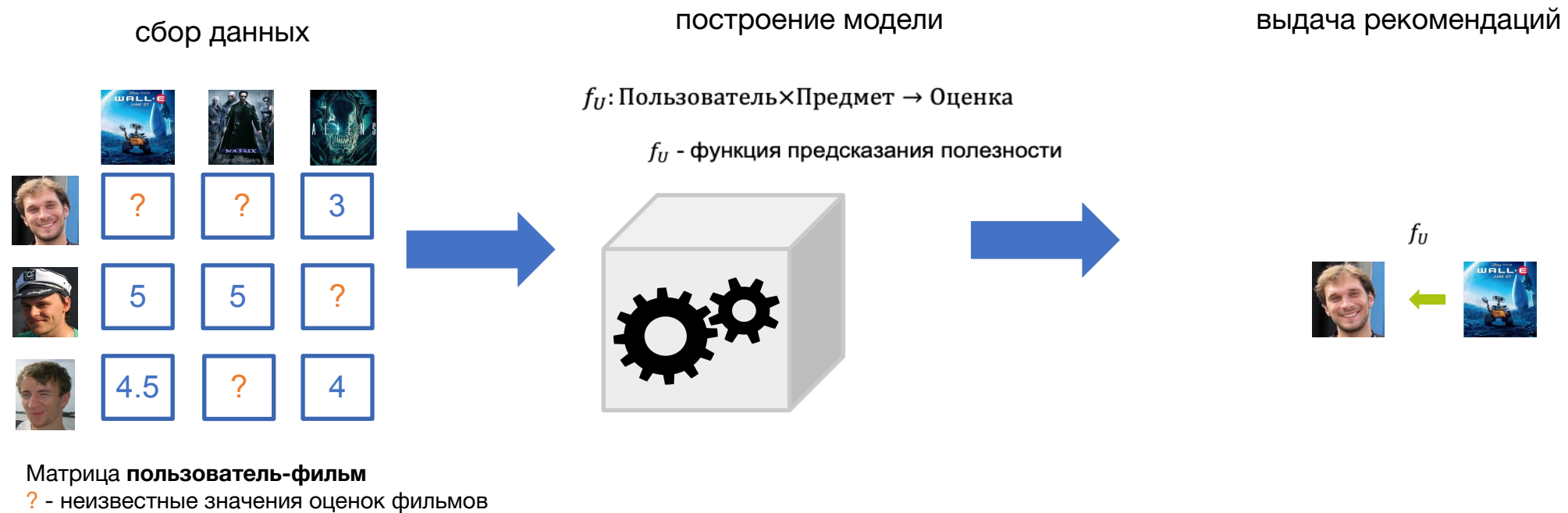


Множественные отношения:
пользователь → действие → место



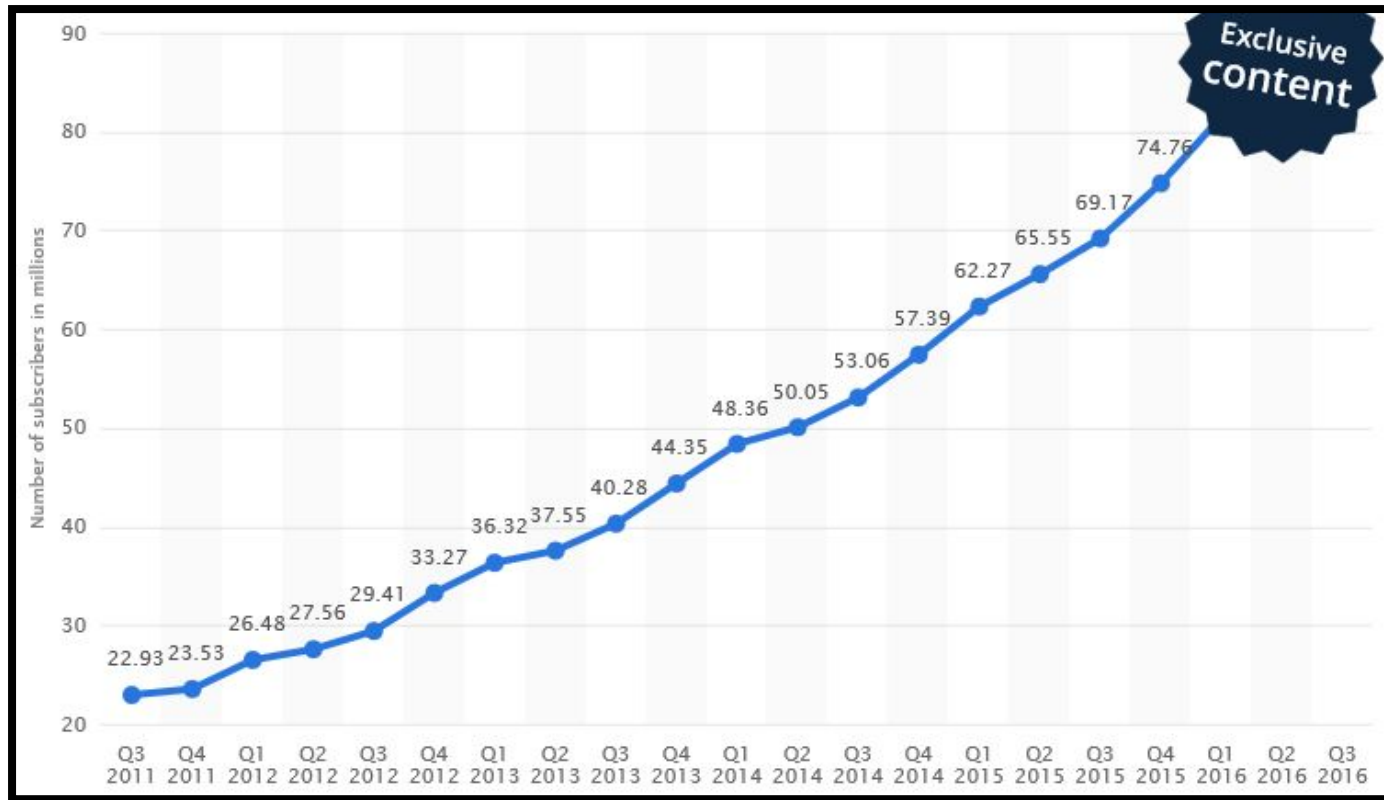
Процесс создания и использования модели рекомендаций

Цель: предсказать пользовательские предпочтения на основе его предыдущих действий.

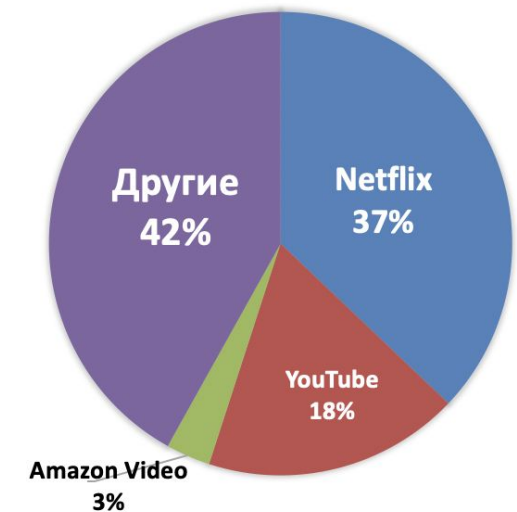


Пример: развлечения

Аудитория Netflix



Распределение интернет трафика



Примеры: финансовый сектор

Возможные сценарии:

- Индивидуальные рекомендации банковских продуктов
- Транзакции пользователей могут быть использованы для угадывания возможных интересов:
 - посетить торговый центр неподалеку
 - приобрести определенные продукты

Что это дает:

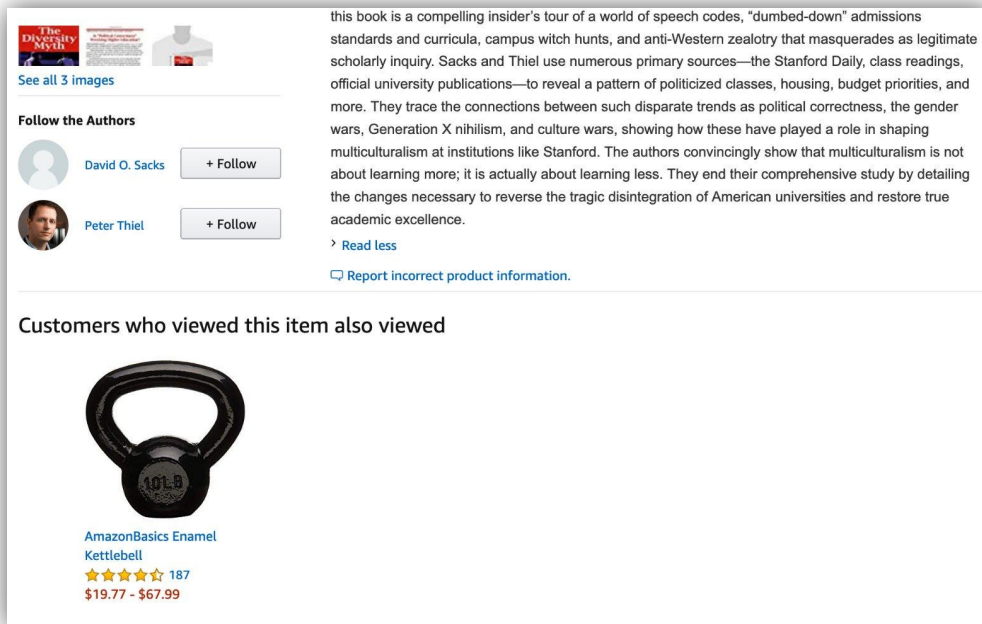
- увеличивает привязанность к банковским услугам
- позволяет удобным образом описать возможные интересы пользователей
- улучшает связность и надежность накопленной информации о пользователях
- дает новые инсайты для принятия бизнес-решений и формирования маркетинговой стратегии



Ведет к увеличению выручки в долгосрочной перспективе

Смешные ошибки рекомендательных систем

“Если вам нравится книга Diversity Myth Питера Тия, то вам может понравиться гиря”



See all 3 images

Follow the Authors

David O. Sacks + Follow

Peter Thiel + Follow

Read less

Report incorrect product information.

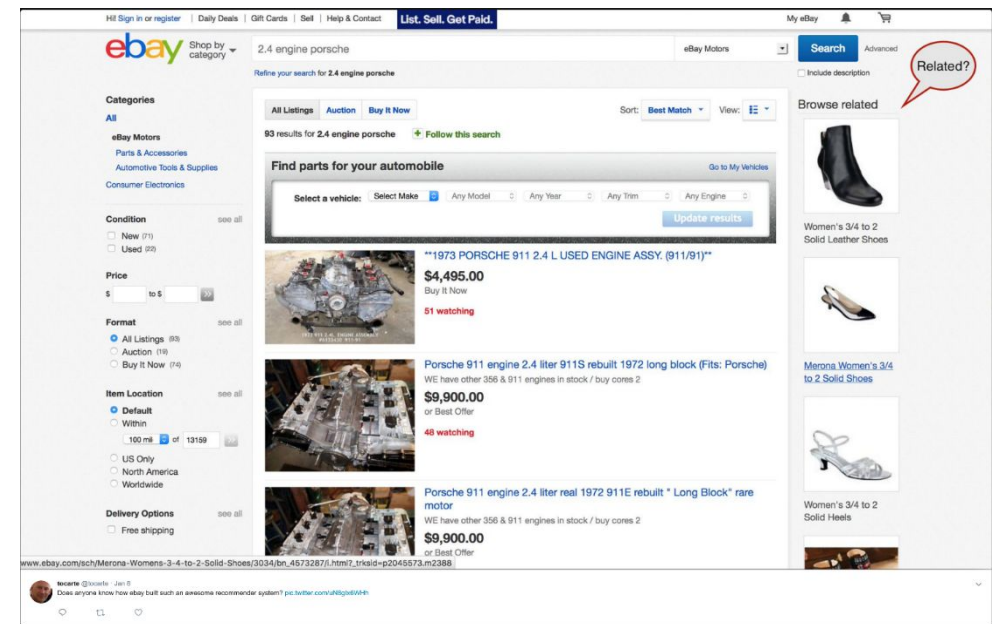
Customers who viewed this item also viewed

AmazonBasics Enamel Kettlebell

★★★★☆ 187

\$19.77 - \$67.99

“Ищете запчасти для двигателя? Вам могут пригодиться женские туфли ”



Hi Sign in or register | Daily Deals | Gift Cards | Sell | Help & Contact | List, Sell, Get Paid.

My eBay

2.4 engine porsche

Shop by category

Categories

ebay Motors

Parts & Accessories

Automotive Tools & Supplies

Consumer Electronics

Condition

New (71)

Used (2)

Price

\$ to \$

Format

All Listings (83)

Auction (14)

Buy It Now (14)

Item Location

Default

Within 100 mi of 13159

US Only

North America

Worldwide

Delivery Options

Free shipping

Find parts for your automobile

Select a vehicle: Select Make Any Model Any Year Any Mile Any Engine

Update results

1973 PORSCHE 911 2.4 L USED ENGINE ASSY. (911/91)*

\$4,495.00

Buy It Now

51 watching

Porsche 911 engine 2.4 liter 911S rebuilt 1972 long block (Fits: Porsche)

WE have other 356 & 911 engines in stock / buy comes 2

\$9,900.00

or Best Offer

48 watching

Porsche 911 engine 2.4 liter real 1972 911E rebuilt "Long Block" rare motor

WE have other 356 & 911 engines in stock / buy comes 2

\$9,900.00

or Best Offer

Browse related

Women's 3/4 to 2 Solid Leather Shoes

Merona Women's 3/4 to 2 Solid Shoes

Women's 3/4 to 2 Solid Heels

Related?

Не такие смешные ошибки рекомендательных систем

Циклическая обратная связь

Радикализация взглядов пользователей
YouTube

Bias

Алгоритмы могут смещены по переменным,
которых нет в данных.



The New York Times @nytimes · Jun 8

Caleb Cain was a college dropout looking for direction. He was then pulled into YouTube's far-right universe, watching thousands of videos filled with conspiracy theories, misogyny and racism. "I was brainwashed."



The Making of a YouTube Radical


Caleb Cain was a college dropout looking for direction. He turned to YouTube, where he was pulled into a world filled with conspiracy theories...

[nytimes.com](https://www.nytimes.com)

<https://habr.com/ru/post/437572/>

<https://www.nytimes.com/interactive/2019/06/08/technology/youtube-radical.html>

<https://www.fast.ai/2019/05/28/google-nyt-mohan/>

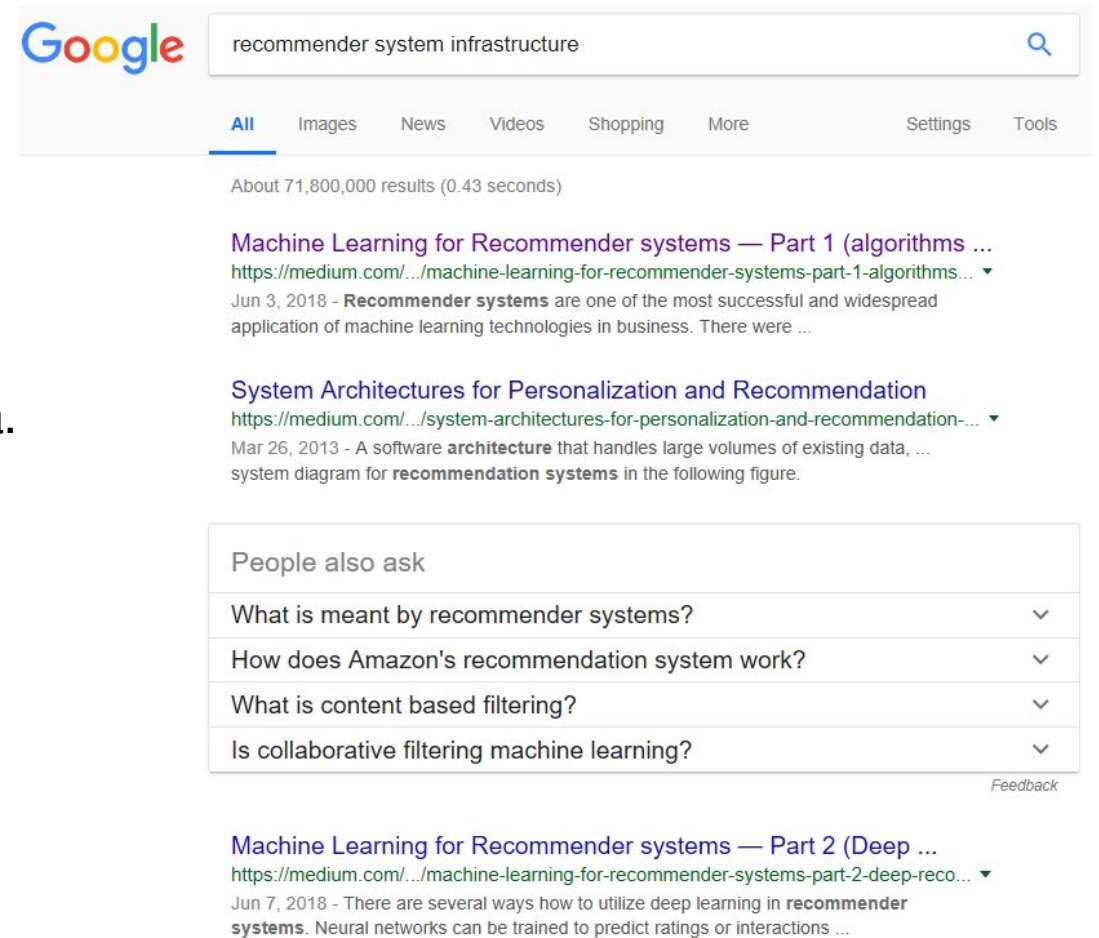


Как оценить качество рекомендательных систем

Поиск vs Рекомендации

Поисковая система использует запрос.

Рекомендательная система – поиск без запроса.



The screenshot shows a Google search interface. The search bar contains the text "recommender system infrastructure". Below the search bar, there are tabs for "All", "Images", "News", "Videos", "Shopping", and "More". The "All" tab is selected. Below the tabs, it says "About 71,800,000 results (0.43 seconds)". There are two search results displayed. The first result is titled "Machine Learning for Recommender systems — Part 1 (algorithms ...)" and includes a URL and a snippet. The second result is titled "System Architectures for Personalization and Recommendation" and also includes a URL and a snippet. Below the search results, there is a section titled "People also ask" with five questions and expandable arrows. At the bottom, there is another search result titled "Machine Learning for Recommender systems — Part 2 (Deep ...)".

Google

recommender system infrastructure

All Images News Videos Shopping More Settings Tools

About 71,800,000 results (0.43 seconds)

Machine Learning for Recommender systems — Part 1 (algorithms ...
<https://medium.com/.../machine-learning-for-recommender-systems-part-1-algorithms...>
Jun 3, 2018 - **Recommender systems** are one of the most successful and widespread application of machine learning technologies in business. There were ...

System Architectures for Personalization and Recommendation
<https://medium.com/.../system-architectures-for-personalization-and-recommendation-...>
Mar 26, 2013 - A software **architecture** that handles large volumes of existing data, ... system diagram for **recommendation systems** in the following figure.

People also ask

- What is meant by recommender systems?
- How does Amazon's recommendation system work?
- What is content based filtering?
- Is collaborative filtering machine learning?

Feedback

Machine Learning for Recommender systems — Part 2 (Deep ...
<https://medium.com/.../machine-learning-for-recommender-systems-part-2-deep-reco...>
Jun 7, 2018 - There are several ways how to utilize deep learning in **recommender systems**. Neural networks can be trained to predict ratings or interactions ...

История Netflix prize: октябрь 2006 – июнь 2009



Условия соревнования: На основе предоставленной компанией информации об оценках пользователей выбранным фильмам, улучшить точность предсказания оценок хотя бы на 10% относительно внутреннего алгоритма Netflix

Приз: \$1,000,000

Соревнование продемонстрировало важность области для индустрии



Интересный факт:
решение победителей не было
использовано в production

Проблемы:

- Умение точно предсказывать оценку \neq умение выдавать релевантные рекомендации

Оценки качества рекомендательных систем



Netflix использовал корень из среднеквадратичной оценки RMSE на оценках.

Проблемы RMSE:

- У ошибок в бизнесе разный вес: сравните 5 vs 3 и 3 vs 1. Так как фильм с оценкой 1 или 3 мы рекомендовать не будем, а с оценкой 5 будем.
- Не отражает бизнес суть задачи!

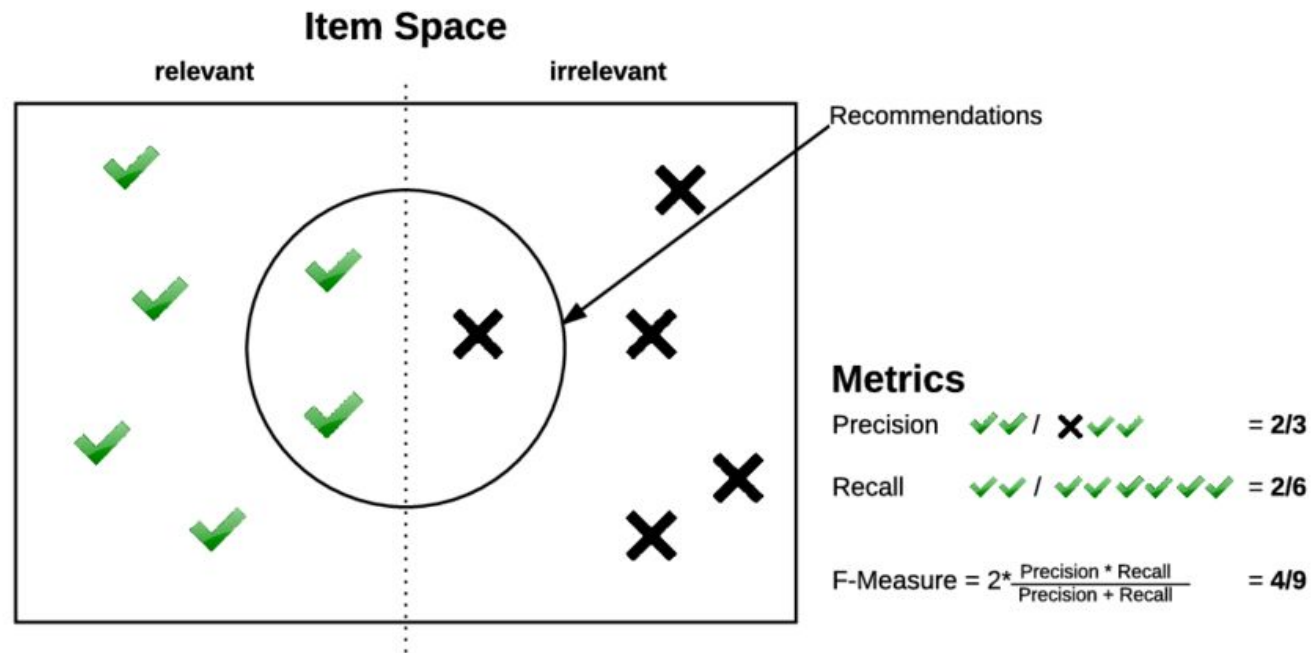
$$RMSE = \sqrt{\sum_{i=1}^n \frac{(\hat{y}_i - y_i)^2}{n}}$$

Точность и специфичность

Precision, точность – доля релевантных для пользователя объектов относительно всех, которые были ему показаны.

Мотивация: пользователю нельзя показать все наши продукты, блок предложений конечен.

Recall, специфичность – доля всех показанных объектов относительно релевантных объектов.



Метрики, учитывающие порядок рекомендаций

- Normalized Discounted Cumulative Gain (NDCG) – штрафует алгоритм за неправильный порядок рекомендаций.
- При этом переставить местами 1 и 2 элемент наказывается сильнее, чем, например 5 и 6.
- Мотивация – чем выше релевантность объекта – тем он должен быть выше.

Если ошибаемся в хвосте рекомендаций – не так страшно.

Дополнительные экзотические метрики

- Разнообразие
- Покрытие запроса
- Новизна
- Неожиданность

Если клиент купил молоко и хлеб можно не рекламировать ему яйца – он с высокой вероятностью их и так купит.

Нужно рекомендовать что-то с более низкой вероятностью покупки, но более разнообразное и интересное.

Бизнес метрики

- Конверсия
- Клики
- Retention
- Прибыль
- LTV - life-time value

У рекомендательной системы много нематематических аспектов:
место расположения, дизайн, и т.д.

Чтобы это корректно оценивать – нужны АВ тесты на пользователях.

АВ тесты часто единственный способ честно оценить бизнес ценность модели.

Sad but true: Такие метрики тяжело прогнозировать.

Метрики



Лучше не
использовать

Основаны на ошибке

RMSE, MAE



Можно
использовать

Основаны на релевантности

precision, recall

F1-score, accuracy

HR (hit rate)



Наибольшая корреляция
с бизнес метриками

Основаны на ранжировании

nDCG (normalized discounted cumulative gain),

MRR (mean reciprocal rank),


ATOP (area under the TOPK-curve)

AUC (area under the ROC-curve)

Измеряем другие аспекты:

- Покрытие (Coverage)
- Новизна (Novelty)
- Неожиданность (Serendipity)
- Разнообразие (Diversity)
- Доверие (Trust)
- Полезность (Utility)

Подробнее: "Evaluating Recommendation Systems",
Guy Shani and Asela Gunawardana, 2011



Алгоритмы построения рекомендательных систем

Данные для построения рекомендательной системы



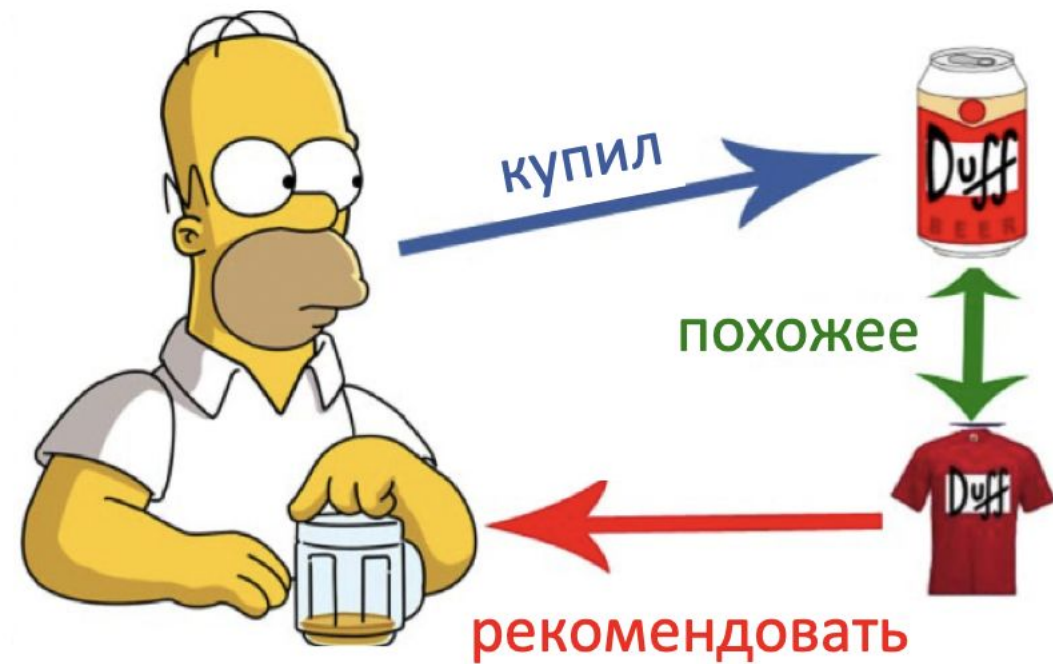
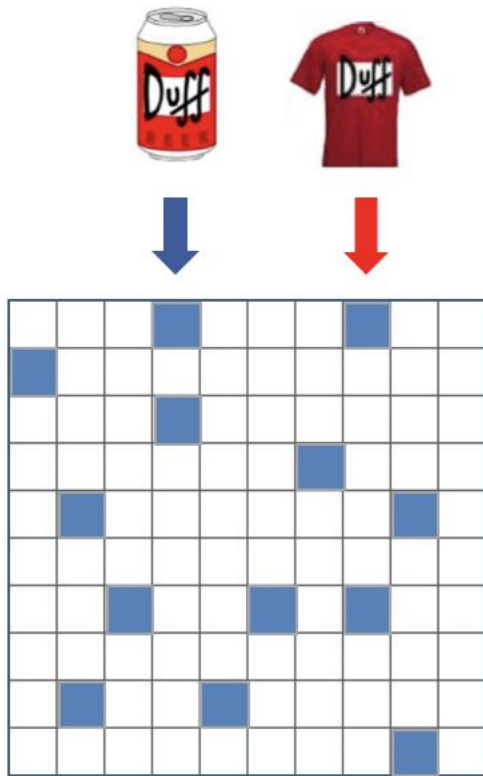
Задача: найти функцию полезности f_U , генерирующую предсказания для любой пары:

f_U : пользователь + предмет \rightarrow релевантность

Если мы умеем считать расстояния между пользователями, то можем использовать *коллаборативную фильтрацию по схожести пользователей (user-based)*

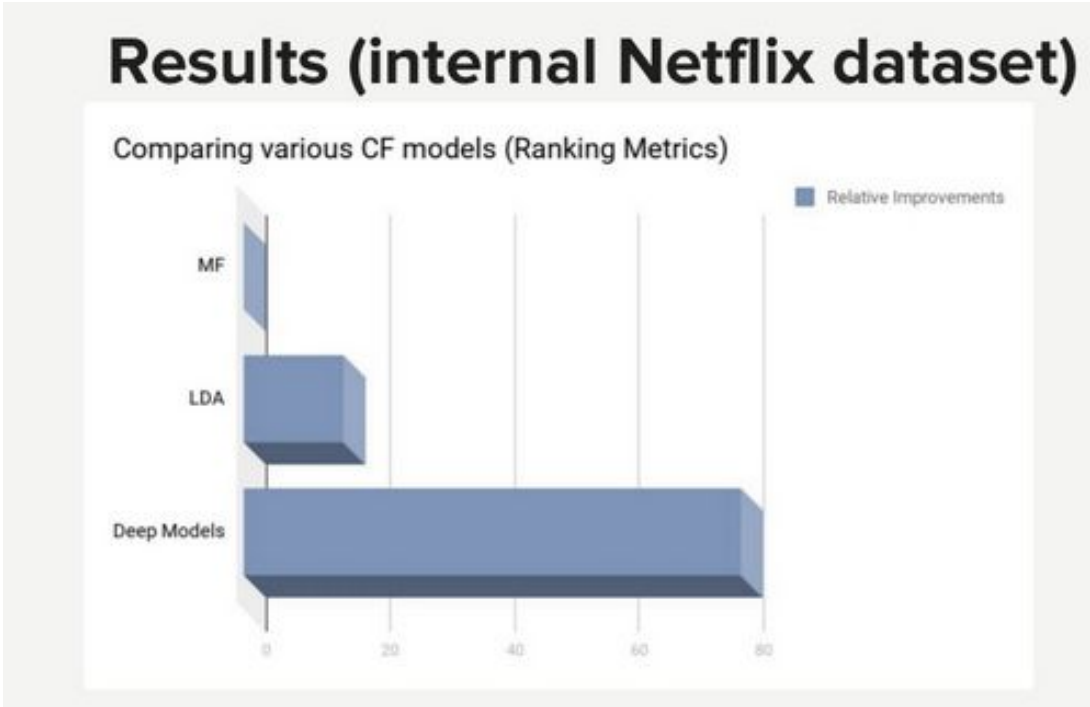


Если мы умеем считать расстояния между предметами, то можем использовать *коллаборативную фильтрацию по схожести предметов (item-based)*



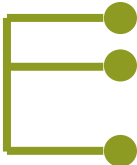
Немного о более сложных методах

Как пишут в статьях



Как все работает на практике

Матричная факторизация у 3 из 4 топовых алгоритмов



Method	movielens			yahoo			pinterest		
	HR [%]	ARHR [%]	NDCG [%]	HR [%]	ARHR [%]	NDCG [%]	HR [%]	ARHR [%]	NDCG [%]
EigenRec	45.21	20.44	26.35	48.12	23.30	29.23	33.81	13.51	18.41
PureSVD	44.14	19.33	25.36	38.68	18.30	22.62	30.97	11.85	16.30
RP3b	34.87	15.02	19.66	41.51	17.82	22.94	27.01	8.07	12.45
SLIM	46.34	21.39	27.28	52.44	26.15	32.35	34.17	13.63	18.57
Mult-DAE	44.06	18.97	24.83	45.37	21.46	27.07	35.03	13.79	18.77
Mult-VAE	44.35	19.50	25.31	45.09	21.22	26.80	35.13	13.73	18.71

Персонализация контента – трудная задача

Пропуски в данных

- 99.99...% неизвестных значений
- смещенность оценки, Missing Not at Random (MNAR)

«длинный хвост»

- 5% предметов могут покрывать 40% всех взаимодействий
- нишевые рекомендации

оценка качества

- Выбор метрики качества, нестандартные метрики (novelty, diversity...)
- проверка качества офлайн vs. A/B тестирование

холодный старт

- неопределенность в том, что рекомендовать
- сложно подобрать наиболее репрезентативные рекомендации

усложнение моделей

- включение признаковой информации
- включение контекста взаимодействия

вычислительная скорость

- вычисление модели за разумное время
- выдача рекомендаций в реальном времени

Холодный старт

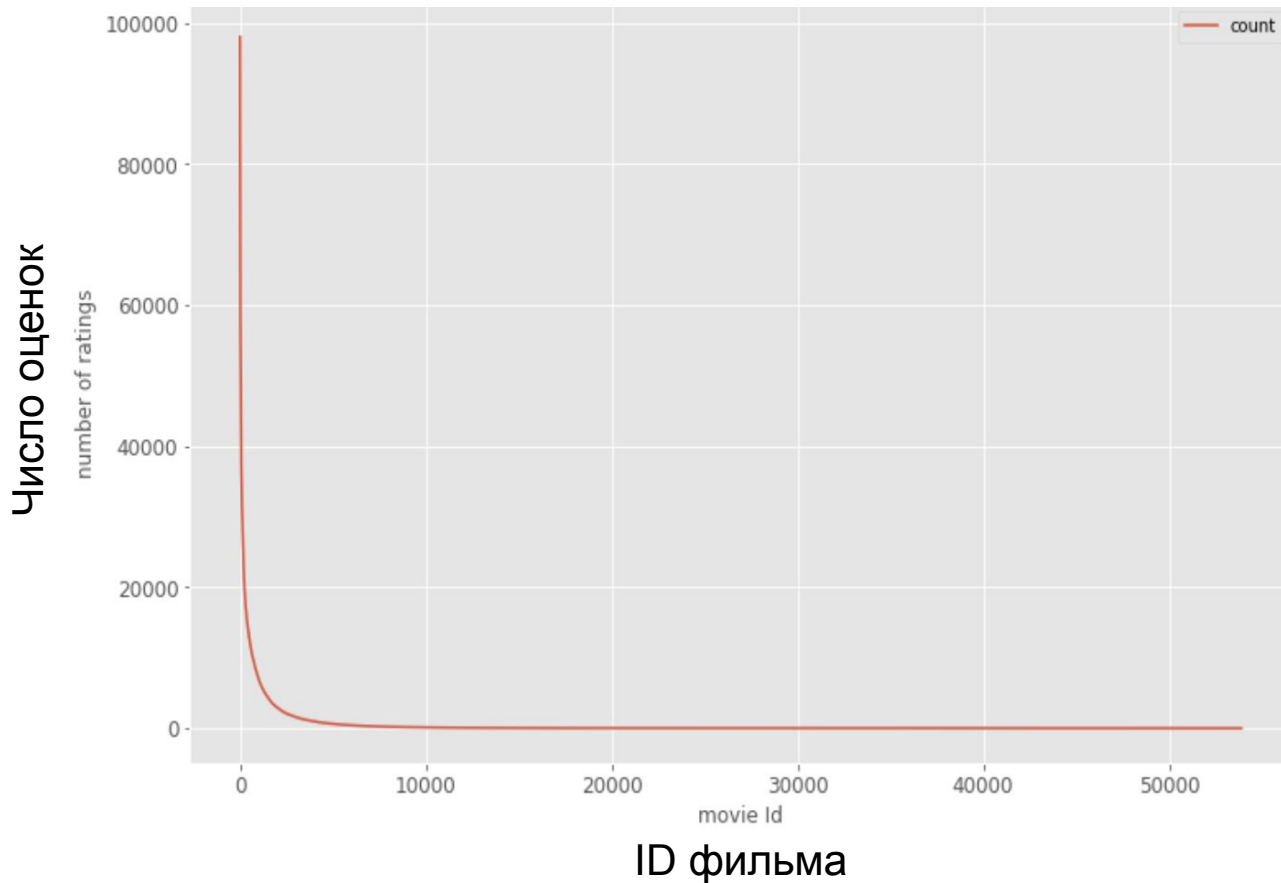
Проблема всех рекомендательных систем:
Первое время нет информации о предпочтениях пользователя.

Решения:

- Использовать профиль «среднего» пользователя
- Рекомендовать самые популярные товары

Длинный хвост

Популярные товары покупают часто.
Большинство товаров покупают редко.



Замкнутый круг: про длинный хвост мало статистики – система его не рекомендует – мало статистики.

Отсутствует новизна предложений.

Гибридные подходы.

Решение:

Продуктовое – создавать категории новинок, иногда рекомендовать 1-2 случайных товара из хвоста.

Гибридные подходы:

Ансамбли из рекомендательных алгоритмов.

Минусы: долго вычисляется, сложно поддерживать

Дополнительные материалы

Кратко и на русском:

- Серия постов С. Николенко на habr.com
- “Как работают рекомендательные системы”, М. Ройзнер, <https://habr.com/ru/company/yandex/blog/241455/>
- Анатомия рекомендательных систем:
Часть 1: <https://habr.com/ru/company/lanit/blog/420499/>
Часть 2: <https://habr.com/ru/company/lanit/blog/421401/>

Кейсы:

- Что происходит в мире рекомендательных систем прямо сейчас <https://youtu.be/8bqgkVmwXi0>
- Рекомендации в Okko <https://habr.com/ru/company/okko/blog/454224/>
- Как рекомендовать музыку, Яндекс <https://habr.com/ru/company/yandex/blog/441586/>
- 10 уроков рекомендательной системы Quora <https://habr.com/ru/company/retailrocket/blog/341346/>

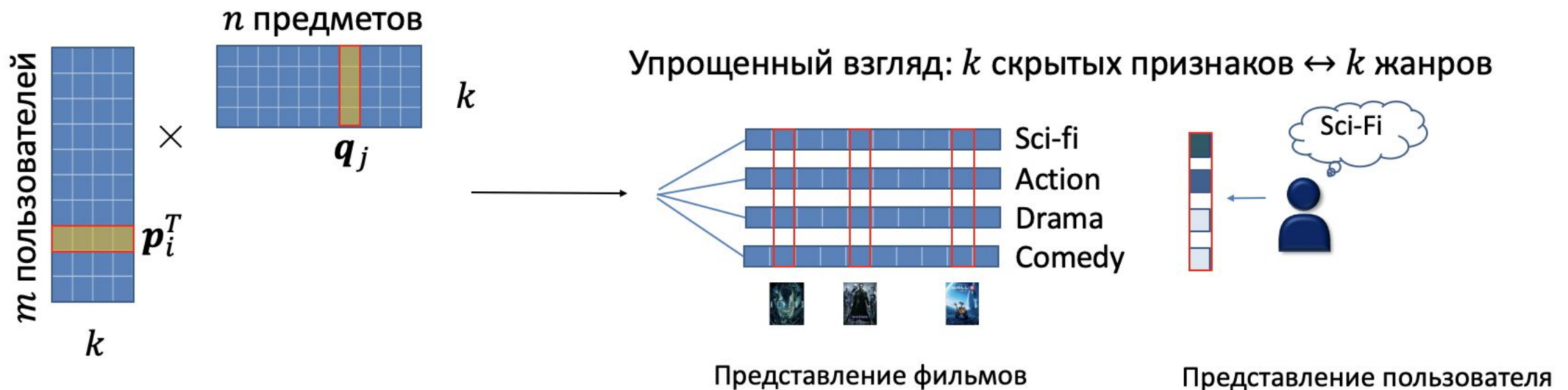
Книжки:

- Recommender Systems Handbook, 2015, 2nd edition; F. Ricci, L. Rokach, B. Shapira
- Recommender Systems. The Textbook, 2016; Charu C. Aggarwal
- Recommender Systems: An Introduction, 2010; D.Jannach, M.Zanker, A.Felfernig, G.Friedrich
- Collaborative Recommendations: Algorithms, Practical Challenges and Applications; S. Berkovsky, I. Cantador and D. Tikk.

Матричная факторизация

Представление пользователей и предметов – способ подсчитать релевантность

Мы построили представление размерности r для пользователей и предметов



Релевантность – расстояние между пользователями и предметами

$$r_{ij} \approx \mathbf{p}_i^T \mathbf{q}_j = \sum_{t=1}^k p_{it} q_{jt}$$

$\mathbf{p}_i = (p_{i1}, p_{i2}, \dots, p_{ik})$ – вектор скрытых признаков пользователя i

$\mathbf{q}_j = (q_{j1}, q_{j2}, \dots, q_{jk})$ – вектор скрытых признаков предмета j

Как получить представления пользователей: матричная факторизация

Основное предположение:

относительно небольшое количество шаблонов поведения + небольшие вариации за счет индивидуальных особенностей

$$\mathcal{L}(A, P, Q) \rightarrow \min_{P, Q} \quad \mathcal{L}(A, P, Q) = \|A - PQ^T\|_2$$

Diagram illustrating matrix factorization. Matrix A (users m by items n) is approximated by the product of matrix P (users m by features k) and matrix Q^T (features k by items n). The element a_{ij} in A is approximated by the dot product of row i in P and column j in Q^T .

Строки матриц P и Q дают представление пользователей и предметов на основе набора скрытых признаков

$$a_{ij} \approx r_{ij} = \mathbf{p}_i^T \mathbf{q}_j = \sum_{t=1}^k p_{it} q_{jt}$$

Релевантности должны быть похожи на элементы исходной матрицы A

Дополнительные слайды

Фреймворки и библиотеки

Frameworks

Surprise

<https://github.com/NicolasHug/Surprise>

Polara

<https://github.com/evfro/polara>

Collaborative Filtering – Apache Spark

<http://spark.apache.org/docs/latest/mllib-collaborative-filtering.html>
(Neighborhood models and MF)

Turi Create (acquired by Apple)

<https://turi.com/learn/userguide/recommender/introduction.html>

Microsoft Azure

<https://github.com/Microsoft/Recommenders>

Полезные библиотеки

Collaborative Filtering for Implicit Feedback Datasets

<https://github.com/benfred/implicit> (fastest)

<https://github.com/quora/qmf> (by Quora)

<https://github.com/MrChrisJohnson/logistic-mf> (as in Spotify)

Factorization Machines

<https://github.com/srendle/libfm>

Другие библиотеки

Нейронные сети

<https://github.com/maciejkula/spotlight>

<https://github.com/MrChrisJohnson/deep-mf>

<https://github.com/songggc/TF-recomm>

<https://github.com/Netflix/vectorflow> (by Netflix)

Билинейные модели

<https://github.com/lyst/lightfm/>

<http://www.recsyswiki.com/wiki/SVDFeature>

MyMediaLite (used to be popular)

<http://www.mymedialite.net>

Модели с латентными факторами

<https://github.com/zhangsi/CisRec>

Простой движок для рекомендаций на основе контента

<https://github.com/grovecoco/content-engine>

Логистическая матричная факторизация

<https://github.com/MrChrisJohnson/implicit-mf>

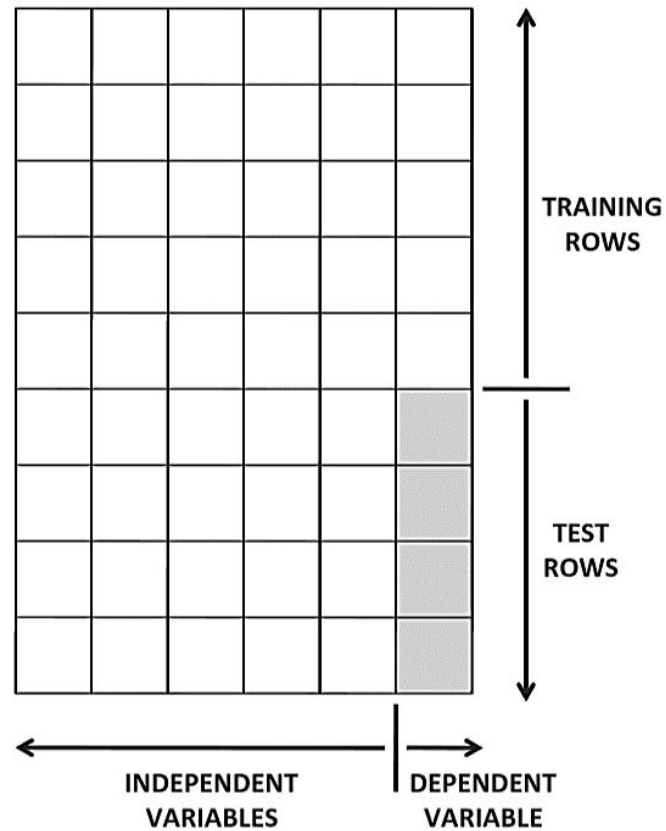
Hermes (поддерживает Spark)

<https://github.com/Lab41/hermes>

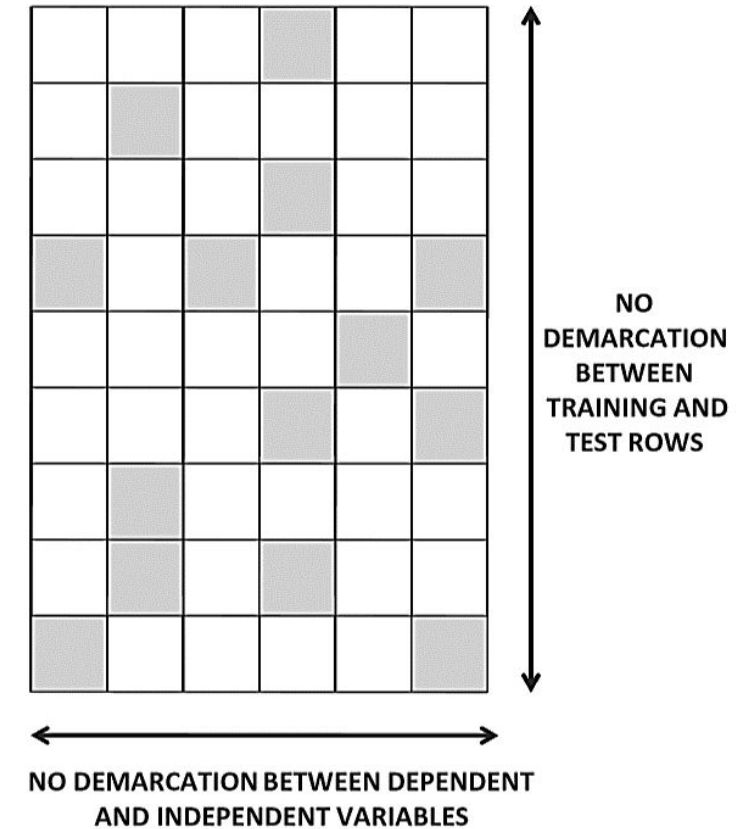
Задачи классификации/регрессии vs задача предсказания предпочтений

Классов в задаче классификации –
немного

Объектов в рекомендательных
системах – очень много



Структура данных
в задаче классификации



Структура данных
в рекомендательных
системах

Разбиение на обучающую и тестовую выборку

