

# Исследование возможности встраивания контекстной информации в алгоритмы коллаборативной фильтрации на основе матричных разложений

Стеценко М. А.    Игнатов Д. И.

Национальный исследовательский университет «Высшая школа экономики»  
Факультет компьютерных наук  
Отделение Прикладной математики и информатики

Москва, 2015

# Оглавление

- ① Введение
- ② Заключение
- ③ Литература

## Аннотация

Наличие контекстной информации является одним из важнейших факторов для построения личных рекомендаций. Однако, классические алгоритмы коллаборативной фильтрации, основанные на матричных разложениях, таких как SVD разложение, используют только информацию о пользователях и предметах и не предоставляют явных методов включения дополнительных факторов. В данной работе будет показан один из методов встраивания контекстной информации в алгоритм, использующий SVD разложение. Для тестирования рассматриваемого метода будет использоваться открытый банк данных MovieLens. База данных содержит пользователей портала MovieLens, каждый из которых оценил не менее 20 фильмов, а так же информацию о каждом фильме.

Context has always been an important factor in personalized Recommender systems. However, standard collaborative filtering algorithms based on matrix factorization rely mainly on user and subject information and don't provide any methods for encapsulating extra data. This work demonstrates such a method based on SVD decomposition. To test results an open data base taken from MovieLens is used. The database provides information about users and movies.

# Введение

- **Рекомендательные системы** – это модели, которые лучше вас знают, что вам хочется.

# Введение

- **Рекомендательные системы** – это модели, которые лучше вас знают, что вам хочется.
  - Netflix - аренда фильмов.

- **Рекомендательные системы** – это модели, которые лучше вас знают, что вам хочется.
  - Netflix - аренда фильмов.
  - Amazon - лидер в области E-Commerce.

- **Рекомендательные системы** – это модели, которые лучше вас знают, что вам хочется.
  - Netflix - аренда фильмов.
  - Amazon - лидер в области E-Commerce.
  - Яндекс.Музыка - стриминг музыки.



- **Рекомендательные системы** – это модели, которые лучше вас знают, что вам хочется.
  - Netflix - аренда фильмов.
  - Amazon - лидер в области E-Commerce.
  - Яндекс.Музыка - стриминг музыки.
- Для множества пользователей  $User$  и предметов  $Objects$  составляем матрицу, каждый элемент матрицы  $a_{ij}$  хранит оценку пользователя  $i$  для предмета  $j$ . Оценки может и не быть, зачастую такие матрицы очень разрежены, много нулей. Теперь задача заключается в заполнении пропусков.

- **Рекомендательные системы** – это модели, которые лучше вас знают, что вам хочется.
  - Netflix - аренда фильмов.
  - Amazon - лидер в области E-Commerce.
  - Яндекс.Музыка - стриминг музыки.
- Для множества пользователей  $User$  и предметов  $Objects$  составляем матрицу, каждый элемент матрицы  $a_{ij}$  хранит оценку пользователя  $i$  для предмета  $j$ . Оценки может и не быть, зачастую такие матрицы очень разрежены, много нулей. Теперь задача заключается в заполнении пропусков.
- **Коллаборативная фильтрация** - метод построения рекомендаций, основываясь на оценках других пользователей.

- Множество переменных, которые влияют на отношение пользователя к предмету и следовательно на рекомендации для этого пользователя, называется **контекстом**. Например, жанр фильма, время покупки товара, темп музыки.
- **Цель работы:** научиться использовать контекстные данные для улучшения результатов работы рекомендательной системы.

## Заключение

- Изучение известных матричных разложений (Singular Value Decomposition, Non-negative matrix factorization).

## Заключение

- Изучение известных матричных разложений (Singular Value Decomposition, Non-negative matrix factorization).
- Поиск вариантов встраивания контекстной информации в эти методы.





## Заключение

- Изучение известных матричных разложений (Singular Value Decomposition, Non-negative matrix factorization).
- Поиск вариантов встраивания контекстной информации в эти методы.
- Реализация системы на языке Python.

## Заключение

- Изучение известных матричных разложений (Singular Value Decomposition, Non-negative matrix factorization).
- Поиск вариантов встраивания контекстной информации в эти методы.
- Реализация системы на языке Python.
- Анализ эффективности в сравнении с обычными моделями, не использующими контекст.

## Литература I

-  Faris Alqadah и др. “Biclustering neighborhood-based collaborative filtering method for top-n recommender systems”. B: *Knowledge and Information Systems* (авг. 2014). ISSN: 0219-1377. DOI: 10.1007/s10115-014-0771-x. URL: <http://link.springer.com/10.1007/s10115-014-0771-x>.
-  Dmitry I Ignatov и др. *Boolean Matrix Factorisation for Collaborative Filtering : An FCA-Based Approach*.
-  Dmitry I Ignatov и др. *Recommendation of Ideas and Antagonists for Crowdsourcing Platform Witology*.
-  “Multiverse Recommendation: N-dimensional Tensor Factorization for Context-aware Collaborative Filtering”. B: *ACM Recommender Systems* (2010). URL: <http://xavier.amatriain.net/pubs/karatzoglu-recsys-2010.pdf>.



Спасибо за внимание!

[makazone@gmail.com](mailto:makazone@gmail.com)