

Практическая работа №14

Collaborative filtering

Цель работы: познакомиться и изучить Collaborative filtering, а также рассмотреть системы рекомендаций, основанные на совместной фильтрации.

Ход работы

Collaborative filtering (коллаборативная фильтрация) - это один из методов построения прогнозов (рекомендаций) в рекомендательных системах. Он использует известные предпочтения (оценки) группы пользователей для прогнозирования неизвестных предпочтений другого пользователя.

Для начала мы выведем два DataFrame, один с данными о фильмах, второй с данными о рейтингах.

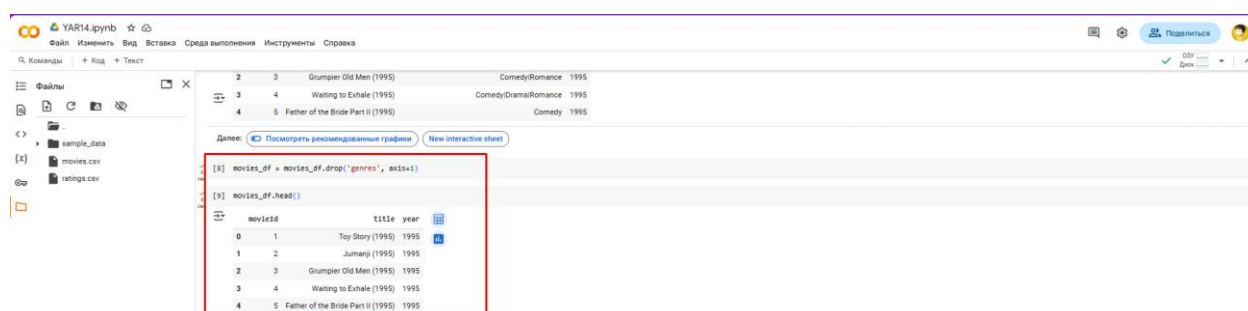


Рисунок 1 – данные о фильмах

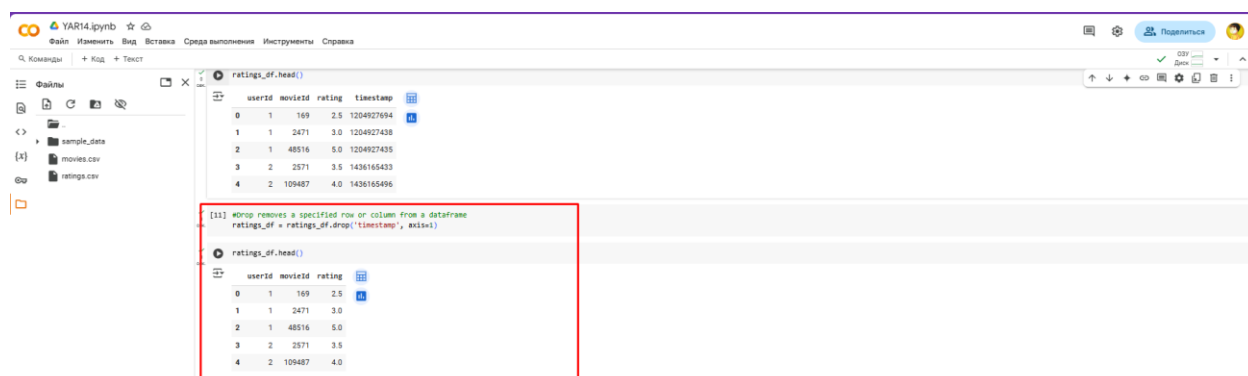
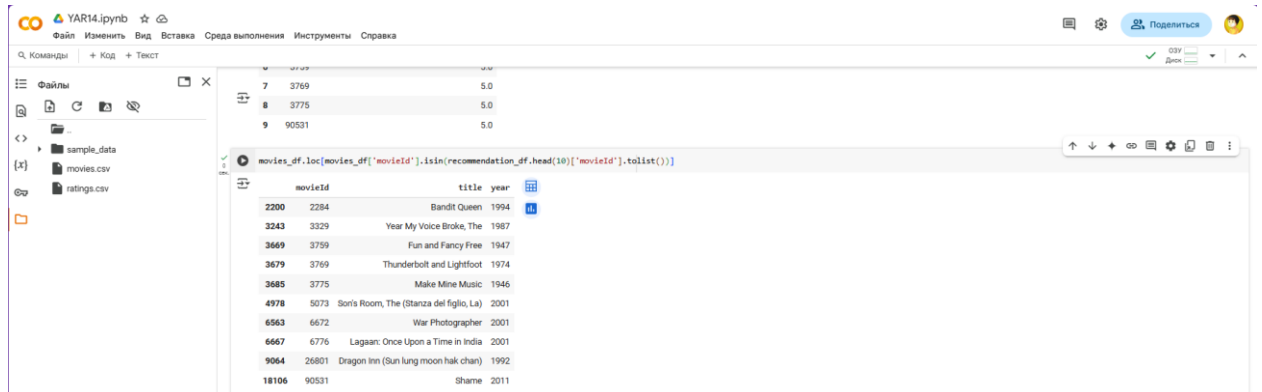


Рисунок 1.1 – данные о рейтингах

Наша задача выбрать пользователя с помощью фильмов, которые он посмотрел. Основываясь на его рейтинге просмотров фильмов, находим X самых популярных соседей.

Затем получаем список просмотренных фильмов пользователя для каждого соседа. Рассчитываем оценку сходства, используя некоторую формулу. И выводим рейтинг фильмов по количеству баллов.



The screenshot shows a Jupyter Notebook window with the title 'YAR14.ipynb'. The left sidebar displays a file explorer with folders 'sample_data' and 'movies.csv', and a file 'ratings.csv'. The main area shows a code cell with the following Python code:

```
movies_df.loc[movies_df['movieId'].isin(recommendation_df.head(10)['movieId'].tolist())]
```

Below the code, a table of movie data is displayed:

movieId	title	year
2200	Bandit Queen	1994
3243	Year My Voice Broke, The	1987
3669	Fun and Fancy Free	1947
3679	Thunderbolt and Lightfoot	1974
3685	Make Mine Music	1946
4978	Son's Room, The (Stanza del figlio, La)	2001
6563	War Photographer	2001
6667	Lagaan: Once Upon a Time in India	2001
9064	Dragon Inn (Sun lung moon hak chan)	1992
18106	Shame	2011

Рисунок 1.2 – рейтинг фильмов по количеству баллов, выведенных алгоритмом

Вывод: в результате практической работы мы познакомились с коллаборативной фильтрацией, рассмотрели системы рекомендаций, и реализовали простую версию такой системы, которая вывела нам рейтинг фильмов по количеству баллов.