

Рекомендательные системы

Студия уроки настоящего

ГБОУ СОШ №255

10 класс

Антипова Ева, Геращенко Соня,
Камилова Ниса, Сердукова Полина



Цель проекта

Ознакомление с рекомендательными системами и их составляющими, применениями. Формирование общего мнения о подобной технологии.

...

Задачи проекта

- Собрать общую информацию по теме рекомендательных систем для дальнейшего изучения
- Проанализировать алгоритмы рекомендательных систем
- Найти первоисточники в интернете с документацией по библиотекам
- Структурировать информацию по теме
- Создать общий документ с описанием алгоритмов и библиотек, используемых для создания рекомендательных систем



Актуальность рекомендательных систем



повышение
удовлетворенности
пользователей

...



увеличение
продажи и
конверсию

...



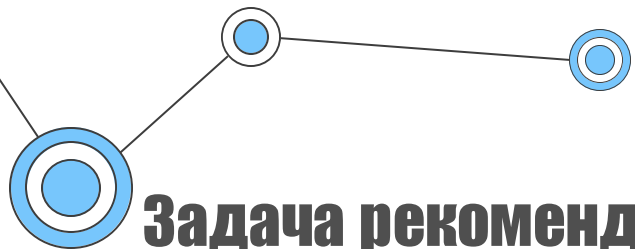
улучшение общей
пользовательской реакции
на предоставляемый
контент

...



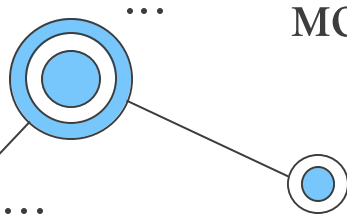
снижение
оттока
пользователей

...



Задача рекомендательной системы -

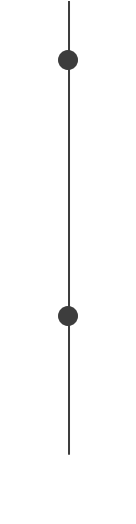
проинформировать
пользователя о товаре,
который ему может
быть наиболее
интересен в данный
момент времени.





01

**Основа
рекомендательных
систем**

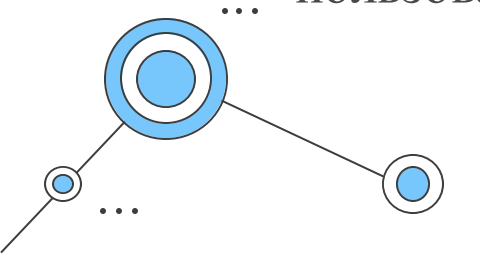




...

Рекомендательные системы -

это системы фильтрации информации, предоставляющие пользователям персонализированную рекомендацию по объектам на основе хранимых сервисом данных о ... пользователях и объектах.



Зачем пользователям рекомендательные системы?

01

Обеспечивают для пользователя
отвечающий его потребностям
контент

02

Обеспечивают экономию
времени на поиск контента

03

Обеспечивают снижение
интенсивности коммуникации
с клиентом



Зачем сервисам рекомендательные системы?



01
...

Увеличивают метрики
удовлетворенности
пользователя

02
...

Позволяют экономить на
коммуникациях

03
...

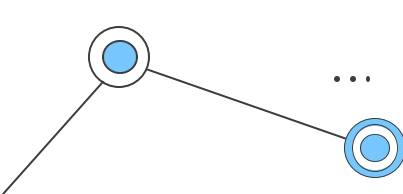
Обеспечивают
дополнительный канал
продаж для услуг



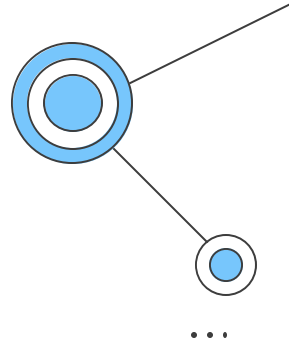
02

Библиотеки





Какие существуют библиотеки?



01 Implicit и LightFM

02 RecTools

03 Pure SVD

04 Surprise

Виды рекомендательных систем

Методы на
основе
контента

Методы
коллаборативной
фильтрации

Гибридные
рекомендательные
системы

Метрические
методы

Методы на основе
матричных
разложений



03

**Коллаборативная
фильтрация**



Коллаборативная фильтрация —

это технология, которая позволяет прогнозировать предпочтения конкретного пользователя интернет-ресурса, сравнивая его интересы с интересами других посетителей ресурса.

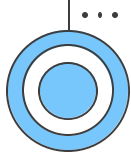


04

**Метрический
метод**

Таблица оценок пользователей

Фильм №	«Бег»	«Идиот»	«Мимино»	«Покровские ворота»	«Римские каникулы»
1	3			7	9
2	8	9	6	4	6
3	9	8	4		5
4			5	7	7
5	8	8		3	

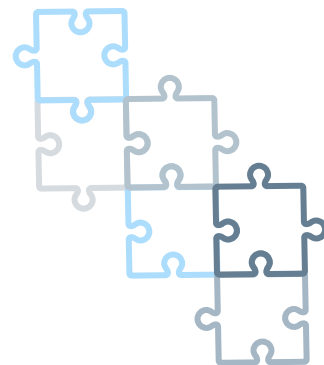


Векторы:

$$\vec{u}_5 = [8, 8, \text{nan}, 3, \text{nan}] \quad \vec{u}_3 = [9, 8, 4, \text{nan}, 5]$$

Формула косинуса для нахождения близости:

$$\cos(u_5, u_3) = \frac{\vec{u}_5 \cdot \vec{u}_3}{|\vec{u}_5| |\vec{u}_3|} = \frac{8 \cdot 9 + 8 \cdot 8}{\sqrt{8^2 + 8^2} \sqrt{9^2 + 8^2}} = 0.998$$



Формулы для выставления оценок на основе метрики близости:

$$\widehat{r}_{ui} = \frac{\sum_{j \in N(u)} r_{ji}}{|N(u)|}$$

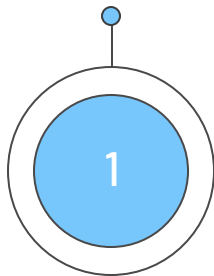
$$\hat{r}_{5, \text{Минимо}} = \frac{\sum_{j \in N[2,3]} r_{j, \text{Минимо}}}{|\{2, 3\}|} = \frac{6 + 4}{2} = 5$$

Пользователь – u
Предсказание - \hat{r}
Объект – i
Настоящая оценка,
которую выставил
пользователь – r

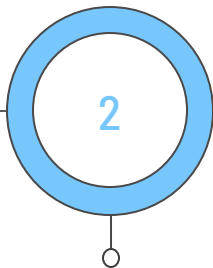


Алгоритм

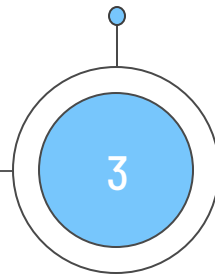
Выбрать метрику
близости, произвести
ее расчет



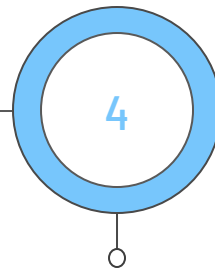
Рассчитать оценки объектов
для пользователя на основе
известных оценок соседей



Выбрать n ближайших
соседей для
пользователя



Отсортировать объекты по
убыванию оценок и показать
пользователю самый
высокорейтинговый



Плюсы

Простота реализации

Интерпретируемость

Низкая вычислительная сложность

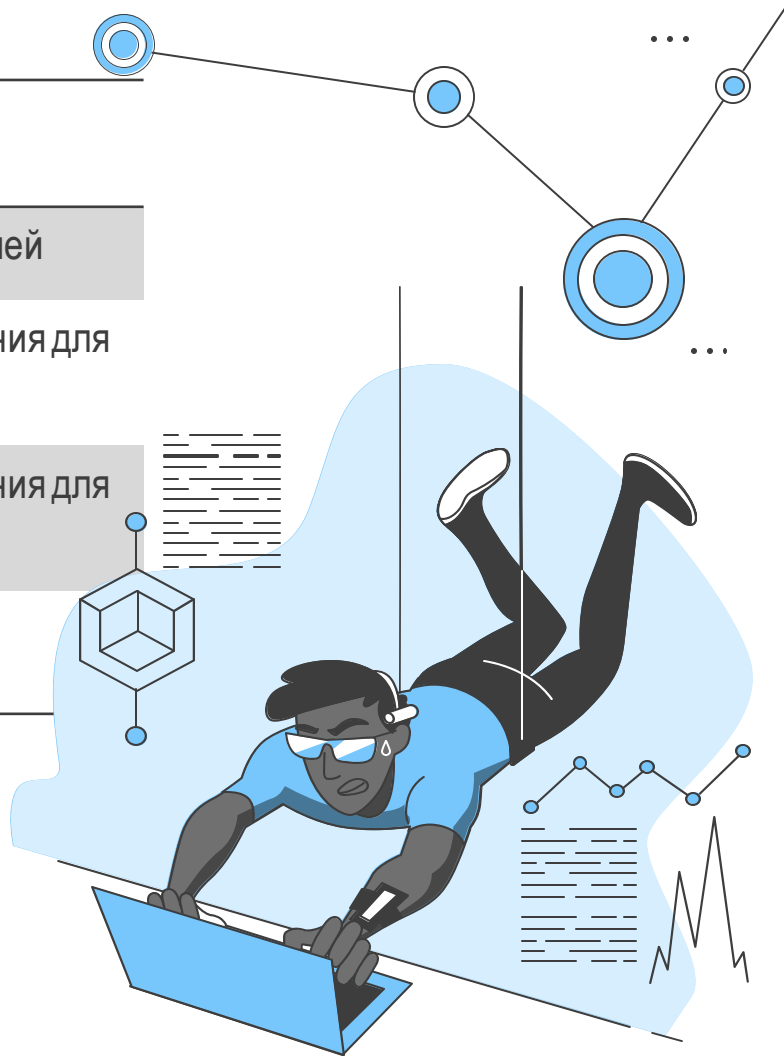
Легкость добавления новых объектов и пользователей

Минусы

Нужны оценки пользователей

Невозможность предсказания для нового пользователя

Невозможность предсказания для нового объекта





05

**Метод на основе
разложения
матрицы**

1. Формула мнимой оценки:

$$\hat{r} = \mu + b_u + b_i$$

μ – средняя оценка объекто

b_u – смещение пользователя

b_i – смещение объекта



2. Формула для вычисления ошибки предсказания:

$$L(u, i) = (r_{ui} - \hat{r}_{ui})^2$$



4. Формула для мнимой оценки с использованием векторов:

$$r_{ui} = \mu + b_u + b_i + q_u^T q_i$$

q_u - вектор пользователя

q_i - вектор объекта



3. Формула для минимизирования ошибки:

$$\sum_{(i,u)} (r_{ui} - \mu - b_u - b_i)^2 \rightarrow \min_{b_u, b_i}$$

\sum - сумма предсказаний по всем фильмам для каждого пользователя



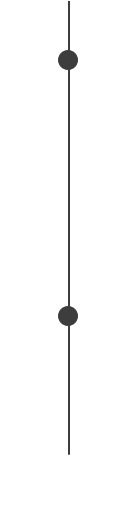
5. Формула для минимизирования ошибки:

$$\sum_{(i,u)} (r_{ui} - \mu - b_u - b_i - q_u^T q_i)^2 \rightarrow \min_{b_u, b_i, q_u, q_i}$$



06

Метрики качества



На уровне машинного обучения используются следующие метрики:

MAE, RMSE

На уровне поиска информации оцениваются следующие метрики:

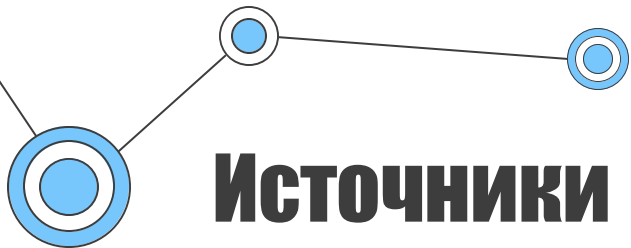
Accuracy, Precision,
Recall, Recall, ROC

На уровне взаимодействия пользователя и компьютера оцениваются следующие метрики:

Diversity, Trust,
Novelty, Serendipity

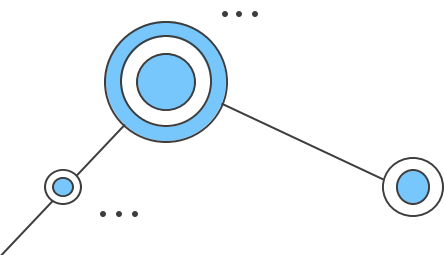
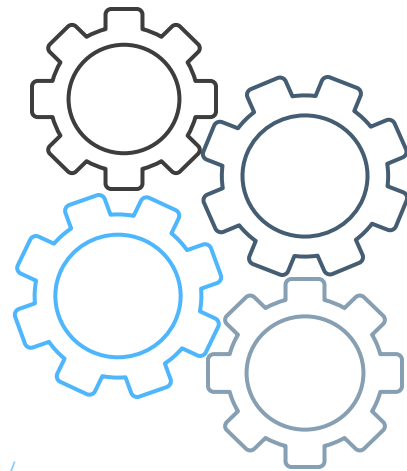
На уровне разработки программного обеспечения оцениваются следующие метрики:

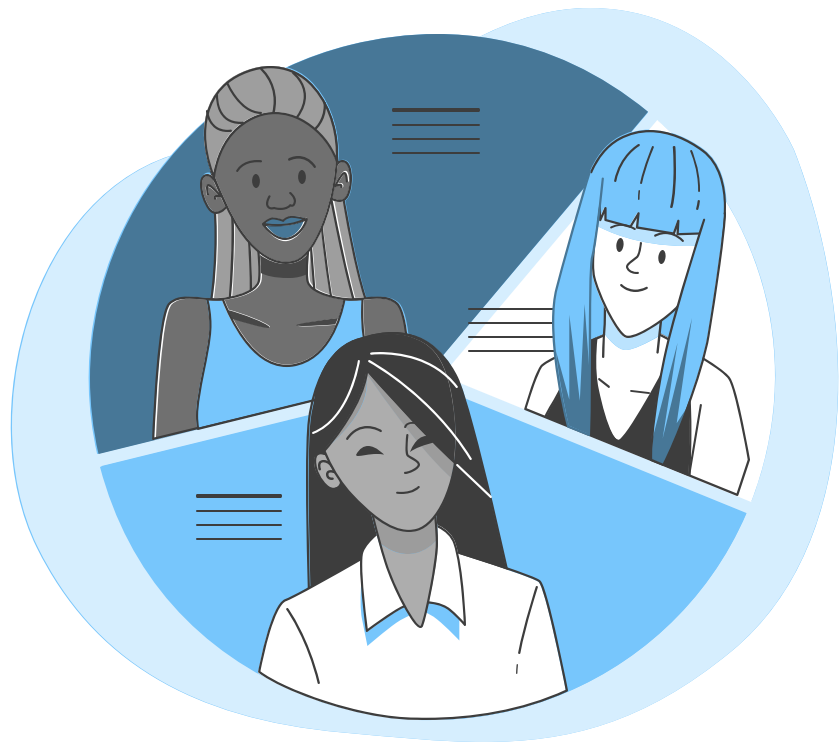
Real-time,
Robustness,
Scalability



Источники

- <https://www.dissercat.com/content/issledovanie-i-razrabotka-algoritmov-rekomendatelnykh-sistem-na-osnove-grafovykh-modelei-dan>
- https://rdc.grfc.ru/2023/05/recommendation_services_methods_and_regulation/
- <https://habr.com/ru/articles/773126/>
- <https://tproger.ru/articles/kak-ustroena-rectools-opensource-biblioteka-dlja-rekomendatelnykh-sistem>





...



**Спасибо за
внимание!**

...

A network diagram consisting of three blue circles connected by lines. The circles are arranged in a line, with the rightmost one being larger and having a white ring around it. There are three dots above the first circle, three dots above the second circle, and three dots below the third circle.