

МОБОД-2022

Рекомендательные сервисы в продакшене

Николай Анохин

8 сентября 2022 г.

Информация о курсе

●oooooo

Зачем нужны рекомендательные сервисы
oooooooooooooooooooo

Что дальше
ooooo

Информация о курсе



Входной опрос



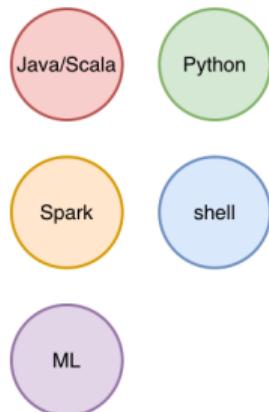
<http://shorturl.at/dosU1>

Николай Анохин

Академический опыт



Навыки



Индустриальный опыт

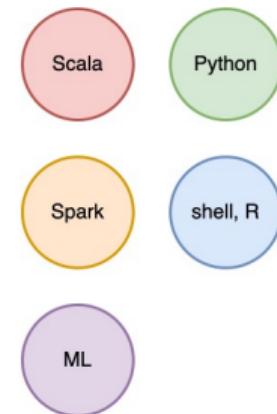


Дарья Никанорова

Академический опыт



Навыки



Индустриальный опыт



Telegram

Чат курса: <https://t.me/+qri5NHAS4Pc3YTky>

- Вопросы вне занятия можно задать в личке или в чате курса (лучше)
- Тегайте нас, чтобы мы не пропустили ваш комментарий в общем потоке сообщений
- Если ответа не последовало в течении 24 часов, то мы, вероятно, не увидели ваше сообщение. Не стесняйтесь его продублировать



Как задать вопрос

- Голосом
- В специально выделенное для этого время
- Перед тем как спросить будущим тоном поставить несколько знаков вопроса

20:23 Саша: ????

20:23 Преподаватель: Ждём вопроса от Саши

20:24 Саша: Какая метрика хорошо работает в задаче рекомендаций?



Если что-то пошло не так

- Пропал голос
- Исчезло изображение
- Плохо слышно
- Любые проблемы другого характера

Сразу пишем в чат много минусов и не ждем других участников. Если вы увидели, что в чате кто-то написал много минусов, а у вас всё хорошо, то поставьте несколько плюсов:

20:24 Петя: - - - - -

20:25 Саша: + + +

20:25 Ольга: + + + + +



Информация о курсе
oooooooo

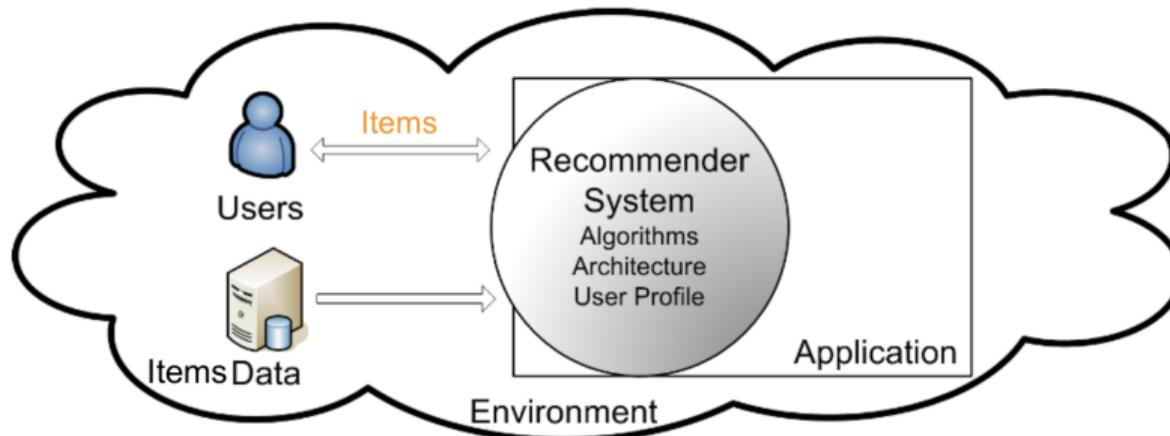
Зачем нужны рекомендательные сервисы
●oooooooooooo

Что дальше
ooooo

Зачем нужны рекомендательные сервисы



Recommender Systems (RS) are software tools and techniques providing suggestions for **items** to be of use to a **user** [RRSK10].



The screenshot shows the Netflix homepage. At the top, there's a navigation bar with links for Home, TV Shows, Movies, New & Popular, and My List. To the right of the navigation are search, Kids, notifications (with 3 notifications), and account settings icons. The main title 'THE TINDER SWINDLER' is displayed prominently in large blue letters. Below the title are two buttons: 'Play' and 'More Info'. A rating of '16+' is shown in the bottom right corner. The background image is a dark, atmospheric scene of a bridge over water at night. Below the main title, there's a section for 'TV Shows' featuring six other series: 'THE WOMAN IN THE HOUSE ACROSS THE STREET FROM THE GIRL IN THE WINDOW', 'OZARK', 'THE SINNER', 'STAY CLOSE', 'After Life', and 'DOWNTON ABBEY'. Each show has its title, a small image, and a 'TOP 10' badge. Below this, there's a 'Trending Now' section with six more series: 'Sweet Magnolias', 'THE TINDER SWINDLER', 'THE POWER OF THE DOG', 'THE CROWN', 'BRIDGERTON', and 'MONEY HEIST'. Each show has its title, a small image, and a 'TOP 10' badge.



The screenshot shows the Netflix homepage. At the top, the title 'THE TINDER SWINDLER' is displayed with a small orange circle containing the number '1' indicating it's the first episode. Below the title are two buttons: 'Play' and 'More Info'. To the right of the title, there are icons for search, kids, notifications (with a count of 3), and account settings. A volume icon and a rating of '16+' are also present. The main background image is a dark, atmospheric scene of a bridge over water at night. Below the main title, there are sections for 'TV Shows' and 'Trending Now', each displaying several movie and TV show thumbnails with their titles and release information.

TV Shows

- N THE WOMAN IN THE HOUSE ACROSS THE STREET FROM THE GIRL IN THE WINDOW TOP 10 NEW EPISODES
- N OZARK TOP 10 NEW EPISODES
- N THE SINNER NEW EPISODES
- N HERMAN CORNELL'S STAY CLOSE NEW EPISODES
- N After Life NEW EPISODES
- DOWNTON ABBEY

Trending Now

- N Sweet Magnolias NEW EPISODES
- N THE TINDER SWINDLER TOP 10
- N THE POWER OF THE DOG
- N THE CROWN
- N BRIDGERTON
- N MONEY HEIST NEW EPISODES



NETFLIX

Home TV Shows Movies New & Popular My List

THE TINDER SWINDLER 1

Play More Info

16+

TV Shows

TOP 10 N THE WOMAN IN THE HOUSE ACROSS THE STREET FROM THE GIRL IN THE WINDOW NEW EPISODES 2 OZARK TOP 10 N THE SINNER NEW EPISODES N HERLÁN CORRÉA'S STAY CLOSE NEW EPISODES N After Life NEW EPISODES DOWNTON ABBEY N

Trending Now

TOP 10 N Sweet Magnolias NEW EPISODES N THE TINDER SWINDLER TOP 10 N THE POWER OF THE DOG N THE CROWN N BRIDGERTON N MONEY HEIST TOP 10 N NEW EPISODES



The screenshot shows the Netflix homepage. At the top, there is a navigation bar with links to Home, TV Shows, Movies, New & Popular, and My List. To the right of the navigation are search, Kids, notifications (with 3 notifications), and account settings icons. The main banner features the TV show 'THE TINDER SWINDLER' with a play button and a 'More Info' button. A large orange circle with the number '1' is overlaid on the banner. Below the banner, there are sections for 'TV Shows' and 'Trending Now'. The 'TV Shows' section includes thumbnails for 'THE WOMAN IN THE HOUSE ACROSS THE STREET FROM THE GIRL IN THE WINDOW', 'OZARK', 'THE SINNER', 'STAY CLOSE', 'After Life', and 'DOWNTON ABBEY'. A large orange circle with the number '2' is overlaid on the 'OZARK' thumbnail. The 'Trending Now' section includes thumbnails for 'Sweet Magnolias', 'THE TINDER SWINDLER', 'THE POWER OF THE DOG', 'THE CROWN', 'BRIDGERTON', and 'MONEY HEIST'. A large orange circle with the number '3' is overlaid on the 'THE TINDER SWINDLER' thumbnail. On the right side of the page, there are additional navigation links for 'Top 10', 'New Releases', 'Watch Party', 'Profile', and 'Help'.



The screenshot shows the Netflix homepage. At the top, there is a navigation bar with links to Home, TV Shows, Movies, New & Popular, and My List. To the right of the navigation are search, Kids, notifications (with 3 notifications), and account settings icons. The main banner features the title 'THE TINDER SWINDLER' in large blue letters, with a small orange circle containing the number '1' in the top right corner. Below the banner are two buttons: 'Play' and 'More Info'. To the right of the banner, there is a volume icon and a rating of '16+'. The page is divided into sections: 'TV Shows' and 'Trending Now'. Under 'TV Shows', there are six thumbnails for different series: 'THE WOMAN ACROSS THE STREET' (ranked 4, with a red 'TOP 10' badge), 'OZARK' (ranked 2, with a red 'TOP 10' badge), 'THE SINNER' (with a red 'NEW EPISODES' button), 'STAY CLOSE' (with a red 'NEW EPISODES' button), 'After Life' (with a red 'NEW EPISODES' button), and 'DOWNTON ABBEY' (with a red 'TOP 10' badge). Each thumbnail has a small orange circle with a number indicating its ranking. Under 'Trending Now', there are seven thumbnails: 'Sweet Magnolias' (ranked 4, with a red 'TOP 10' badge), 'THE TINDER SWINDLER' (ranked 3, with a red 'TOP 10' badge), 'THE POWER OF THE DOG', 'THE CROWN', 'BRIDGERTON', and 'MONEY HEIST' (with a red 'TOP 10' badge and a red 'NEW EPISODES' button). Each thumbnail also has a small orange circle with a number indicating its ranking.



The screenshot shows the Netflix homepage with several recommended TV shows highlighted by orange circles with numbers:

- THE TINDER SWINDLER** (1) - Top recommendation, currently playing.
- THE WOMAN ACROSS THE STREET** (4) - Top 10, new episodes available.
- OZARK** (2) - Top 10, new episodes available.
- THE SINNER** - New episodes available.
- STAY CLOSE** - New episodes available.
- After Life** - New episodes available.
- DOWNTON ABBEY** - New episodes available.
- Sweet Magnolias** (4) - Trending Now, new episodes available.
- THE TINDER SWINDLER** (3) - Trending Now, new episodes available.
- THE POWER OF THE DOG** - Trending Now.
- THE CROWN** - Trending Now.
- BRIDGERTON** - Trending Now.
- MONEY HEIST** (10) - Trending Now, new episodes available.

Other visible elements include the Netflix logo, navigation links (Home, TV Shows, Movies, New & Popular, My List), search bar, Kids mode, notifications, and a 16+ rating indicator.



The screenshot shows the Netflix homepage with several recommended shows highlighted by orange circles with numbers:

- THE TINDER SWINDLER** (1) - Top recommendation, main banner.
- THE WOMAN ACROSS THE STREET FROM THE GIRL IN THE WINDOW** (4) - TV Show, Trending Now.
- OZARK** (2) - TV Show, Trending Now.
- THE SINNER** - TV Show, Trending Now.
- STAY CLOSE** - TV Show, Trending Now.
- After Life** - TV Show, Trending Now.
- DOWNTON ABBEY** - TV Show, Trending Now.
- Sweet Magnolias** (4) - TV Show, Trending Now.
- THE TINDER SWINDLER** (3) - TV Show, Trending Now.
- THE POWER OF THE DOG** - TV Show, Trending Now.
- THE CROWN** - TV Show, Trending Now.
- BRIDGERTON** - TV Show, Trending Now.
- MONEY HEIST** (16+) - TV Show, Trending Now.

Other visible elements include the Netflix logo, navigation links (Home, TV Shows, Movies, New & Popular, My List), search bar, Kids mode, notifications, and account settings.



RS в индустрии и науке

Компании

Amazon.com, YouTube, Netflix, Yahoo, Tripadvisor, Spotify, Booking.com

Конференции

RecSys, KDD, WSDM, SIGIR

Книги

Recommender Systems Handbook [RRSK10], Mining of Massive datasets [LRU14]



Зачем RS бизнесу

- Увеличить продажи
- Продвигать более разнообразные айтемы
- Улучшить пользовательский опыт
- Добиться большей лояльности
- Лучше понимать пользователей



Зачем RS пользователям

- Найти лучший айтем
- Найти **все** подходящие айтемы
- Найти последовательность или набор айтемов
- Залипнуть
- Помочь другим сделать выбор



Зачем RS инженерам



- Делать высоконагруженный отказоустойчивый сервис :D
- Анализировать большие данные
- Окунуться в волшебный мир матана машинного обучения
- Объективно измерять результат своей работы
- Узнать много нового о людях
- Все это за зарплату

RS vs другие задачи ML [SBE⁺21]

- RS ориентированы на продакшен
- Наблюдаемые данные очень разреженные
- Отсутствующие данные – missing not-at-random
- Отсутствующие данные – либо ненаблюдаемые позитивные, либо негативные
- Рекомендательные сервисы живут в петле обратной связи



Пройдя этот курс, вы...

- Разработаете свой рекомендательный сервис (почти) с нуля
- Научитесь использовать Spark для нужд RS
- Сможете выбирать правильные инструменты под конкретную задачу
- Узнаете о проблемах, возникающих в боевых RS, и научитесь их решать
- Будете в курсе SOTA моделей и задач RS, над которыми работают ученые
- Подготовитесь к собеседованию по RS на junior позицию



Информация о курсе
oooooooo

Зачем нужны рекомендательные сервисы
oooooooooooooooooooo

Что дальше
●oooo

Что дальше



Программа модуля

Неделя	Тема	Квиз	Семинар	Домашка
1	Рекомендательные сервисы в продакшене	✓	✓	
2	Метрики и базовые подходы	✓	✓	
3	Классические алгоритмы 1	✓	✓	
4	Классические алгоритмы 2	✓	✓	✓
5	Нейросетевые рекомендеры	✓	✓	
6	Нерешенные проблемы и новые направления	✓	✓	
7	Рекомендации и Reinforcement Learning	✓	✓	



Оценка $\in [0, 1]$

- Идеальное выполнение домашки = 0.6 (нижняя граница “тройки”)
- Идеальное выполнение квизов и домашки = 0.9 (нижняя граница “пятерки”)
- Выберем трех самых активных слушателей и накинем им по 0.1
- Дополнительные баллы можно получить на зачете





Литература I

-  Jure Leskovec, Anand Rajaraman, and Jeffrey David Ullman, *Mining of massive datasets*, 2nd ed., Cambridge University Press, USA, 2014.
-  Francesco Ricci, Lior Rokach, Bracha Shapira, and Paul B. Kantor, *Recommender systems handbook*, 1st ed., Springer-Verlag, Berlin, Heidelberg, 2010.
-  Harald Steck, Linas Baltrunas, Ehtsham Elahi, Dawen Liang, Yves Raimond, and Justin Basilico, *Deep learning for recommender systems: A netflix case study*, AI Magazine 42 (2021), no. 3, 7–18.

