

The background is a dark blue gradient. It features several vertical white lines of varying lengths. Scattered throughout are small squares in three colors: light blue, orange, and pink. Some squares are solid, while others are outlined.

# Проект курса SkillFactory Рекомендательная система

Трек ML-инженер

# Цели и задачи проекта

Содержание:

1. Цель.
2. Постановка задачи.
3. Выбор технических метрик качества.
4. Описание данных.
5. Факторы.
6. Выбор модели.
7. Результат.

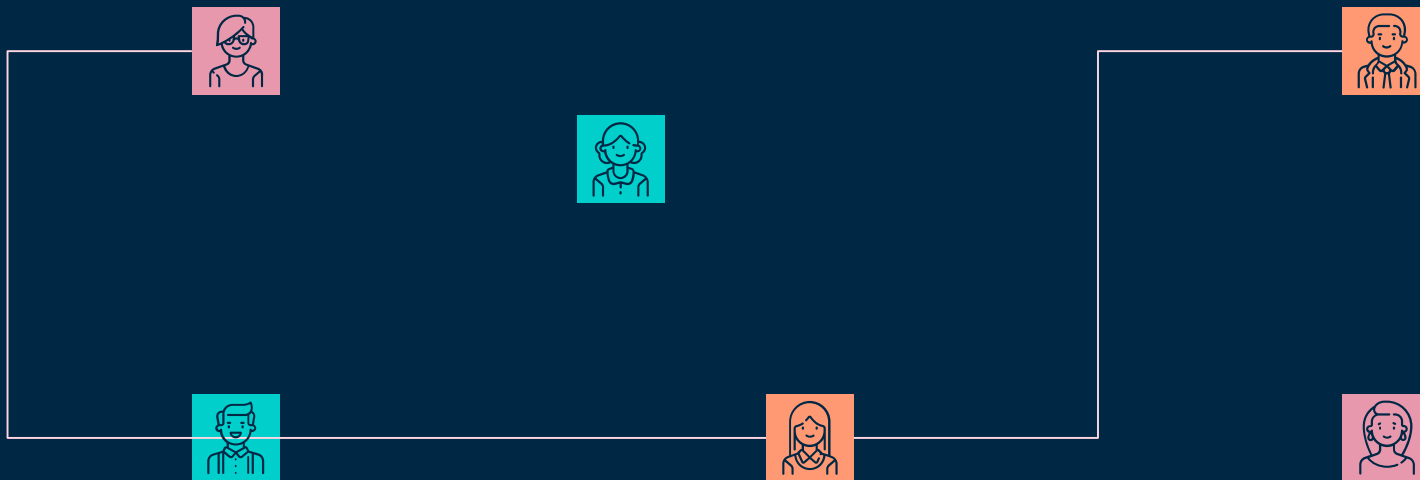
# Цель

Повысить прибыль от допродаж в интернет-магазине на 20%.



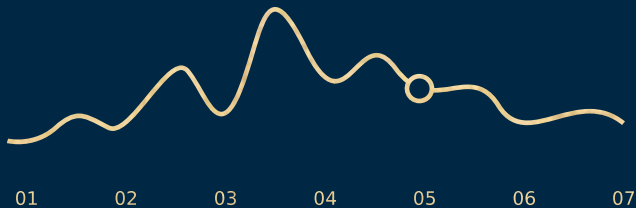
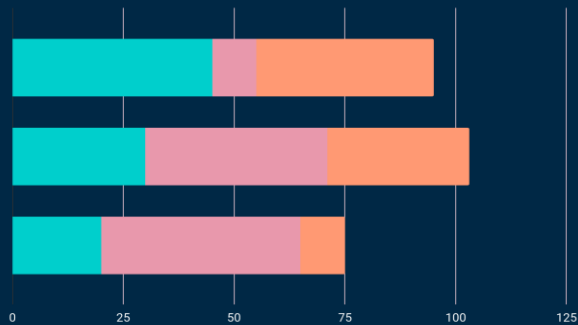
# Задача

Построить рекомендательную систему на основе предоставленных данных.

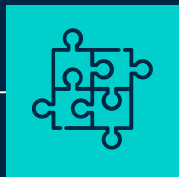


# Метрика

*Precision@3* - показывает, какую долю из того, что мы порекомендовали, купил пользователь



# Данные



01

events

датасет с событиями

| # | Column        | Dtype   |
|---|---------------|---------|
| 0 | timestamp     | int64   |
| 1 | visitorid     | int64   |
| 2 | event         | object  |
| 3 | itemid        | int64   |
| 4 | transactionid | float64 |



02

category\_tree

файл с деревом категорий

| # | Column     | Dtype   |
|---|------------|---------|
| 0 | categoryid | int64   |
| 1 | parentid   | float64 |



03

item\_properties

файл со свойствами товаров

| # | Column    | Dtype  |
|---|-----------|--------|
| 0 | timestamp | int64  |
| 1 | itemid    | int64  |
| 2 | property  | object |
| 3 | value     | object |

# Факторы

User-Item

- visitorid
- itemid
- rating (оценка на основе action:  
    'view': 1,  
    'addtocart': 2,  
    'transaction': 3)



# Модель

В проекте используется метод колоборативной фильтрации, учитывающий средние оценки каждого пользователя (surprise.`KNNWithMeans`).





# Результат

На тестовых данных достигли результата:

- RMSE: 0.6361
- Precision@3: 6.2%

