



ТИНЬКОФФ
ОБРАЗОВАНИЕ

Тинькофф - кэшбеки

Создание рекомендательной системы кэшбеков

Тимошкин Максим МГУ 6 курс

Никишин Андрей СПбПУ 4 курс



Какие стояли задачи?

- ✓ Создание Телеграм-бота
- ✓ Развертывание бота
- ✓ Рекомендательная система кэшбеков
- ✓ Оценка качества



Немного о рекомендательных системах



Существуют два основных типа рекомендательных систем: коллаборативные и контентные.



Коллаборативные системы рекомендаций анализируют взаимодействия пользователей с товарами, чтобы предоставить рекомендации на основе схожих пользовательских предпочтений.



Контентные системы анализируют характеристики айтемов и стремятся на их основе рекомендовать похожие айтемы.

Холодный старт

В начальный момент **моделька**
ничего не знает о пользователе :(

Фильтр пузыря

Ограничение **разнообразия**
рекомендаций

Exploration vs Exploitation

Если мы знаем, что нравится
пользователю, должны ли мы
рекомендовать **одно и то же** ?

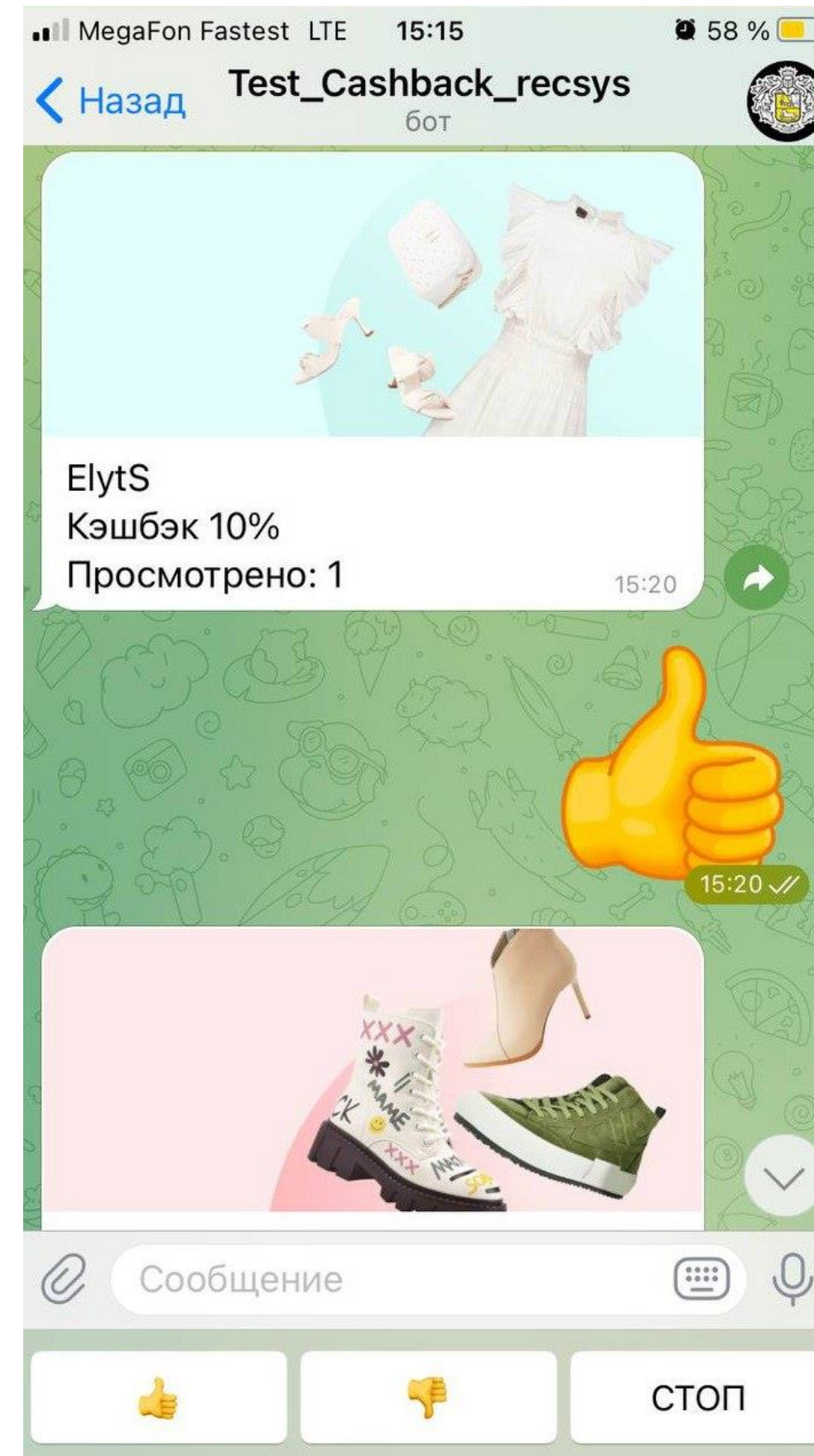
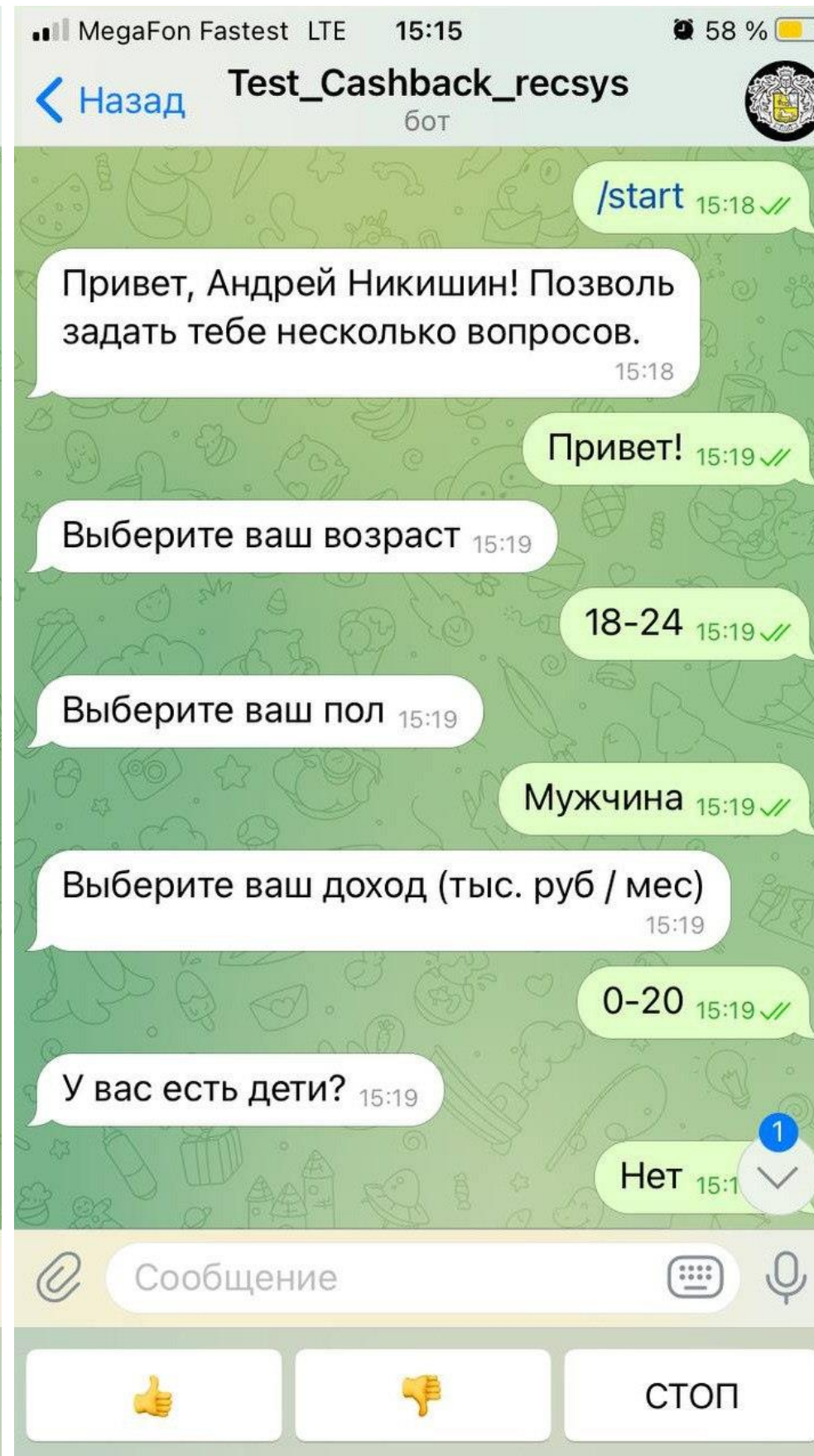
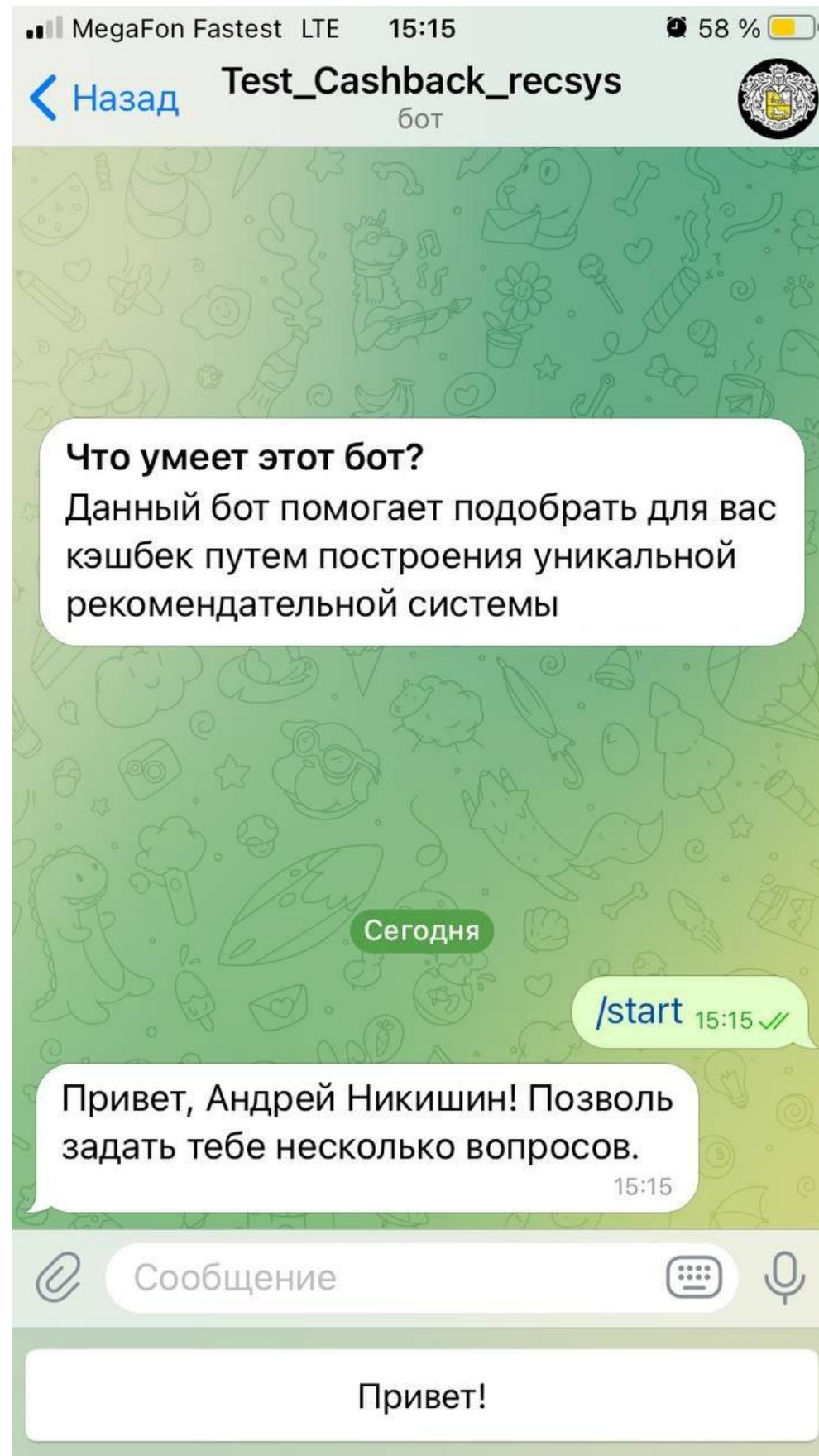
Основные проблемы **любых** рекомендательных систем



Roadmap to Tinkoff



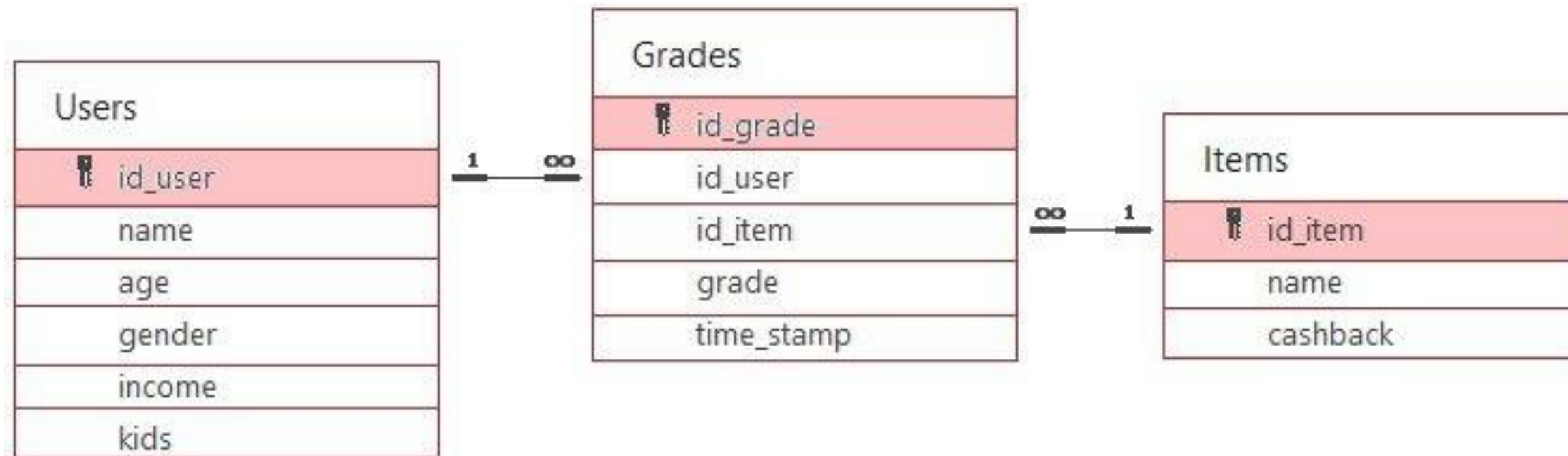
Создание ТГ-бота



1. Telebot
2. Стартовые вопросы
3. Обработка ввода
4. Режим лайк-дизлайк

Создание базы данных

1. Архитектура БД



2. Выбор СУБД

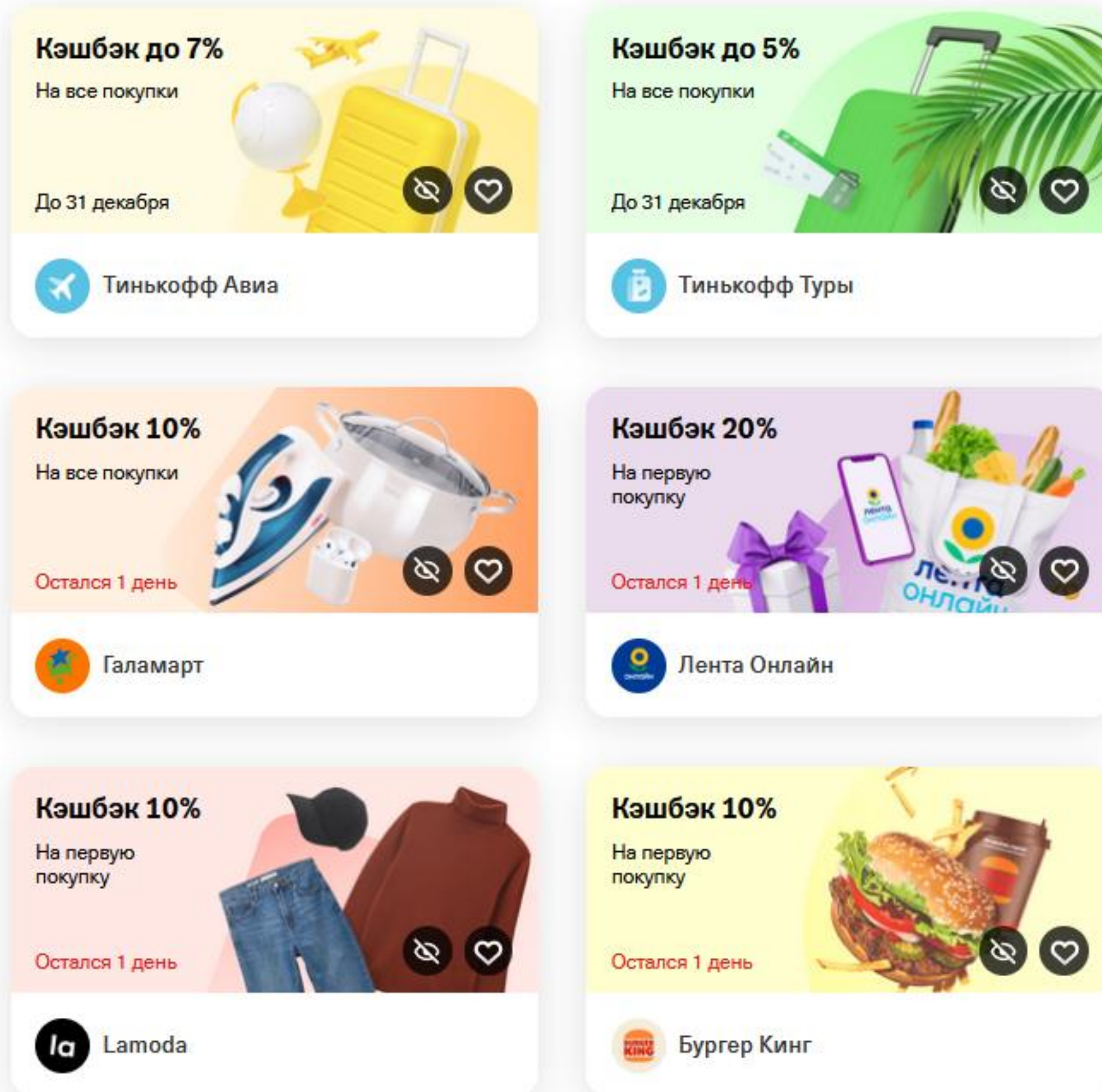
Преимущества SQLite

- Простота использования.
- Низкие требования к ресурсам.



Парсинг кэшбеков

Рекомендуем



1. Парсинг HTML с помощью регулярных выражений;
2. Автоматическое скачивание всех изображений кэшбеков;
3. Автоматическое получение информации (название, процент) с кэшбеков;
4. Заполнение айтемов в БД.

Формирование матрицы объекты-признаки

	Item_1	Item_2	Item_M	User_Features	
User_1	9 1 0 5 2 5 7 2 4	8 7 6 4 0 1 9 2 6	6 2 9 0 0 0 6 9 0	5 5 8 0 3 4 9 9 5	3 6 4 3 5 6 0 0 0
	2 6 9 1 8 6 3 6 1	4 9 5 2 9 3 9 5 1	9 3 7 8 4 3 0 1 9	3 9 1 3 0 7 2 0 0	2 7 5 0 4 0 7 9 7
	1 4 2 7 7 7 5 2 6	8 7 4 3 9 7 2 2 7	4 9 6 7 4 7 7 9 6	7 1 5 3 6 8 2 2 5	3 0 8 4 9 4 8 8 6
	6 5 1 3 1 7 7 9 4	6 1 2 7 2 9 4 3 1	2 5 5 8 4 7 0 4 0	8 0 4 7 6 1 0 7 0	0 7 6 8 8 7 5 5 7
User_2	3 8 7 8 5 7 8 9 4	3 6 4 1 4 0 4 1 7	5 6 8 1 6 8 0 0 4	0 7 5 6 1 4 8 3 6	8 5 6 3 1 0 7 2 5
	9 8 5 8 0 1 9 5 2	0 2 1 3 1 5 9 6 9	1 0 4 6 7 9 9 4 5	1 7 7 9 0 2 0 1 9	2 4 1 7 7 3 6 3 7
	9 5 4 0 1 3 9 1 7	4 3 7 4 7 0 7 6 8	0 6 4 4 6 4 0 4 0	2 3 1 5 1 1 5 4 4	8 3 0 7 8 0 5 9 7
	5 8 0 6 7 4 3 0 1	9 4 5 0 7 2 8 9 0	6 2 7 9 7 5 2 7 2	0 4 4 0 5 1 7 1 4	2 8 6 4 5 3 6 3 0
User_3	1 6 9 6 4 4 1 6 8	6 0 0 3 8 5 0 3 4	2 3 1 4 4 4 7 7 5	3 2 1 2 1 6 4 6 8	6 9 0 6 3 8 3 8 2
	0 3 2 7 7 8 5 7 7	8 0 8 6 0 8 4 9 0	4 0 3 3 9 0 5 9 7	8 3 9 2 8 5 7 8 3	4 6 8 7 1 9 1 2 9
	1 8 8 2 4 1 8 9 2	3 8 4 6 5 5 2 5 8	SYSTEM FAILURE	4 1 0 3 8 9 0 8 4	2 6 1 2 5 4 0 9 2
	1 0 8 0 5 8 7 3 9	7 1 2 7 8 9 7 9 8	3 2 3 9 8 0 3 6 0	5 2 8 9 0 0 8 9 6	9 0 5 3 3 6 3 3 6
...	4 4 5 3 3 4 3 5 3	9 5 1 5 2 0 4 7 2	5 1 9 8 7 8 2 4 4	3 4 0 4 0 9 1 4 6	2 6 2 1 3 0 2 1 0
	7 4 1 2 9 4 0 7 0	0 1 8 5 8 3 8 1 6	8 7 0 0 5 2 4 7 9	4 2 1 7 4 9 1 0 3	8 6 9 1 3 6 0 0 5
	3 3 6 3 1 8 9 0 9	7 2 5 0 0 3 6 3 7	7 5 9 7 8 5 6 5 3	3 4 3 4 6 4 3 6 6	7 9 9 2 9 1 8 4 1
	4 9 1 3 6 0 8 5 3	1 8 3 2 3 2 5 9 2	4 4 2 0 7 6 4 8 8	6 2 4 9 4 5 6 9 5	3 9 2 9 3 7 9 7 9
User_N	8 5 8 1 0 3 6 5 1	2 8 5 6 3 6 5 2 5	6 0 4 0 3 3 3 3 3	0 3 4 4 4 4 2 2 0	6 4 8 7 7 1 0 7 8
	6 6 9 8 3 2 0 7 8	1 8 7 2 1 5 7 8 0	8 9 3 3 7 1 3 4 9	7 7 9 6 1 1 2 5 6	4 2 9 2 4 5 1 5 2
	2 7 7 4 7 8 0 4 3	1 6 4 3 8 2 7 3 0	3 2 2 2 0 3 6 0 6	2 1 9 9 9 4 5 9 4	2 9 3 7 6 4 6 0 3
	0 8 5 9 5 6 2 1 4	4 8 0 9 2 8 4 5 7	8 8 7 1 9 9 5 0 1	7 3 8 3 6 4 2 9 6	4 5 1 1 3 8 8 8 2
	4 5 3 9 3 8 9 6 7	8 0 2 2 4 9 3 9 0	6 8 2 0 1 4 2 3 7	6 4 8 7 3 8 3 4 8	4 3 4 6 4 3 6 1 3
	5 3 4 9 0 3 0 1 9	2 2 7 9 0 4 6 6 4	7 5 8 5 4 3 6 7 4	1 0 6 2 2 3 8 5 7	1 2 3 4 7 8 6 4 3
	6 4 5 5 5 4 4 8 9	7 2 5 5 0 3 1 0 9	3 1 3 3 7 8 0 7 4	4 3 1 4 1 4 6 3 7	4 4 0 2 4 3 5 7 4
	3 4 8 7 0 8 0 0 9	9 8 1 8 0 0 5 8 4	6 6 9 5 6 4 0 4 3	1 6 7 4 6 3 0 4 4	8 6 2 6 6 7 9 3 5
	8 5 5 8 4 1 8 1 5	7 9 4 2 4 7 1 9 2	7 1 4 9 0 5 0 9 9	8 6 0 1 4 4 7 9 5	4 3 1 9 4 7 5 2 3

1. One-Hot-Encoding признаков юзеров:

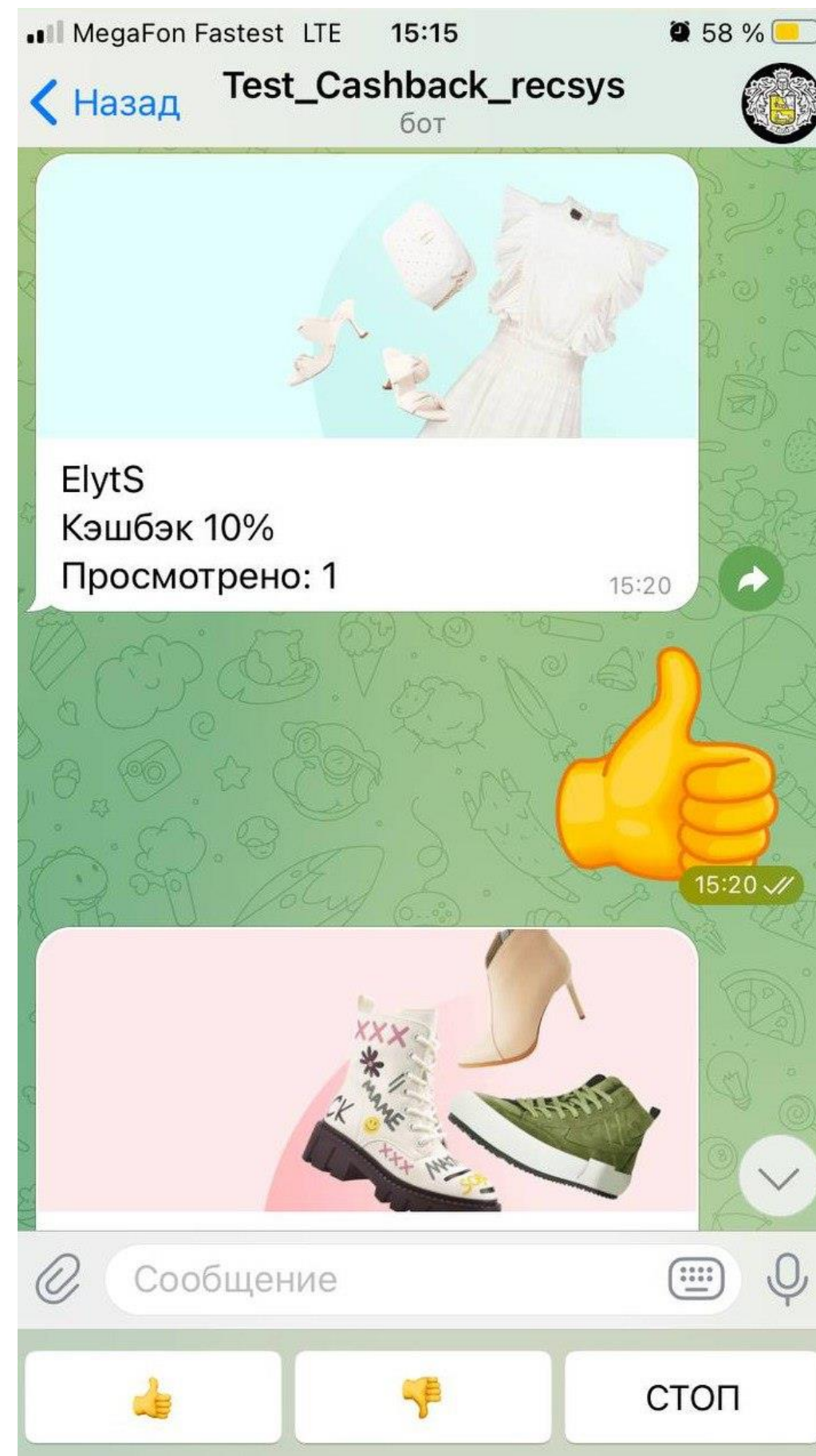
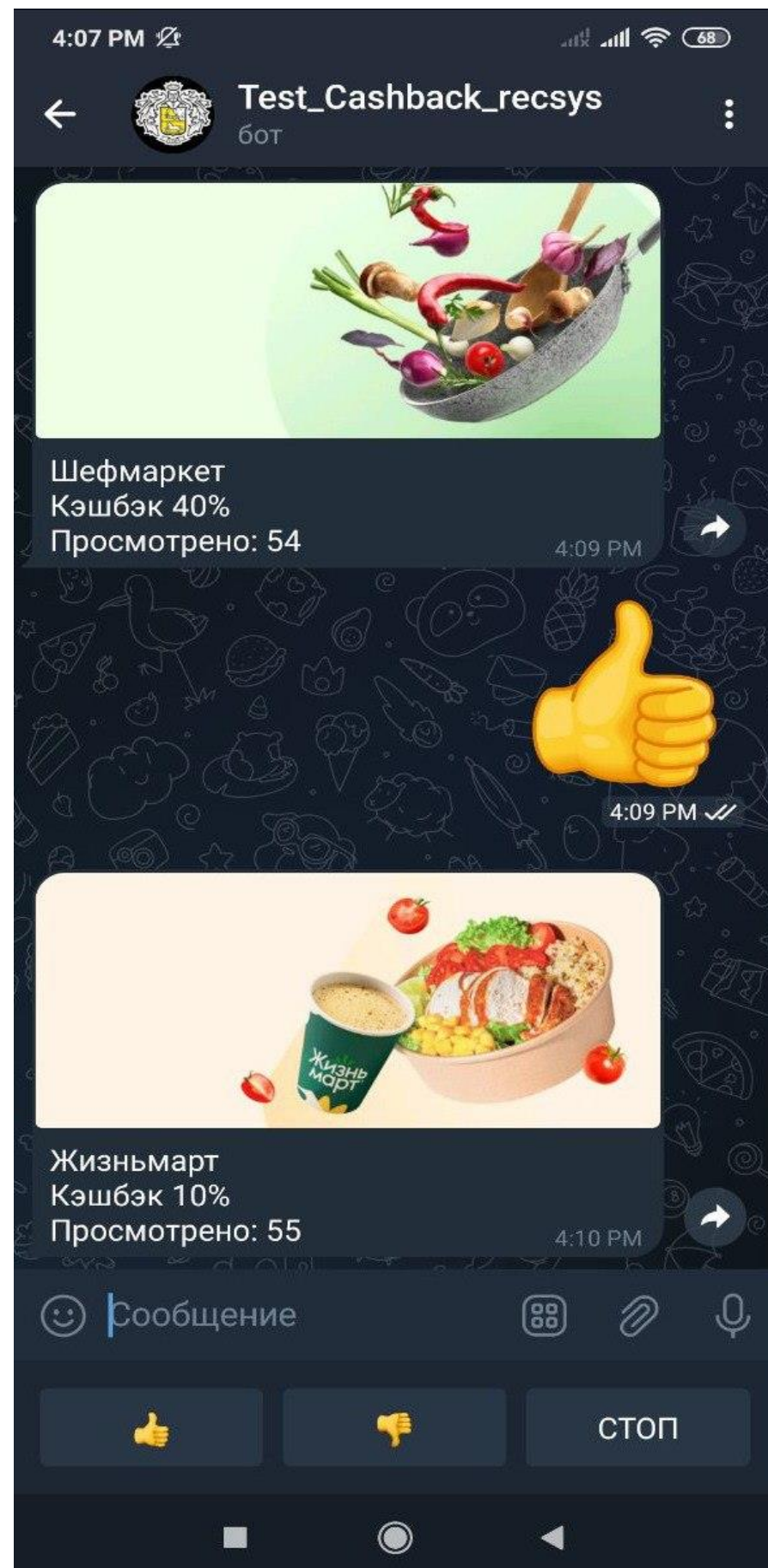
- возрастная группа
- пол
- заработок
- наличие детей

2. Оценки айтемов юзерами (-1, 0, +1)

Roadmap to Tinkoff



Тестирование на рандомных рекомендациях



Первым 15
пользователям даются
случайные рекомендации

Реализация индивидуальных рекомендаций



1. Основной метод:

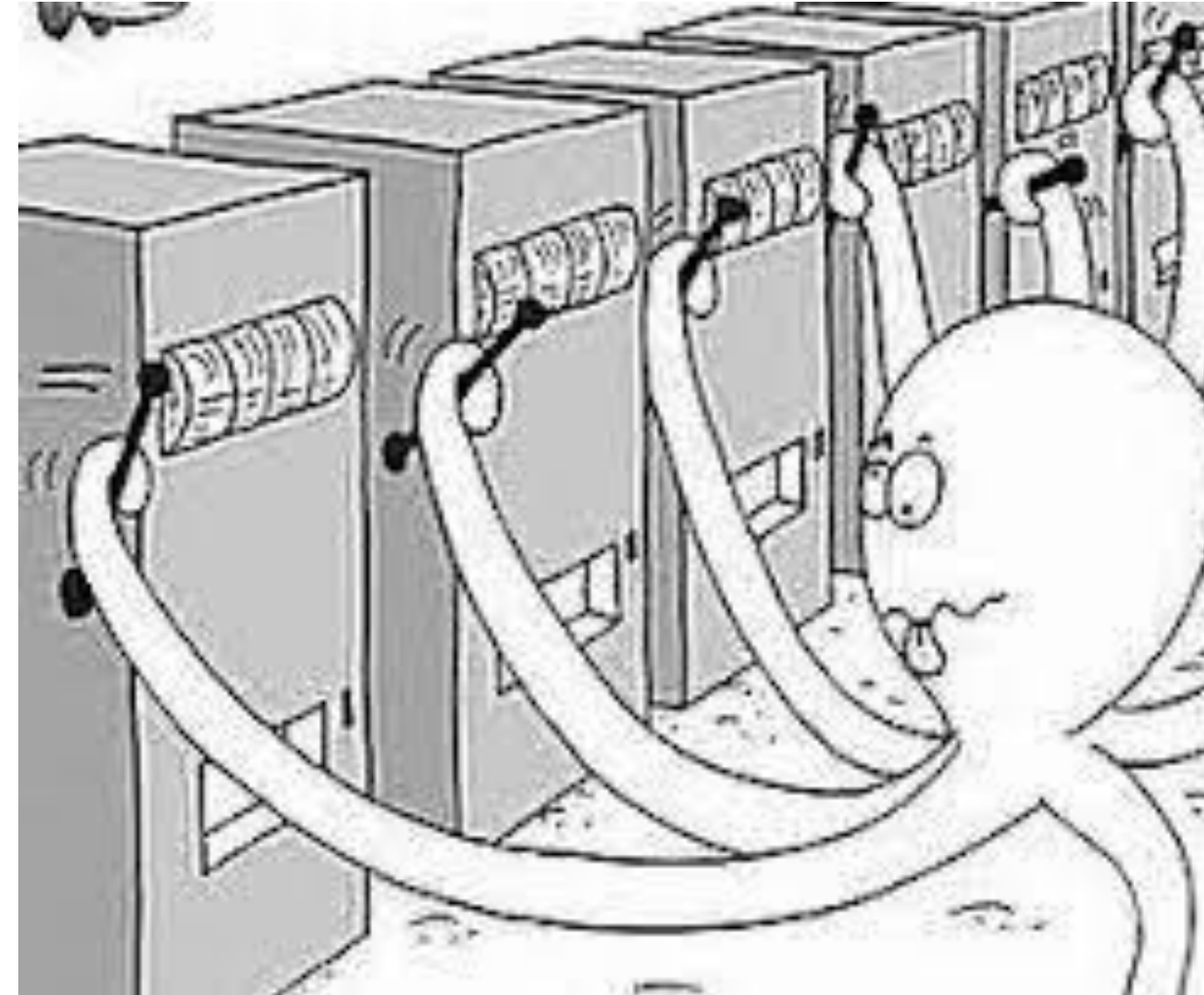
Item2Item KNN;

2. Применение **многоруких бандитов**
(эпсилон жадная стратегия);



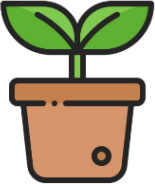










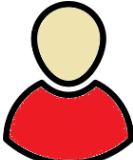













3. Добавление **категориальных юзеров**

Применение **многоруких** бандитов

item $\left\{ \begin{array}{ll} \text{Best KNN item} & \text{with probability } 1-\epsilon \\ \text{Random item} & \text{with probability } \epsilon \end{array} \right.$



Добавление категориальных юзеров



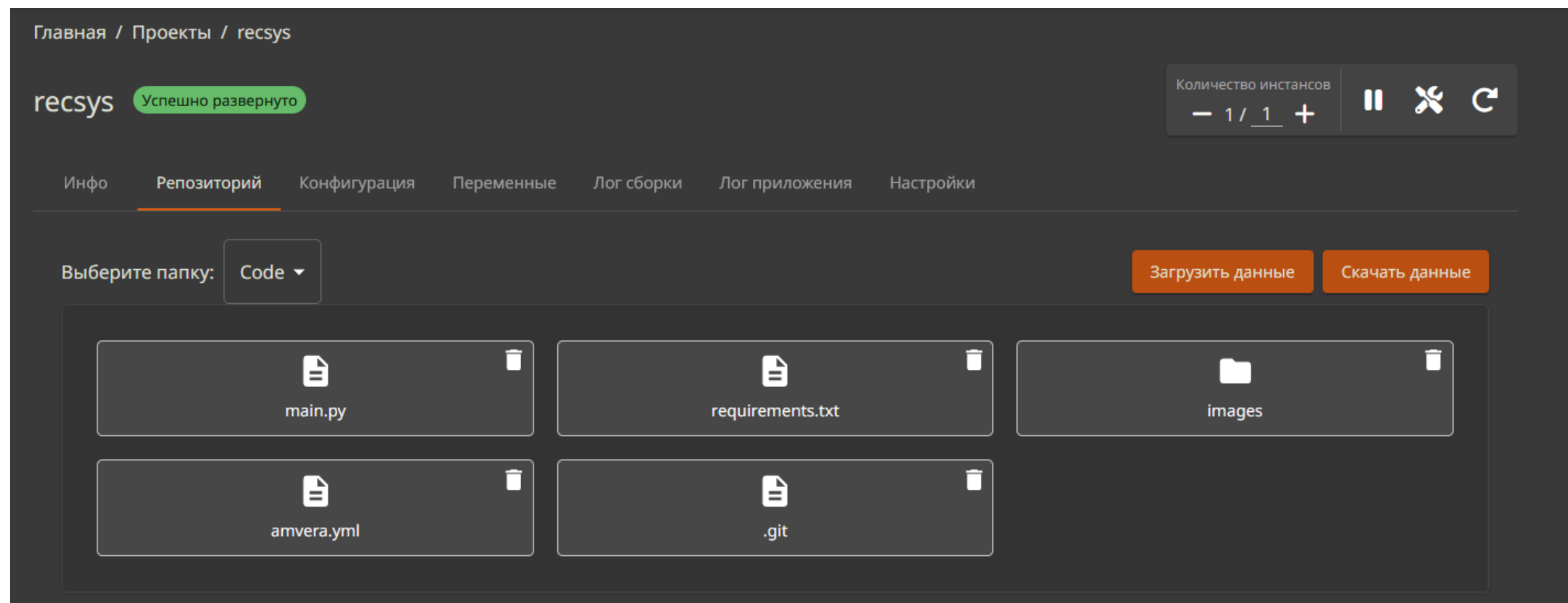
Roadmap to Tinkoff



Развертывание на бесплатном сервере

```
100%|██████████| 202/202 [00:00<00:00, 2258.58it/s]
2023-10-13T14:25:58.03837014Z stdout F user: 1990884208. recommended item: 57.
2023-10-13T14:25:58.131075864Z stderr F
0%|          | 0/202 [00:00<?, ?it/s]
100%|██████████| 202/202 [00:00<00:00, 2184.96it/s]
2023-10-13T14:26:03.140692419Z stdout F user: 1990884208. recommended item: 26.
2023-10-13T14:26:03.140728151Z stdout F user: 1990884208. random item: 83.
2023-10-13T14:26:03.231835021Z stderr F
0%|          | 0/202 [00:00<?, ?it/s]
100%|██████████| 202/202 [00:00<00:00, 89466.67it/s]
2023-10-13T14:27:02.638426116Z stdout F user: 1526882500. recommended item: 178.
2023-10-13T14:27:02.638462672Z stdout F user: 1526882500. random item: 7.
2023-10-13T14:27:02.732194042Z stderr F
0%|          | 0/202 [00:00<?, ?it/s]
100%|██████████| 202/202 [00:00<00:00, 2161.18it/s]
2023-10-13T14:27:17.73018252Z stdout F user: 1526882500. recommended item: 23.
2023-10-13T14:27:17.732932455Z stderr F
0%|          | 0/202 [00:00<?, ?it/s]
100%|██████████| 202/202 [00:00<00:00, 80559.99it/s]
```

1. Выбор сервера - [Amvera](#)
2. Развертывание Бота на сервере
3. Разбор ПОЧЕМУ БД НЕ ОБНОВЛЯЕТСЯ, а бот работает... [и исправление!](#)



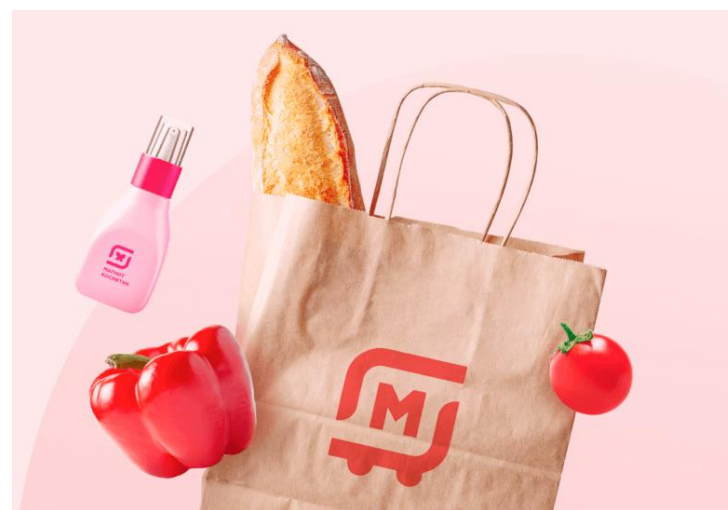
Сбор данных



ТОП-5 **ЛУЧШИХ** предложений КЭШБЕКОВ



Перекресток
87%



Магнит Доставка
83%



ЕАптека
82%

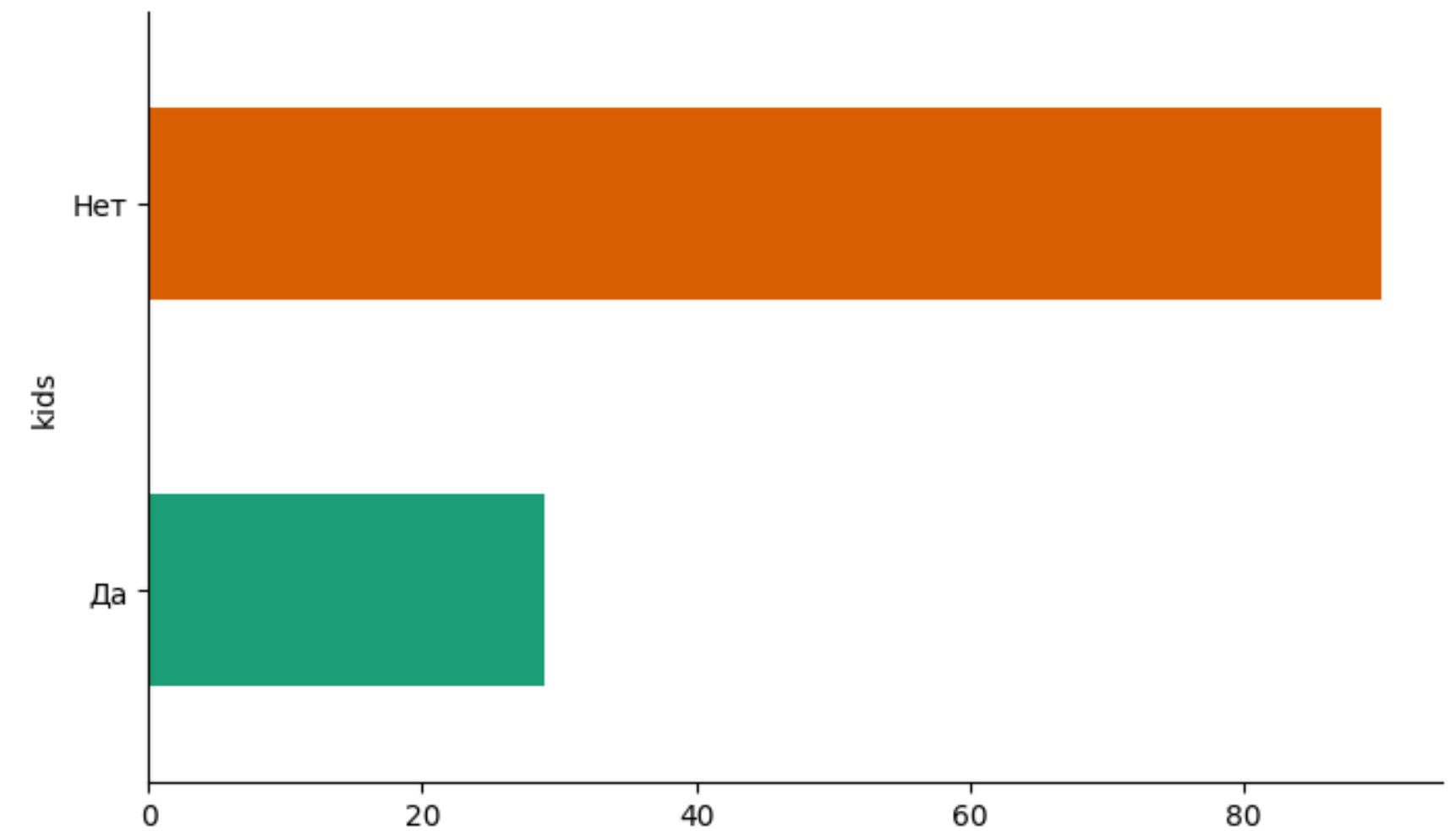
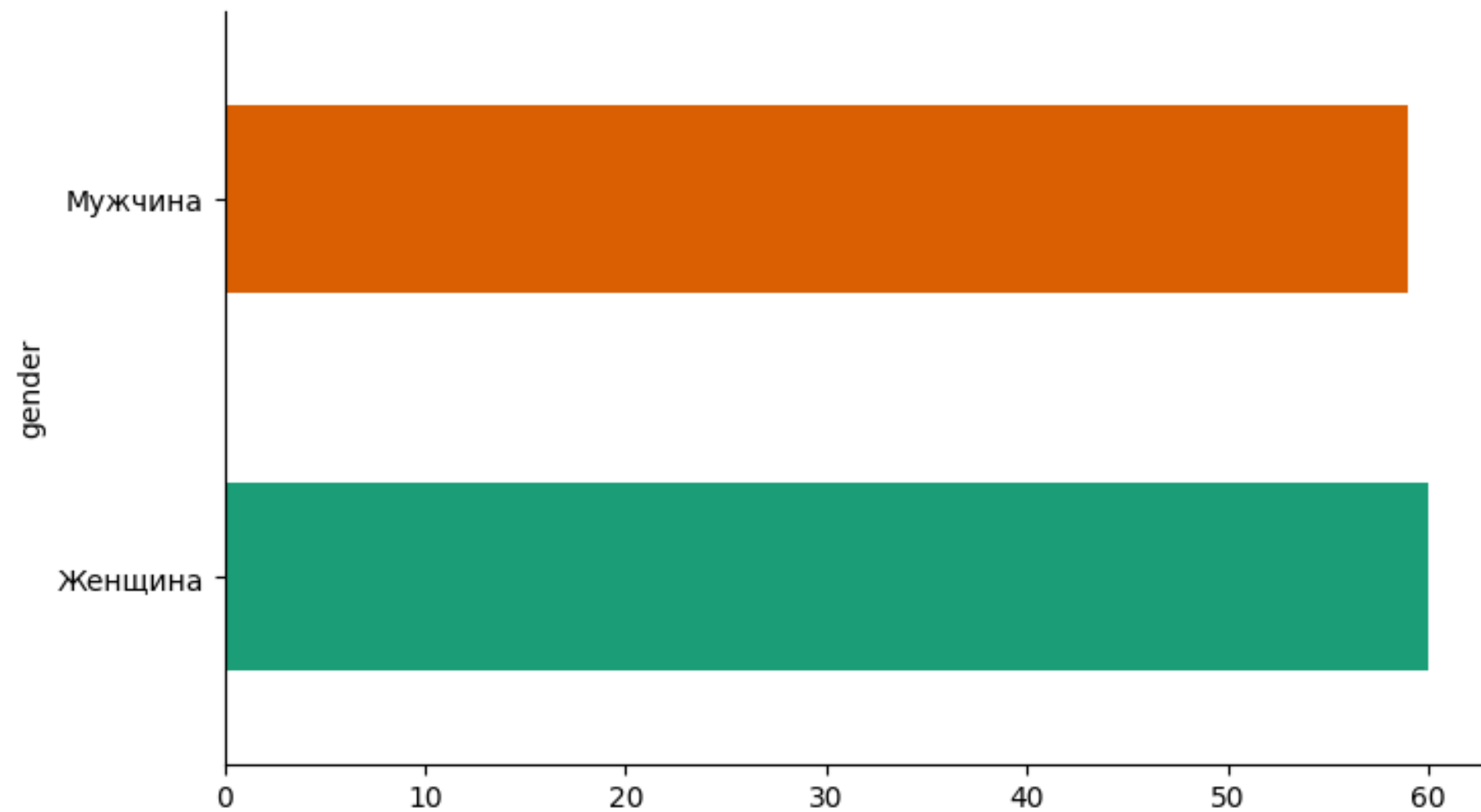


Лента Онлайн
82%

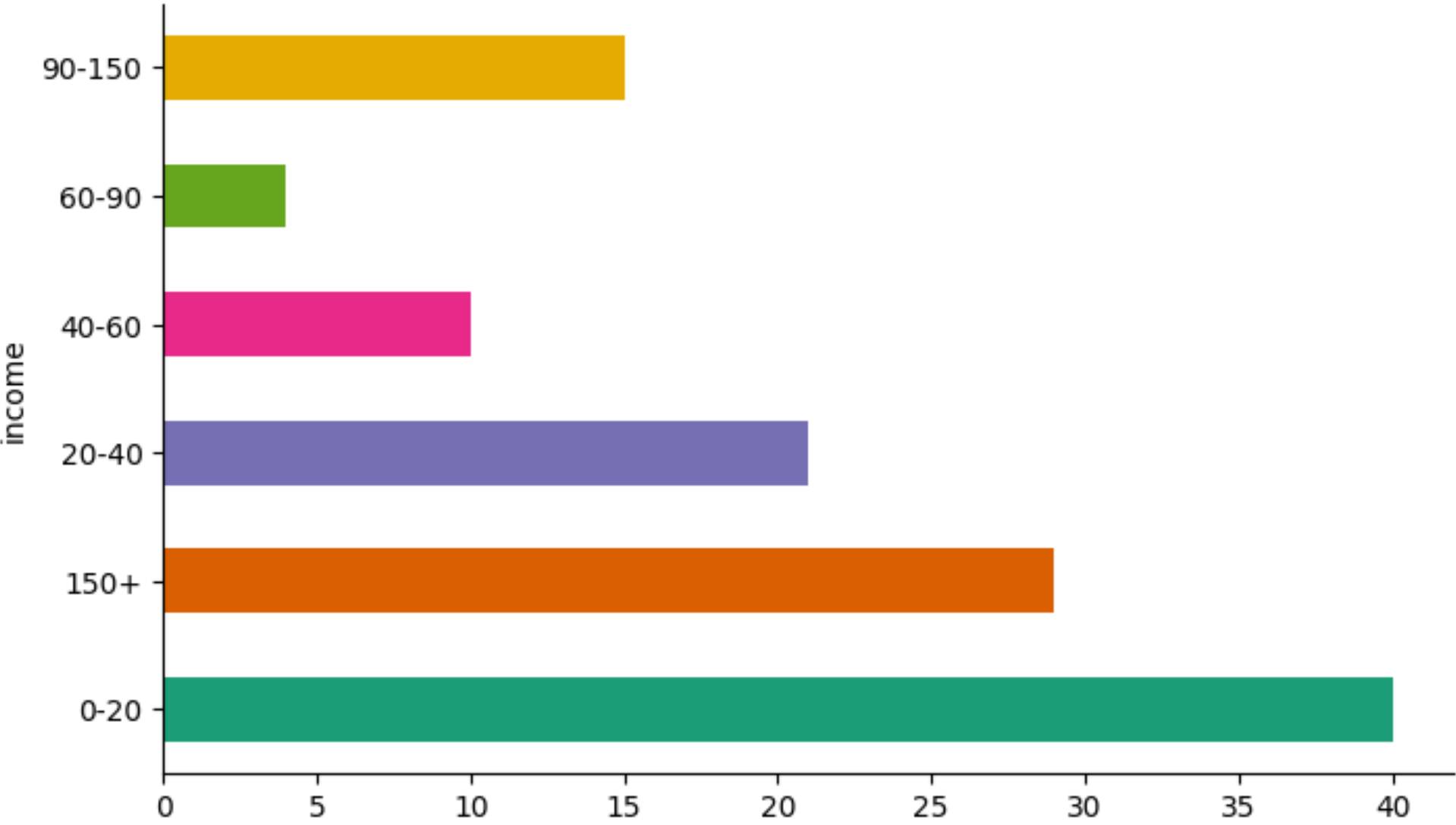
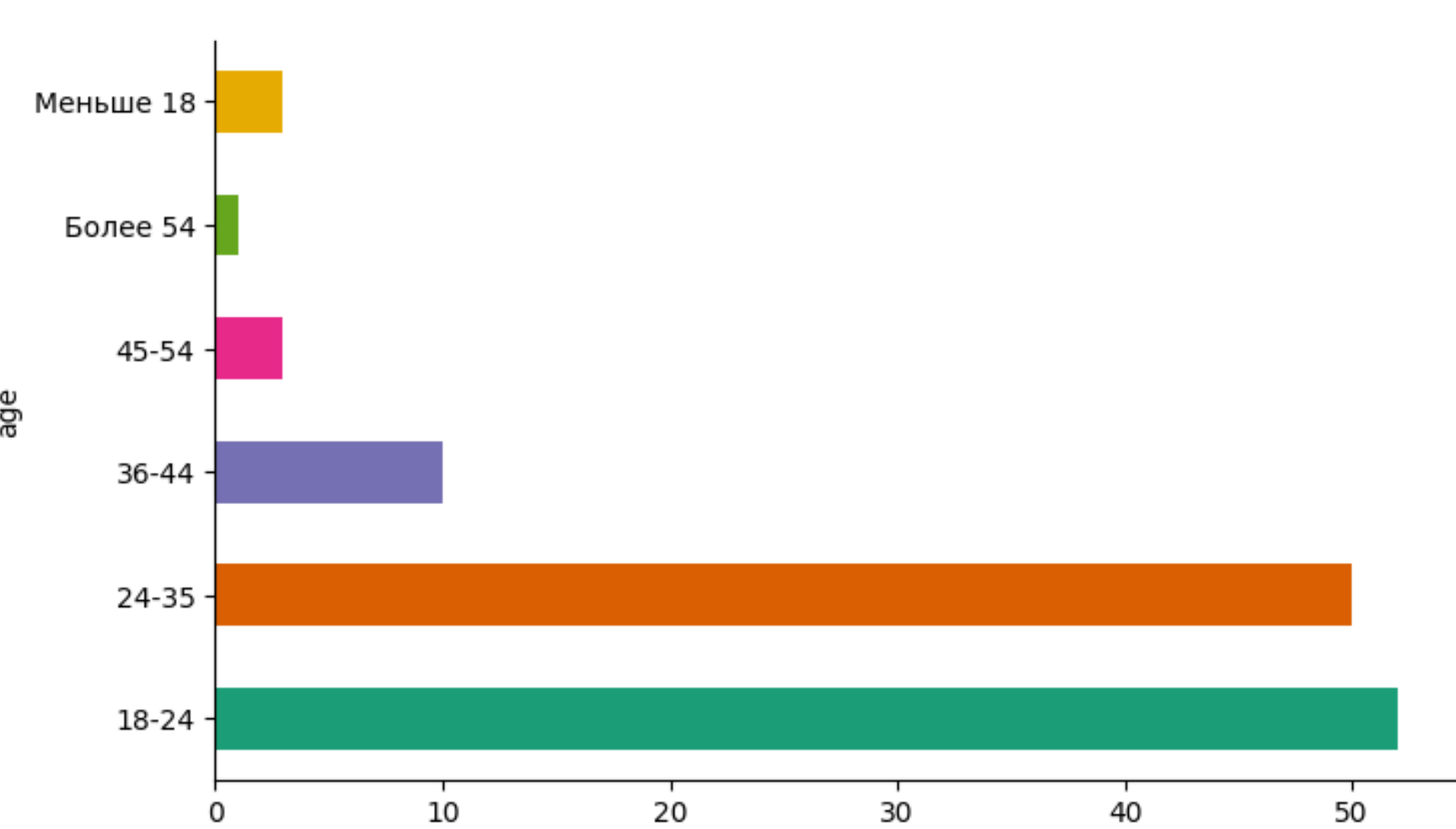


Вкусвилл
81%

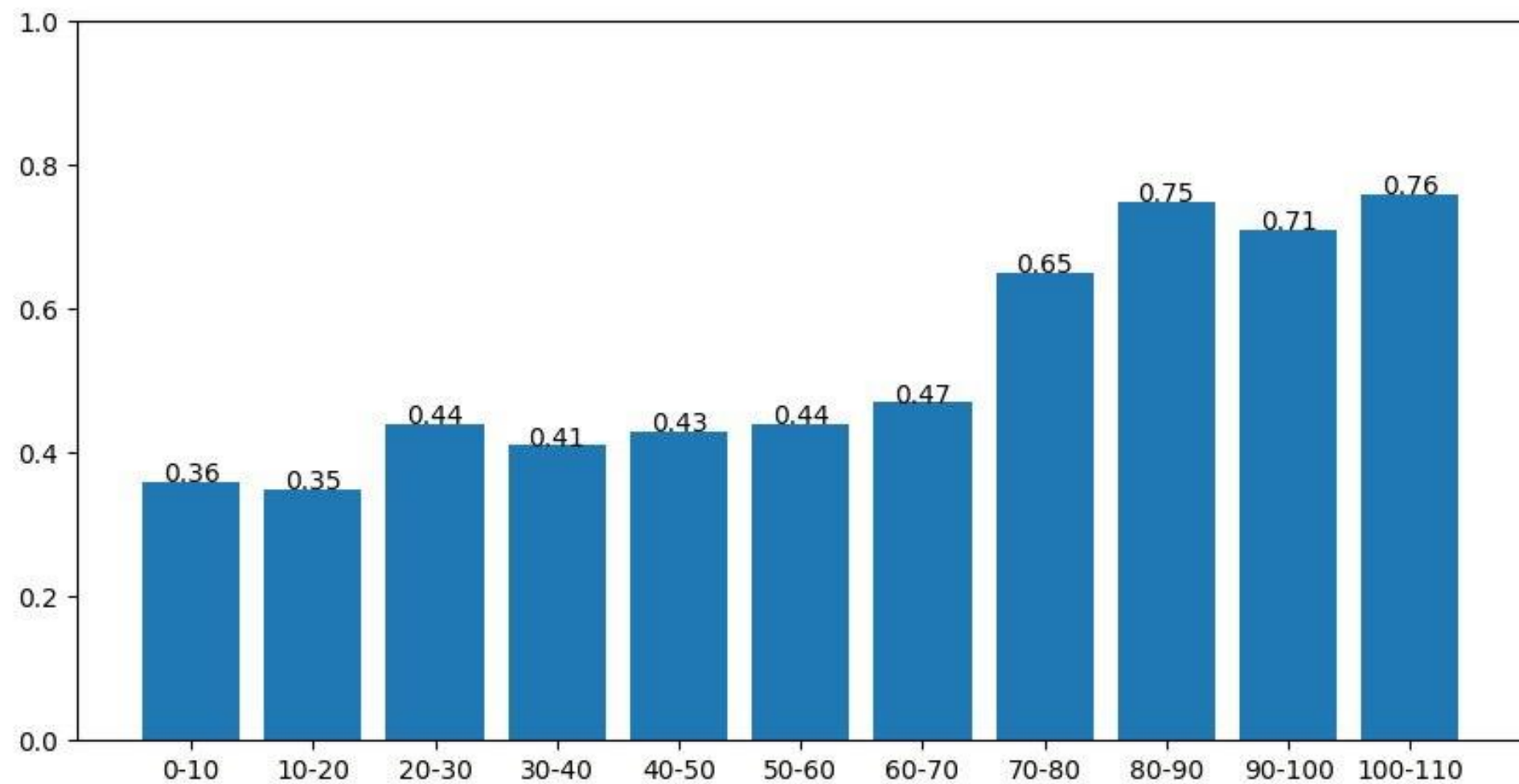
Статистика по нашей выборке



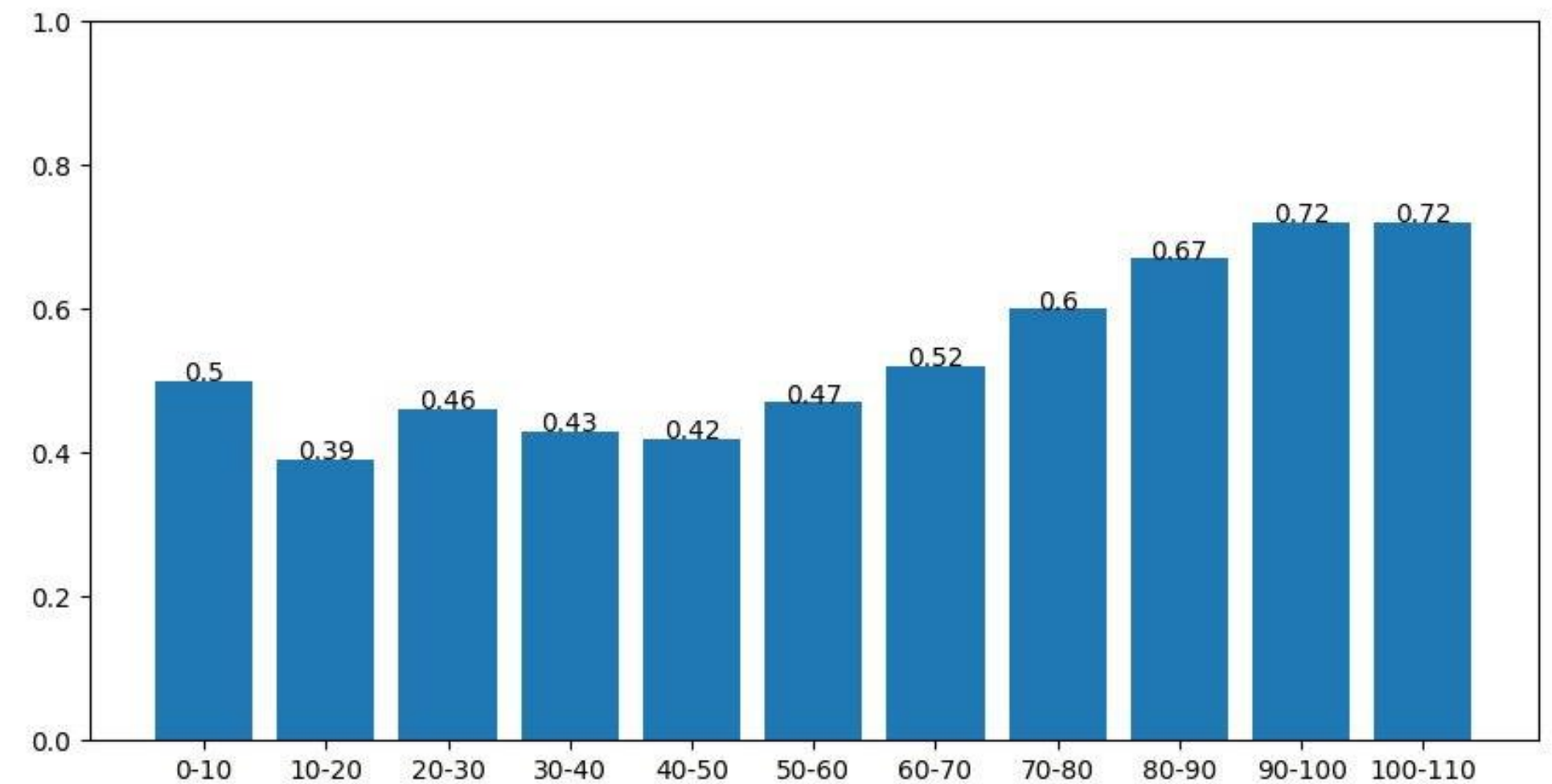
Статистика по обучающей выборке



Доли лайкнутых рекомендаций усредненных по последовательным группам пользователей



Усреднение оценок юзеров для
5 – 35 рекомендаций



Усреднение оценок юзеров по всем
просмотренным рекомендациям

Отзывы о боте

Интересный чат, все доступно

ХОРОШИЕ ПРЕДЛОЖЕНИЯ, Я БЫ ВОСПОЛЬЗОВАЛАСЬ!
МНОГО ВЫГОДЫ!

более точные попадания

Бот понравился, много разных кэшбеков нашел для себя.
Даже тех которых нет на моей карте (не отображаются)

На первый взгляд удобный бот, быстро
предлагает варианты. Хотелось бы наверно
разделение на каталоги еда одежда и тд
раздельно

Интересный бот. Можно выбирать скидки и кешбеки и
примерно знать какие из них и где будут

Всё удобно. Хорошие предложения.
Большие проценты кэшбэка

Суммаризация (отзывов СНАТ GPT3.5)

1. Большинство пользователей считают чат-бот интересным и удобным.
2. Отмечается более точное попадание в интересы пользователей и подбор персонального кэшбека.
3. Есть положительные комментарии относительно предложений и выгоды, которую пользователи получают.
4. В некоторых отзывах упоминается недостаток в виде отсутствия описания к названию бренда или невозможности определения категории товаров.
5. Некоторые отзывы указывают на проблемы с открытием ссылок и неясностью некоторых предложений.
6. Много отзывов, где пользователи высоко оценивают уровень кэшбека.
7. Кое-кто отмечает ограничения, связанные с доступными магазинами в их городе.

Выводы

1. Получилось создать MVP модель рекомендательной системы кэшбеков от Тинькофф
2. Качество модели и количество лайкнутых кэшбеков увеличивается с приходом новых пользователей
3. Положительные отзывы от пользователей



@TESTCASHBACKRECS
YSBOT



ТИНЬКОФФ
ОБРАЗОВАНИЕ

**Спасибо
за внимание!**

Он такой один!

