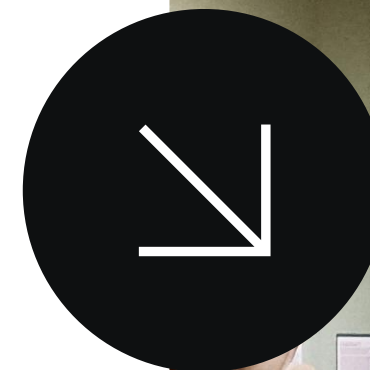
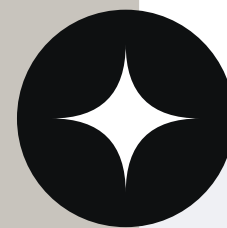


Оценка тональности отзывов о музее

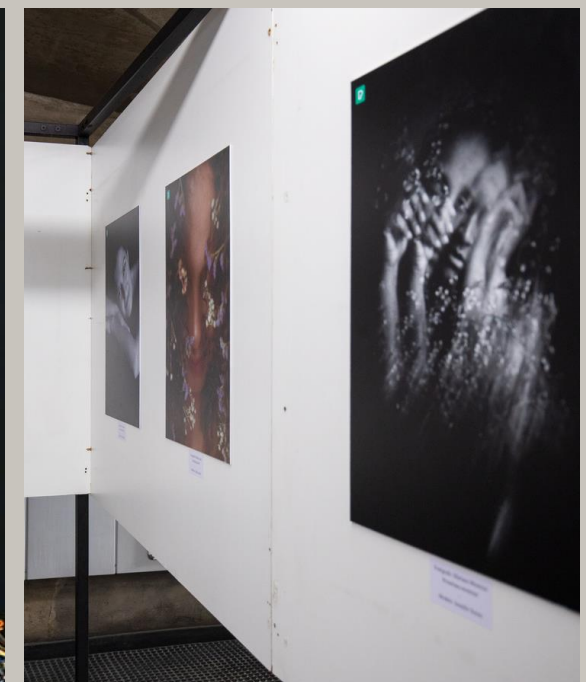
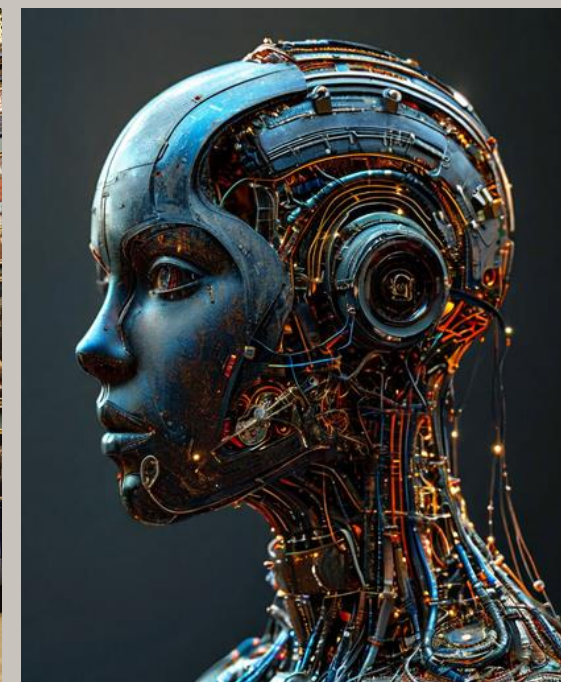
Команда Solvers
Светова Виктория
Мастеров Дмитрий

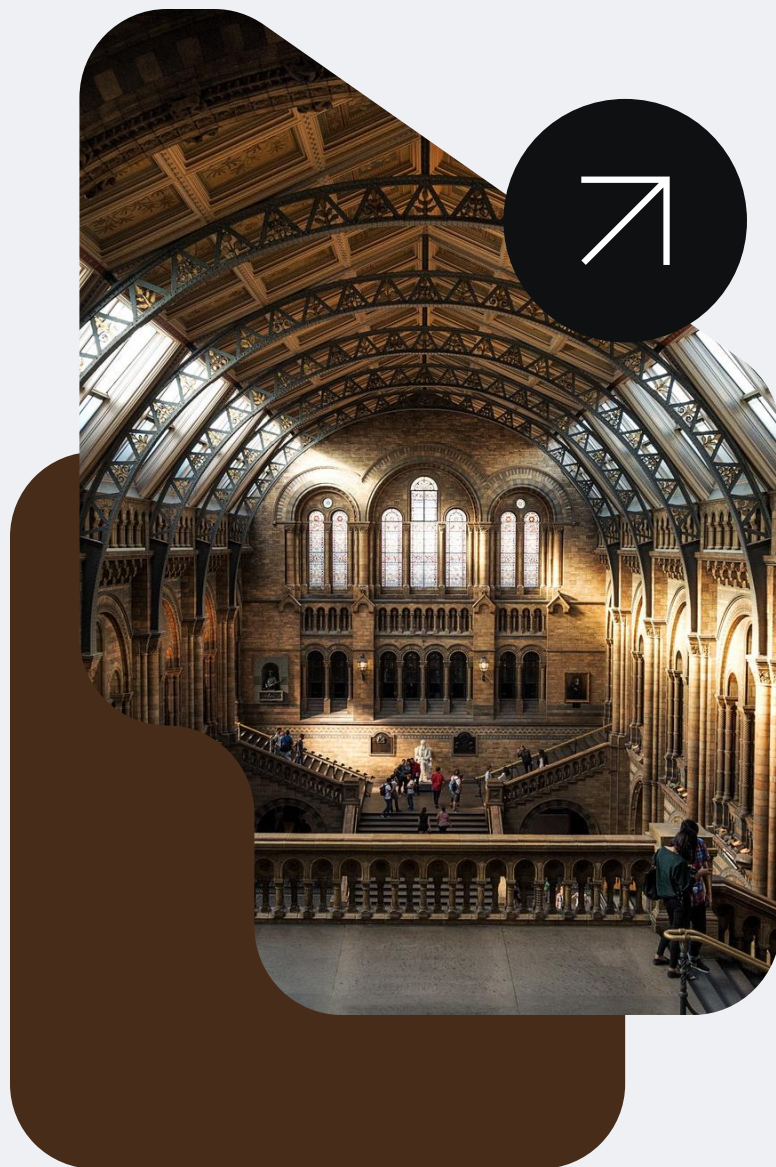


Введение



Машинное обучение
помогает музеям
автоматизировать анализ
отзывов посетителей,
улучшая качество
обслуживания.





Актуальность

- Улучшение пользовательского опыта в музеях
- Помощь в принятии решений со стороны музея для повышения комфорта посетителей.
- Неверное соотношение отзывов и количества звезд в Яндекс картах
- Автоматизация рутинных процессов
- Выявление ключевых тенденций и проблем

Потеря негативной информации из-за высокой оценки



Екатерина В.

Знаток города 13 уровня

[Подписаться](#)

★★★★★ 3 декабря 2019

Сам музей очень понравился, доброжелательные приветливые сотрудники. Очень интересные экскурсии, много ценных экспонатов...Но.. нынешнее состояние музея удручает, разбитые лестницы, грязные стены и потолки, старый потёртый линолеум, в общем, ремонт требуется однозначно!



Дмитрий Март

Знаток города 12 уровня

[Подписаться](#)

★★★★★ 31 октября 2019

Интересное место. Но, как я понимаю, нуждается в финансовой поддержке.



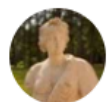
Юлия

Знаток города 10 уровня

[Подписаться](#)

★★★★★ 25 сентября 2021

Была в музее с экскурсией. Очень познавательно. Но здание требует ремонта(



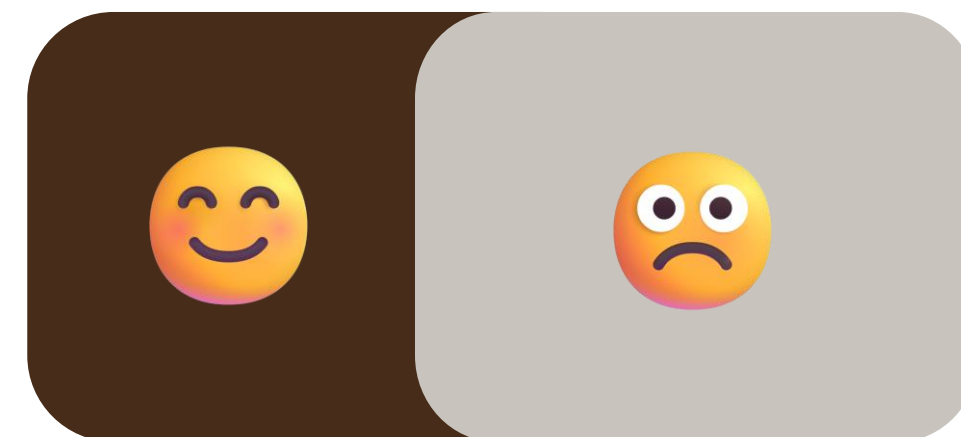
Любовь Тарасова

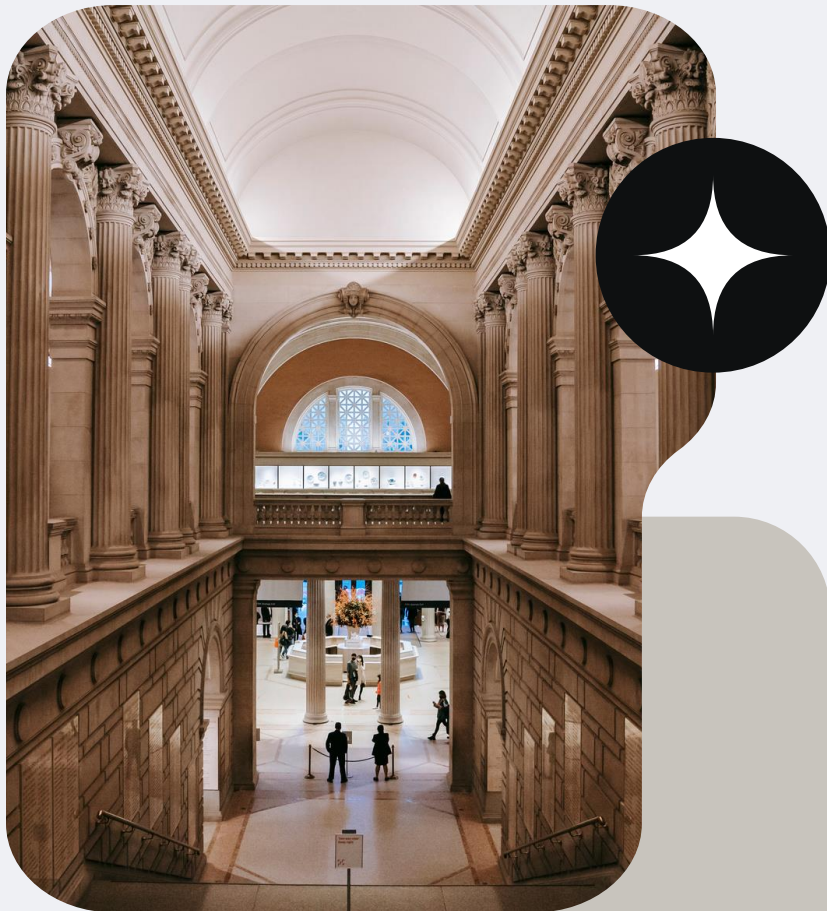
Знаток города 18 уровня

[Подписаться](#)

★★★★★ 23 апреля 2023

Очень интересный и важный музей. Но экспозиция уместается в достаточно тесном пространстве, требуется ремонт. Посещение только с экскурсией.





Цель проекта

Разработать модель для классификации отзывов о музее на положительные и отрицательные.

Постановка задач

1. Сбор данных с помощью парсинга Яндекс карт
2. Произвести разметку отзывов на 0(негативный) и 1(позитивный).
3. Проанализировать полученный датасет с размеченными данными.
4. Дообучить готовую модель на составленном датасете.
5. Разработать telegram бота для взаимодействия с моделью

Данные

Получение отзывов с помощью парсинга сайта

Разметка тональности: 0 – негативный, 1 – позитивный

Получаем итоговый датасет:

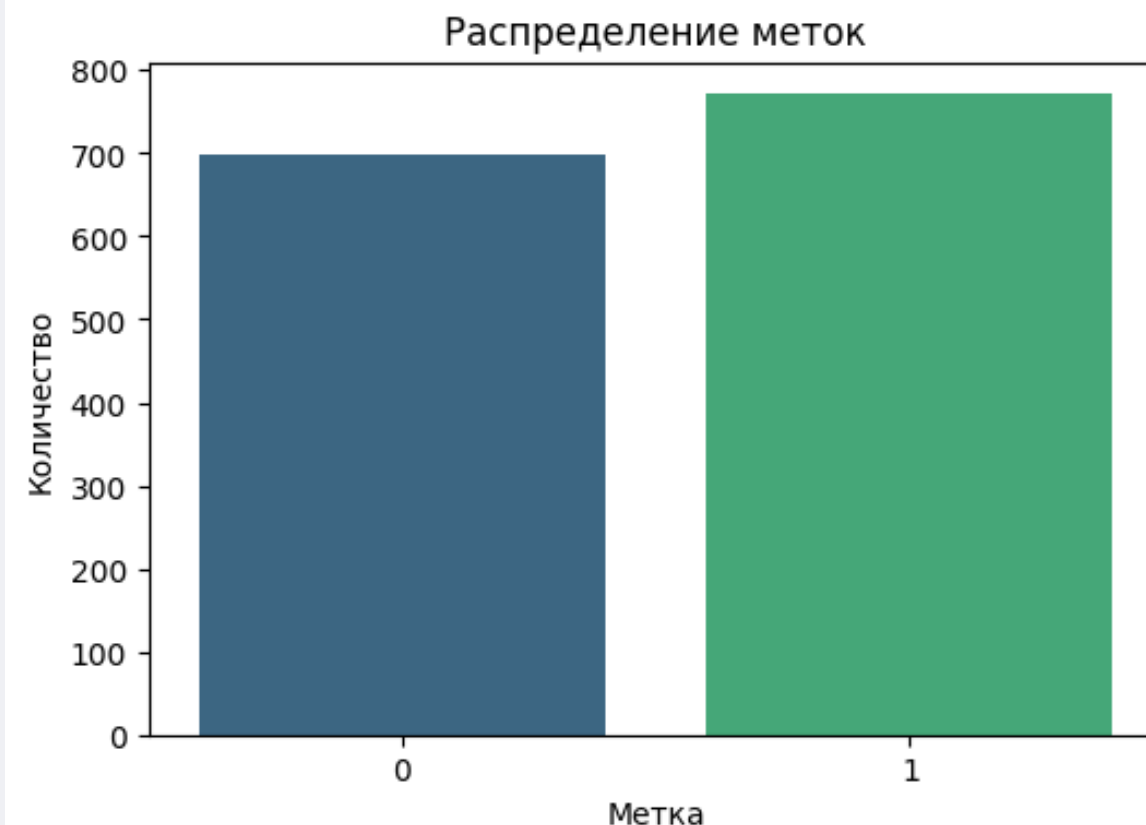
Index

Review

label

Информация о данных:

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 1470 entries, 0 to 1469  
Data columns (total 3 columns):  
#   Column   Non-Null Count  Dtype    
---  ---      -  
0    Index    1470 non-null   int64    
1    Review   1470 non-null   object   
2    label    1470 non-null   int64    
dtypes: int64(2), object(1)  
memory usage: 34.6+ KB
```



Распределение меток:

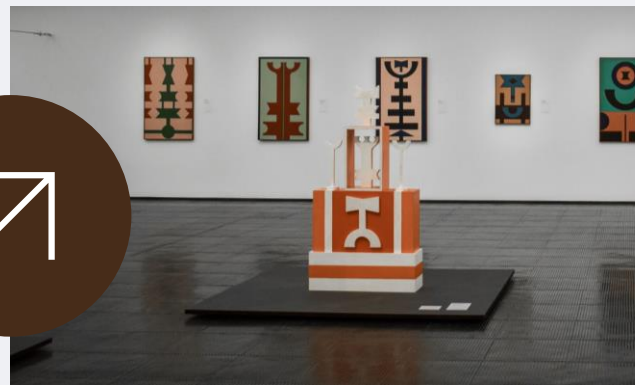
```
label  
1    771  
0    699  
Name: count, dtype: int64
```


Выбор модели

cointegrated/rubert-tiny-sentiment-balanced была выбрана по нескольким основным причинам:

Адаптация к русскоязычным текстам

Готовность к дообучению



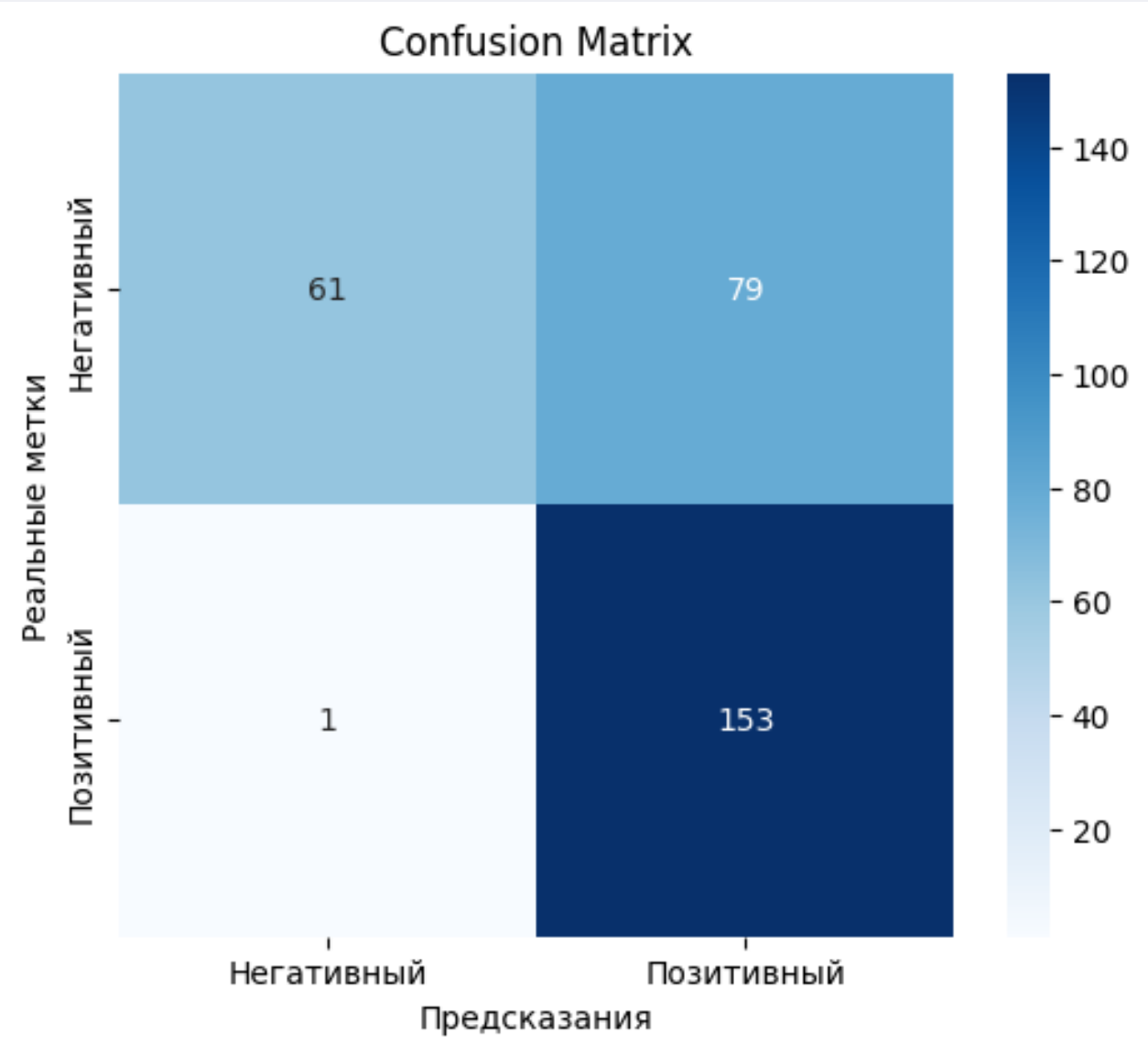
Модель 1

cointegrated/rubert-tiny-sentiment-balanced

Accuracy: 0.73

F1-score: 0.79

Модель 1	precision	recall	f1-score
0	0,98	0,44	0,60
1	0,66	0,99	0,79



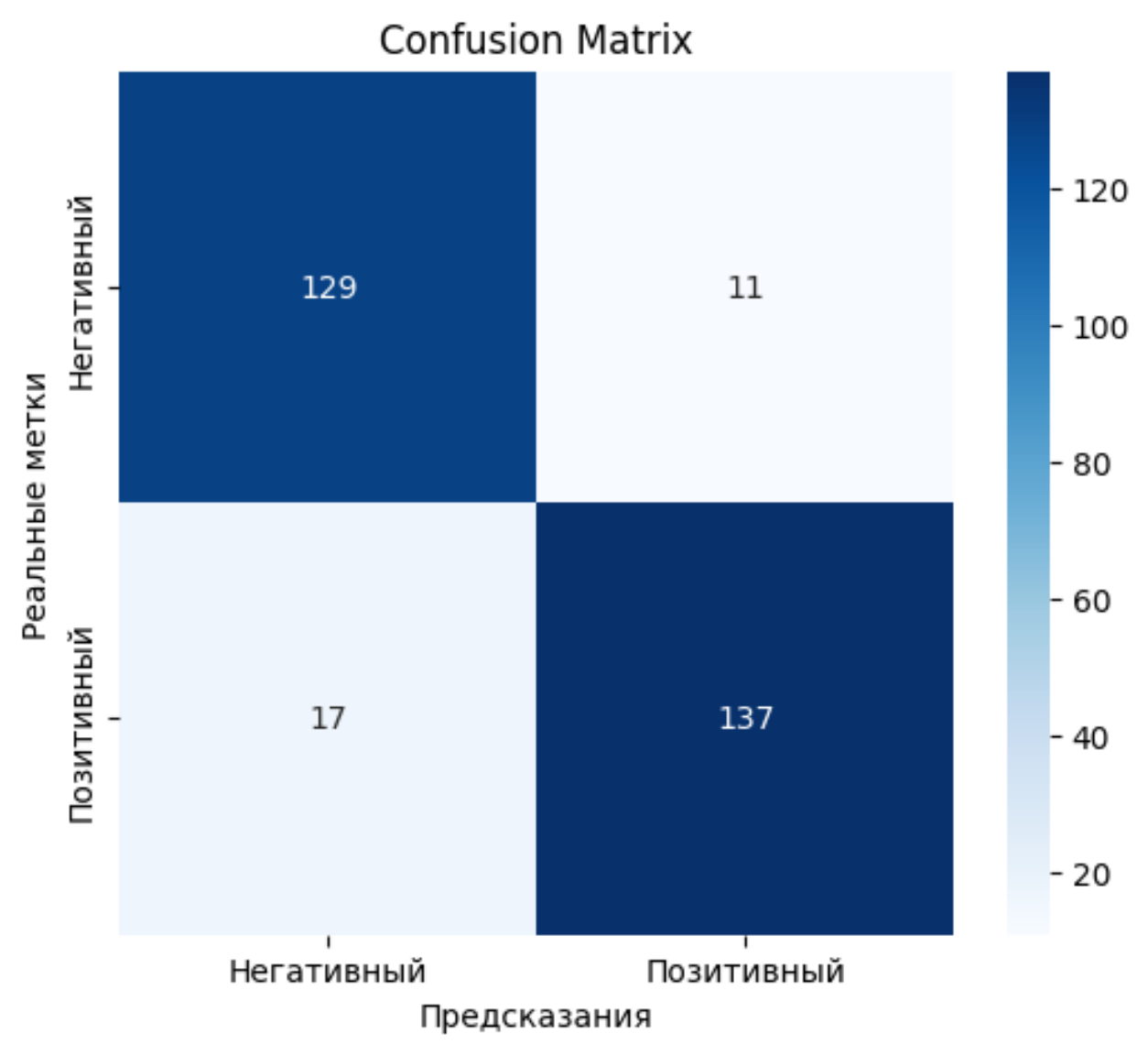
Модель 2

дообучение cointegrated/rubert-tiny-sentiment-balanced

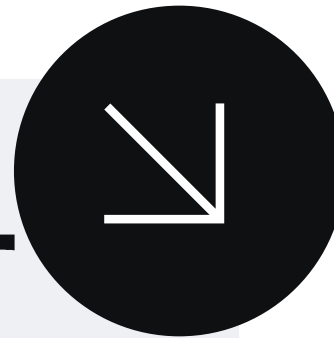
Accuracy: 0.91

F1 Score: 0.91

Модель 2	precision	recall	f1-score
0	0,88	0,92	0,90
1	0,93	0,89	0,91

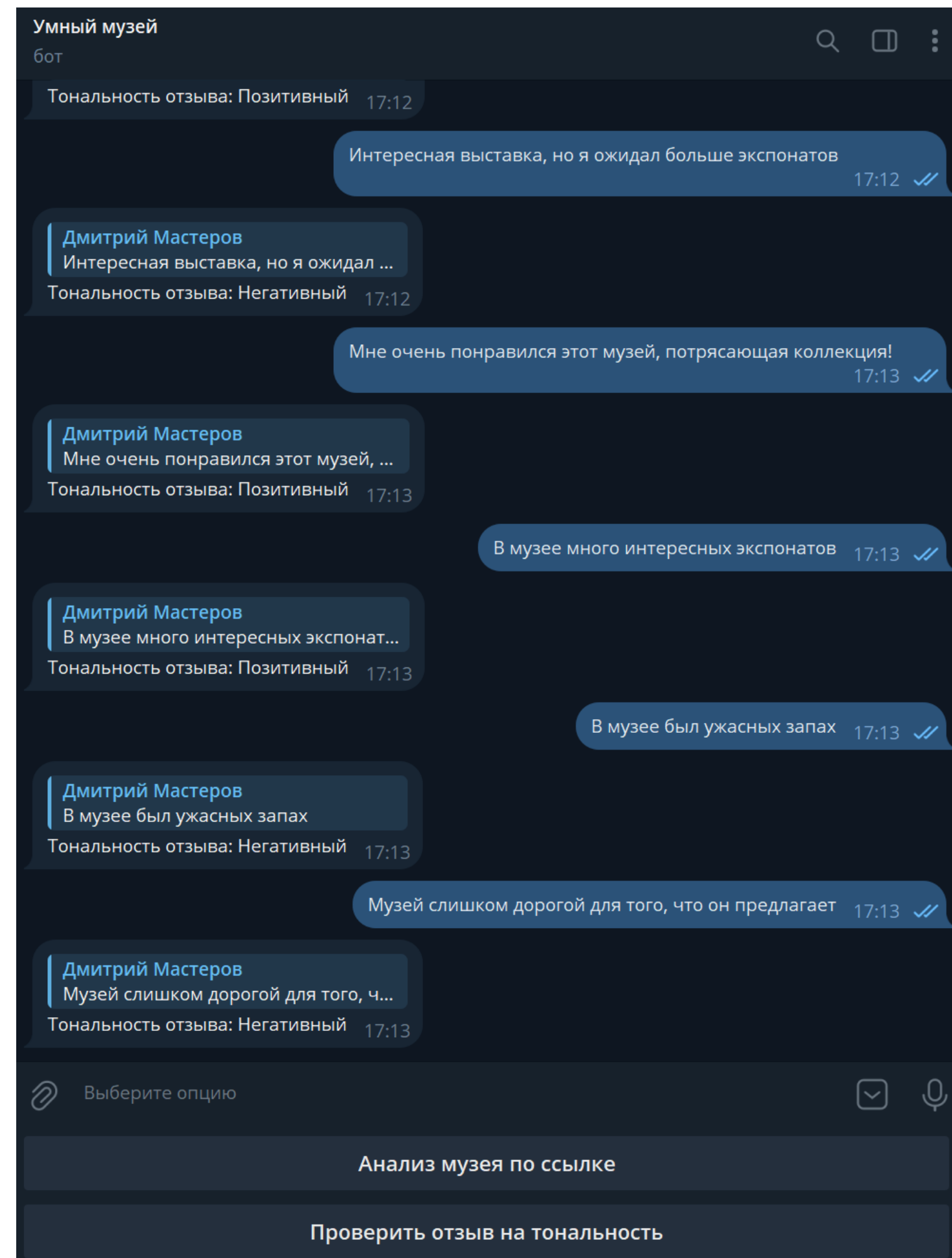


Telegram-бот



Два режима работы:

1. Проверить отзыв на тональность – на вход принимает текст отзыва. Выдает тональность отзыва (позитивный или негативный)
2. Анализ музея по ссылке – проводит подробный анализ музея по предоставленной ссылке на Яндекс картах



Телеграм-бот



Отзыв и оценка из Яндекс карт



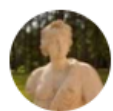
Екатерина В.

Знаток города 13 уровня

[Подписаться](#)

★★★★★ 3 декабря 2019

Сам музей очень понравился, доброжелательные приветливые сотрудники. Очень интересные экскурсии, много ценных экспонатов...Но.. нынешнее состояние музея удручает, разбитые лестницы, грязные стены и потолки, старый потёртый линолеум, в общем, ремонт требуется однозначно!



Любовь Тарасова

Знаток города 18 уровня

[Подписаться](#)

★★★★★ 23 апреля 2023

Очень интересный и важный музей. Но экспозиция уместается в достаточно тесном пространстве, требуется ремонт. Посещение только с экскурсией.

Дообученная модель rubert

Сам музей очень понравился, доброжелательные приветливые сотрудники. Очень интересные экскурсии, много ценных экспонатов...Но.. нынешнее состояние музея удручает, разбитые лестницы, грязные стены и потолки, старый потёртый линолеум, в общем, ремонт требуется однозначно! 17:16 ✓✓

Дмитрий Мастеров

Сам музей очень понравился, добр...

Тональность отзыва: Негативный 17:16

Очень интересный и важный музей. Но экспозиция уместается в достаточно тесном пространстве, требуется ремонт. Посещение только с экскурсией. 17:19 ✓✓

Дмитрий Мастеров

Очень интересный и важный музей...

Тональность отзыва: Негативный 17:20

Telegram-бот

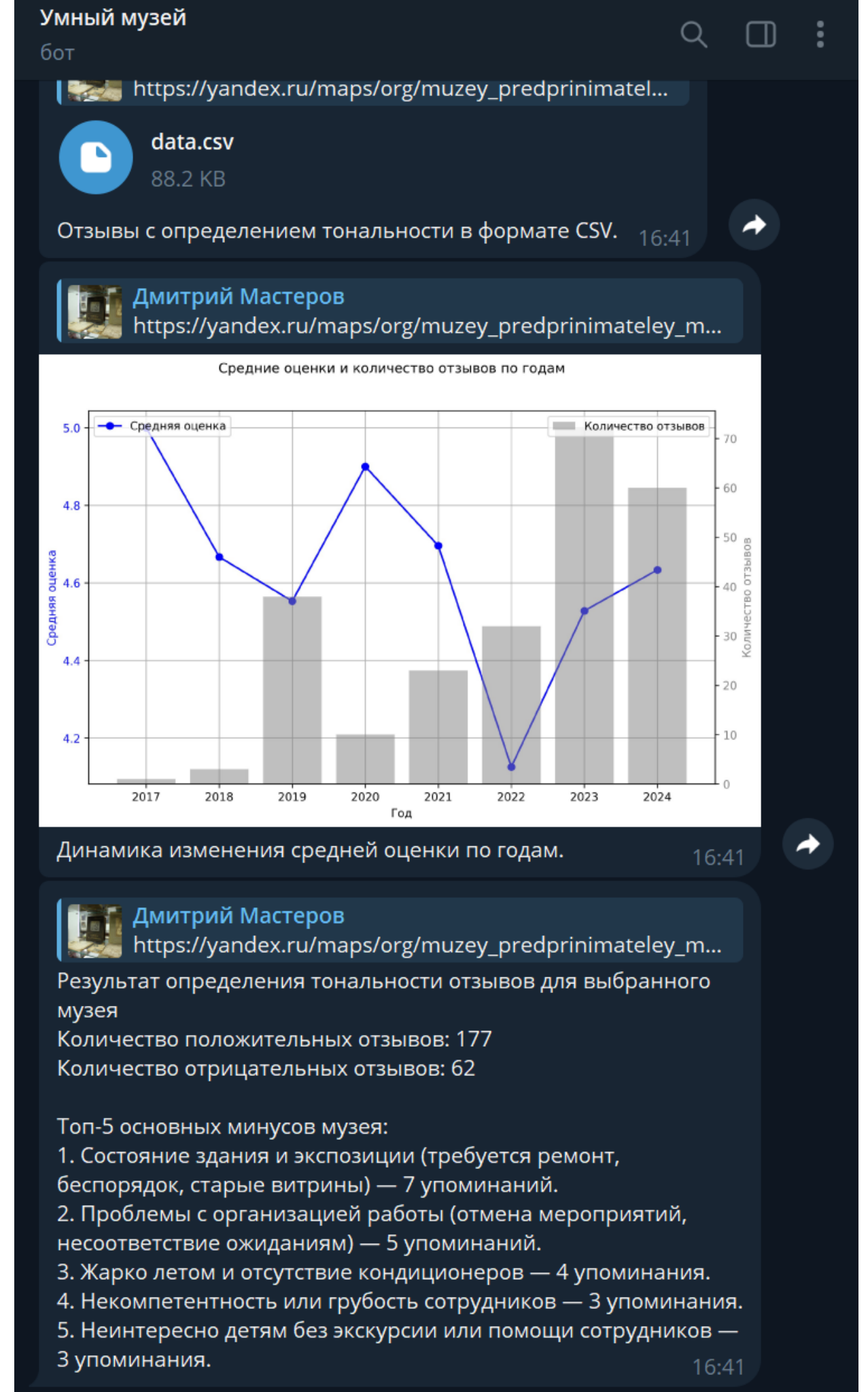
Подробный анализ музея по предоставленной ссылке из Яндекс карт

Парсинг отзывов и оценок

Определение тональности

Выделение ключевых минусов

Анализ динамики оценок



Возможные области применения

1. Улучшение качества обслуживания:

- Выявление часто упоминаемых проблем в отзывах, таких как долгие очереди, неудобный график работы или недостаток удобств.
- Принятие решений для повышения комфорта посетителей.

2. Мониторинг репутации:

- Своевременное реагирование на негативные отзывы для предотвращения ухудшения имиджа музея.
 - Отслеживание динамики восприятия музея и влияния маркетинговых кампаний.
-

Бизнес-модель проекта

Целевая аудитория: музеи, галереи, культурные центры, разработчики маркетинговых решений

Платформа анализа данных:

- Подписка для музеев и культурных учреждений, предоставляющая доступ к инструментам анализа отзывов, включая полярность, частотный анализ тем и динамику мнений.

Бот-помощник для музеев:

Функционал:

- Автоматическая классификация отзывов на положительные и отрицательные.
- Выявление ключевых тем и тенденций в отзывах посетителей.
- Уведомления о негативных отзывах с рекомендациями для реакции.

Монетизация через:

- Лицензирование бота для использования музеями и галереями.
- Интеграция в туристические платформы для анализа отзывов о достопримечательностях.
- Сотрудничество с компаниями, занимающимися улучшением клиентского опыта, для предоставления аналитических данных.