ОЛИМПИАДА ШКОЛЬНИКОВ «ШАГ В БУДУЩЕЕ»

НАУЧНО-ОБРАЗОВАТЕЛЬНОЕ СОРЕВНОВАНИЕ

«ШАГ В БУДУЩЕЕ, МОСКВА»

регистрационный номер: 36129

Секция: Информационные технологии (ИУ7)

SpaceAlien

Автор: Елькин Кирилл Валерьевич

Ученик 11 “Б” класса ГБОУ школа №1467

Научный руководитель:

Старший преподаватель в МГТУ им. Н.Э. Баумана

Волкова Лилия Леонидовна

Москва, 2024

SpaceAlien

Аннотация

Название: SpaceAlien

Цель работы: Разработать сервис для исследования и поиска различной информации.

Методы и материалы:

Серверная часть проекта была реализована на Python с помощью фреймворка Django. Пользовательский интерфейс реализован на HTML, CSS и JavaScript. Модель искусственного интеллекта реализована на основе нейронных сетей с применение архитектуру Google BERT.

Результаты:

Проект находится на стадии разработки, но уже сейчас обладает базовым функционалом. Модель искусственного интеллекта на основе BERT была обучена и показала высокую точность. Архитектура модели также показала высокую эффективность. На данный момент модель способна классифицировать 8 классов. Реализован базовый пользовательский интерфейс и серверная часть сервиса с использованием фреймворка Django. На сайте реализованы основные разделы.

Перспективы:

Планируется обучить модель искусственного интеллекта на новом наборе данных. Модель получит возможность давать определения словам и предоставлять разную информацию. Будет произведена настройка алгоритмов для обработки естественного языка. Также будет продолжена настройка параметров модели. Кроме того, планируется доработать дизайн сервиса. Будет реализован раздел “Лекции”, который будет содержать лекции нашего времени и прошлого столетия. Планируется реализовать поиск среди информации, представленной на сайте. Будет доработан раздел “Статьи”.

Оглавление

Введение

Наука играет важную роль в различных аспектах нашей жизни. Она оказывает влияние на все сферы жизни общества, способствуя их развитию и прогрессу. Кроме того, наука даёт возможность найти ответы на вопросы или пути решения актуальных проблем. На основе научных знаний можно строить новые теории, гипотезы или ставить эксперименты, которые позволят лучше изучить наш мир.

Технологией, которая позволит упростить работу с информации, - искусственный интеллект. Это одна из самых актуальных технологий на данный момент. Он используется в различных областях: науке, медицине, экономике и так далее. На данный момент технология искусственного интеллекта находится на активной стадии развития. Каждый новый проект вносит свои коррективы и инновации в текущую технологию, что, в свою очередь, ведёт к созданию новых продуктов и услуг. В нашей стране искусственным интеллектом занимаются небольшое количество компаний, что открывает большие перспективы для продвижения данной технологии.

1 Актуальность работы

Наука в наше время содержит огромное количество знаний, поэтому важно сделать её доступной для большого количества людей. Данный проект поможет людям найти различную информацию, лекции и фотографии, которые сделаны с помощью специальной аппаратуры. Также данный сервис поможет улучшить взаимосвязь между научным сообществом и людьми. Кроме того, каждый желающий сможет изучить научные достижения прошлых столетий.

Также важно принимать во внимание, что количество информации, которую необходимо обрабатывать, анализировать и структурировать, в мире постоянно увеличивается. Одним из возможных решений является технология искусственного интеллекта, которая сможет решить данную проблему. Кроме того, искусственный интеллект сможет сделать информацию более доступной и полезной для большого числа людей. Он сможет упростить задачу поиска

ответов на конкретные вопросы пользователя. Также в ходе разработки искусственного интеллекта данные будут систематизированы, что позволит проще совершать различные действия с ними.

2 Цели и задачи работы

Цель работы - разработать сервис для исследования и поиска различной информации.

Задачи работы:

1. Определить архитектуру сервиса;
2. Разработать серверную часть сервиса;
3. Разработать веб-приложение для взаимодействия с пользователями;
4. Разработать принципы хранения и записи данных;
5. Определить архитектуру искусственного интеллекта;
6. Произвести обучение и оптимизирование модели искусственного интеллекта.

3 Этапы

3.1 Первый этап

Этап планирование, который продлился с 1 января 2024 года по 11 января 2024 года, заключался в том, что была продумана концепция сервиса. И определена архитектура сервиса. В качестве фреймворка был выбран Django [3], так как он обладает рядом преимуществ, по сравнению с другими фреймворками: высокая масштабируемость, безопасность и производительность.

Далее, создана базовая структура сервиса на основе Django (Рис. 1). Проект был разделён на два приложения: *main*, которое отвечает за основной функционал сервиса, и *ai*, отвечающее за взаимодействие с искусственным интеллектом. Это было сделано для того, чтобы упростить разработку и поддержку проекта и сделать проект структурированным. Также определены

цели и задачи, произведён расчёт ресурсов, необходимых для корректной работы сервиса.

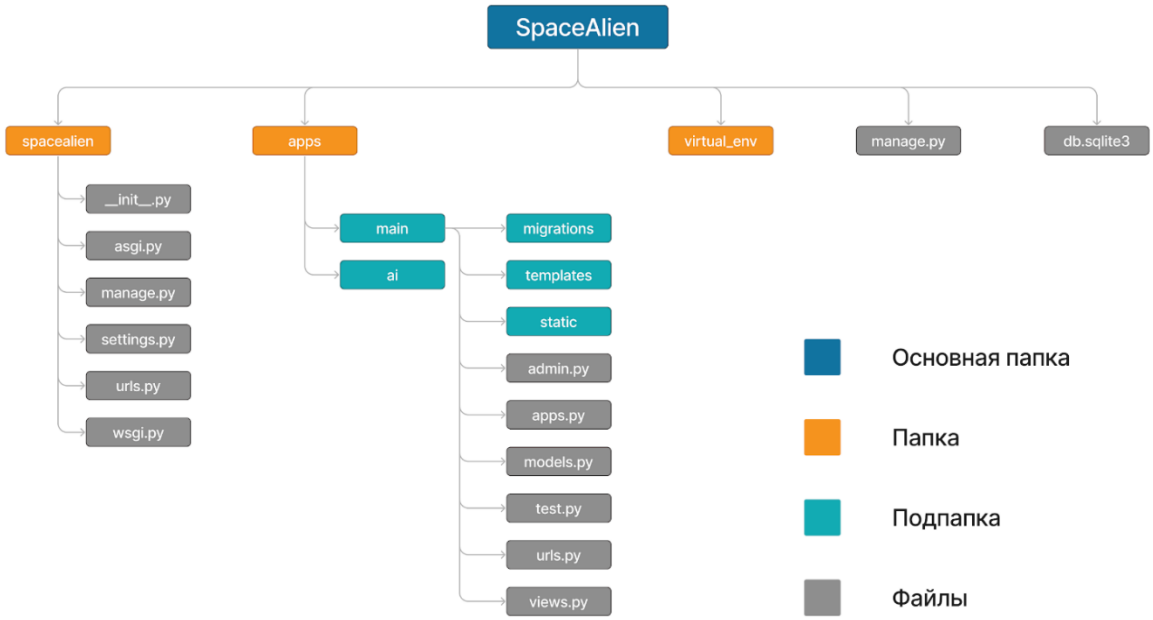

Рисунок 1 — *Структура сервиса*

3.2 Второй этап

Этап разработка дизайна, продлившийся с 12 января 2024 года по 18 января 2024 года, заключался в проектировании пользовательского интерфейса сервиса. Продумана и создана концепция интерфейса. Определены основные цвета пользовательского интерфейса (Рис. 2). Создан макет визуальной составляющей сервиса. Проектирование происходило в приложение Figma – это графический редактор для совместного проектирования сайтов, приложений и других дизайнерских продуктов. Преимуществом использования данного редактора является то, что хранение файлов происходит в облачном хранилище - это позволяет вести разработку с разных устройств без предварительного копирования файлов.

Далее, согласно ранее созданному макету, был написан пользовательский интерфейс с помощью языка разметки (HTML) и таблицы стилей (CSS). Для того чтобы улучшить взаимодействие со страницей и взаимодействовать с серверной часть сервиса, был использован JavaScript. Кроме того, на этом этапе начался сбор данных для искусственного интеллекта и фотографий космических объектов для раздела “Галерея”.

Рисунок 2 — *Основные цвета*

3.3 Третий этап

Этап разработки искусственного интеллекта, продлился с 19 января 2024 года по 16 февраля 2024 года. В качестве основы для искусственного интеллекта была выбрана модельная архитектура машинного обучения BERT[1] [8, 11]. Данная модель была выбрана по нескольким причинам: во-первых, данная модель обладает открытым исходным кодом, что позволит производить тонкую настройку модели и контролировать каждый её элемент, во-вторых, BERT демонстрирует более высокую точность по сравнению с другими моделями. Кроме того, BERT помогает обрабатывать языки и понимать контекст. Также было произведено оптимизирование и доработка архитектуры BERT для задачи

[1] BERT (Bidirectional Encoder Representations from Transformers) — языковая модель, разработанная компанией Google и основанная на архитектуре трансформеров, предназначенная для предобучения языковых представлений с целью их последующего применения в широком спектре задач обработки естественного языка.

многоклассовой классификации. Определены основные параметры модели, представленные на рисунке 3.

Параметры	Модель
Слои (Layers)	12
Скрытые слои (Hidden units)	768
Attention heads	12

Рисунок 3 — *Основные параметры модели искусственного интеллекта*

Важно отметить, что была создана определённая логика обработки поступающего запроса: основной моделью является модифицированная модель BERT, которая, классифицирует информацию на 8 классов, далее, с помощью алгоритмов обработки естественного языка или моделей искусственного интеллекта, которые также построены на основе BERT, определяется дополнительная информация из запроса пользователя, например, определение города, в котором пользователь хочет узнать погоду. Таким образом, данный подход позволяет постепенно анализировать данные и делать акцент только на нужные параметры запроса. Принцип работы искусственного интеллекта представлен на рисунке 4.

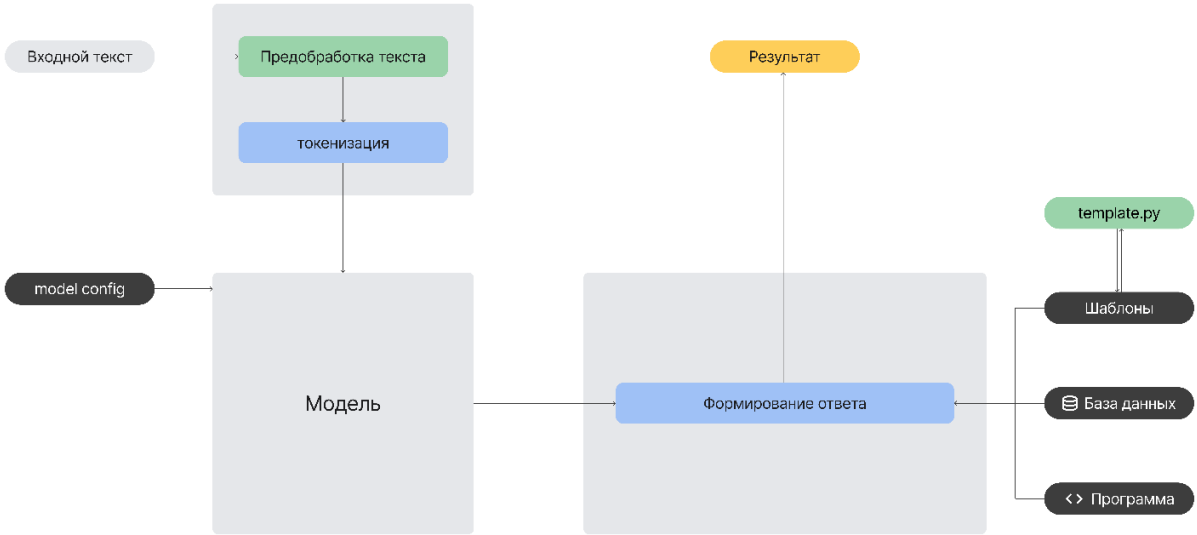

Рисунок 4 — *Принцип работы искусственного интеллекта*

Далее, произведена подготовка данных для обучения искусственного интеллекта: данные были разбиты на тестовую и обучающую выборку (рисунки 5-7). Большое количество тестовых данных обуславливается тем, чтобы корректно оценить модель искусственного интеллекта.

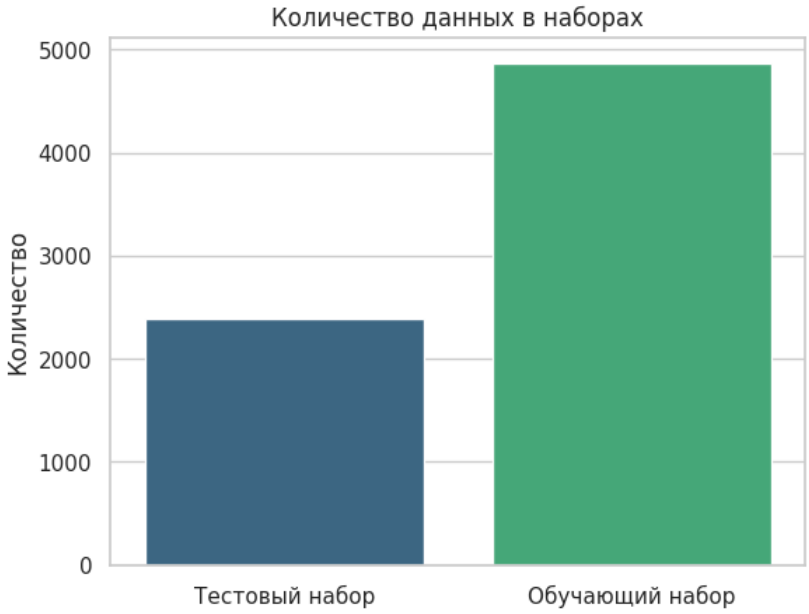

Рисунок 5 — *Количество данных*

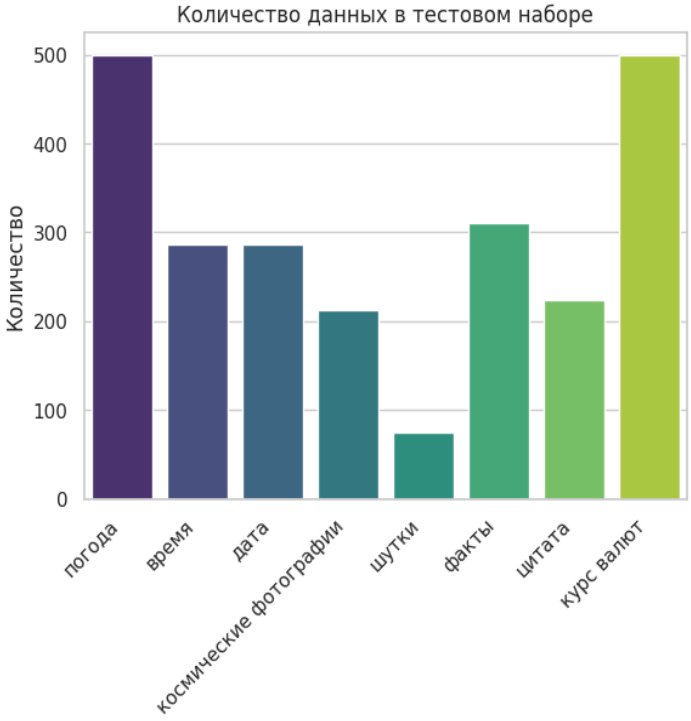

Рисунок 6 — *Распределение тестовых данных по классам*

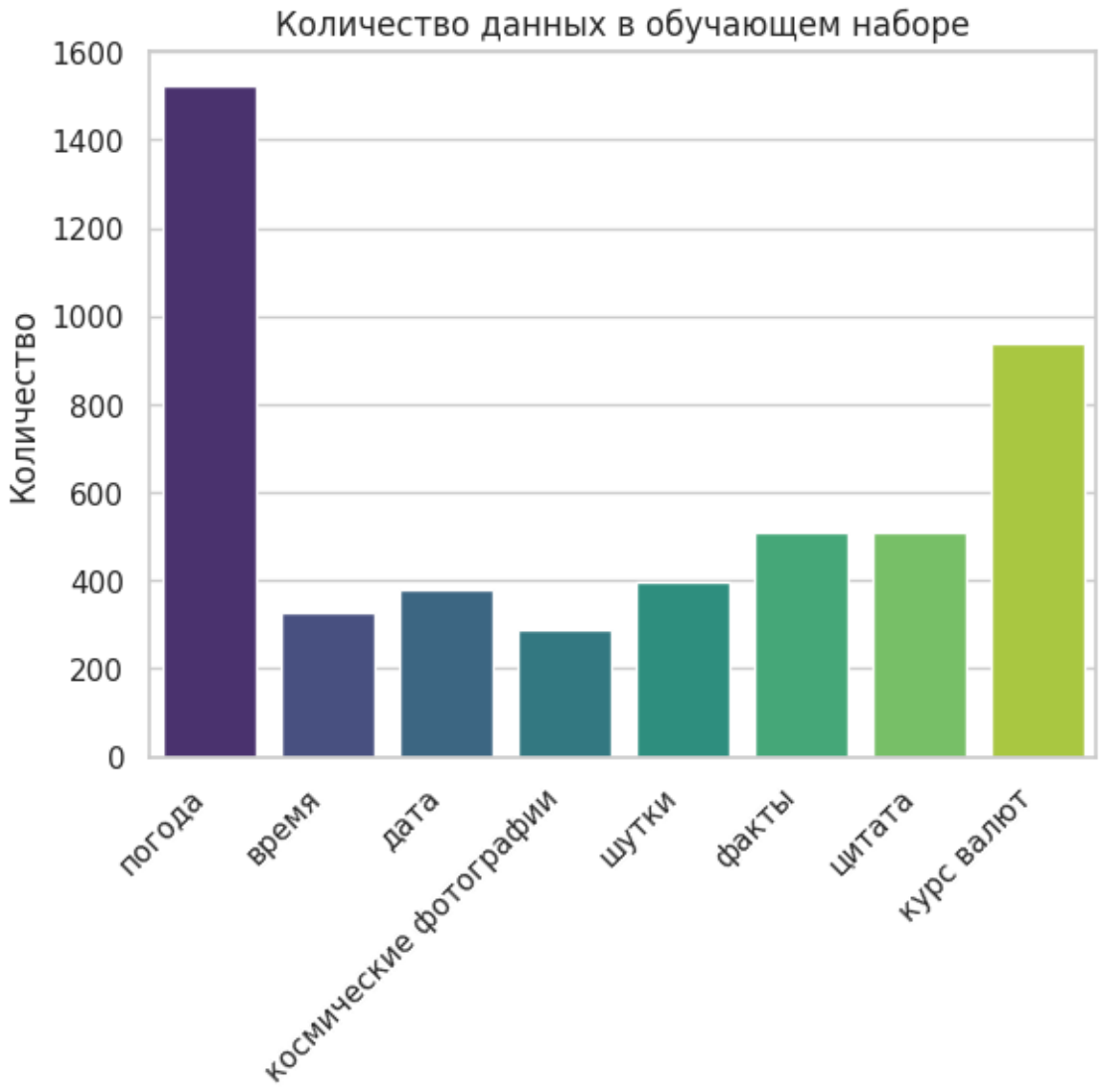

Рисунок 7 — *Распределение обучающих данных по классам*

После этого было произведено обучение модели на начальном наборе данных с использованием CPU[2] Intel Core I5-8300H @ 2.30 ГГц. Точность модели в зависимости от эпохи обучения представлено на рисунке 8. Точность определения искусственным интеллектом принадлежности к определённому классу входного запроса представлено на рисунке 9. Подробная информация об обучении модели искусственного интеллекта представлена в приложении Б.

[2] Central processing unit (CPU) - Центральный процессор

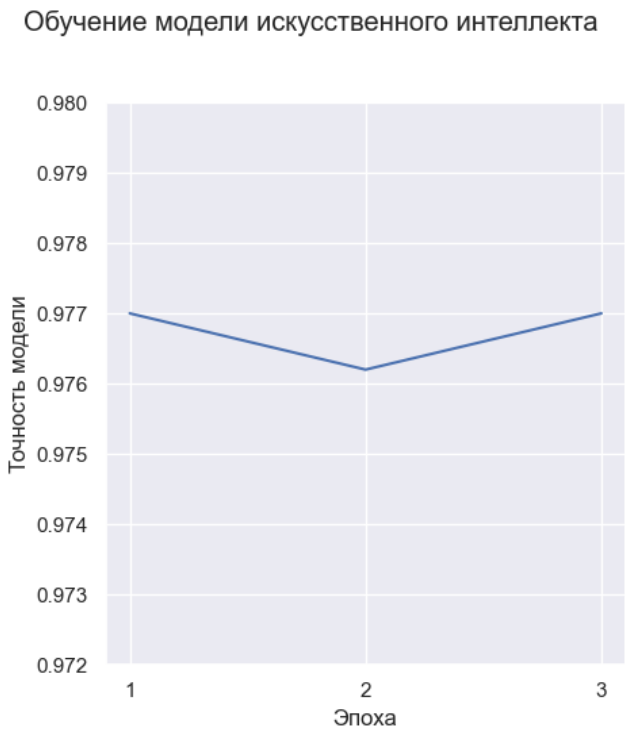

Рисунок 8 — *Точность модели в зависимости от эпохи обучения*

Класс	Точность
weather	0.98
time	0.94
date	0.95
space_image	1.00
fact	0.98
joke	0.91
quote	1.00
currency	1.00

Рисунок 9 — *Точность определения искусственным интеллектом класса входного текста*

3.4 Четвертый этап

Этап разработка серверной части сервиса, продлился с 17 февраля 2024 года по 6 марта 2024 года. Написана схема URL-адресов и обработка запросов. Также создана единая база данных для хранения данных, необходимых для работы сервера. Для всех изображений была создана копия в формате WebP[3]. Важно отметить, что при переходе на страницу, которая содержит полную информацию о фотографии, загружается оригинальная версия фотографии без потери качества. Формат WebP был выбран по нескольким причинам: во-первых, при просмотре списка всех доступных изображений не требует высокое качество, во-вторых, при уменьшении размера изображения повышается скорость загрузки страницы, в-третьих, по сравнению с JPEG [4]форматом, WebP продемонстрировал большую степень сжатия при низких потерях качества изображения. Сравнение размера изображений в форматах WebP с JPEG представлено на рисунке 10.

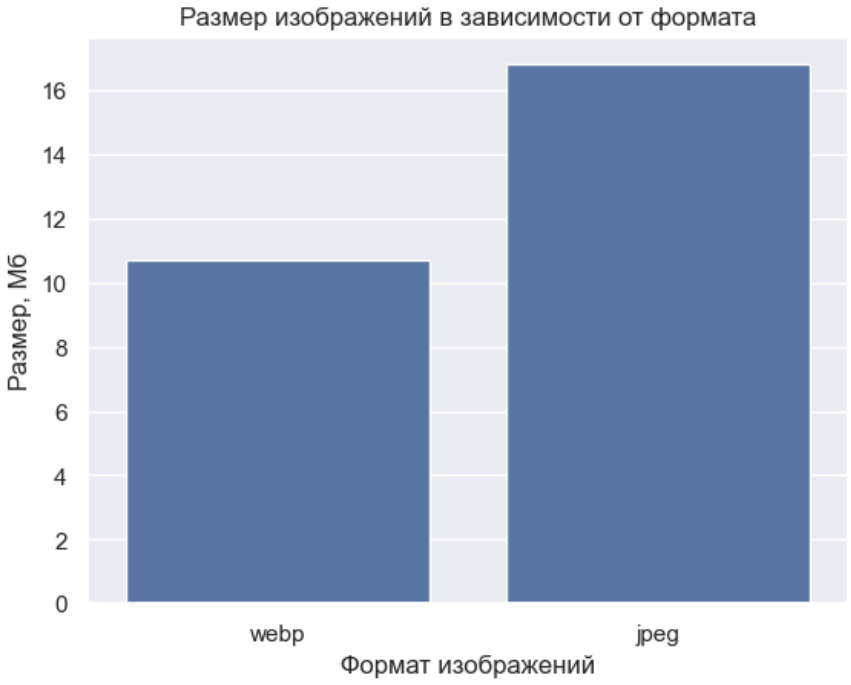

Рисунок 10 — *Сравнение размеров всех изображений в формате WebP и JPEG*

[3] WebP — формат сжатия изображений как с потерями, так и без потерь, предложенный компанией Google в 2010 году.

[4] JPEG (Joint Photographic Experts Group) — один из популярных растровых графических форматов, применяемый для хранения фотографий и подобных им изображений.

В качестве базы была выбрана база данных Sqlite3 [10], так как сервис не обладает большим количество данных и не требуется большая скорость обработки запросов. Кроме того, Sqlite3 не требует отдельного серверного

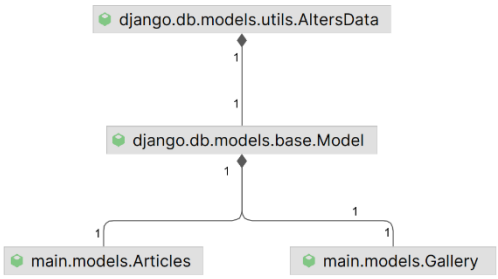

Рисунок 11 — *Архитектура базы данных в приложении "main"*

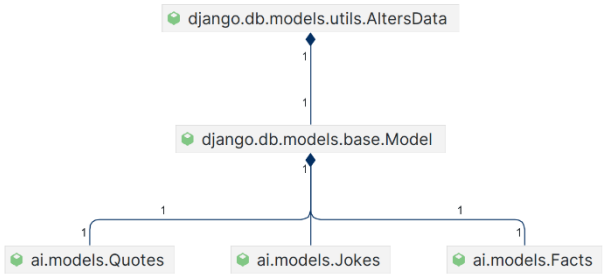

Рисунок 12 — *Архитектура базы данных в приложении "ai"*

процесса, что позволит сэкономить время на настройку базы данных. Для описания структуры таблиц в базе данных использовались стандартные ресурсы Django.

Кроме того, произведена настройка основных параметров проекта. В частности, была настроена криптографическая подпись для того, что обнаруживать любое вмешательство в передаваемые данные и защитить от подделки данные, хранящиеся в скрытых полях формы. Также была разработана программа для создания шаблонов и последующего конвертирования их в HTML код.

4 Используемые библиотеки

Для реализации серверной части сервиса и модели искусственного интеллекта использовались библиотеки Python. Основные из них представлены в таблице 1.

Таблица 1 — Основные используемы библиотеки Python

Название библиотеки	**Версия**	**Описание**
Requests [13]	2.31.0	HTTP-клиентская библиотека для

		языка программирования Python.
NumPy [7]	1.26.3	Фундаментальный пакет для научных вычислений на языке Python.
sklearn [2]	1.4.0	Python-библиотека для машинного обучения
Django [3]	5.0.1	Фреймворк для веб-приложений на языке Python
Pandas [1]	2.2.0	Библиотека в Python для работы с данными.
TensorFlow [5]	2.15.0	Фреймворк для глубокого машинного обучения
Transformers [12]	4.36.2	Transformers предоставляет API для быстрой загрузки и использования предварительно обученных моделей, которые находятся на сайте huggingface.co
PyTorch [6]	2.1.2+cpu	Фреймворк машинного обучения для языка Python с открытым исходным кодом, созданный на базе Torch.
Spacy [9]	3.7.2	Библиотека для продвинутого естественного языка, написанная на языках программирования Python и Cython.
Sqlite3 [10]	2.6.0	DB-API 2.0 интерфейс для SQLite баз данных.

5 Результаты

На данный момент проект находится на активной стадии разработки, но уже сейчас обладает базовым функционалом, который позволяет пользователю взаимодействовать с сервисом. Кроме того, модель искусственного интеллекта на основе BERT была обучена и показала высокую точность на первоначальном наборе данных. Архитектура модели также показала высокую эффективность. На данный момент модель способна классифицировать 8 классов. Также реализован базовый пользовательский интерфейс и серверная часть сервиса с использованием фреймворка Django. На сайте реализован раздел “Галерея”, главная страница и раздел “Чат” для взаимодействия с искусственным интеллектом. Собрана начальная коллекция изображений с полной информацией для раздела “Галерея”. Изображения страниц сайт можно найти в приложении А.

6 Экономические расчёты

Таблица 2 — Экономические расчёты

№	Название	Стоимость, рублей
1.	Облачная среда разработки	1500
2.	Покупка доменного имени	800
	ИТОГО	**2300**

7 Перспективы

В ближайшее время планируется обучить модель искусственного интеллекта на новом наборе данных, который будет содержать больше разнообразной информации и увеличит количество доступных классов для классификации. Модель получит возможность давать определения словам и предоставлять разную информацию, например, о животном или о планете. Будет произведена настройка алгоритмов для обработки естественного языка, чтобы повысить их эффективность. Также будет продолжена настройка параметров модели.

В частности, хочется отметить, что планируется доработать дизайн сервиса, чтобы пользовательский интерфейс корректно отображался на всех видах устройств и был удобен в использовании. Будущий дизайн главной страницы сервиса представлен на рисунке.

Также будет реализован раздел “Лекции”, который будет содержать лекции нашего времени и прошлого столетия, и доработан раздел “Статьи”. Планируется реализовать поиск среди информации, представленной на сайте.

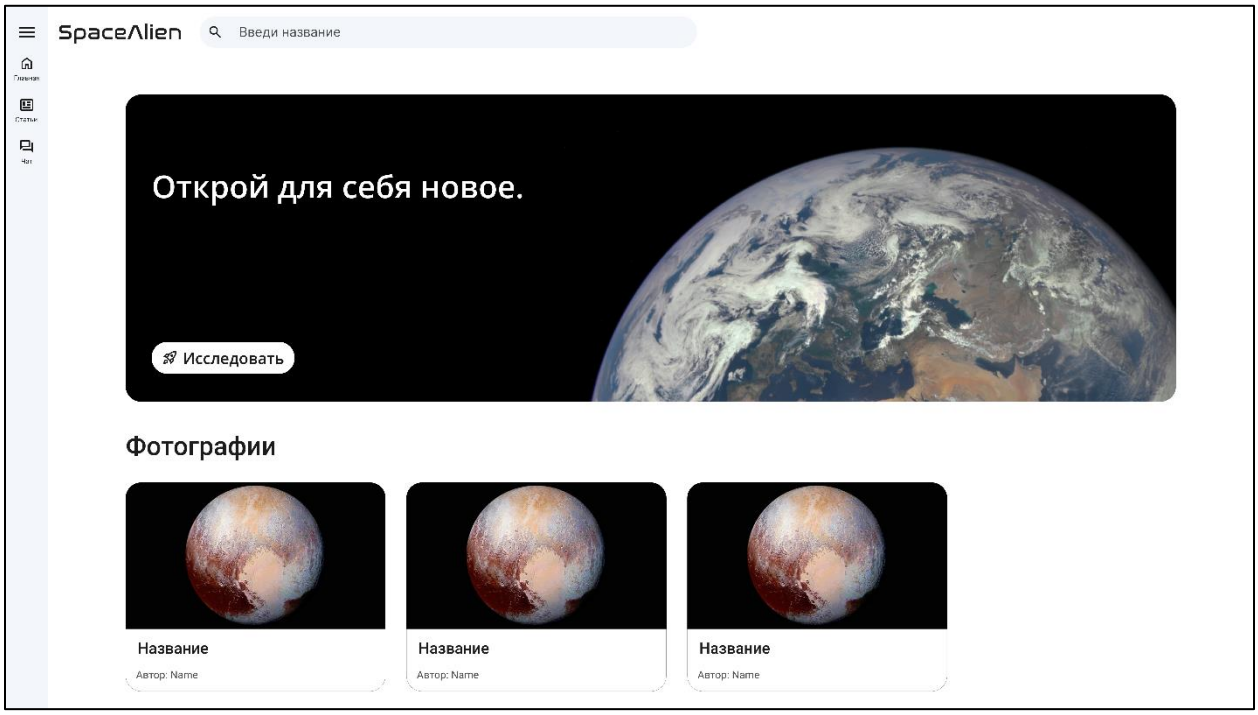

Рисунок 13 — *Новый дизайн главной страницы сервиса*

8 Список литературы

1. Pandas / [Электронный ресурс] // pandas: [сайт]. — URL: https://pandas.pydata.org/ (дата обращения: 06.03.2024).
2. Scikit-learn / [Электронный ресурс] // scikit-learn: [сайт]. — URL: https://scikit-learn.org/stable/ (дата обращения: 06.03.2024).
3. Django / [Электронный ресурс] // django: [сайт]. — URL: https://www.djangoproject.com/ (дата обращения: 06.03.2024).
4. Bert / [Электронный ресурс] // bert: [сайт]. — URL: https://github.com/google-research/bert (дата обращения: 06.03.2024).

5. Tensorflow/ [Электронный ресурс] // tensorflow: [сайт]. — URL: https://www.tensorflow.org/ (дата обращения: 06.03.2024).
6. Pytorch/ [Электронный ресурс] // pytorch: [сайт]. — URL: https://pytorch.org/ (дата обращения: 06.03.2024).
7. Numpy/ [Электронный ресурс] // numpy: [сайт]. — URL: https://numpy.org/ (дата обращения: 06.03.2024).
8. BERT - In Depth Understanding/ [Электронный ресурс] // BERT: [сайт]. — URL: https://www.kaggle.com/code/mdfahimreshm/bert-in-depth-understanding (дата обращения: 06.03.2024).
9. Spacy [Электронный ресурс] // spaCy: [сайт]. — URL: https://spacy.io/ (дата обращения: 06.03.2024).
10. SQLite3 [Электронный ресурс] // SQLite3: [сайт]. — URL: https://docs.python.org/3/library/sqlite3.html (дата обращения: 06.03.2024).
11. Google/bert [Электронный ресурс] // github: [сайт]. — URL: https://github.com/google-research/bert (дата обращения: 06.03.2024).
12. Transformers [Электронный ресурс] // huggingface: [сайт]. — URL: https://huggingface.co/docs/transformers/index (дата обращения: 06.03.2024).
13. Requests [Электронный ресурс] // Requests: [сайт]. — URL: https://requests.readthedocs.io/en/latest/ (дата обращения: 06.03.2024).

Приложение А. Изображение пользовательского интерфейса

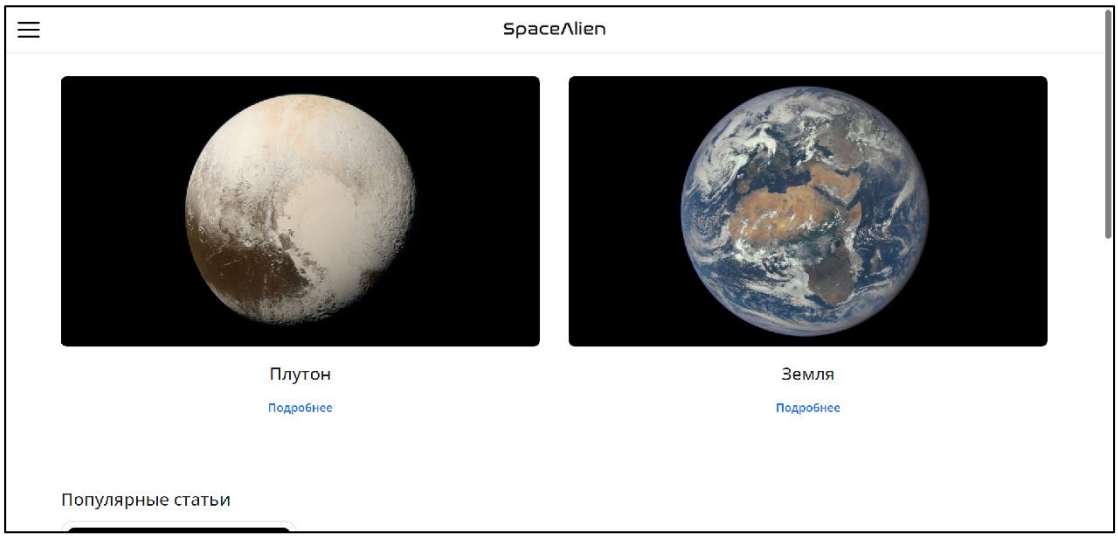

Рисунок 14 — *Главная страница*

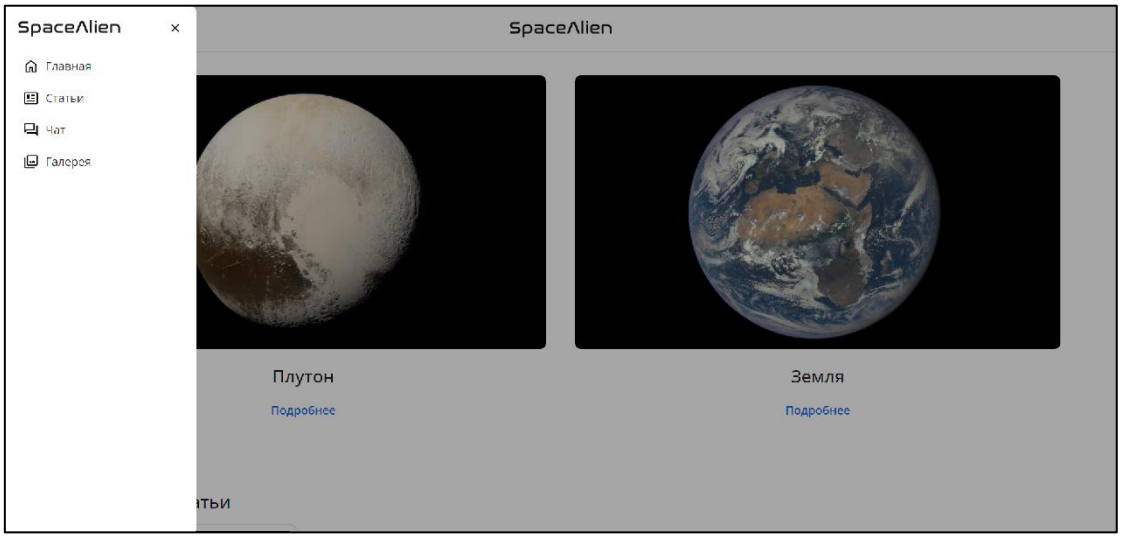

Рисунок 15 — *Меню сайта*

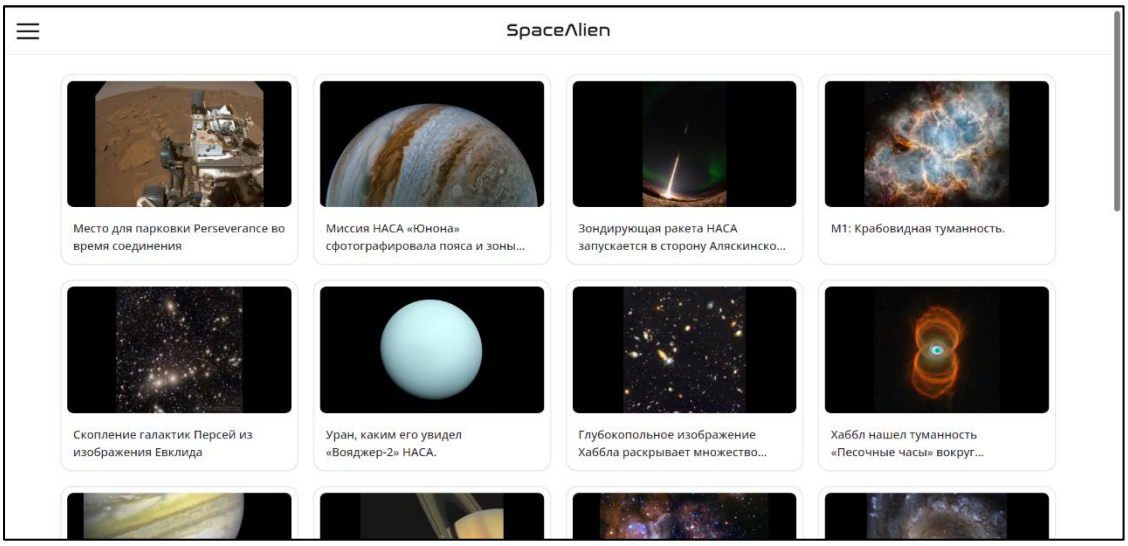

Рисунок 16 — *Страница “Галерея”*

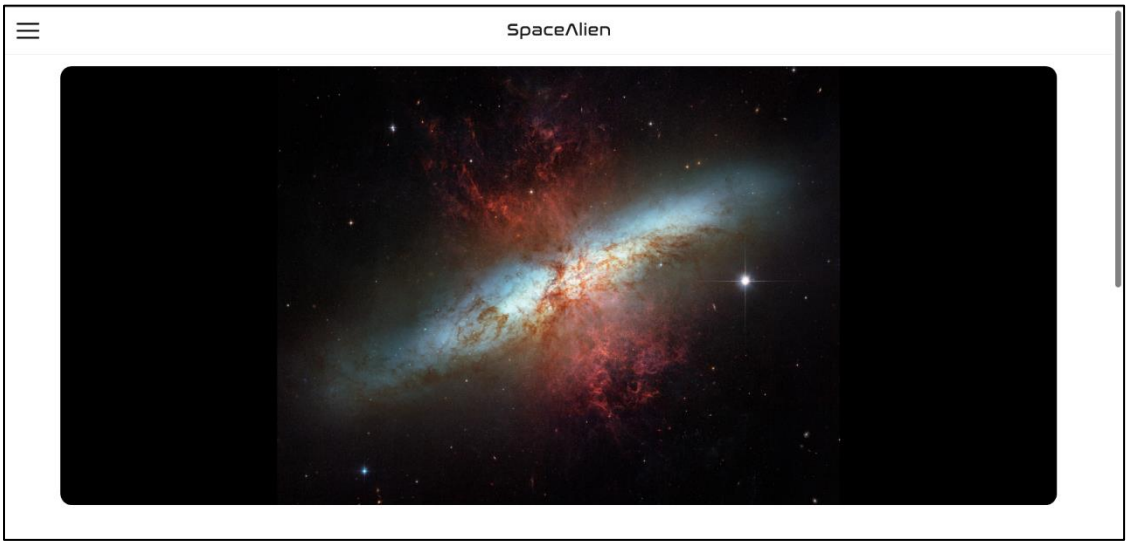

Рисунок 17 — *Страница просмотра фотографии. Часть 1*

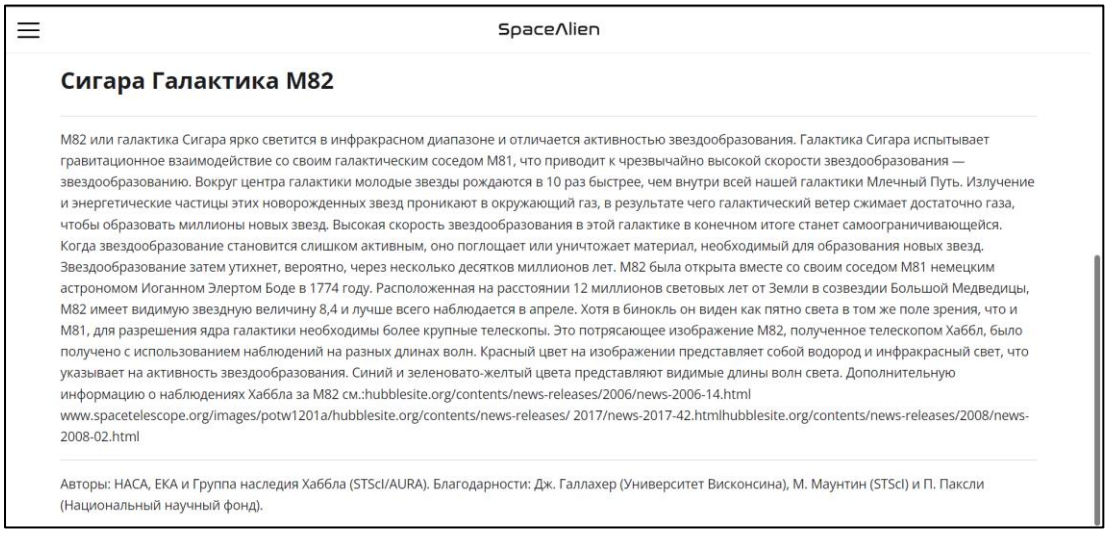

Рисунок 18 — *Страница просмотра фотографии. Часть 2*

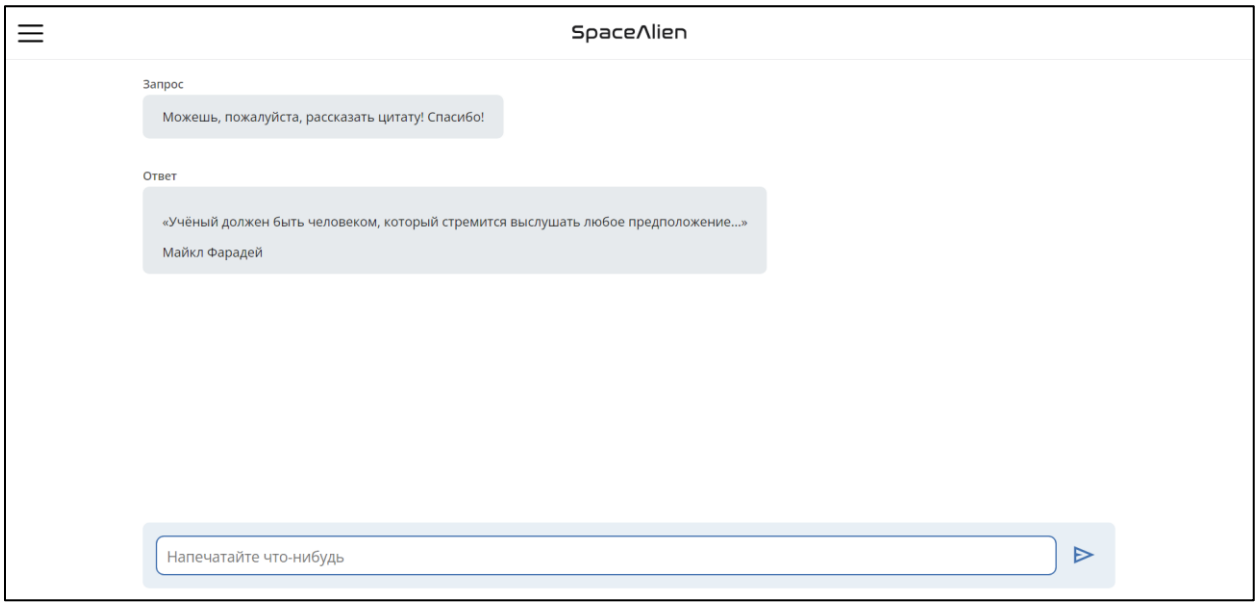

Рисунок 19 — *Страница для взаимодействия с искусственным интеллектом*

Приложение Б. Подробная информация об обучении модели искусственного интеллекта

```
Epoch 1/3
Validation Accuracy: 0.9770
              precision    recall  f1-score   support

           0       1.00      0.99      1.00       500
           1       0.95      0.90      0.92       287
           2       0.90      0.95      0.93       286
           3       0.99      1.00      0.99       212
           4       0.94      0.97      0.95        75
           5       0.99      0.98      0.99       311
           6       1.00      0.99      1.00       224
           7       1.00      1.00      1.00       500

    accuracy                           0.98      2395
   macro avg       0.97      0.97      0.97      2395
weighted avg       0.98      0.98      0.98      2395

Epoch 2/3
Validation Accuracy: 0.9762
              precision    recall  f1-score   support

           0       0.97      1.00      0.98       500
           1       0.96      0.90      0.93       287
           2       0.96      0.96      0.96       286
           3       1.00      1.00      1.00       212
           4       0.84      0.99      0.91        75
           5       0.99      0.96      0.97       311
           6       1.00      0.99      1.00       224
           7       1.00      1.00      1.00       500

    accuracy                           0.98      2395
```

```
   macro avg       0.96      0.97      0.97      2395
weighted avg       0.98      0.98      0.98      2395

Epoch 3/3
Validation Accuracy: 0.9770
              precision    recall  f1-score   support

           0       0.98      1.00      0.99       500
           1       0.94      0.91      0.92       287
           2       0.95      0.94      0.95       286
           3       1.00      1.00      1.00       212
           4       0.91      0.97      0.94        75
           5       0.98      0.98      0.98       311
           6       1.00      0.99      1.00       224
           7       1.00      1.00      1.00       500

    accuracy                           0.98      2395
   macro avg       0.97      0.97      0.97      2395
weighted avg       0.98      0.98      0.98      2395
```