**Программная документация: Анализатор тональности текста**

## **1. Оглавление**

1. **Введение**
2. **Описание системы**
3. **API документация**
4. **Руководство пользователя**
5. **Требования к системе**
6. **Установка и запуск**
7. **Примеры использования**

## **2. Введение**

Данная документация описывает работу **веб-приложения для анализа тональности текста** и соответствующего **REST API**.  
Система позволяет:

* Определять эмоциональную окраску текста (позитивная/негативная).
* Выводить статистику по предобработанным данным.
* Работать как через веб-интерфейс, так и через API.

## **3. Описание системы**

### **3.1. Компоненты**

1. **FastAPI-сервер** (api.py) – обрабатывает запросы, возвращает предсказания модели.
2. **Streamlit-интерфейс** (app.py) – визуализирует результаты анализа.
3. **ML-модель** (model\_rf.pkl) – Random Forest, обученный на размеченных данных.
4. **Векторизатор** (vectorizer.pkl) – TF-IDF для преобразования текста в числовой формат.

### **3.2. Технологии**

* **Python 3.9+**
* **FastAPI** (API сервер)
* **Streamlit** (веб-интерфейс)
* **Scikit-learn** (ML модель)
* **Pandas, Matplotlib, WordCloud** (анализ и визуализация)

## **4. API документация**

### **4.1. Эндпоинты**

#### POST /predict

Анализ тональности текста.

{

"text": "текст который приходит от пользователя"

}

Ответ

{

"sentiment": "positive | negative",

"confidence": 0.95,

"probabilities": {

"positive": 0.95,

"negative": 0.05

}

}

#### GET /model\_info

**Назначение**: Получение информации о модели.  
**Ответ**:

{

"model\_type": "RandomForestClassifier",

"vectorizer\_type": "TF-IDF",

"features": 10000

}

## **5. Руководство пользователя**

### **5.1. Веб-интерфейс (Streamlit)**

#### **Вкладка "Text Scan"**

1. Введите текст в поле ввода.
2. Нажмите **"EXECUTE SCAN"**.
3. Получите результат:
   * **Sentiment** (positive/negative)
   * **Confidence** (уверенность модели)
   * **График распределения вероятностей**

#### **Вкладка "Data Matrix"**

1. **Общая статистика**:
   * Количество записей
   * Средняя длина текста
   * Доля позитивных отзывов
2. **Графики**:
   * Распределение длины текста
   * Облако слов

## **6. Примеры использования**

### **6.1. Через веб-интерфейс**

