



Модель машинного обучения для определения тональности коротких комментариев

Яковлев Е.А.



Структура работы

1. Введение

- Актуальность темы
- Цели и задачи проекта
- Обзор существующих методов анализа тональности
- Календарь

2. Обзор литературы

- Список источников-материалов

3. Методика исследования

- Подготовка данных (сбор, очистка, разметка)
- Выбор модели и архитектуры (например, наивный Байес, LSTM, CNN, трансформеры)
- Описание алгоритмов и инструментов

4. Рабочая модель

- Предоставление исходного кода
- Предоставление данных работы модели
- Предоставление эффективности работы модели

5. Заключение

- Основные выводы
- Практическое значение работы

6. Доп материалы

Введение

Определение тональности (позитив/негатив) в коротких текстах напримере отзывов.

Выполнение проекта по определению тональности коротких текстов даст мне практический опыт в программировании на Python и работе с PyTorch. Также поможет понять эффективность современных методов машинного обучения для анализа эмоций в текстах. Для меня важно получить ответ на вопрос, насколько точно ИИ может распознавать позитивные и негативные отзывы. Данная тема напрямую связано с моей повседневной жизнью и семьёй, поскольку мы часто ориентируемся на отзывы в интернете при выборе товаров и услуг.

Для общества проект полезен, так как автоматический анализ тональности помогает быстро оценивать общественное мнение, улучшать коммуникацию между клиентами и организациями, использоваться в школах и социальных проектах. В России этим вопросом занимаются научные организации, например, лаборатории Высшей школы экономики, РЭУ им. Г.В. Плеханова и Университета Тюмени. Они исследуют методы анализа тональности для русского языка, стремясь улучшить точность и применять эти технологии в разных сферах. Мой проект предлагает доступное решение, которое может применяться для образовательных и бытовых целей.

Научная актуальность подтверждается исследованиями в области обработки естественного языка и анализа тональности, где применяются современные нейросетевые методы, включая трансформеры и BERT-модели. Мой проект отличается фокусом на практическом использовании PyTorch для создания моделей, адаптированных под русскоязычные тексты, и экспериментами с их оптимизацией. Таким образом, он вносит вклад в развитие доступных и эффективных инструментов анализа эмоциональной окраски текста в русском языке.

Календарь

Дата	Этапы работы
23.10.2025	Формулирование календаря работы над проектом, структуры проекта и списка литературы
До 31.10.2025	Сбор и подготовка данных - начало обучение и подговки модели
До 11.11.2025	Завершение обучения модели и создание структуры работы, формулирование глав и их содержания
До 2.12.2025	Оформление итоговой текстовой работы
До 16.12.2025	Оформление презентации и выступления по итогам проектной работы



Статья разработка методов анализа тональности публицистических текстов на русском языке [Электронный ресурс]. - URL: <https://rscf.ru/project/23-21-00495/> (дата обращения: 23.10.2025).

Статья новое нейросетевое решение для анализа тональности текста на естественном языке [Электронный ресурс]. - URL: <https://www.rea.ru/news/25616-uchnyie-reu-razrabotali-novoe-neyrosetevoe-reshenie-dlya-analiza-tonalnosti-teksta-na-estestvennom-yazyike> (дата обращения: 23.10.2025).

Официальная документация фреймворка Pytorch [Электронный ресурс]. - URL: <https://docs.pytorch.org/docs/stable/index.html> (дата обращения: 23.10.2025).

Официальный сайт фреймворка Pytorch [Электронный ресурс]. - URL: <https://pytorch.org/> (дата обращения: 23.10.2025).

Официальный сайт HuggingFace (Статьи, дата-сеты)[Электронный ресурс]. - URL: <https://huggingface.co/> (дата обращения: 23.10.2025)

Официальный сайт Kaggle (Статьи, дата-сеты)[Электронный ресурс]. - URL: <https://www.kaggle.com/> (дата обращения: 23.10.2025)