

1 Проект 1: Классификация отзывов на фильмы

1.1 Категория

Классификация

1.2 Описание

Проект направлен на классификацию отзывов на фильмы как позитивные или негативные с использованием анализа тональности. Используются методы обработки естественного языка (NLP) и алгоритмы машинного обучения.

1.3 План реализации

- 1. Сбор данных:** Использовать датасет IMDb или аналогичный.
- 2. Предобработка текста:** Токенизация, удаление стоп-слов, стемминг/лемматизация.
- 3. Векторизация текста:** TF-IDF или Word2Vec.
- 4. Модель:** Логистическая регрессия или нейронная сеть.
- 5. Оценка:** Метрики precision, recall, F1-score.