План работы

Этап 1: Анализ источников

Проведение обзора научных статей, книг и интернет-ресурсов, посвященных теме выделения ключевых слов. Изучение различных методов и подходов к выделению ключевых слов.

Этап 2: Определение и выбор методов

Анализ существующих алгоритмов выделения ключевых слов. Выбор наиболее подходящих методов для реализации и тестирования (например, TF-IDF, TextRank, BERT и др.).

Этап 3: Реализация алгоритмов

Реализация выбранных методов на языке Python. Использование библиотек для обработки текста, таких как NLTK, SpaCy и других. Написание понятного и структурированного кода с комментариями для облегчения понимания.

Этап 4: Тестирование и оценка качества

Проведение экспериментов на текстовых данных с использованием реализованных алгоритмов. Оценка производительности алгоритмов по критериям точности, полноты, времени обработки и сложности. Сравнение результатов выделения ключевых слов с ручными результатами.

Этап 5: Оптимизация алгоритмов

Настройка параметров выбранных методов для улучшения производительности. Проведение дополнительных тестов после оптимизации для оценки улучшений.

Этап 6: Анализ и выводы

Подведение итогов работы, обсуждение полученных результатов. Оценка преимуществ и недостатков реализованных методов. Формулировка выводов относительно эффективности алгоритмов и возможных направлений дальнейшего исследования.

Этап 7: Подготовка отчета и презентации

Оформление итоговой работы в соответствии с требованиями. Подготовка презентации для защиты проекта, включающей основные результаты и выводы. Включение в приложение полного кода реализованных алгоритмов и дополнительных материалов, таких как графики и таблицы с результатами.

Таким образом, работа будет включать теоретическое исследование, практическую реализацию и сравнение методов выделения ключевых слов с использованием современных инструментов обработки текстовых данных.