

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Макаренко Елена Николаевна

Должность: Ректор

Дата подписания: 29.07.2022 17:52:52

Уникальный идентификатор документа:
c098bc0c1041cb2a4cf926cf171d6715d99a6ae00adc8e27b55cbe1e2dbd7c78

ЛАБОРАТОРНАЯ РАБОТА № 1

Построение векторной модели текста и классификация длинных текстов

Лабораторная работа посвящена изучению алгоритмов векторизации текста и получению навыков практического применения этих алгоритмов для решения задач классификации и кластеризации текстов.

Работа включает в себя:

- изучение алгоритмов, необходимых для выполнения задания;
- изучение архитектуры нейронных сетей, необходимых для выполнения задания;
- изучение библиотек, содержащих реализацию требуемых алгоритмов;
- написание программы, решающей поставленную задачу;
- составление отчета о проделанной работе с полным описанием программы и полученных результатов.

Варианты заданий:

- бинарная классификация спам сообщений (сравнить эффективность использования методов векторизации текста: мешок слов и CBOW);
- бинарная классификация спам сообщений (сравнить эффективность использования методов векторизации текста: прямое кодирование и skip-gram);
- поиск в корпусе документов, наиболее релевантных заданной фразе (использовать меру TF-IDF);
- тематическая классификация документов (использовать меру TF-IDF);
- кластеризация документов корпуса (использовать меру TF-IDF).

Студент может предложить свой собственный вариант.

Лабораторная работа проводится в аудитории, оборудованной в соответствии с рабочей программой дисциплины.

Основанием для допуска к лабораторной работе является знание теоретического материала.

Процесс выполнения лабораторной работы документируется с помощью текстового редактора MS Word, полученные сведения служат основой для формирования отчета о выполнении лабораторной работы.

Защита отчета о выполнении лабораторной работы сопровождается демонстрацией полученных результатов, соответствующих теоретических знаний и ответов на дополнительные вопросы преподавателя по лабораторной работе.

ЛАБОРАТОРНАЯ РАБОТА № 2

Статистические языковые модели

Лабораторная работа посвящена изучению статистических методов обработки языка и их использованию для решения задачи генерации и восстановления текстов.

Работа включает в себя:

- изучение алгоритмов, необходимых для выполнения задания;
- изучение архитектуры нейронных сетей, необходимых для выполнения задания;
- изучение библиотек, содержащих реализацию требуемых алгоритмов;
- написание программы, решающей поставленную задачу;
- составление отчета о проделанной работе с полным описанием программы и полученных результатов.

Варианты заданий:

- генерация текста на заданную тему с использованием рекуррентной нейронной сети;
- генерация текста на заданную тему с использованием рекуррентной нейронной сети с контекстными признаками;
- восстановление пропущенных фрагментов текста с использованием рекуррентной нейронной сети;
- восстановление пропущенных фрагментов текста с использованием рекуррентной нейронной сети с контекстными признаками.

Студент может предложить свой собственный вариант.

Лабораторная работа проводится в аудитории, оборудованной в соответствии с рабочей программой дисциплины. Основанием для допуска к лабораторной работе являются:

- знания теоретического материала;
- отсутствие задолженности по предыдущей лабораторной работе.

Процесс выполнения лабораторной работы документируется с помощью текстового редактора MS Word, полученные сведения служат основой для формирования отчета о выполнении лабораторной работы.

Защита отчета о выполнении лабораторной работы сопровождается демонстрацией полученных результатов, теоретических знаний и ответов на дополнительные вопросы преподавателя по лабораторной работе.

ЛАБОРАТОРНАЯ РАБОТА № 3
Распознавание структуры коротких текстов
и преобразование последовательностей

Лабораторная работа заключается в выделении в тексте сегментов или коротких подпоследовательностей и соотнесения их с заданными категориями. На вход поступает короткий текст (чаще всего это отдельное предложение). Для каждого токена нужно предсказать класс.

Работа включает в себя:

- изучение алгоритмов, необходимых для выполнения задания;
- изучение архитектуры нейронных сетей, необходимых для выполнения задания;
- изучение библиотек, содержащих реализацию требуемых алгоритмов;
- написание программы, решающей поставленную задачу;
- составление отчета о проделанной работе с полным описанием программы и полученных результатов.

Варианты заданий:

- найти в заданном тексте имя и локацию главного героя;
- в заданном предложении определить части речи;
- поиск именованных сущностей в тексте;
- генерация аннотаций, то есть сжатого описания содержимого более длинного текста.

Лабораторная работа проводится в аудитории, оборудованной в соответствии с рабочей программой дисциплины. Основанием для допуска к лабораторной работе являются:

- знания теоретического материала;
- отсутствие задолженностей по предыдущим лабораторным работам.

Процесс выполнения лабораторной работы документируется с помощью текстового редактора MS Word, полученные сведения служат основой для формирования отчета о выполнении лабораторной работы.

Защита отчета о выполнении лабораторной работы сопровождается демонстрацией полученных результатов, теоретических знаний и ответов на дополнительные вопросы преподавателя по лабораторной работе.