**ЛАБОРАТОРНАЯ РАБОТА No1**

Тема лабораторной работы: тестирование документации.

**1. Выбор документации**

Я выбрала для тестирования документацию к моей курсовой работе, посвященной применению методов машинного обучения для нормализации слов текста.

Эта работа, представленная в формате Word, детально описывает требования к проектированию и разработке моделей для обучения, а также структуру базы данных, используемой в проекте.

Код программы, написанный на языке Python, представлен в Jupyter Notebook, что позволяет наглядно демонстрировать процесс обработки и нормализации текста.

**2. Анализ и составление отчета**

**a. Цель работы**

Целью данной курсовой работы является разработка и реализация методов нормализации слов текста с использованием технологий машинного обучения. Проект нацелен на обучение моделей, способных эффективно преобразовывать различные формы слов в их каноническую (нормализованную) форму, что является ключевым этапом в обработке естественного языка. Нормализация слов позволяет улучшить качество и точность анализа текста, что особенно важно для задач автоматической обработки текстов, таких как поиск информации, машинный перевод и другие.

**b. Описание тестируемой документации**

Курсовая работа состоит из документа Word, количество страниц 37, включая приложение.

Код программы прикреплен в качестве пункта «Приложение», а также представлен в Jupyter Notebook.

Тестируемая документация включает следующие разделы:

1. Введение
2. Теоретическая часть
   1. Основные методы нормализации текста
   2. Используемые технологии
   3. Описание алгоритмов машинного обучения
3. Обработка данных
   1. Сбор и очистка данных
   2. Удаление стоп-слов
   3. Стемминг и лемматизация
4. Разработка и обучение модели
   1. Применение алгоритмов машинного обучения
   2. Повышение эффективности модели
   3. Анализ результатов
5. Заключение
6. Список литературы
7. Приложение

Код программы, размещенный в Jupyter Notebook, сопровождается комментариями, поясняющими каждый шаг процесса нормализации текста.

**c. Описание критериев качества тестируемой документации**

Документация будет оцениваться по следующим критериям:

1. Полнота:

Каждый элемент функциональности должен быть представлен в документе в требуемом объеме.

1. Однозначность:

Интерпретация написанного в документе должна быть однозначной, чтобы обеспечить ясность и понимание своих собственных инструкций и требований.

1. Непротиворечивость:

Документ проверяется на наличие конфликтных требований.

1. Актуальность:

Документация и реально разработанный программный продукт соответствуют друг другу.

1. Структурированность:

Структурные элементы документации позволяют пользователю осуществлять поиск нужной информации.

1. Тестируемость:

Описанная функциональность должна быть проверяема на финальном этапе разработки.

1. Соответствие стандартам:

Тестируемая документация должна быть выполнена в соответствии со стандартами (ГОСТ) и ТЗ.

**d. Описание и обоснование метода тестирования документации**  
 Для тестирования документации я выбрала **метод рецензирования.** Этот метод включает тщательный анализ документации с моей стороны для выявления потенциальных проблем и несоответствий. Рецензирование позволяет мне получить всесторонний взгляд на качество документации, что помогает выявить различные проблемы.

Такой выбор обусловлен тем, что я самостоятельно выполняю всю работу без участия других специалистов, и рецензирование не требует создания дополнительных материалов, что делает его наиболее подходящим для данной курсовой работы.

**e. Список несоответствий в документации критериям качества**

1. Полнота

* Тема курсового проекта полностью раскрыта, подробно рассмотрены цель, задачи и все аспекты данной темы.

1. Однозначность

* Пункт 3: Термин "токенизация" не объяснен должным образом.
* Рекомендация: включить определение термина и его отличие от стемминга и лемматизации.

1. Непротиворечивость

* Пункт 3.3: в разделе "Анализ результатов" указано, что точность модель после Grid Search равна 0.93, одна в итоговых графиках точность 0.91
* Рекомендация: обновить документацию для согласования с итоговыми графиками.

1. Актуальность

* Пункт 3.2: в разделе "Анализ результатов" приведены устаревшие данные тестирования.
* Рекомендация: провести актуальные тесты и обновить результаты.

1. Структурированность

* Структура курсового проекта четкая, имеет подзаголовки и структурированный текст, что облегчает понимание целей и задач работы.

1. Тестируемость

* Пункт 5.5: отсутствует оценка времени выполнения для каждой модели
* Рекомендация: провести эксперименты и измерить время обучения каждой модели

1. Соответствие стандартам:

* Текст работы не содержит ссылки на используемую литературу согласно ГОСТ.
* Работа полностью соответствует ТЗ
* Рекомендация: включить в работу ссылки используемой литературы согласно ГОСТ.

**Заключение**

В ходе тестирования документации к курсовой работе по применению методов машинного обучения для нормализации слов текста были выявлены как положительные аспекты, так и некоторые несоответствия.

Положительные моменты включают полноту раскрытия темы работы, структурированность документации и четкость изложения основных концепций.

Однако, были выявлены некоторые недочеты, такие как отсутствие объяснения определенных терминов, противоречия в представленных данных и несоответствие стандартам.

В целом, тестирование документации позволило выявить области для улучшения и обеспечить более высокое качество курсовой работы.

**Список используемых источников**

1. Учебное пособие «Основы управления качеством программных средств»
2. ГОСТ 28806-90 «КАЧЕСТВО ПРОГРАММНЫХ СРЕДСТВ. Термины и определения (Software quality. Terms and definitions)» [1].
3. ГОСТ 28195-89 «ОЦЕНКА КАЧЕСТВА ПРОГРАММНЫХ СРЕДСТВ. Общие положения (Quality control of software systems. General principles)» [2].
4. ГОСТ Р ИСО/МЭК 25000. Требования и оценка качества систем и программного обеспечения (SQuaRE). Модели качества систем и программных продуктов [3].