**ЛАБОРАТОРНАЯ РАБОТА №2**

**Тема работы: классификация видов тестирования.**

**1. Выбор программного проекта**

Выбираем программный проект из курсовой работы, где реализована система классификации текстов на примере анализа тональности отзывов фильмов.

**2. Анализ проекта и составление отчета**

**a. Цель работы** Целью данной лабораторной работы является анализ и оценка разработанного программного проекта, включающего в себя систему классификации текстов с применением методов машинного обучения.

**b. Описание программного проекта** Проект представляет собой систему классификации текстов, предназначенную для анализа тональности отзывов фильмов. В ходе выполнения проекта использованы различные методы машинного обучения, такие как наивный байесовский классификатор, метод опорных векторов (SVM), логистическая регрессия, деревья решений и случайные леса, а также нейронные сети и глубокое обучение. В проекте проведены этапы предобработки текстов, включающие токенизацию, нормализацию, исключение стоп-слов и преобразование текстов в числовые векторы.

**c. Элементы классификации и виды тестирования**

* **Виды тестирования**
  + **Модульное тестирование**: Проверка отдельных модулей системы, таких как токенизация и нормализация текстов.
  + **Интеграционное тестирование**: Проверка взаимодействия различных модулей, например, этапа предобработки данных и этапа обучения моделей.
  + **Системное тестирование**: Комплексное тестирование всей системы классификации для обеспечения корректной работы всех компонентов.
  + **Приемочное тестирование**: Проверка соответствия системы требованиям и ожиданиям заказчика, в данном случае — корректная классификация отзывов по тональности.
* **Типы тестирования**
  + **Функциональное тестирование**: Проверка функциональных требований системы, таких как корректная классификация текста.
  + **Нефункциональное тестирование**: Оценка производительности, масштабируемости и надежности системы.
  + **Тестирование безопасности**: Проверка защиты данных и предотвращения несанкционированного доступа.
* **Методы тестирования**
  + **Динамическое тестирование**: Тестирование, проводимое путем выполнения кода, например, проверка точности классификации на тестовом наборе данных.
  + **Статическое тестирование**: Тестирование без выполнения кода, например, обзор кода и анализ архитектуры.
* **Уровни тестирования**
  + **Низкий уровень (unit)**: Модульное тестирование отдельных функций и компонентов.
  + **Средний уровень (integration)**: Интеграционное тестирование, проверка взаимодействия модулей.
  + **Высокий уровень (system)**: Системное тестирование всей системы.

**d. Примеры дефектов, характерных для каждого вида тестирования**

* **Модульное тестирование**: Выявление ошибок в функции токенизации, таких как неправильное разделение слов или некорректная нормализация.
* **Интеграционное тестирование**: Обнаружение проблем в передаче данных между модулями предобработки и обучения модели, например, потеря данных при преобразовании.
* **Системное тестирование**: Ошибки в классификации, когда система неверно определяет тональность текста.
* **Приемочное тестирование**: Несоответствие системы требованиям, например, недостаточно высокая точность классификации для использования в реальных условиях.

**e. Выводы по работе**

Проект представляет собой эффективное решение для классификации текстов с использованием методов машинного обучения. Разработанная система успешно справляется с задачей анализа тональности отзывов фильмов, однако для улучшения качества классификации возможно применение дополнительных методов предобработки данных и более сложных моделей машинного обучения.

**f. Список использованных источников**

* Попов Д.С. «Машинное обучение в задачах классификации текстов». Курсовая работа, Финансовый университет при Правительстве Российской Федерации, Москва, 2024.
* Горячева, А. В., & Гурьянов, В. В. (2020). Обзор методов машинного обучения для классификации текстов. *Информационно-управляющие системы*, 1, 23-29.
* Лавров, С. М. (2021). Машинное обучение и искусственный интеллект: теоретические и практические аспекты. *Программные продукты и системы*, 4, 5-13.