

Документация проекта

Краткое описание задачи

Современные системы для разработки документов и научные журналы сталкиваются с рядом вызовов при работе с математическими формулами. Формулы являются важным элементом, который необходимо легко редактировать, сохранять и интегрировать в другие модули. Также в научной среде важно проверять формулы на совпадения и защищать авторство математических конструкций.

Текущие инструменты либо ограничены функционально, либо требуют значительных ресурсов для интеграции, что создает барьеры для пользователей и увеличивает трудозатраты на создание, проверку и использование формул в цифровом виде.

Цель проекта

Создать универсальное решение, которое будет поддерживать:

- Удобный интерфейс для создания формул.
- Инструмент проверки формул на совпадения, в том числе для задач антиплагиата.

Содержание решения

Разработка web-интерфейса редактора формул

- Рабочая область для создания формул.
- Возможность вставки формул в текст и записи формул как отдельных объектов.
- Поддержка стандартных математических элементов: дробь, корень, интеграл, логарифм, возведение в степень и т.д.
- Расширяемость библиотеки математических блоков.

Разработка web-интерфейса анализа формул

- Рабочая область для ввода и импортирования формул для анализа.
- Поддержка стандартных математических элементов: дробь, корень, интеграл, логарифм, возведение в степень и т.д.
- Возможность расширения анализируемых конструкций формата LaTeX.

Обработка формул

- Преобразование формул в строковый формат LaTeX для хранения в БД и передачи в JSON.
- Обратное преобразование из строки LaTeX в визуальное представление в редакторе.

Анализ формул

- Индексирование записей формул в строковом формате LaTeX для сопоставления.
- Отображение процента совпадений и выделение совпадающих частей формул в визуальном представлении.

Выходные данные

- Web-интерфейс для ввода и редактирования формул.
- Web-интерфейс для анализа формул.
- Экспорт формул в строковый формат LaTeX и обратно.

- Обработка формул в строковом формате LaTeX.
-

Требования к результату разработки

Функционал редактора формул

- Создание и редактирование формул пользователем в web-интерфейсе.
- Конвертация формул в строковый формат LaTeX.
- Обратная конвертация строки LaTeX в визуальную формулу.
- Наличие библиотеки математических элементов с возможностью ее расширения.

Функционал анализа формул

- Ввод и импортирование формул в графическом формате и строковом формате LaTeX.
- Индексирование и обработка записей формул в строковом формате LaTeX.
- Отображение процента совпадений и выделение совпадающих частей формул.
- Работа с базой данных формул в строковом формате LaTeX.

Реализация интерфейса

- Удобство для пользователя.
- Простота записи и редактирования формул.

Код и документация

- Оптимизация кода.
 - Комментарии в исходниках.
 - Соответствие документации кода требованиям проекта.
-

Стек технологий

Frontend

- **HTML, CSS, JavaScript**: Для создания пользовательского интерфейса.

Backend

- **Flask**: Для реализации серверной части приложения и обработки запросов.
- **scikit-learn**: Для реализации анализа формул, включая вычисление схожести и построение моделей.
- **all-MiniLM-L6-v2**: Модель для обработки текстовых данных, включая сравнение математических формул.

Хранение данных

- **PostgreSQL**: Для хранения формул и связанных данных в реляционной базе данных.
 - **JSON**: Для обмена данными между фронтендом и бэкендом, а также для хранения промежуточных данных.
-

Диаграммы архитектуры

Архитектура системы

- **Frontend**: Пользователь взаимодействует с редактором формул через интерфейс, созданный с использованием HTML, CSS и JavaScript.
- **Backend**: Flask обеспечивает связь между фронтендом и обработкой данных.
- **База данных**: PostgreSQL хранит формулы в строковом формате LaTeX и данные о пользователях.

- Модель анализа: scikit-learn и all-MiniLM-L6-v2 используются для вычисления схожести формул.

Схема базы данных

- Таблица "Formulas": содержит идентификатор, строку формата LaTeX, дату создания.

Сценарии использования (Use Cases)

Создание и редактирование формулы

1. Пользователь открывает редактор формул.
2. Выбирает нужные элементы: дроби, интегралы, степени.
3. Редактирует формулу через визуальный интерфейс.
4. Сохраняет формулу в базе данных.

Анализ формул на совпадение

1. Пользователь импортирует формулы через интерфейс (LaTeX).
2. Запускается анализ: формулы индексируются и сравниваются с базой данных.
3. Выводится процент совпадений с подсветкой одинаковых частей.

Примеры формул

- Пример 1: $a^2 + b^2 = c^2$ (LaTeX: `a^2 + b^2 = c^2`).
- Пример 2: $\int_a^b f(x) dx$ (LaTeX: `\int_a^b f(x) dx`).
- Пример 3: $\sqrt{a^2 + b^2}$ (LaTeX: `\sqrt{a^2 + b^2}`).

Заключение

Данное решение должно предоставить удобный и функциональный инструмент для работы с математическими формулами в web-среде. Реализация учитывает как удобство пользователя, так и технические аспекты для разработчиков, включая гибкость расширения библиотеки блоков и поддержку формата LaTeX.