

Условие задачи:

Новый классификатор продукции на основе стандартных товарных единиц (СТЕ) на Портале поставщиков помогает заказчикам просматривать все предложения поставщиков - позиции электронного каталога СТЕ, которые удовлетворяют критериям поиска, а также сравнивать характеристики продукции и цены по конкретному типу продукции.

Участникам необходимо разработать систему, которая оповещает поставщика о возможных ошибках при заведении новых СТЕ, а модератор информирует о конкретных выявленных несоответствиях.

Раздел «Создать ofertу и СТЕ» для поставщика:

- Уведомление о возможной ошибке в СТЕ: текст, картинки, категории, характеристики (см. типовые ошибки из статусов обработки).

Раздел модератора:

- Уведомление о возможной ошибке: текст, картинки, категории, характеристики (название, единица измерения, значение, формат);
- Вероятность итоговой некорректности для ранжирования и распределения заявок между модераторами.

Система поддержки и принятия решений:

- Оповещает пользователя о возможных ошибках при заведении новых СТЕ;
- Для модераторов выводит информацию о вероятных ошибках и позволяет ранжировать список новых СТЕ по этому показателю.

Важно разработать систему, которая оповещает поставщика и модератора о возможных ошибках при заведении новых СТЕ. Организаторы не накладывают требований на создание интерфейса. На этапе защиты необходимо будет продемонстрировать работоспособность системы.

Описание бизнес-процесса: Модератор в ручном режиме производит проверку новых СТЕ.

Для успешной подачи заявки на СТЕ необходимо выполнить следующие условия:

- изображение должно полностью соответствовать текстовому описанию;
- изображение товара/работы/услуги должно полностью соответствовать требованиям портала (без использования логотипов, оттисков печатей, водяных знаков, QR-кода, ссылок на веб-ресурсы, а также иных документов);
- новые СТЕ не должны дублировать те, которые уже размещены на портале;

- проверять характеристики на полноту и точность данных;
- наличие стоп-слов (опечатки, неприемлемый контент, недопустимая продукция и т.д.).

Бизнес-метрика: уменьшение времени между подачей заявки и ее опубликованием на портале.

Данные:

sku_images - изображения СТЕ, где название файла соответствует полю “Изображение” в таблице SkuChangeRequests_200k_2.xlsx

SkuChangeRequests.csv - таблица с заявками и информацией по ним.

- Номер заявки;
- Форма заявки на СТЕ полная и упрощенная (характеристики заполняет модератор) - обязательно указывается наименование, необязательно модель, ссылка на сайт производителя, описание;
- Наименование;
- Единица измерения;
- Изображение;
- Классификация ГОСТ/ТУ;
- Описание - текстовое поле с описанием товара;
- Статус;
- Причина отказа;
- Комментарий - От модератора, в случае отказа или возврата на доработку;
- Ид оферты;
- Наименование оферты;
- Артикул оферты - Уникальный, произвольный идентификатор конкретного поставщика;
- Регион поставки;
- Срок поставки в днях «от»;
- Срок поставки в днях «до»;
- Доступное количество «от»;
- Доступное количество «до»;
- Исходные характеристики - набор качественных и количественных характеристик со значениями;
- Категория оферты;
- Категория справочника;
- Вид продукции;
- Количество эталонных утвержденных характеристик в категории;
- Количество использованных поставщиком эталонных утвержденных характеристик.

Требования к решениям:

1. Реальный кейс в процессе защиты. Кейс содержит ошибку в изображении и тексте. Участники получают файл с данными и изображение

Наименование *

Вид продукции *

Единица измерения *

Изображения

Классификация ГОСТ/ТУ

Описание

Похожие STE

Соответствий не найдено. Попробуйте изменить наименование

Перенесите файлы сюда или нажмите мышкой...

Поиск...

ХАРАКТЕРИСТИКИ

Наименование	Значение	Ед. измерения
--------------	----------	---------------

Проверить и отправить заявку на STE

Рис. 1 - Интерфейс, в котором создаётся STE

2. Код выложен в публичный репозиторий на GitHub. В README.md описана структура проекта и методология. Презентация отдельным файлом в формате .pdf или Power Point.

Этапы защиты:

Этап 1: Защита проектов (презентация до 5 мин) + реальный кейс параллельно демонстрирует второй участник команды. Максимальное кол-во баллов - 100.

Этап 2: Защита топ 5 решений (презентация до 10 мин, 10 минут на вопросы жюри). Максимальное кол-во баллов - 100. Баллы суммируются с баллами этапа 1.

Критерии оценки:

1. Работоспособность прототипа - 20
2. Оригинальность идеи - 20
3. Масштабируемость - 20
4. Уровень понимания участников сферы тендерных закупок - 20
5. Применение передовых методов машинного и глубокого обучения (ИИ) - 20

Полезные ссылки:

- <https://colab.research.google.com> – Среда разработки от Google
- <https://scikit-learn.org/> – Python-библиотека для машинного обучения
- https://radimrehurek.com/gensim_3.8.3/index.html – Python-библиотека для работы с текстом и создания текстовых моделей
- <https://fasttext.cc/> – Python-библиотека для создания текстовых моделей
- <https://docs.python.org/3.8/> – Документация Python
- <https://www.anaconda.com/> – Дистрибутив языков Python и R, включающий библиотеки для машинного обучения
- <https://jupyter.org/> – Среда разработки для анализа данных