# Введение

**Заказчик**: UniGoTo

**Команда**: Большой Ух

**Цель**:

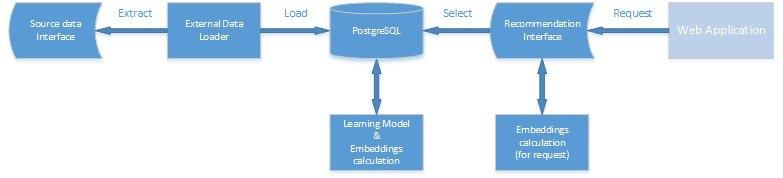
Реализация рекомендательной системы для выдачи персонального ранжированного списка направлений обучения в высших учебных заведениях на основе данных, вводимых пользователем рекомендательной системы.

**Задачи**:

* организовать сервер для разработки, тестирования и презентации решения;
* реализовать и запустить загрузчик данных;
* выбрать и реализовать модель для рекомендательной системы;
* спроектировать и реализовать интерфейс для взаимодействия с рекомендательной системой;
* реализовать web-приложение для взаимодействия пользователя с рекомендательной системой.

# Описание решения

## Концепция IT-решения



Решение состоит из:

* загрузчика данных внешнего интерфейса в базу данных;
* выгрузки из базы данных отфильтрованных данных для обучения модели и формирования эмбеддингов по тексту отфильтрованных записей в базе данных;
* интеграционного интерфейса для получения рекомендаций на основе входных данных, который формирует вектор по тексту запроса и получает из базы список похожих записей в исходных данных и направлений с ними ассоциированных;
* web-приложения для работы пользователя с рекомендательной системой.

## План реализации

1. Загрузка данных
2. Анализ данных
3. Подготовка прототипа модели
4. Дизайн архитектуры сервиса
5. Имплементация сервиса
6. Покрытие тестами
7. Подготовка документации

## Технологии, Инструменты, Алгоритмы

* **AirFlow:** для оркестрации операций обработки данных.
* **PyTorch + SentenceTransformers:** для преобразования текстовых описания абитуриентов в векторы признаков.
* **Docker Compose:** для организации развертывания и взаимодействия составных частей рекомендательной системы.
* **PostgreSQL:** для хранения исходных данных и поиска ближайших векторов (kNN).
* **FastAPI:** интеграционный интерфейс рекомендательной системы.
* **Streamlit:** web-приложение для работы пользователя с рекомендательной системой.

# Практическая ценность и применимость

Значительная часть выпускников школ испытывает затруднения в выборе образовательной траектории для приобретения необходимых будущих профессиональных навыков; многие из них согласятся со взглядом родителей на выбор той или иной профессии. Такой выбор зачастую не удовлетворяет внутренним запросам будущего студента: разность интересов с сокурсниками, направленность на развитие в рамках научного сообщества или прикладные знания, выражения себя через спорт, искусство, изобретательство и т.д. К тому же, в настоящее время постоянно появляются новые, перспективные направления, которые еще не успели получить широкую известность и могут быть упущены. Наш сервис призван помочь абитуриенту в выборе направления обучения, которому он посвятит следующие несколько лет. Сердцем нашего сервиса является рекомендательная система, которая и поможет определиться с будущей профессией посредством формирования рекомендаций выбора вуза и факультета максимально близко к интересам будущего студента для комфортного обучения в группе единомышленников и с учетом доступности учебных программ в вузах домашнего региона, чтобы позволить сохранить возможность тесного общения с друзьями и родственниками.

# Команда и план действий

## Команда

* Туомас Эдвард Витальевич – Software Developer
* Анисимова Татьяна Александровна – ML Engineer
* Голубев Артем Алексеевич – Business Analyst
* Гуков Алексей Александрович – MLOps Engineer
* Колотий Вячеслав Витальевич - PM
* Ротерман Виктор Андреевич – Test Automation Engineer

## Перечень и распределение работ

* Sprint 1: 28.11.2023 - 1.12.2023
* Sprint 2: 2.12.2023 - 6.12.2023
* Sprint 3: 7.12.2023 - 15.12.2023

| **Item** | **mh** | **Sprint** | **Assignee** |
| --- | --- | --- | --- |
| Создание сервера и доступ для команды | 2 | 1 | Вячеслав |
| Имплементация загрузчика данных | 2 | 1 | Алексей |
| Анализ данных и определение политики очистки данных | 2 | 1 | Артем |
| Анализ справочника вузов и его соответствия внешним данным. | 5 | 2 | Виктор |
| Проверка модели на эмбеддингах (Google Colab) | 10 | 2 | Татьяна |
| **Сервис географических координат** |  |  |  |
| Исследовать и выбрать сервис получения географических координат по адресу | 4 | 2 | Эдвард |
| Имплементация класса для получения координат городов. | 4 | 3 | Эдвард |
| Дополнение справочника вузов их географическими координатами. | 3 | 3 | Эдвард |
| Исследование GRPC. Выбор GRPC или REST | 2 | 1 | Артем |
| Имплементация скрипта вычисления векторов по исходным данным | 4 | 2 | Алексей |
| Имплементация тестов для валидации метрик качества модели | 3 | 3 | Виктор |
| **Имплементация сервиса рекомендательной системы** |  |  |  |
| Имплементация интеграционного интерфейса | 0.5 | 2 | Алексей |
| Имплементация валидации и очистки данных полученных от входного интерфейса | 0.5 | 3 | Артем |
| Имплементация получения списка рекомендованных направлений | 0.5 | 2 | Алексей |
| Имплементация поиска географически близких вузов, имеющих полученные направления | 1 | 3 | Эдвард |
| Имплементация тестов для сервиса рекомендательной системы | 4 | 3 | Виктор |
| Оформление составных частей решения в виде докер-контейнеров. Написание и отладка конфигурации docker compose. | 2 | 2 | Алексей |
| Настройка пайплайна (загрузка, вычисление векторов для новых данных, тестирование-валидация, вывод в эксплуатацию) | 2 | 2 | Алексей |
| Имплементация веб-приложения для работы пользователя с рекомендательной системой | 4 | 2 | Артем |
| Добавить линтер на репозиторий в Github | 1 | 2 | Артем |
| Подготовка документации по первому деливераблу | 8 | 2 | Вячеслав |
| Оформление документации в git | 4 | 3 | Вячеслав |
| Подготовка презентации для финального выступления и демонстрации | 3 | 3 | Вячеслав |

## 

## Риски

| **№** | **Риск** | **Опция на случай реализации риска** |
| --- | --- | --- |
| 1 | Дефицит времени или сложность реализации GRPC-интерфейса | Реализовать REST-интерфейс с использованием FastAPI |
| 2 | Web-приложение для работы пользователя с рекомендательной системой | Отказаться от имплементации веб-приложения. Оставить только интеграционный интерфейс. |
| 3 | Недоступность API внешних данных, невалидные данные | Настойчиво просить менторов связаться с бизнес-заказчиком |

# Заключение

**Практическая ценность:**

* контейнеризированное решение, готовое для установки и использования;
* автоматическое пополнение базы новыми данными;
* использование самостоятельно реализованного геокодера для поиска по выбранному региону и учета географической близости пользователя и вузов;
* хранение исходных данных для возможного переиспользования новыми моделями в случае расширения ансамбля моделей или изменения подхода формирования рекомендаций;
* рекомендации на основе саморепрезентации пользователя.

**Планируемые улучшения:**

* увеличение количества моделей до 3-х и объединение их в ансамбль для получения более точных результатов;
* использование специализированных БД для хранения векторных данных Milvus (<https://github.com/milvus-io/milvus/>);
* подготовка набора данных и обучение supervised deep learning модели, а именно GNN-модели для учета гетерогенной саморепрезентации пользователя системы.