**Задание 2.1. ВСР Создать раздаточный материал для защиты выпускной квалификационной работы**

**Тема:**  
Использование машинного обучения для анализа эффективности цифровизации учебного процесса

**1. Краткое описание проекта**

Проект посвящён анализу анкетных данных студентов с помощью методов машинного обучения для оценки эффективности цифровизации учебного процесса. В работе реализован комплексный подход, включающий кластеризацию, классификацию и генерацию персонализированных рекомендаций по профилю цифровой активности.

**2. Архитектура программного решения**

* Сбор и хранение данных: Анкетные данные студентов (цифровые навыки, использование онлайн-ресурсов, вовлечённость).
* Модуль предобработки: Очистка, нормализация, кодирование признаков.
* Кластеризация:
  + Методы: UMAP (снижение размерности), Fuzzy C-Means, агломеративная кластеризация, метрика Говера.
  + Цель — выделить однородные группы (профили) студентов по цифровому поведению.
* Профилирование:
  + Автоматический анализ ключевых характеристик и визуализация различий между профилями (radar chart, barplot, таблицы).
* Классификация:
  + Обученная модель (случайный лес) автоматически определяет цифровой профиль нового студента.
* Генерация рекомендаций:
  + Интеграция с GigaChat API: на основании профиля система автоматически формирует персонализированную текстовую рекомендацию для каждого студента.
* Визуализация:
  + Scatter plot кластеров, radar chart профилей, barplot различий, скриншоты интерфейса.
* Пользовательский интерфейс:
  + Управление анализом, просмотр профилей, генерация и получение рекомендаций.

**3. Используемые алгоритмы**

* Кластеризация: UMAP, Fuzzy C-Means, агломеративная кластеризация (на матрице Говера).
* Классификация: случайный лес, решающее дерево, нейронная сеть (MLP).
* Оценка качества: Silhouette Score, Davies-Bouldin Index, F1-score.
* Генерация текста: языковая модель GigaChat (ИИ).
* Вспомогательные библиотеки: Pandas, scikit-learn, Matplotlib.

**4. Иллюстрации (скриншоты и схемы — по описанию)**

* Диаграмма рассеяния (scatter plot UMAP): График распределения студентов по кластерам.
* Таблица профилей: Краткая характеристика и численность студентов в каждом профиле.
* Radar chart: Визуальное сравнение выраженности цифровых навыков по профилям.
* Barplot различий: Топовые признаки, отличающие кластеры.
* Скриншоты интерфейса: Окно кластеризации, окно персональных рекомендаций.
* Блок-схема архитектуры: Последовательность: данные → предобработка → кластеризация → профилирование/классификация → генерация рекомендаций → вывод пользователю.

**5. Основные результаты**

* Обнаружено 3 основных профиля студентов по цифровой активности:
  1. Цифровые энтузиасты — высокая вовлечённость, использование современных технологий, самостоятельное обучение онлайн.
  2. Умеренно вовлечённые — используют цифровые инструменты только по необходимости, сочетают офлайн и онлайн.
  3. Традиционалисты — преимущественно классические подходы, низкая вовлечённость в цифровую среду.
* Классификация новых студентов по профилю осуществляется автоматически и с высокой точностью.
* Генерация персональных рекомендаций на базе ИИ — для каждого профиля предлагаются шаги по развитию цифровых навыков и адаптации к современному образовательному процессу.
* Практическая польза:
  1. Возможность мониторинга цифровизации на уровне факультета.
  2. Индивидуальная поддержка студентов по результатам анализа.
  3. Автоматизация аналитики и обратной связи.

**Контакт для обратной связи и вопросов:**  
Воложанин Владислав Олегович, студент 4 курса РГПУ им. А. И. Герцена