**Задание к итоговому проекту**

**Задачи**

**1. Задачи регрессии (оценка целевого признака из области действительных чисел). Построить не менее 3 моделей (линейная, дерево и любая другая на выбор).**

**2. Задачи классификации и кластеризации. С использованием одного и того же набора данных решить задачу классификации и кластеризации (при кластеризации удалить целевой признак – класс).**

**Для задачи классификации построить не менее 3 моделей (логистическая регрессия, дерево и любая другая на выбор). Для задачи кластеризации использовать методы К-средних и DBSAN.**

0. Выбрать задачу.

1. Подобрать и скачать набор данных.

2. Подписать блокнот, описать используемый набор данных, указать ссылку на источник.

3. Сформулировать решаемую задачу, выбрать целевую переменную для задачи классификации или регрессии.

4. Загрузить данные в датафрейм.

5. Вывести статистическую информацию о наборе данных.

6. Вывести названия столбцов и строк.

7. Заменить или удалить пропущенные значения.

8. Заменить категориальные данные количественными.

9. Выбрать важные признаки с помощью корреляционного анализа (оценить их влияние на значение целевого признака в задачах регрессии и классификации).

10. Выполнить любые другие действия по конструированию признаков на ваше усмотрение.

11. Визуализировать данные с помощью Matplotlib и Seaborn.

12. Выбрать и обосновать выбор метрик.

13. Построить модели и обучить их. Подбор значений гиперпараметров можно выполнить вручную либо с использованием библиотек Optuna, GridSearch и т.д.

14. Рассчитать значения метрик.

15. Проанализировать полученные результаты обучения, выбрав наилучшую модель. Обосновать выбор.

16. Оформить отчёт.