Отчет по лабораторной работе 3 (Вариант 4)

Задание1

Цель работы:

Построение диаграммы рассеяния для визуализации взаимосвязи признаков mean radius и mean texture из датасета breast\_cancer с цветовой маркировкой классов.

Задачи:

1. Загрузить датасет breast\_cancer из библиотеки sklearn.

2. Выбрать столбцы mean radius (X) и mean texture (Y).

3. Визуализировать данные с цветовым разделением по классам (target).

Инструменты и алгоритмы:

Язык: Python

Библиотеки:

matplotlib.pyplot для визуализации,

sklearn.datasets для загрузки данных.

Алгоритм:

1. Загрузка данных → load\_breast\_cancer().

2. Выбор столбцов 0 (mean radius) и 1 (mean texture).

3. Построение scatter plot с параметрами:

c=target (цвета классов),

cmap='coolwarm' (палитра),

s=8 (размер точек).

Ошибки и исправления:

Проблема: неверное обращение к столбцам по индексам.

Решение: сверка с feature\_names для точного выбора (mean radius — 0, mean texture — 1).

Проблема: слишком большие точки на графике (перекрывают друг друга).

Решение: подбор размера (s=8).

Выводы

Построена диаграмма рассеяния, показывающая разделение классов (0/1) по признакам mean radius и mean texture. Визуально подтверждена корреляция: чем больше radius и texture, тем выше вероятность класса 0 (злокачественные, на графике обозначенны синим цветом).

Задание2

Проблемы с загрузкой Grunfeld, предоставлю решение позже