Wejściówka lab1-lab2 28.10.2022 gr. IO 5.13

- 1. Wczytaj dane z pliku surgical.csv. Zbiór powinien zawierać 54 wiersze i 9 kolumn. Zmienna zależna oznaczona jest jako "y" i oznacza ona czas przeżycia pacjenta po operacji wątroby. Znajdź kolumnę o największym współczynniku zmienności (współczynnik zmienności=średnia/odchylenie standardowe). Wyświetl nazwę tej kolumny.
- 2. Stwórz dwa wykresy na jednym obrazie za pomocą fig, ax=plt.subplots(...). Na pierwszym rysunku ma znaleźć się histogram (hist) zmiennej y, a na drugim rysunku wykres boxplot tej zmiennej.
- 3. Podziel zbiór na testowy i uczący za pomocą X_train, X_test, y_train, y_test=train_test_split(X,y, test_size=0.2, random_state=2022, shuffle=False) Wytrenuj na zbiorze uczącym model regresji liniowej. Zapisz jego współczynniki (wagi) w zmiennej weights. Sprawdź, która zmienna ma najsilniejszy wpływ na y, wyświetl jej nazwę i wagę w modelu.
- 4. Dokonaj predykcji z modelu na zbiorze testowym. Oblicz i wyświetl wartości trzech znanych Ci miar jakości pracy modelu.

Potrzebne importy:

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

from sklearn.linear_model import LinearRegression

from sklearn.model_selection import train_test_split

from sklearn.metrics import mean_absolute_percentage_error, mean_squared_error, mean_absolute_error