

# WSI Lab4

Jakub Śledź

WSI - ćwiczenie 4. Regresja i klasyfikacja

## Dane:

Dane są wczytywane, dyskretyzują wejście poprzez konwersję wieku pacjenta na lata a następnie kategoryzują wartości ciągłe przy użyciu listy progów

```
age_bins = [40, 50, 60]           # kategorie wieku
height_bins = [155, 165, 175]     # kategorie wzrostu
weight_bins = [60, 80, 100]       # kategorie wagi
ap_hi_bins = [120, 140, 160]      # ciśnienie skurczowe
ap_lo_bins = [80, 90, 100]        # ciśnienie rozkurczowe
```

tworząc dla listy  $[b_1, b_2, b_3, \dots]$  kategorie  $(-\infty, b_1], (b_1, b_2], \dots, (b_N, \infty)$

Dane dzielone są na zbiór uczący na podstawie którego tworzy się drzewa o wielkości 0.6 danych wejściowych, zbiór walidacyjny służący ustaleniu optymalnej głębokości drzewa (0.2 wielkości wejścia) i zbiór testowy (0.2 wejścia).

## Wyniki:

Testowałem głębokości drzew z zakresu  $<1;20$ ). Najczęściej optymalna głębokość drzewa to 3, czasem przez losowość danych pojawiała się wartość 4. Trafność działania algorytmu na zbiorach walidacyjnym i testowym powtarzalnie wynosi 0.72.