

Projekt: Rozpoznawanie chorób serca

S. Hoa Nguyen

1 Źródło danych

<https://archive.ics.uci.edu/ml/datasets/heart+Disease>

2 Opis danych

Dane są zbierane przez cztery szpitale kardiologii w Węgrach, Szwajcarii i Stanach Zjednoczonych. Opisują one objawy chorobowe i profile pacjentów. Każdy pacjent jest opisywany 76 atrybutami (wśród których jest 14 istotnych, popatrz źródło danych). Jest 5 grup chorób (0 - zdrowe, 1,2,3,4 - kolejne stadium choroby). Baza danych zawiera ponad 300 przykładów.

3 Cel badania

Tworzyć *klasyfikator binarny* do identyfikacji czy pacjent jest zdrowy (etykieta 0) czy chory (etykiety 1,2,3 lub 4).

4 Model testowania

- Zbiór uczący: 80% wylosowanych przykładów.
- Zbiór testowy: 20% reszta.
- Miara oceny klasyfikatora: Trafność klasyfikacji na zbiorze testowym (procent przykładów testowych dobrze sklasyfikowanych).

5 Metoda i implementacja

- a) Zaprojektować jednokierunkową sieć neuronową do klasyfikacji pacjentów.
- b) Przetestować parametry uczenia sieci (liczba warstw, liczba neuronów, współczynnik uczenia).
- c) Wybierać optymalną sieć.
- d) Podać trafność klasyfikatora na zbiorze testowym.