

Należy zaimplementować naiwny klasyfikator Bayes'a. Proszę przetestować klasyfikator na jednym ze zbiorów danych podanych poniżej:

- <https://www.kaggle.com/competitions/titanic> Proszę przewidzieć czy pasażer przeżył katastrofę na podstawie 5 wybranych atrybutów (różnych od id). Atrybuty rzeczywiste można skategoryzować, np. atrybut wiek można zamienić na atrybut, którego wartości to {dziecko, dorosły, osoba starsza}. Jaka jest jakość stworzonego klasyfikatora? Czy jest niedouczenie, albo przeuczenie? Jak to zmierzyć skoro zbiór testowy nie jest otagowany? Proszę obliczyć jakie prawdopodobieństwo przeżycia miała osoba pod warunkiem, że:
  - a) Była mężczyzną.
  - b) Była kobietą.
  - c) Była chłopcem.
  - d) Była dziewczynką.Klasyfikator niekoniecznie musi zwracać prawdopodobieństwo warunkowe! Proszę zinterpretować uzyskane wyniki.
- Snake. Do uzyskania maksymalnej liczby punktów musi on uzyskiwać średnio co najmniej 20 punktów, wszystkie chwytty są dozwolone.

Raport oraz pliki proszę spakować do pliku o nazwie WSI-7-NAZWISKO-IMIE.zip i przesać na adres grzegorz.rypec.dokt@pw.edu.pl.

Data oddania tego ćwiczenia to 11.06. Należy do tego dnia przesać do mnie brakujące pliki z zaległych ćwiczeń. Będą to ostatnie zajęcia.