

Ćwiczenie DD

Drzewa decyzyjne

Część teoretyczna

Wykład na temat drzew decyzyjnych.

Zadania pomocnicze

Ćwiczenie wykorzystuje algorytmy drzew decyzyjnych opisane w rozdz. 8.3 książki G. James, D. Witten, T. Hastie, R. Tibshirani, J. Taylor: An Introduction to Statistical Learning with Applications in Python (<https://www.statlearning.com/>). Zapoznaj się z tym materiałem.

Zadania do wykonania

Według Światowej Organizacji Zdrowia (WHO) udar jest drugą najczęstszą przyczyną zgonów na świecie i odpowiada za około 11% wszystkich zgonów.

Zadanie polega na budowie klasyfikatora opartego na drzewach decyzyjnych do przewidywania, czy pacjent może doznać udaru na podstawie parametrów wejściowych, takich jak płeć, wiek, schorzenia i przebyte choroby oraz palenie tytoniu. Każdy wiersz danych zawiera istotne informacje o pacjencie.

Źródło danych <https://www.kaggle.com/datasets/fedesoriano/stroke-prediction-dataset>

...