

Metody klasyfikacji w R - pRojekt

Zadanie

Podpowiedź do zadania - do danych z pakietu `mlbench` można dostać się w następujący sposób:

```
library(mlbench)
data("PimaIndiansDiabetes")
```

Na podstawie danych `PimaIndiansDiabetes` z pakietu `mlbench`:

- a) Wczytaj dane oraz podziel zbiór na uczący i testowy w proporcji 75% do 25%.
- b) Zbuduj dwa modele prognozujące zmienną `diabetes`:
 - drzewo klasyfikacyjne z wszystkimi zmiennymi objaśniającymi przy domyślnych ustawieniach,
 - las losowy z wszystkimi zmiennymi objaśniającymi (ustaw liczbę wykorzystanych drzew na 200).
- c) Sprawdź która ze zmiennych objaśniających ma największy wpływ na zmienną prognozowaną?
- d) Na podstawie zbioru testowego, policz `Accuracy`, `Recall` oraz `False Negative Rate (FNR)`. Oceń, który model lepiej prognozuje zmienną `diabetes` ze względu na `FNR`?
- e) Dla lepszego modelu policz `AUC` oraz narysuj `ROC`, a także krzywą `lift`.