

# Metoda MARS (Multivariate Adaptive Regression Splines)

## Zadania:

1. Przetestuj działanie metody MARS (Multivariate Adaptive Regression Splines) na danych symulacyjnych. Implementacja metody MARS znajduje się w pakiecie `earth`. Rozważamy następujące schematy symulacji:
  - (a) Przykład 1:
    - Liczność próbki  $n = 1000$ , liczba zmiennych  $p = 50$ ,
    - $X_1 \sim U(0, 4)$ ,  $X_2, \dots, X_{50} \sim N(0, 1)$ ,  $\epsilon \sim N(0, 0.1)$ ,
    - Zmienna odpowiedzi  $Y = \sqrt{X_1} + \epsilon$ .
  - (b) Przykład 2:
    - Liczność próbki  $n = 1000$ , liczba zmiennych  $p = 50$ ,
    - $X_1 \sim U(0, 4)$ ,  $X_2, \dots, X_{50} \sim N(0, 1)$ ,  $\epsilon \sim N(0, 0.1)$ ,
    - Zmienna odpowiedzi  $Y = X_1^2 + \epsilon$ .
  - (c) Przykład 3:
    - Liczność próbki  $n = 1000$ , liczba zmiennych  $p = 50$ ,
    - $X_1, \dots, X_{50} \sim N(0, 1)$ ,  $\epsilon \sim N(0, 0.1)$ ,
    - Zmienna odpowiedzi  $Y = (X_1 - 0)_+ + (X_1 - 1)_+ + \epsilon$ .
  - (d) Przykład 4:
    - Liczność próbki  $n = 1000$ , liczba zmiennych  $p = 50$ ,
    - $X_1 \sim U(0, 4)$ ,  $X_2, \dots, X_{50} \sim N(0, 1)$ ,  $\epsilon \sim N(0, 0.1)$ ,
    - Zmienna odpowiedzi  $Y = \sin(X_1) + \epsilon$ .
  - (e) Przykład 5:
    - Liczność próbki  $n = 1000$ , liczba zmiennych  $p = 50$ ,
    - $X_1, \dots, X_{50} \sim N(0, 1)$ ,  $\epsilon \sim N(0, 0.1)$ ,
    - Zmienna odpowiedzi  $Y = I(X_1 < 0)$ .
2. Wybierz dowolny zbiór dotyczący regresji z repozytorium UCI <http://archive.ics.uci.edu/ml/index.php>. Porównaj działanie 3 metod: MARS, regresja liniowa oraz drzewo regresyjne stosując krosvalidację oraz obliczając błąd predykcji (RMSE).