

# Michał Broś, Radosław Barszczak

## Zadanie 5: programowanie genetyczne i regresja symboliczna

Termin realizacji: 12 maja 2025

Zadanie do oddania przez MS Teams. Do oddania: kod oraz krótkie sprawozdanie w PDF (można na przykład przy użyciu `quarto render notebook.ipynb --to pdf`).

### Na 3.0

Do realizacji:

1. Zmodyfikuj przykład `pysr_demo.ipynb` tak, aby uczył się funkcji  $f(x) = 2.2 \sin(x_0 + 2x_1) - x_5^2 - 3$  której dziedziną jest  $R^6$ . Uczenie ma się odbywać w oparciu o 200 wylosowanych z dziedziny próbek (między -5 a 5).
2. Zanotuj wzory trzech rozwiązań o najwyższej wartości `score` oraz rozwiązanie `best` dla następujących zestawów ustawień:
  - a. `binary_operators=["+", "*"], unary_operators=["cos", "exp", "sin"], maxsize=20,`
  - b. `binary_operators=["+", "*", "-", "^"], unary_operators=["cos", "exp", "sin", "log"], maxsize=30,` (dodaj ograniczenie dla argumentów operatora "^").
  - c. `binary_operators=["+", "*", "-", "^"], unary_operators=["exp", "sin"], maxsize=15.`
3. Powtórz eksperymenty z zadania na 3.0 po dodaniu szumu do próbek z funkcji  $f$  (rozkład normalny o średniej 0 i odchyleniu standardowym 0.5)

### Na 4.0

Do realizacji:

1. Punkty z zadania na 3.0.
2. Dodaj do porównania dopasowanie oparte o próbki losowane w szerszym zakresie (między -15 a 15) oraz wyższy poziom szumu (odchylenie standardowe równe 2 oraz 5).

### Na 5.0

Do realizacji:

1. Punkty z zadania na 4.0.

2. Zamień funkcję  $f$  na  $f(x) = 2.2 \sin(x_0 + 2x_1) - x_5^2 - p(l, x_0)$  gdzie  $p(i)$  oznacza  $i$ -tą liczbę pierwszą. Uwzględnij  $p$  jako dodatkowy operator unarny analogicznie do przykładu "Julia packages and types" z notatnika `pysr_demo.ipynb`. Powtórz eksperymenty opisane w zadaniach na 3.0 i 4.0.

## Rozwiązania na 3.0

### Ad. 2.1

Trzy najlepsze rozwiązania (według score):

1.  $x_5 * (x_5 * -1.1931057)$  (score: 2.0312, loss: 6.758767)
2.  $((x_5 * x_5) * -0.9815544) + -3.1631014$  (score: 0.5600, loss: 2.205166)
3.  $((x_1 * (\cos(x_0) * (\cos(x_1 * 2.3874586) * 0.52057564))) + (x_5 * x_5)) + 3.1727533) * -0.990387$  (score: 0.1085, loss: 1.715046)

Najlepsze rozwiązanie (best):

1.  $x_5 x_5 (-0.9815544) - 3.1631014$

### Ad. 2.2

Trzy najlepsze rozwiązania (według score):

1.  $-3.0127609 - (x_5 * x_5)$  (score: 3.1427, loss: 2.224043)
2.  $(-3.055794 - (x_5 * x_5)) + \sin((x_0 + x_1) + x_1)$  (score: 1.1397, loss: 0.664082)
3.  $\cos(x_5) - 10.987488$  (score: 0.0721, loss: 51.521240)

Najlepsze rozwiązanie (best):

1.  $-x_5 * x_5 + \sin(x_0 + x_1 + x_1) - 3.055794$

### Ad. 2.3

Trzy najlepsze rozwiązania (według score):

1.  $-3.0128005 - (x_5 * x_5)$  (score: 3.2011, loss: 2.224043)
2.  $-3.0069795 - ((x_5 * x_5) - \sin(x_0 + (x_1 + x_1)))$  (score: 1.1433, loss: 0.661699)
3.  $(-3.0832508 - (x_5 * x_5)) - \sin(x_1 + x_1)$  (score: 0.0350, loss: 2.118636)

Najlepsze rozwiązanie (best):

1.  $-(x_5 * x_5 - \sin(x_0 + x_1 + x_1)) - 3.0069795$

## Z szumem

### Ad. 3.1

Trzy najlepsze rozwiązania (według score):

1.  $x_5 * (x_5 * -1.1879785)$  (score: 2.0183, loss: 6.855120)
2.  $((x_5 * -0.9798253) * x_5) + -3.1116424$  (score: 0.5148, loss: 2.448527)
3.  $((\sin(x_1) * \sin(x_0 * \sin(x_1))) * -0.54738796) + ((x_5 * (x_5 * -0.9749486)) + -2.9543612)$  (score: 0.0753, loss: 2.152109)

Najlepsze rozwiązanie (best):

1.  $x_5(-0.9798253)x_5 - 3.1116424$

## Ad. 3.2

Trzy najlepsze rozwiązania (według score):

1.  $-2.9472342 - (x_5 * x_5)$  (score: 3.0386, loss: 2.471086)
2.  $\cos(x_5) + -10.922029$  (score: 0.0692, loss: 51.586185)
3.  $((x_5 * ((x_5 + 0.038573753) * -0.98144275)) - 2.5140498) + \cos(\sin(-0.27121347 - ((x_1 + ((x_5 * -0.27121347) * 0.005949329)) * -2.0556355)) - 3.843657)$  (score: 0.0206, loss: 2.251434)

Najlepsze rozwiązanie (best):

1.  $-x_5 * x_5 - 2.9472342$

## Ad. 3.3

Trzy najlepsze rozwiązania (według score):

1.  $-2.947309 - (x_5 * x_5)$  (score: 3.0967, loss: 2.471086)
2.  $(x_5 * ((x_5 * -0.9783928) + (x_4 * 0.036650464))) - 3.130253$  (score: 0.0149, loss: 2.362945)
3.  $-11.091976 - \sin(x_0)$  (score: 0.0111, loss: 54.672173)

Najlepsze rozwiązanie (best):

1.  $-x_5 * x_5 - 2.947309$

# Rozwiązania na 4.0

## Ad. 2.1

Trzy najlepsze rozwiązania (według score):

1.  $x_5 * (x_5 * -1.0195214)$  (score: 3.0744, loss: 9.260938)
2.  $(x_5 * (x_5 * -0.99853265)) + -2.8212833$  (score: 0.2484, loss: 5.635047)
3.  $((x_5 * x_5) * -0.99825096) + (\sin(x_5 * -6.664151) + -2.9023483)$  (score: 0.0427, loss: 5.344251)

Najlepsze rozwiązanie (best):

1.  $x_5x_5(-0.99853265) - 2.8212833$

## Ad. 2.2

Trzy najlepsze rozwiązania (według score):

1.  $-3.0794163 - (x5 * x5)$  (score: 3.3044, loss: 6.661105)
2.  $\sin(\exp(x1 + x5) * 15.111193) - ((x5 * x5) - -3.177719)$  (score: 0.0233, loss: 6.174113)
3.  $(\sin(x2 + 1.0673308) - (x5 * x5)) - 3.1319969$  (score: 0.0144, loss: 6.461315) Najlepsze rozwiązanie (best):
4.  $-x5*x5 - 3.0794163$

## Ad. 2.3

Trzy najlepsze rozwiązania (według score):

1.  $-3.1119175 - (x5 * x5)$  (score: 3.2863, loss: 6.531839)
2.  $(\sin((x1 + x0) + x1) - (x5 * x5)) - 3.10501$  (score: 0.2079, loss: 4.961884)
3.  $(\sin(\exp(x3 - -1.3543315)) - (x5 * x5)) + -3.292466$  (score: 0.0326, loss: 6.108549) Najlepsze rozwiązanie (best):
4.  $-x5*x5 - 3.1119175$

## Rozwiązania z szumem:

### Ad. 3.1

Trzy najlepsze rozwiązania (według score):

1.  $(x5 * -1.025387) * x5$  (score: 4.8739, loss: 31.651186)
2.  $(x5 * (x5 * -1.0022084)) + -3.001038$  (score: 0.0648, loss: 27.806509)
3.  $((((x5 * x5) + \cos(x0)) * -1.0036274) + -2.832869)$  (score: 0.0289, loss: 26.962204) Najlepsze rozwiązanie (best):
4.  $x5(-1.025387)x5$

### Ad. 3.2

Trzy najlepsze rozwiązania (według score):

1.  $-3.1801066 - (x5 * x5)$  (score: 2.4977, loss: 27.126000)
2.  $x5 + -69.9174$  (score: 0.0211, loss: 4007.538000)
3.  $(\cos(x3) - (x5 * x5)) - 3.2211828$  (score: 0.0207, loss: 26.358912) Najlepsze rozwiązanie (best):
4.  $-x5*x5 - 3.1801066$

### Ad. 3.3

Trzy najlepsze rozwiązania (według score):

1.  $-3.267118 - (x5 * x5)$  (score: 2.5022, loss: 25.626910)
2.  $(\sin(x2 * -4.076018) + -3.2835457) - (x5 * x5)$  (score: 0.0326, loss: 24.419416)

3.  $\sin((x_2 * -4.093544) + \sin(x_1)) + (-3.280562 - (x_5 * x_5))$  (score: 0.0147, loss: 23.912691)  
Najlepsze rozwiązanie (best):
4.  $-x_5 * x_5 - 3.267118$

## Rozwiązania na 5.0

=== Eksperyment: zakres (-5, 5), szum 0, zestaw operatorów 1 ===

Trzy najlepsze rozwiązania (według score):

1.  $p(x_0) * -1.0122799$  (score: 6.4995, loss: 104.812355)
2.  $(p(x_0) + (x_5 * x_5)) * -0.9996523$  (score: 1.3654, loss: 3.599497)
3.  $(p(x_0) + 7.303805) * -1.0047319$  (score: 0.3203, loss: 55.234028) Najlepsze rozwiązanie (best):

4.  $(x_5 x_5 + \text{sympy\_p}(x_0))(-0.9996523)$

=== Eksperyment: zakres (-5, 5), szum 0, zestaw operatorów 2 ===

Trzy najlepsze rozwiązania (według score):

1.  $-8.170675 - p(x_0)$  (score: 7.1995, loss: 63.663666)
2.  $(0.4298719 - (x_5 * x_5)) - p(x_0)$  (score: 2.8412, loss: 3.658588)
3.  $((0.4777713 - (x_5 * x_5)) - \cos((x_1 + x_1) - ((x_0 - -1.5376787) * -1.0096959))) - p(x_0)$  (score: 0.2006, loss: 1.953838) Najlepsze rozwiązanie (best):

4.  $-x_5 x_5 - \text{sympy\_p}(x_0) - \cos(x_1 + x_1 - (-1.0096959)(x_0 - 1 * (-1.5376787))) + 0.4777713$

=== Eksperyment: zakres (-5, 5), szum 0, zestaw operatorów 3 ===

Trzy najlepsze rozwiązania (według score):

1.  $-7.5788875 - p(x_0)$  (score: 7.4884, loss: 60.613644)
2.  $0.3522636 - ((x_5 * x_5) + p(x_0))$  (score: 2.9474, loss: 3.140937)
3.  $((x_5 * x_5) * -0.98932993) - p(x_0) - (x_0 * -0.17667101)$  (score: 0.0229, loss: 2.969951)  
Najlepsze rozwiązanie (best):

4.  $0.3522636 - (x_5 * x_5 + \text{sympy\_p}(x_0))$

=== Eksperyment: zakres (-5, 5), szum 0.5, zestaw operatorów 1 ===

Trzy najlepsze rozwiązania (według score):

1.  $p(x_0) * -1.0080464$  (score: 6.5105, loss: 108.903275)
2.  $((x_5 * x_5) + p(x_0)) * -0.99949193$  (score: 1.3106, loss: 3.906000)
3.  $(p(x_0) + 7.764764) * -1.0000557$  (score: 0.3533, loss: 53.720340) Najlepsze rozwiązanie (best):

$$4. (x5x5 + \text{sympy\_p}(x0))(-0.99949193)$$

=== Eksperyment: zakres (-5, 5), szum 0.5, zestaw operatorów 2 ===

Trzy najlepsze rozwiązania (według score):

1.  $-7.22119 - p(x0)$  (score: 7.5070, loss: 57.798237)
2.  $(0.32242456 - (x5 * x5)) - p(x0)$  (score: 2.6298, loss: 3.808837)
3.  $((0.023105383 - x5) * x5) - p(x0 - (2.7450666 ^ (x0 + -4.9263973)))$  (score: 0.2118, loss: 2.977677) Najlepsze rozwiązanie (best):

$$4. x5*(0.023105383 - x5) - \text{sympy\_p}(-2.7450666**(x0 - 4.9263973) + x0)$$

=== Eksperyment: zakres (-5, 5), szum 0.5, zestaw operatorów 3 ===

Trzy najlepsze rozwiązania (według score):

1.  $-7.0982428 - p(x0)$  (score: 7.6804, loss: 50.200020)
2.  $(0.3705284 - (x5 * x5)) - p(x0)$  (score: 2.6491, loss: 3.466925)
3.  $((x0 * 0.23814917) - (x5 * x5)) - p(x0)$  (score: 0.0527, loss: 3.120271) Najlepsze rozwiązanie (best):

$$4. -x5*x5 - \text{sympy\_p}(x0) + 0.3705284$$

=== Eksperyment: zakres (-15, 15), szum 2, zestaw operatorów 1 ===

Trzy najlepsze rozwiązania (według score):

1.  $(p(x0) + (x5 * x5)) * -0.998894$  (score: 3.0639, loss: 8.606525)
2.  $p(x0) * -1.0744728$  (score: 1.2139, loss: 9228.832000)
3.  $(p(x0) + 76.35268) * -0.97736406$  (score: 0.4249, loss: 3945.543200) Najlepsze rozwiązanie (best):

$$4. (x5x5 + \text{sympy\_p}(x0))(-0.998894)$$

=== Eksperyment: zakres (-15, 15), szum 2, zestaw operatorów 2 ===

Trzy najlepsze rozwiązania (według score):

1.  $1.149591 - ((x5 * x5) + p(x0))$  (score: 3.0068, loss: 10.170794)
2.  $-69.831894 - p(x0)$  (score: 1.9482, loss: 4235.213000)
3.  $0.26773232 - (p(x0 * 0.9493936) + (x5 * x5))$  (score: 0.1580, loss: 7.414894) Najlepsze rozwiązanie (best):

$$4. 1.149591 - (x5*x5 + \text{sympy\_p}(x0))$$

=== Eksperyment: zakres (-15, 15), szum 2, zestaw operatorów 3 ===

Trzy najlepsze rozwiązania (według score):

1.  $(0.538977 - p(x_0)) - (x_5 * x_5)$  (score: 3.0536, loss: 9.427520)
2.  $-76.56597 - p(x_0)$  (score: 2.0283, loss: 4368.440400)
3.  $0.10952565 - ((x_5 * x_5) + p(x_0 * 0.9568246))$  (score: 0.1262, loss: 7.324997) Najlepsze rozwiązanie (best):

$$4. -x_5 * x_5 - \text{sympy\_p}(x_0) + 0.538977$$

=== Eksperyment: zakres (-15, 15), szum 5, zestaw operatorów 1 ===

Trzy najlepsze rozwiązania (według score):

1.  $(p(x_0) + (x_5 * x_5)) * -0.9997934$  (score: 2.4460, loss: 31.277485)
2.  $p(x_0) * -1.1294073$  (score: 1.6818, loss: 8413.479000)
3.  $(p(x_0) * -1.0474375) + -67.211334$  (score: 0.3513, loss: 4166.950000) Najlepsze rozwiązanie (best):

$$4. (x_5 x_5 + \text{sympy\_p}(x_0))(-0.9997934)$$

=== Eksperyment: zakres (-15, 15), szum 5, zestaw operatorów 2 ===

Trzy najlepsze rozwiązania (według score):

1.  $-72.15584 - p(x_0)$  (score: 2.7191, loss: 4264.712000)
2.  $(0.36088267 - (x_5 * x_5)) - p(x_0)$  (score: 2.4159, loss: 32.750587)
3.  $(-0.3260557 - p(x_0 * 0.95008445)) - (x_5 * x_5)$  (score: 0.0526, loss: 30.938730) Najlepsze rozwiązanie (best):

$$4. -x_5 * x_5 - \text{sympy\_p}(x_0) + 0.36088267$$

=== Eksperyment: zakres (-15, 15), szum 5, zestaw operatorów 3 ===

Trzy najlepsze rozwiązania (według score):

1.  $(0.61306316 - p(x_0)) - (x_5 * x_5)$  (score: 2.4935, loss: 30.214243)
2.  $-78.33836 - p(x_0)$  (score: 2.3448, loss: 4541.064000)
3.  $(0.7217275 - p(x_0 - -0.03897474)) - (x_5 * x_5)$  (score: 0.0149, loss: 29.765318) Najlepsze rozwiązanie (best):

$$4. -x_5 * x_5 - \text{sympy\_p}(x_0) + 0.61306316$$