[Running] python -u "/root/PROJEKT\_NEXTFLOW/ML\_1.py"

Statystyki danych:

        Tissue\_1   Tissue\_2   Tissue\_3  ...   Tissue\_8   Tissue\_9  Tissue\_10

count  50.000000  50.000000  50.000000  ...  50.000000  50.000000  50.000000

mean    9.650018  10.507886   9.594654  ...   9.753132  10.681131   9.582222

std     4.662643   4.777863   4.953694  ...   4.328336   4.509407   5.430526

min     0.406144   0.196161   0.337211  ...   0.201649   1.483088   0.062155

25%     5.926023   7.283070   6.277186  ...   7.009722   7.371066   6.159043

50%     9.812685  10.259566   9.876692  ...   9.849409  10.493319   9.169161

75%    11.891990  14.195351  13.383086  ...  12.246496  12.601090  13.034496

max    21.573293  19.261391  20.610781  ...  19.483965  25.394404  29.263657

[8 rows x 10 columns]

Rozmiar całego zbioru: (50, 12)

Rozmiary zbiorów:

X\_train: (24, 11), y\_train: (24,)

X\_val:   (11, 11), y\_val:   (11,)

X\_test:  (15, 11), y\_test:  (15,)

Przykład X\_train:

    Gene\_ID   Tissue\_1   Tissue\_2  ...   Tissue\_8   Tissue\_9  Tissue\_10

43  Gene\_44   4.083707   0.196161  ...   3.099493   1.483088   9.722262

49  Gene\_50   9.960137  17.399721  ...   9.048307   5.621909   3.086001

18  Gene\_19  13.128337   5.714212  ...   9.635855   5.766031   2.425764

46  Gene\_47  19.823626  10.176318  ...  13.787539   7.347494   7.120909

29  Gene\_30   8.959389   7.534995  ...  11.536498  14.064311  13.148144

[5 rows x 11 columns]

Przykład y\_train:

43    transcription\_factor

49                receptor

18      structural\_protein

46                  enzyme

29    transcription\_factor

Name: Gene\_Function, dtype: object