

Klaudia Pakos

IO, rok 3, 267561

## PODSTAWY SZTUCZNEJ INTELIGENCJI

### SPRAWOZDANIE NR 2

# BUDOWA I DZIAŁANIE SIECI WIELOWARSTWOWEJ TYPU FEEDFORWARD

## 1. Cel ćwiczenia

Celem ćwiczenia jest poznanie budowy i działania wielowarstwowych sieci neuronowych poprzez uczenie kształtu funkcji matematycznej z użyciem algorytmu wstecznej propagacji błędu.

## 2. Zadania do wykonania

- Wygenerowanie danych uczących i testujących dla funkcji Rastrigin 3D dla danych wejściowych z przedziałów od -2 do 2
- Przygotowanie (implementacja lub wykorzystanie gotowych narzędzi) wielowarstwowej sieci oraz algorytmu wstecznej propagacji błędu.
- Uczenie sieci dla różnych współczynników uczenia (np. 0.5, 0.1, 0.01) i bezwładności (np. 0, 0.5, 1).
- Testowanie sieci.

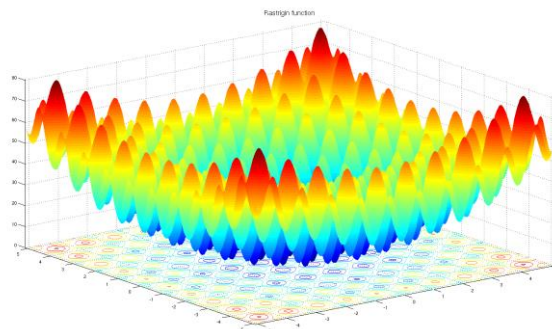
## 3. Opis budowy sieci

Celem budowanej sieci jest rozpoznawanie funkcji rastrigin. Funkcja ta w naszym przypadku przyjmuje za wejście współrzędne x, y z przedziału <-2; 2> i zwraca współrzędną z.

Wzór funkcji rastrigin:

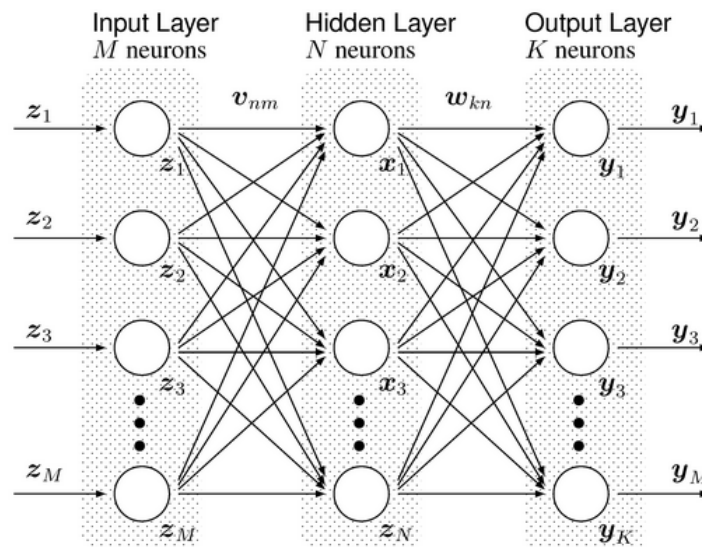
$$f(x) = 10 * n + \sum_{i=1}^n [x_i^2 - 10 * \cos(2\pi x_i)]$$

Rastrigin ma minimum globalne w punkcie (0, 0, 0).



Sieć wielowarstwowa składa się z warstwy wejściowej (Input Layer), co najmniej jednej warstwy ukrytej (Hidden Layer) oraz warstwy wyjściowej (Output Layer). Warstwy ukryte służą do przetwarzania sygnałów w sieci neuronowej.

Zastosowany typ **feedforward** oznacza, że w sieci istnieje z góry określony kierunek przepływu danych – dane przechodzą od warstwy wejściowej przez wszystkie warstwy ukryte kończąc na warstwie wyjściowej (dane nie mogą się „cofać” w użytej sieci). Każda z warstw jest powiązana tylko z warstwą poprzednią i następną. Dane wyjściowe każdego neuronu w jednej warstwie są jednocześnie danymi wejściowymi dla neuronów w kolejnej warstwie na zasadzie każdy z każdym. Sygnał wyjściowy nie jest dzielony, więc jest podawany taki sam na wejścia wszystkich neuronów kolejnej warstwy. Natomiast neurony w jednej warstwie nie są ze sobą w żaden sposób połączone.



Cykl uczenia metodą wstecznej propagacji błędów (backpropagation) składa się z następujących etapów:

1. Wyznaczenie odpowiedzi neuronów warstwy wyjściowej oraz warstw ukrytych na zadany sygnał wejściowy.
2. Wyznaczenie błędu popełnianego przez neurony znajdujące się w warstwie wyjściowej i przesłanie go w kierunku warstwy wejściowej.
3. Adaptacja wag.

#### 4. Przebieg zadania-listing kodu

```
%funkcja Rastrigin 3D- implementacja
function wynik = rastrigin(x)
    if x==0
        wynik=0;
    else
        x1=x;
        A=10;
```

```

        n=100;
        dx=(5.12-x)/n;
        wynik=A*n;
        for i=1:1:n
            x=x1+(i*dx);
            wynik=wynik+x^2-A*cos(2*pi*x);
        end
    end
end

close all; clear all; clc;

%Wygenerowanie danych wejściowych (-2.0, -1.9 ... 1.9 2.0)
wejście=zeros(1);
m=-2;
for i=1:41
    wejście(i)=m;
    m=m+0.1;
end
%wygenerowanie wartości funkcji Rastrigin 3D dla danych
wejściowych
wyjście=zeros(1);
for i=1:41
    wyjście(i)=rastrigin(wejście(i));
end

%Inicjalizacja wartości testowych
test = zeros(1);
k=-1.15;
for i=1:10
    test(i)=k;
    k=k+0.25;
end

%Sprawdzenie poprawnej wartości funkcji dla wartości testowych
test_wynik = zeros(1);

for i=1:10
    test_wynik(i)=rastrigin(test(i));
end

%Tworzenie sieci z 4 warstwami ukrytymi
net = feedforwardnet(4);

%Użycie algorytmu wstecznej propagacji
net.trainFcn = 'traingd';

%Dobranie współczynników dla funkcji
%współczynnik uczenia
net.trainParam.lr = 0.5;
% współczynnik bezwładności

```

```

net.trainParam.mc = 0;

%Trenowanie sieci
net = train(net, wejscie, wyjscie);

%Inizjalizacja tablicy wynikowej
wynik = zeros(size(net));

%Wywołanie funkcji Rastrigin dla liczb z przedziału [-2,2]
for i = 1:41
    wynik(i) = sim(net, wejscie(i)); %Testowanie sieci
end

for i = 1:41
    fprintf('Dla %i : Wartość funkcji %i, wartość podana przez
sieć: %i \n',wejscie(i), wyjscie(i), wynik(i));
end

test_wynik_siec = zeros(1);

for i=1:10
    test_wynik_siec(i)=sim(net, test(i));
end
fprintf('\n -----TEST-----\n');
for i = 1:10
    fprintf('Dla %i : Wartość funkcji %i, wartość podana przez
sieć: %i \n',test(i), test_wynik(i), test_wynik_siec(i));
end

```

## 5. Przedstawienie wyników

Dokonano sprawdzenia działania sieci dla współczynników uczenia (0.5, 0.1, 0.01- pierwszy wiersz) i współczynników bezwładności (0, 0.5, 1- drugi wiersz).

		0.5	0.1	0.01	0.5	0.1	0.01	0.5	0.1	0.01
Argumen ty	Wartość funkcji	0	0	0	0.5	0.5	0.5	1	1	1
-2	1663,31 4	2620,01 7	2227,92 9	2269,66 7	2302,82 2	1723,16 9	2177,61 4	1367,84 6	2120,16 3	2432,51 8
-1,9	1679,50 7	2592,09 7	2246,58 8	2182,38 2	2233,13 3	1787,83	2185,26 7	1346,07 1	2180,34 4	2346,70 4
-1,8	1690,11 3	2567,96	2262,52 4	2107,24 6	2173,12 2	1843,47	2192,08 8	1327,54 5	2232,10 9	2272,55 5
-1,7	1692,68 7	2549,53 3	2274,31 6	2050,50 2	2127,75 7	1885,45 1	2197,69 3	1313,98	2271,12 5	2216,01 1
-1,6	1688,03 5	2536,6	2281,89 4	2011,82 9	2096,75 9	1913,98 7	2202,36 5	1305,53 4	2297,56 9	2176,45 3

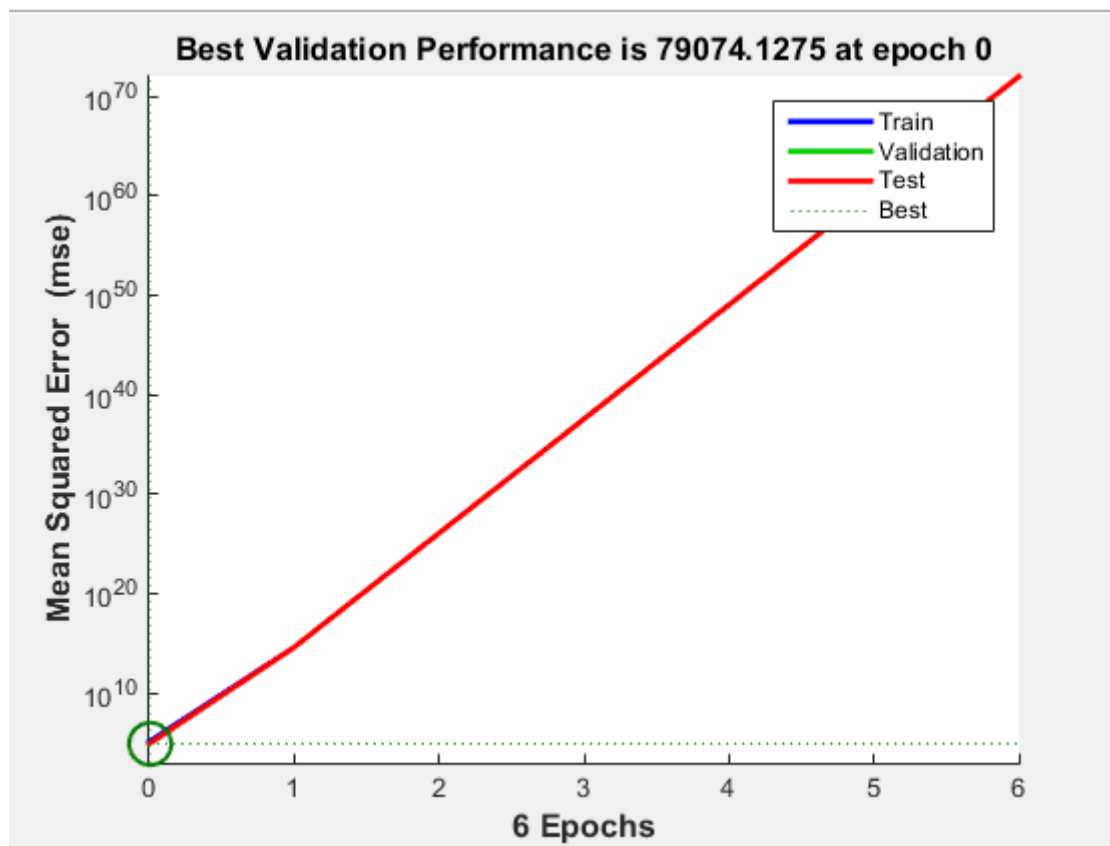
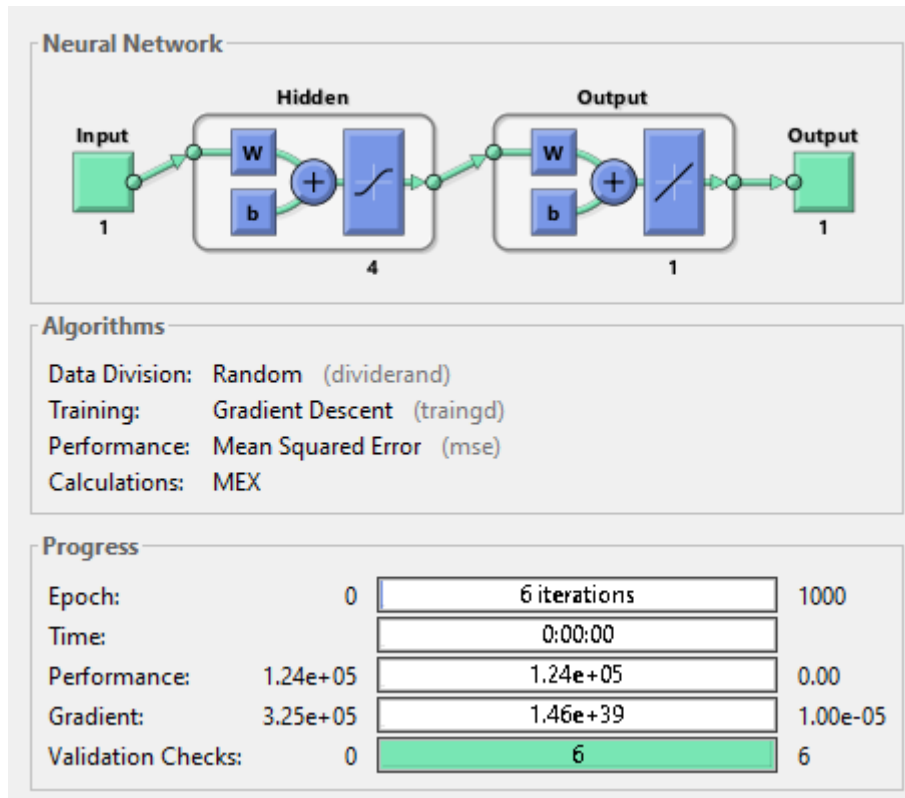
-1,5	1680,00 4	2527,76 8	2285,83 9	1987,44 9	2077,07	1931,84 1	2206,86 1	1301,66 4	2313,97 5	2149,65 6
-1,4	1674,09 9	2521,36 2	2286,66 1	1973,13 3	2065,24 4	1942,08	2212,28 2	1302	2323,13 1	2130,58 5
-1,3	1675,42 5	2515,80 1	2284,39 3	1965,62 6	2058,56 6	1947,00 8	2220,14 8	1306,88 8	2327,07 2	2114,56 1
-1,2	1686,72 1	2509,57 4	2278,39 2	1962,93 4	2055,25 3	1947,92 6	2232,62 3	1317,69 8	2326,83 2	2097,23 5
-1,1	1707,24	2501,07 8	2267,23 8	1964,16	2054,30 6	1945,23 6	2252,78 7	1337,02	2322,51 4	2074,12 7
-1	1732,92 3	2488,53 6	2248,80 2	1969,22	2055,27 6	1938,64 7	2284,63 8	1368,58 2	2313,49 3	2040,36 4
-0,9	1757,84 9	2470,24 8	2220,80 9	1978,43 2	2058,02 6	1927,55 9	2332,25 4	1416,31 8	2298,84 6	1991,32 9
-0,8	1776,44 1	2445,41 9	2182,24 5	1991,86 6	2062,43 1	1911,75 3	2397,47 4	1481,99 8	2278,21 1	1924,85 6
-0,7	1785,61 5	2415,33 8	2135,27 6	2008,55 7	2068,06 1	1892,25	2476,71 1	1561,96 9	2252,88 5	1844,36 3
-0,6	1786,04 2	2383,71 4	2085,82 3	2026,24 1	2074,05 8	1871,59 4	2559,99 8	1646,19 9	2226,17 2	1759,74 3
-0,5	1782,01 9	2355,08 5	2041,09 4	2042,22	2079,41 7	1852,83 1	2635,15 6	1722,47 6	2202,07	1683,11 4
-0,4	1779,95 2	2332,43 5	2005,85 2	2054,68	2083,43 1	1837,95 3	2694,07 1	1782,76 8	2183,25 4	1622,42 1
-0,3	1785,97 1	2316,25 1	1980,97 4	2063,20 2	2085,84 4	1827,29 6	2735,10 5	1825,70 3	2170,32 2	1578,92 5
-0,2	1803,57 3	2305,31 7	1964,72 5	2068,25 3	2086,64 1	1820,06	2760,89 8	1854,43	2162,52 4	1549,30 9
-0,1	1832,18 4	2297,86 4	1954,61 2	2070,47 7	2085,74 8	1815,07 1	2775,15 9	1873,52 7	2158,82 5	1528,72 2
0,00E+00	1867,22 4	2292,19 8	1948,44 9	2070,23 3	2082,79	1811,18 7	2780,73 7	1887,27 9	2158,57 9	1512,43 4
0,1	1901,71 5	2286,81	1944,64 7	2067,39 5	2076,90 7	1807,37 3	2778,94 6	1899,31 8	2161,80 2	1496,09 6
0,2	1928,86	2280,25 7	1942,13 4	2061,26 6	2066,59 8	1802,60 9	2769,48 2	1912,89 8	2169,27 2	1475,34 4
0,3	1944,65 7	2271,01 5	1940,19 7	2050,58 7	2049,66 6	1795,79 9	2750,59 7	1931,23 8	2182,50 3	1445,41 9
0,4	1949,54 9	2257,55 3	1938,35 3	2033,78 9	2023,54 5	1785,82 5	2719,71 3	1957,45 7	2203,41 5	1401,42 9
0,5	1948,43 7	2238,82 2	1936,30 3	2009,75 6	1986,44 5	1771,92 8	2674,95 6	1993,66 9	2233,36 6	1340,03 5
0,6	1948,98 6	2215,18 6	1933,95 3	1979,11 6	1939,31 3	1754,40 3	2617,65 8	2039,21 5	2271,52	1262,53 5
0,7	1958,80 3	2189,06 4	1931,42 9	1945,12 2	1887,19 4	1735,09 2	2554,05 4	2089,45 1	2313,73 2	1177,01 1
0,8	1982,51 1	2164,14 6	1928,98 4	1912,66 1	1837,70 8	1716,80 5	2493,45 6	2137,24 8	2353,76 1	1095,78 9
0,9	2019,84 2	2143,42 1	1926,80 3	1885,68 6	1797,12 1	1701,86 8	2443,45 6	2176,79 8	2386,45 7	1028,98 4
1	2065,50 8	2127,84 8	1924,85 7	1865,49 6	1767,74 9	1691,16 5	2406,74	2206,13 9	2409,84 4	980,221 7
1,1	2110,99 9	2116,60 3	1922,89 2	1851,07 1	1748,57 3	1684,37	2381,79 3	2226,65 4	2424,6	947,590 3

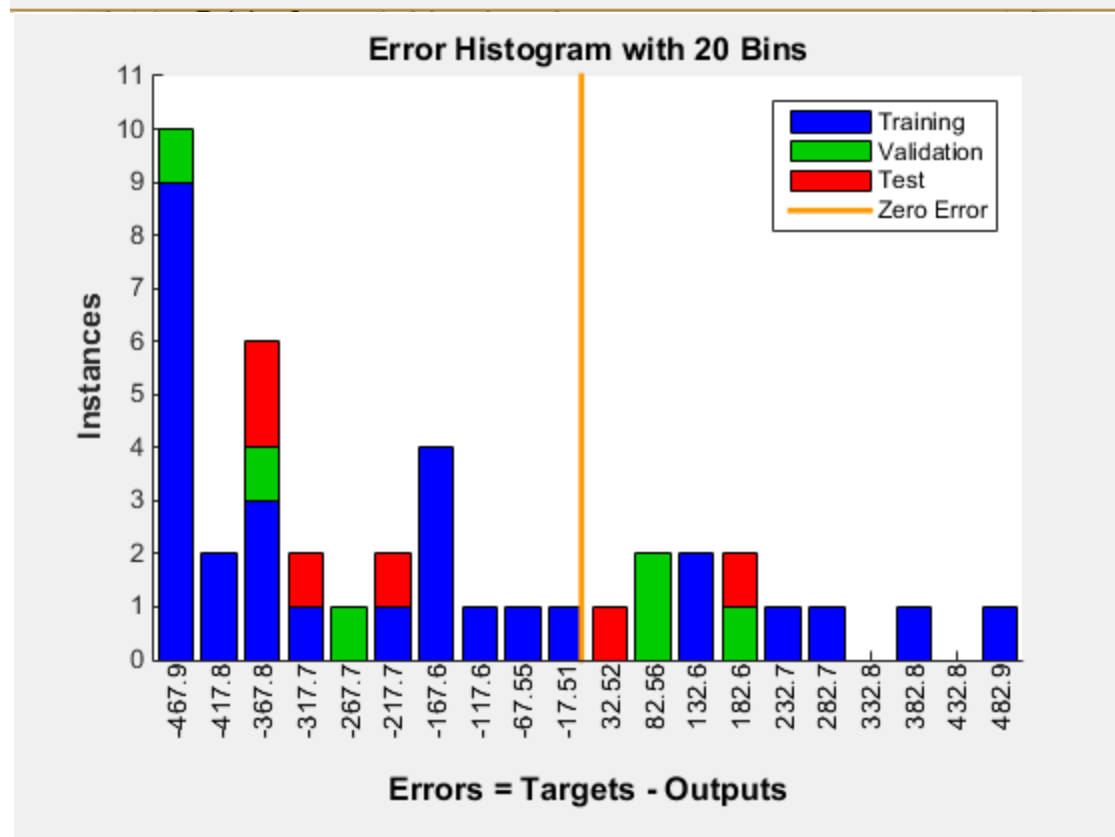
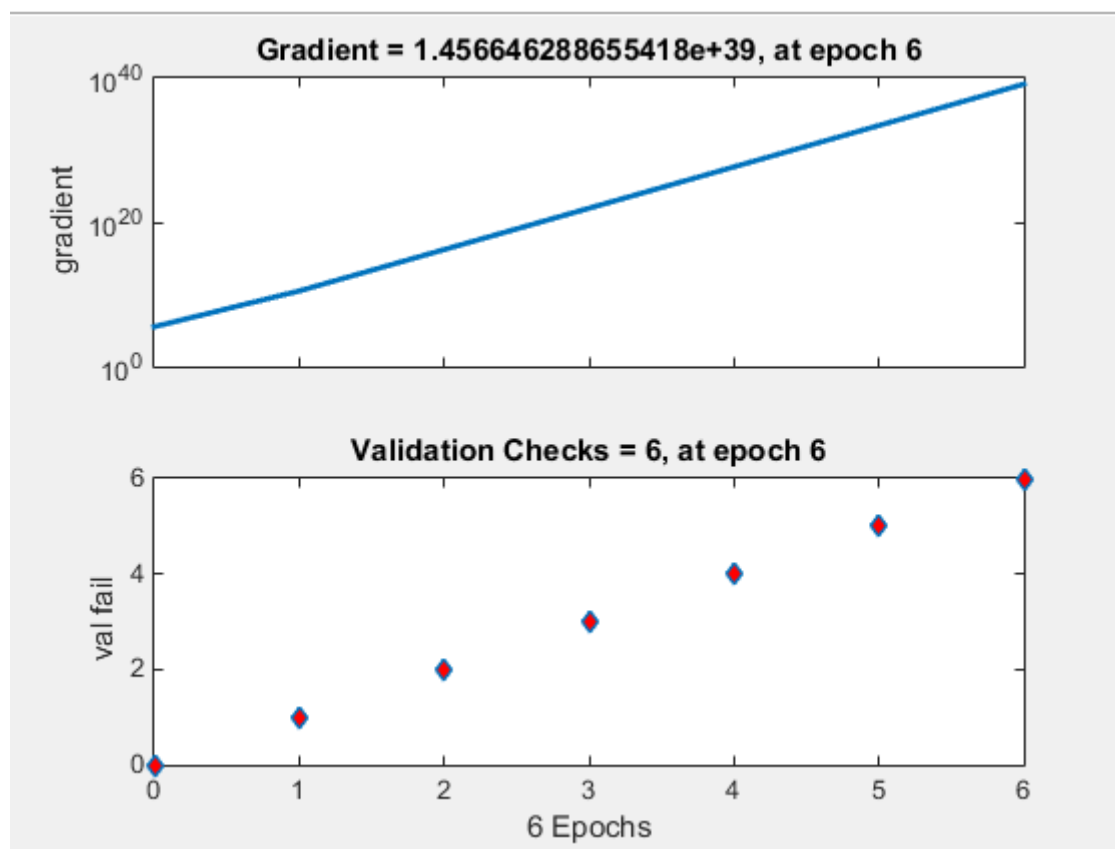
1,2	2147,69 2	2107,99 7	1920,45 5	1840,29 2	1737,27 6	1680,72 1	2365,30 9	2241,24 1	2432,34 8	926,899 2
1,3	2170,16	2100,07 7	1916,86 4	1830,75	1731,75 7	1679,6	2353,90 6	2253,05 7	2434,36 2	914,037 6
1,4	2178,41 9	2090,71 4	1911,09 5	1819,89 6	1730,75 2	1680,80 7	2344,69 1	2265,20 8	2430,95 2	905,825 8
1,5	2178,20 6	2077,38 4	1901,60 1	1804,80 2	1733,95 3	1684,64 9	2335,15 1	2280,98 2	2421,23 8	900,003 6
1,6	2179,07 8	2056,92 1	1886,18	1781,86 9	1741,93 7	1691,95 6	2322,87 7	2304,18 4	2403,10 6	894,954 1
1,7	2190,95	2025,62 4	1862,10 5	1746,93	1755,91 5	1703,94 2	2305,46 9	2339,09 1	2373,52 1	889,469 6
1,8	2220,29 6	1980,29 2	1826,97 7	1696,39 6	1777,09 3	1721,72 9	2280,97 5	2389,34 3	2329,69 2	882,705 4
1,9	2267,46	1920,55 2	1780,55 8	1629,83 6	1805,45 5	1745,37 9	2249,04 7	2455,41 8	2271,45 5	874,367 1
2	2326,15 1	1851,27 9	1726,67 6	1552,67	1838,54 9	1772,89 6	2212,18 4	2531,96 8	2203,71 1	864,958 4

Dokonano również testu sieci, czego wyniki przedstawiono poniżej.

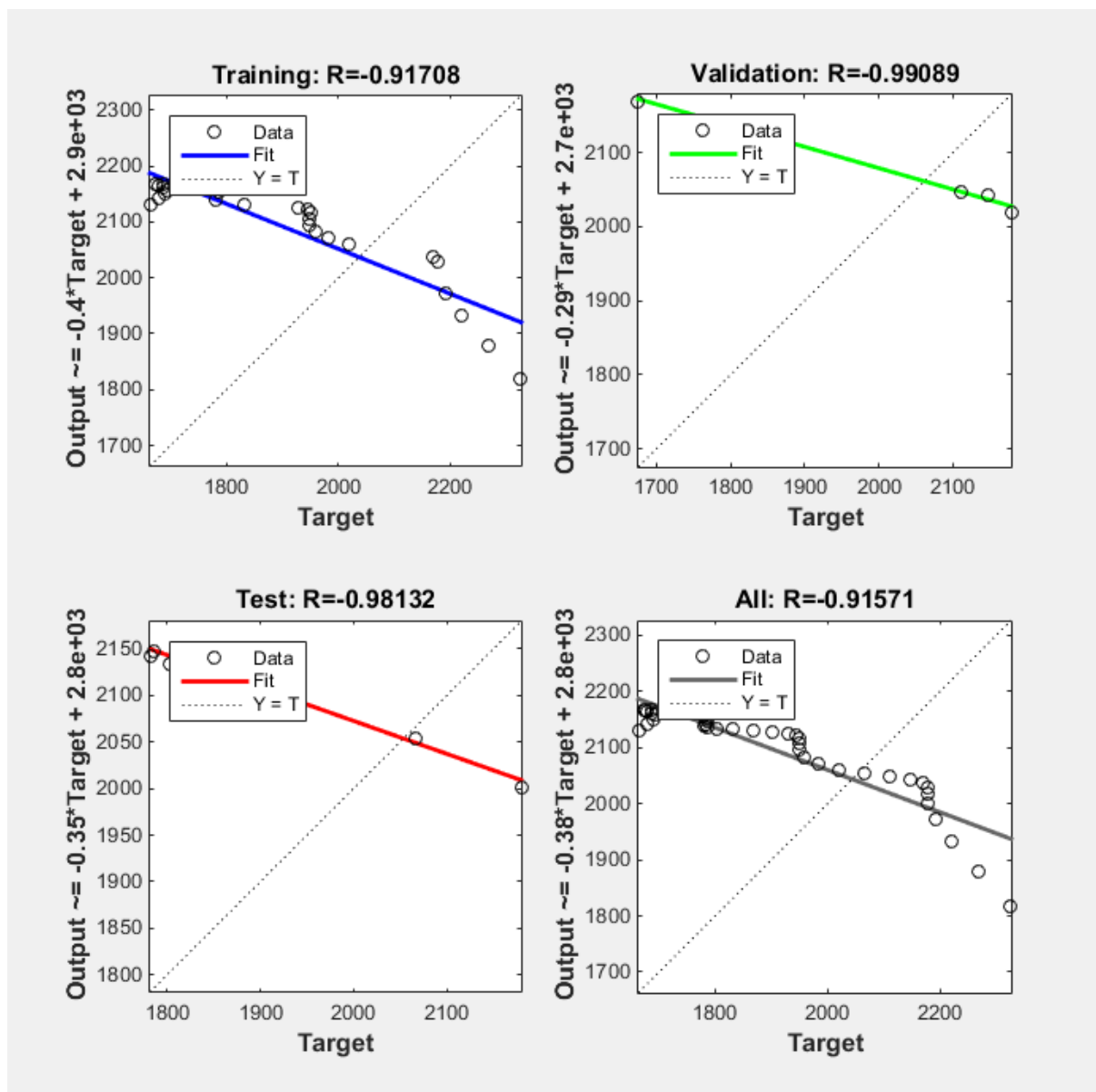
Argumenty	Wartość funkcji	0.5	0.1	0.01	0.5	0.1	0.01	0.5	0.1	0.01
		0	0	0	0.5	0.5	0.5	1	1	1
-1,15	1696,015	2505,717	2273,579	1963,078	2054,522	1947,037	2241,512	1326,084	2325,208	2086,7
-0,9	1757,849	2470,248	2220,809	1978,432	2058,026	1927,559	2332,254	1416,318	2298,846	1991,329
-0,65	1786,727	2399,424	2110,392	2017,45	2071,079	1881,868	2518,638	1604,344	2239,443	1801,782
-0,4	1779,952	2332,435	2005,852	2054,68	2083,431	1837,953	2694,071	1782,768	2183,254	1622,421
-0,15	1816,707	2301,266	1959,051	2069,686	2086,42	1817,359	2769,26	1864,899	2160,216	1538,193
0,1	1901,715	2286,81	1944,647	2067,395	2076,907	1807,373	2778,946	1899,318	2161,802	1496,096
0,35	1948,252	2264,898	1939,286	2043,047	2037,905	1791,271	2736,815	1943,189	2191,884	1425,459
0,6	1948,986	2215,186	1933,953	1979,116	1939,313	1754,403	2617,658	2039,215	2271,52	1262,535
0,85	1999,699	2153,149	1927,857	1898,34	1816,059	1708,826	2466,84	2158,274	2371,228	1060,195
1,1	2110,999	2116,603	1922,892	1851,071	1748,573	1684,37	2381,793	2226,654	2424,6	947,5903

Przykładowe wyniki działania programu dla współczynnika uczenia 0.5 i współczynnika bezwładności 1.









## 6. Analiza błędów

Obliczono błąd względny dla każdego z otrzymanych wyników, gdzie dla wartości uczących otrzymano:

Argumenty	0.5	0.1	0.01	0.5	0.1	0.01	0.5	0.1	0.01
	0	0	0	0.5	0.5	0.5	1	1	1
-2	58%	34%	36%	38%	4%	31%	18%	27%	46%
-1,9	54%	34%	30%	33%	6%	30%	20%	30%	40%
-1,8	52%	34%	25%	29%	9%	30%	21%	32%	34%
-1,7	51%	34%	21%	26%	11%	30%	22%	34%	31%

-1,6	50%	35%	19%	24%	13%	30%	23%	36%	29%
-1,5	50%	36%	18%	24%	15%	31%	23%	38%	28%
-1,4	51%	37%	18%	23%	16%	32%	22%	39%	27%
-1,3	50%	36%	17%	23%	16%	33%	22%	39%	26%
-1,2	49%	35%	16%	22%	15%	32%	22%	38%	24%
-1,1	46%	33%	15%	20%	14%	32%	22%	36%	21%
-1	44%	30%	14%	19%	12%	32%	21%	34%	18%
-0,9	41%	26%	13%	17%	10%	33%	19%	31%	13%
-0,8	38%	23%	12%	16%	8%	35%	17%	28%	8%
-0,7	35%	20%	12%	16%	6%	39%	13%	26%	3%
-0,6	33%	17%	13%	16%	5%	43%	8%	25%	1%
-0,5	32%	15%	15%	17%	4%	48%	3%	24%	6%
-0,4	31%	13%	15%	17%	3%	51%	0%	23%	9%
-0,3	30%	11%	16%	17%	2%	53%	2%	22%	12%
-0,2	28%	9%	15%	16%	1%	53%	3%	20%	14%
-0,1	25%	7%	13%	14%	1%	51%	2%	18%	17%
0,00E+00	23%	4%	11%	12%	3%	49%	1%	16%	19%
0,1	20%	2%	9%	9%	5%	46%	0%	14%	21%
0,2	18%	1%	7%	7%	7%	44%	1%	12%	24%
0,3	17%	0%	5%	5%	8%	41%	1%	12%	26%
0,4	16%	1%	4%	4%	8%	40%	0%	13%	28%
0,5	15%	1%	3%	2%	9%	37%	2%	15%	31%
0,6	14%	1%	2%	0%	10%	34%	5%	17%	35%
0,7	12%	1%	1%	4%	11%	30%	7%	18%	40%
0,8	9%	3%	4%	7%	13%	26%	8%	19%	45%
0,9	6%	5%	7%	11%	16%	21%	8%	18%	49%
1	3%	7%	10%	14%	18%	17%	7%	17%	53%
1,1	0%	9%	12%	17%	20%	13%	5%	15%	55%
1,2	2%	11%	14%	19%	22%	10%	4%	13%	57%
1,3	3%	12%	16%	20%	23%	8%	4%	12%	58%
1,4	4%	12%	16%	21%	23%	8%	4%	12%	58%
1,5	5%	13%	17%	20%	23%	7%	5%	11%	59%
1,6	6%	13%	18%	20%	22%	7%	6%	10%	59%
1,7	8%	15%	20%	20%	22%	5%	7%	8%	59%
1,8	11%	18%	24%	20%	22%	3%	8%	5%	60%
1,9	15%	21%	28%	20%	23%	1%	8%	0%	61%
2	20%	26%	33%	21%	24%	5%	9%	5%	63%

A dla wartości testowych otrzymano:

Argumenty	0.5	0.1	0.01	0.5	0.1	0.01	0.5	0.1	0.01
	0	0	0	0.5	0.5	0.5	1	1	1
-1,15	48%	34%	16%	21%	15%	32%	22%	37%	23%
-0,9	41%	26%	13%	17%	10%	33%	19%	31%	13%
-0,65	34%	18%	13%	16%	5%	41%	10%	25%	1%
-0,4	31%	13%	15%	17%	3%	51%	0%	23%	9%
-0,15	27%	8%	14%	15%	0%	52%	3%	19%	15%
0,1	20%	2%	9%	9%	5%	46%	0%	14%	21%
0,35	16%	0%	5%	5%	8%	40%	0%	13%	27%
0,6	14%	1%	2%	0%	10%	34%	5%	17%	35%
0,85	8%	4%	5%	9%	15%	23%	8%	19%	47%
1,1	0%	9%	12%	17%	20%	13%	5%	15%	55%

## 7. Wnioski

Podczas tworzenia sieci wielowarstwowej trzeba zwrócić największą uwagę na jej strukturę oraz na współczynnik uczenia. Sieci wielowarstwowe posiadające większą liczbę neuronów oraz warstw ukrytych na ogół uczą się bardziej efektywnie od sieci z ich mniejszą ilością. Przeprowadzone ćwiczenie pokazało, że niezależnie od współczynnika uczenia i bezwładności sieć potrzebowała do nauki tylko 6 epok. Różne natomiast otrzymano wyniki nauki dla tych współczynników. Najgorzej wypadają sieci neuronowe z wysokim współczynnikiem uczenia i niskim współczynnikiem bezwładności oraz wysokim współczynnikiem bezwładności i niskim współczynnikiem uczenia. Uzyskują największe błędy sięgające nawet 60%. Kluczem do sukcesu jest zatem dobry wybór tych współczynników. W analizowanym przykładzie funkcji Rastrigin 3D najlepiej sprawdził się model ze współczynnikiem uczenia 0.5 i współczynnikiem bezwładności 1. Parametry uczenia pozwalają nam sterować wynikami końcowymi i mają główny wpływ na wygląd i wartości otrzymanych funkcji.