Kamil Majchrowicz
Grupa II
Inżynieria Obliczeniowa
Podstawy Sztucznej Inteligencji

Sprawozdanie 3

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:

a) Wygenerowanie danych uczących i testujących dla funkcji Rastrigin 3D dla danych wejściowych z przedziałów od -2 do 2.

b) Przygotowanie (implementacja lub wykorzystanie gotowych narzędzi) wielowarstwowej sieci oraz algorytmu wstecznej propagacji błędu.

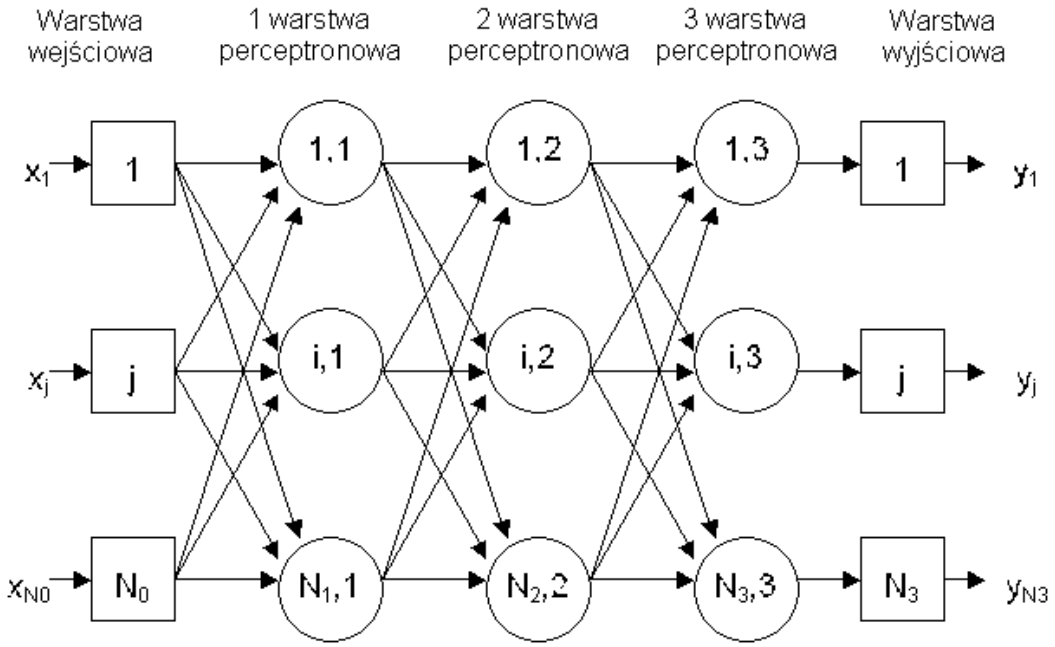
c) Uczenie sieci dla różnych współczynników uczenia (np. 0.5, 0.1, 0.01) i bezwładności (np. 0, 0.5, 1).

d) Testowanie sieci.

3. Opis pojęć:

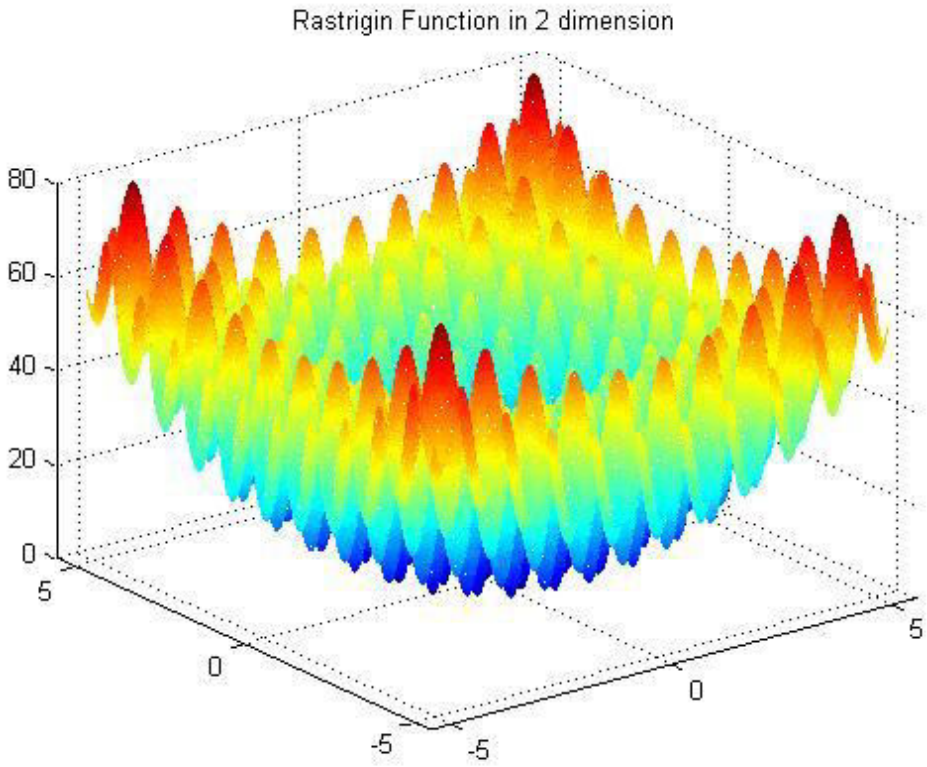
Perceptron wielowarstwowy – najpopularniejszy typ sztucznych sieci neuronowych. Sieć tego typu składa się zwykle z jednej warstwy wejściowej, kilku warstw ukrytych oraz jednej warstwy wyjściowej. Warstwy ukryte składają się najczęściej z neuronów McCullocha-Pittsa. Ustalenie właściwej liczby warstw ukrytych oraz liczby neuronów znajdujących się w poszczególnych warstwach jest trudnym zagadnieniem, które musi rozwiązać twórca sieci neuronowej. Warstwa wyjściowa może składać się z neuronów liniowych (w przypadku regresji) lub neuronów nieliniowych (w przypadku klasyfikacji). Trenowanie sieci typu MLP możliwe jest dzięki zastosowaniu metody wstecznej propagacji błędów. Perceptron wielowarstwowy w przeciwieństwie do perceptronu jednowarstwowego może być wykorzystywany do klasyfikowania zbiorów, które nie są liniowo separowalne. Sieć MLP w swojej podstawowej wersji jest siecią, w której nie ma sprzężenia zwrotnego, w przeciwieństwie do sieci zwanych sieciami rekurencyjnymi. Na bazie sieci MLP zbudowane są splotowe sieci neuronowe, służące do rozpoznawania obrazów.

Przykład sieci neuronowej o 3 warstwach:



Propagacja wsteczna – podstawowy algorytm uczenia nadzorowanego wielowarstwowych, jednokierunkowych sieci neuronowych. Podaje on przepis na zmianę wag dowolnych połączeń elementów przetwarzających rozmieszczonych w sąsiednich warstwach sieci. Oparty jest on na minimalizacji sumy kwadratów błędów (lub innej funkcji błędu) uczenia z wykorzystaniem optymalizacyjnej metody największego spadku. Dzięki zastosowaniu specyficznego sposobu propagowania błędów uczenia sieci powstałych na jej wyjściu, tj. przesyłania ich od warstwy wyjściowej do wejściowej, algorytm propagacji wstecznej stał się jednym z najskuteczniejszych algorytmów uczenia sieci.

Funkcja rastrign- przyjmuje na wejściu współrzędne x i y z przedziału <-2,2> a na wyjściu zwraca



z.

Wzór funkcji rastrign:

$$f(\mathbf{x}) = An + \sum_{i=1}^{n} \left[x_i^2 - A\cos(2\pi x_i) \right]$$

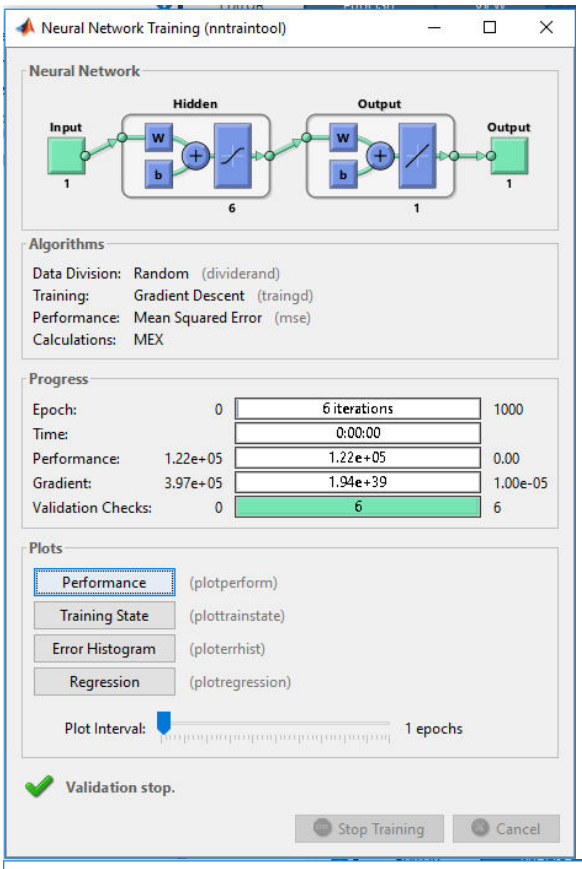
Gdzie:

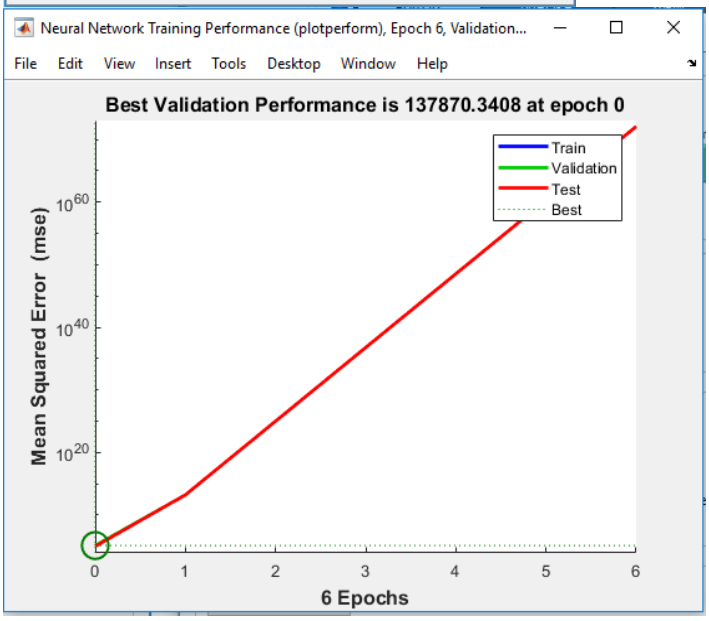
$$A = 10$$
$$x_i \in [-5.12, 5.12]$$

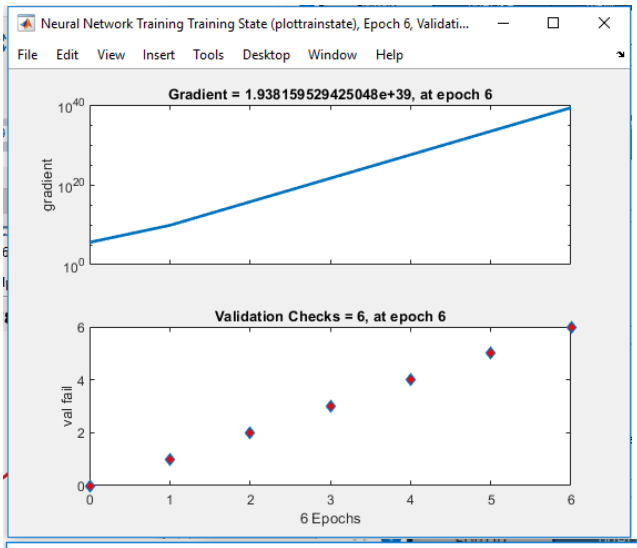
a minimum lokalne zawarte jest w punkcie (0,0,0).

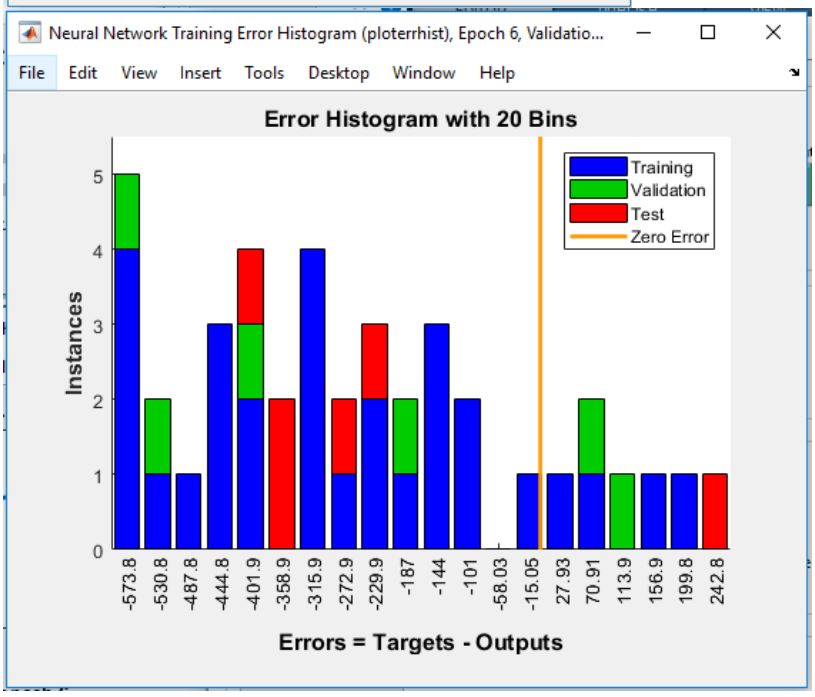
4. Analiza Wyników

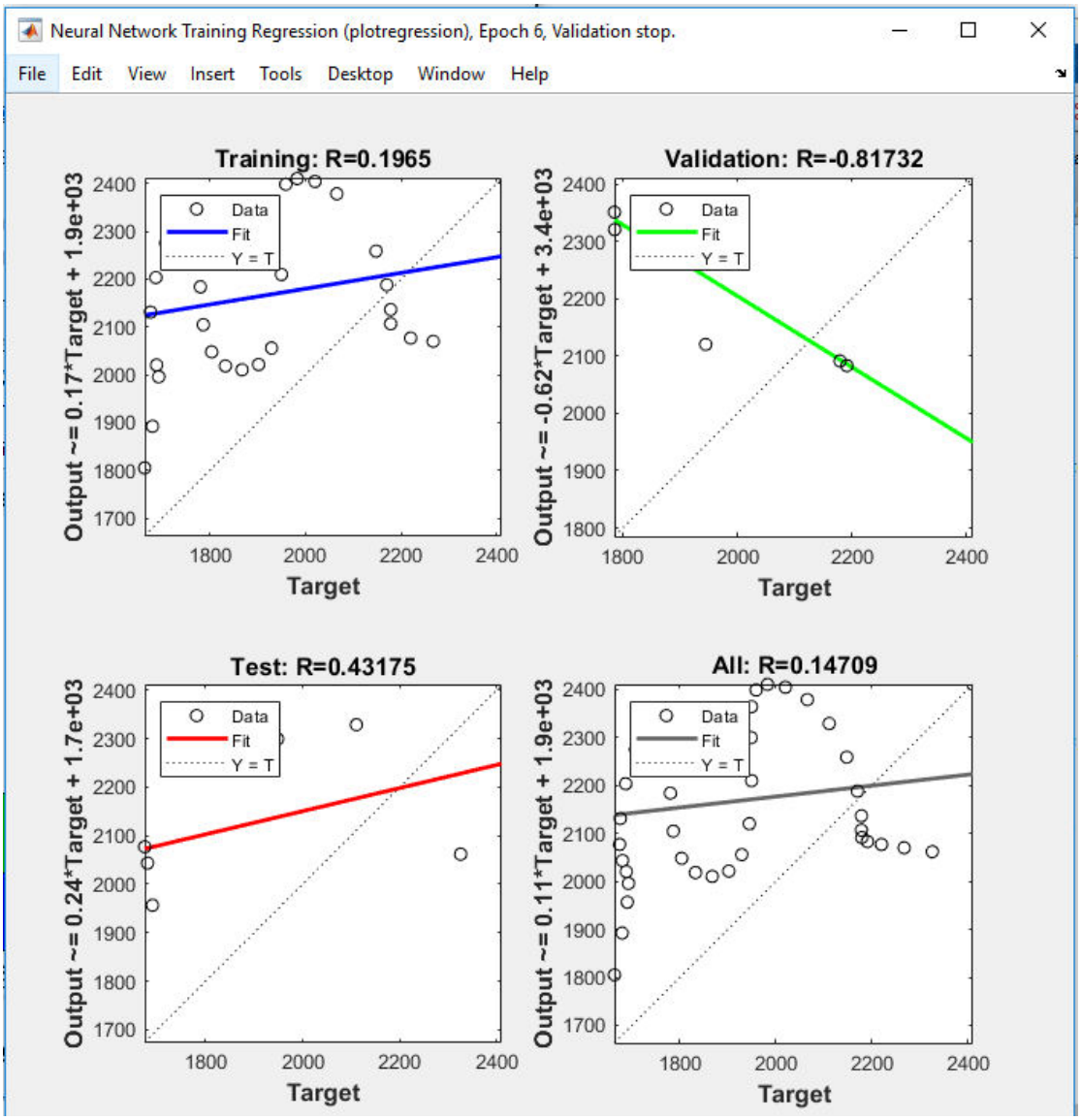
Sieć z 6 warstwami
Współczynnik uczenia - 0.5
Współczynnik bezwładności - 0


Neural Network Training (nntraintool)
Neural Network
Input
Hidden
Output
Output
W
b
1
6
1
1
Algorithms
Data Division: Random (dividerand)
Training: Gradient Descent (traingd)
Performance: Mean Squared Error (mse)
Calculations: MEX
Progress
Epoch: 0 6 iterations 1000
Time: 0:00:00
Performance: 1.22e+05 1.22e+05 0.00
Gradient: 3.97e+05 1.94e+39 1.00e-05
Validation Checks: 0 6 6
Plots
Performance (plotperform)
Training State (plottrainstate)
Error Histogram (ploterrhist)
Regression (plotregression)
Plot Interval: 1 epochs
Validation stop.
Stop Training
Cancel

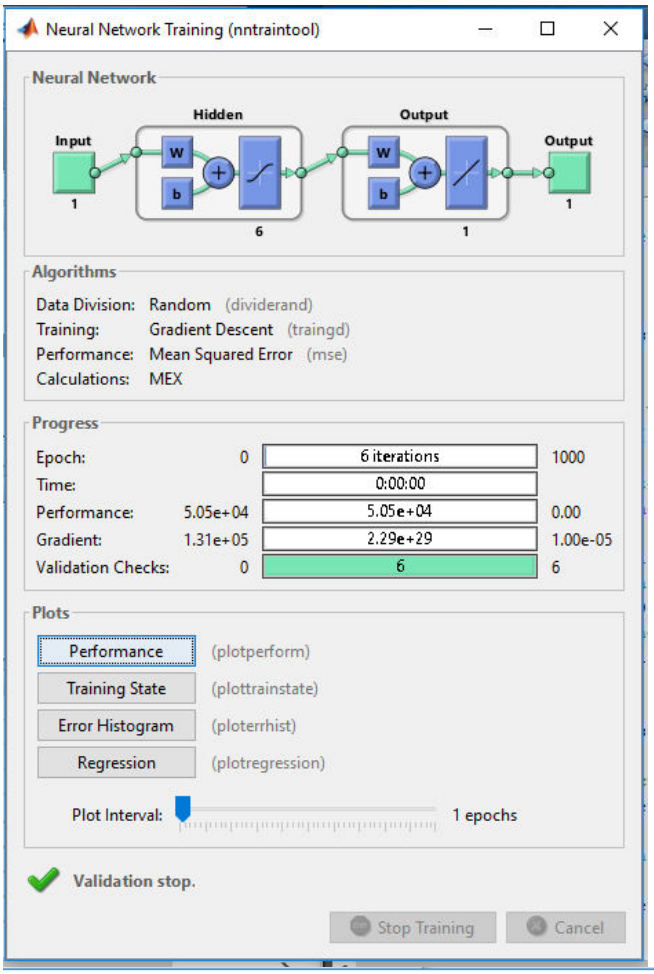

Neural Network Training Performance (plotperform), Epoch 6, Validation...
File Edit View Insert Tools Desktop Window Help
Best Validation Performance is 137870.3408 at epoch 0
Train
Validation
Test
Best
Mean Squared Error (mse)
10^60
10^40
10^20
0
1
2
3
4
5
6
6 Epochs

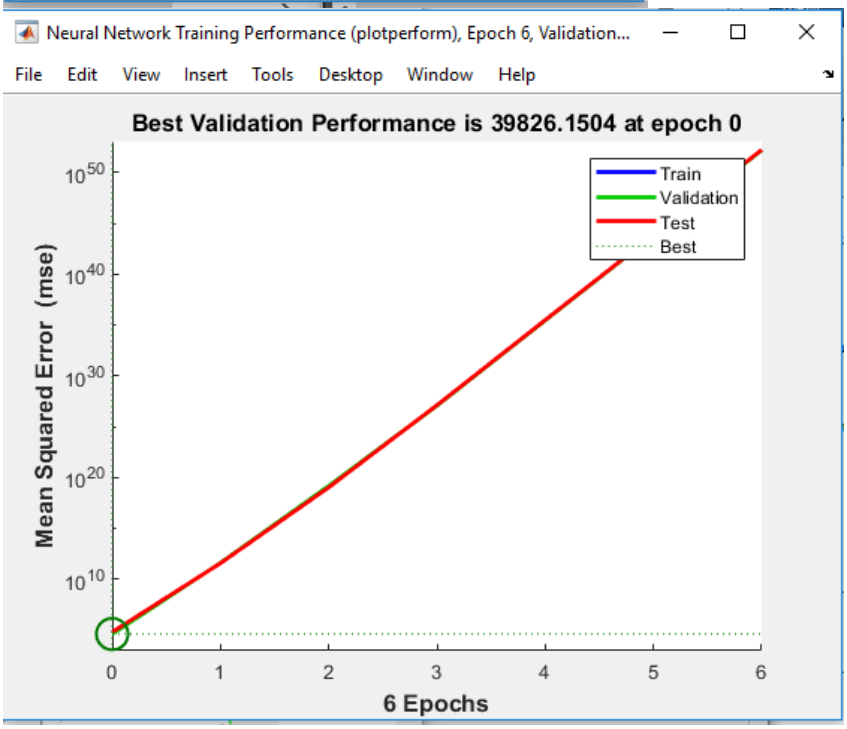

Neural Network Training Training State (plottrainstate), Epoch 6, Validati...
File Edit View Insert Tools Desktop Window Help
Gradient = 1.938159529425048e+39, at epoch 6
gradient
Validation Checks = 6, at epoch 6
val fail
6 Epochs

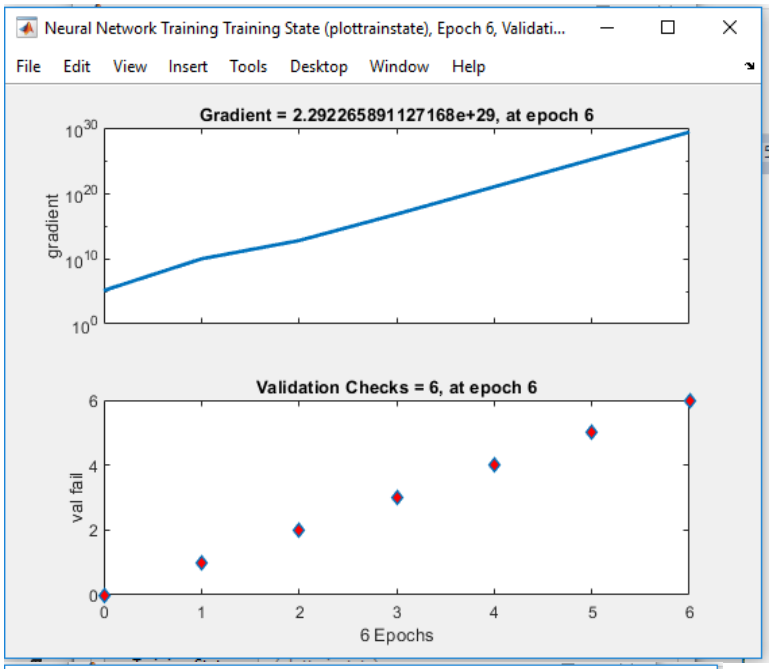

Neural Network Training Error Histogram (ploterrhist), Epoch 6, Validatio...
File Edit View Insert Tools Desktop Window Help
Error Histogram with 20 Bins
Training
Validation
Test
Zero Error
Instances
-573.8
-530.8
-487.8
-444.8
-401.9
-358.9
-315.9
-272.9
-229.9
-187
-144
-101
-58.03
-15.05
27.93
70.91
113.9
156.9
199.8
242.8
Errors = Targets - Outputs

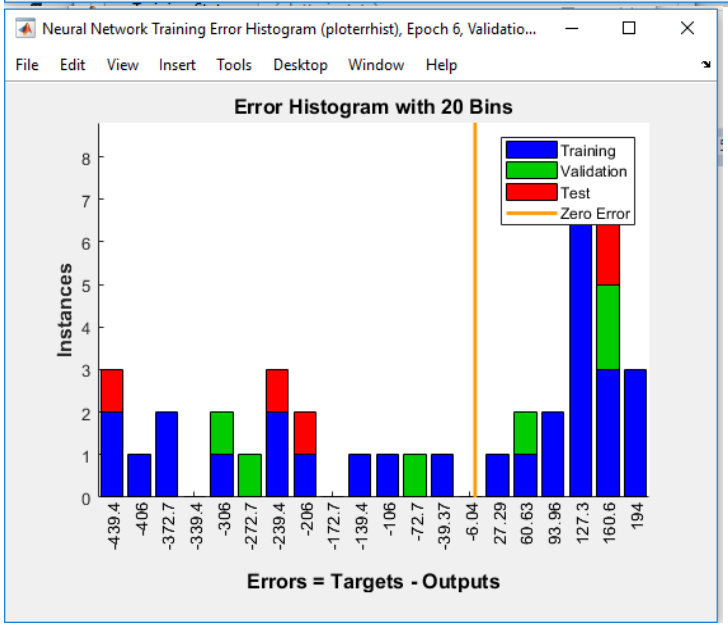


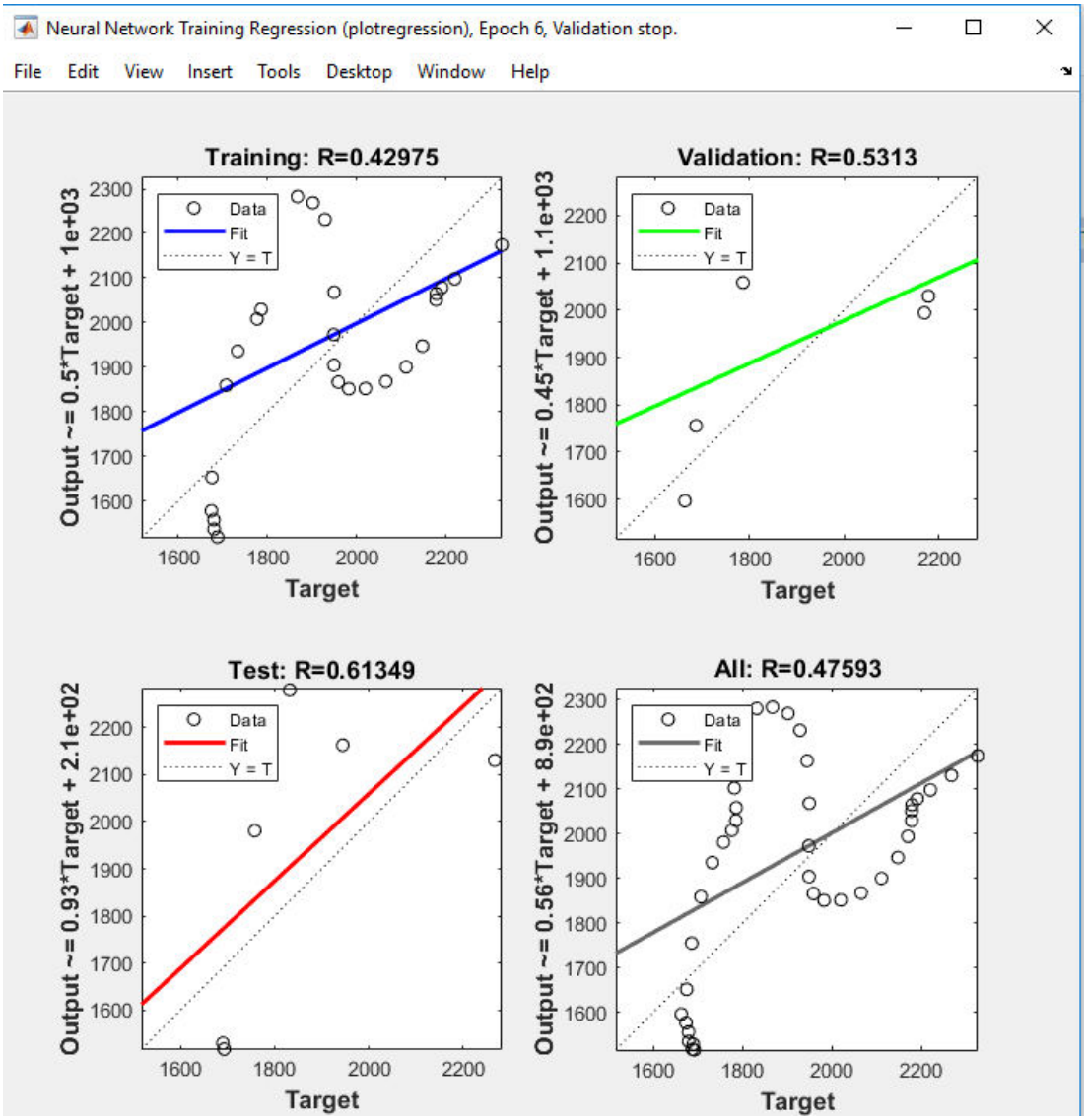
Współczynnik uczenia - 0.01
Współczynnik bezwładności - 1


Neural Network Training (nntraintool)
Neural Network
Input
Hidden
Output
Output
W
b
1
6
1
1
Algorithms
Data Division: Random (dividerand)
Training: Gradient Descent (traingd)
Performance: Mean Squared Error (mse)
Calculations: MEX
Progress
Epoch: 0 6 iterations 1000
Time: 0:00:00
Performance: 5.05e+04 5.05e+04 0.00
Gradient: 1.31e+05 2.29e+29 1.00e-05
Validation Checks: 0 6 6
Plots
Performance (plotperform)
Training State (plottrainstate)
Error Histogram (ploterrhist)
Regression (plotregression)
Plot Interval: 1 epochs
Validation stop.
Stop Training
Cancel


Neural Network Training Performance (plotperform), Epoch 6, Validation...
File Edit View Insert Tools Desktop Window Help
Best Validation Performance is 39826.1504 at epoch 0
Train
Validation
Test
Best
Mean Squared Error (mse)
10^50
10^40
10^30
10^20
10^10
0 1 2 3 4 5 6
6 Epochs


Neural Network Training Training State (plottrainstate), Epoch 6, Validati...
File Edit View Insert Tools Desktop Window Help
Gradient = 2.2922658911271 68e+29, at epoch 6
gradient
Validation Checks = 6, at epoch 6
val fail
6 Epochs


Neural Network Training Error Histogram (ploterrhist), Epoch 6, Validatio...
File Edit View Insert Tools Desktop Window Help
Error Histogram with 20 Bins
Training
Validation
Test
Zero Error
Instances
Errors = Targets - Outputs



5. Kod:

```
close all; clear all; clc;



%Dane wejściowe
wejscie=zeros(1);
m=-2;
for i=1:41
    wejscie(i)=m;
    m=m+0.1;
end
%Wartosci funkci rastrigin
wyjscie=zeros(1);
for i=1:41
    wyjscie(i)=rastrigin(wejscie(i));
```

```
end

%Wartości testowe
wart_test = zeros(1);
k=-1.15;
for i=1:10
    wart_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(wart_test(i));
end

%Tworzenie sieci, w nawiasie liczba warstw
net = feedforwardnet(6);

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


% współczynnik bezwładności
net.trainParam.mc = 1;
%współczynnik uczenia
net.trainParam.lr = 0.001;


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

%Tablica wynikowa
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('Funkcja %i: \n Wartość: %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, wart_test(i));
end
fprintf('\n -----TEST------\n');
for i = 1:10
fprintf('Funkcja %i: \n Wartość: %i, wartość podana przez sieć: %i \n',test(i),
test_wynik(i), test_wynik_siec(i));
```

```
end

wynik2=wynik';
test_wynik_siec2=test_wynik_siec';

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
```

6. Wnioski:

- Najgorsze wyniki osiągają sieci których współczynnik uczenia jest równy współczynnikowi bezwładności
- W wypadku nie używania algorytmu propagacji danych należy zachowywać duże różnice między wartościami współczynników
- Wartości współczynników uczenia i bezwładności znacząco wpływają na otrzymywane wyniki
- Sieci wielowarstwowe z algorytmem bezwładności działają gorzej niż te bez algorytmu
- Algorytm propagacji danych znacznie przyspiesza proces uczenia sieci
- Przy zastosowaniu coraz to bardziej skomplikowanych sieci wielowarstwowych można osiągać dokładniejsze wyniki operacji