

Michał Gawlik, IO

Podstawy Sztucznej Inteligencji

Sprawozdanie 3

Celem Projektu było poznanie budowy i działania wielowarstwowych sieci neuronowych.

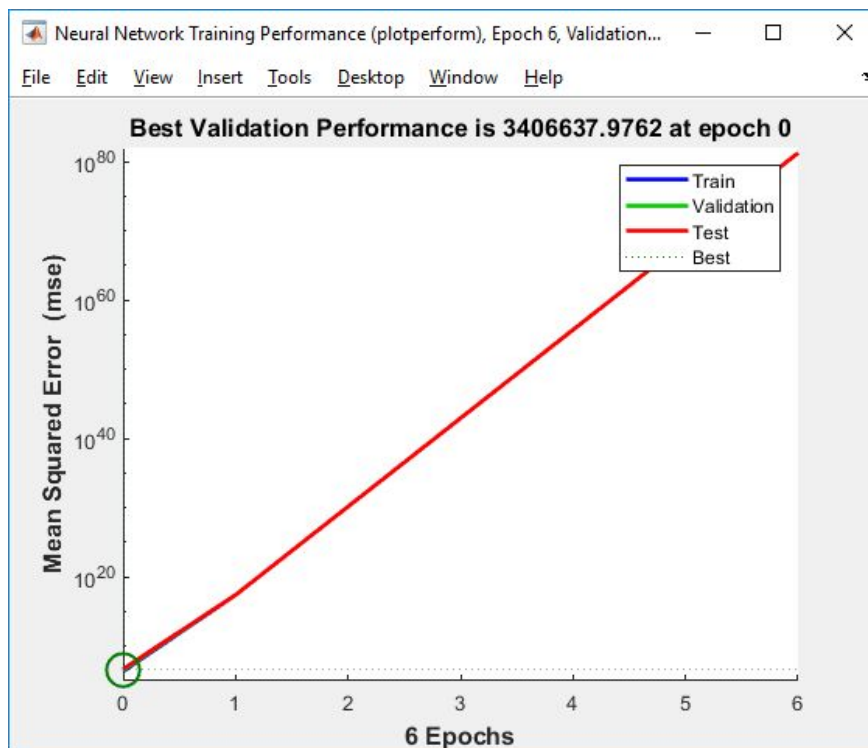
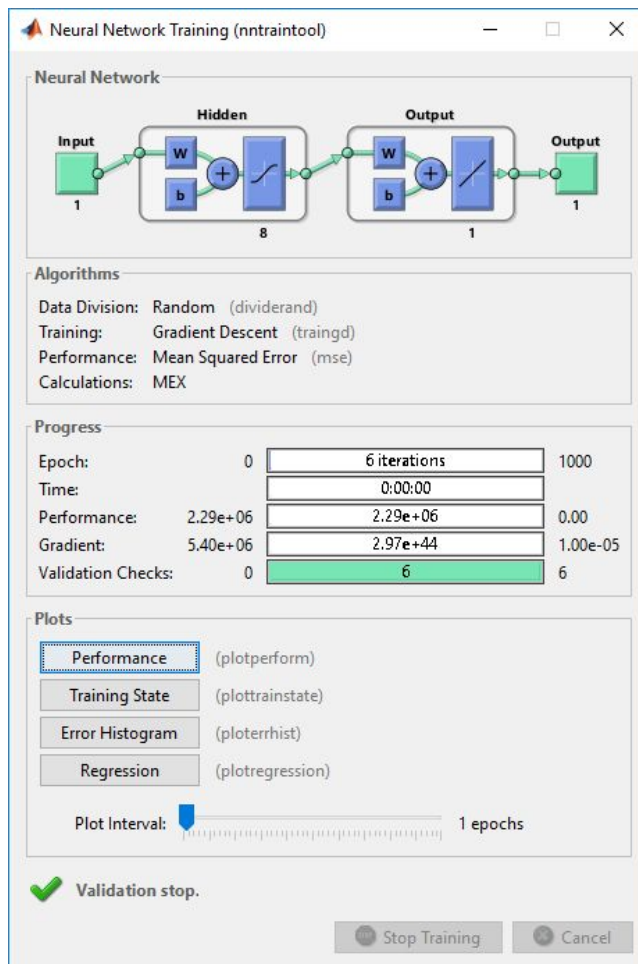
Sieci wielowarstwowe to połączone ze sobą warstwy neuronów. Posiadają one conajmniej dwie warstwy: wyjściową i wejściową. Do tego mogą dojść warstwy ukryte, które są między nimi. Ilość neuronów w warstwach nie musi być taka sama. Odpowiedni warstwy mogą łączyć się tylko z sąsiednimi warstwami. Neurony w sąsiednich warstwach są połączone na zasadzie każdy z każdym.

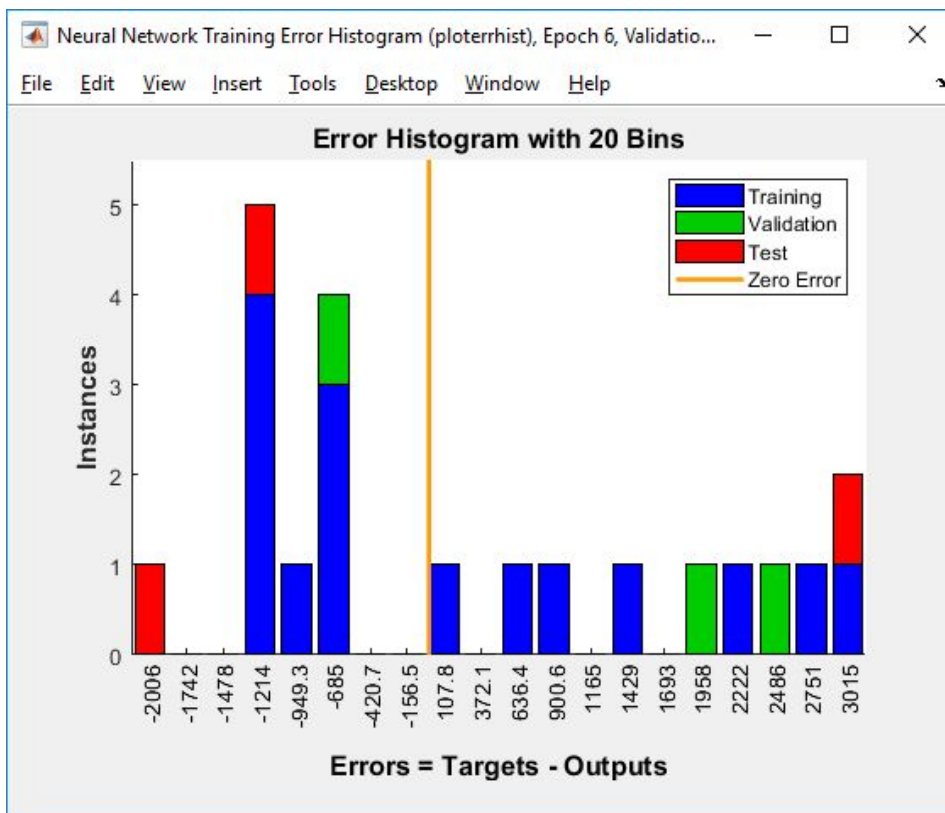
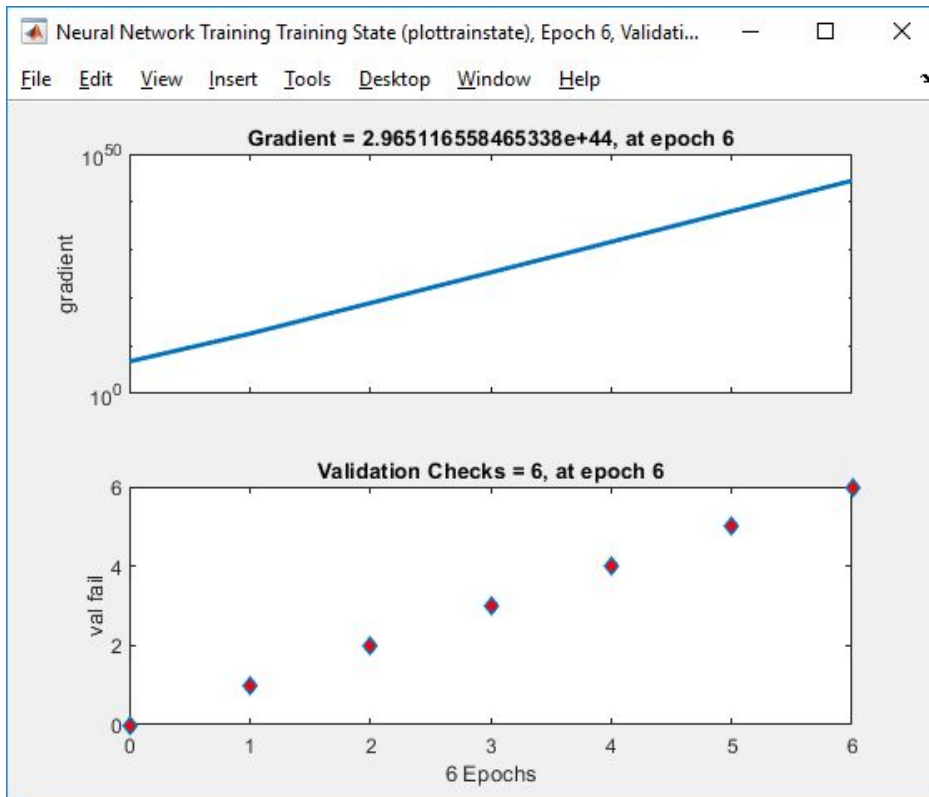
Algorytm wczesnej propagacji błędów określa strategię doboru wag w sieci wielowarstwowej. Wykorzystuje w tym celu gradientowe metody optymalizacji.

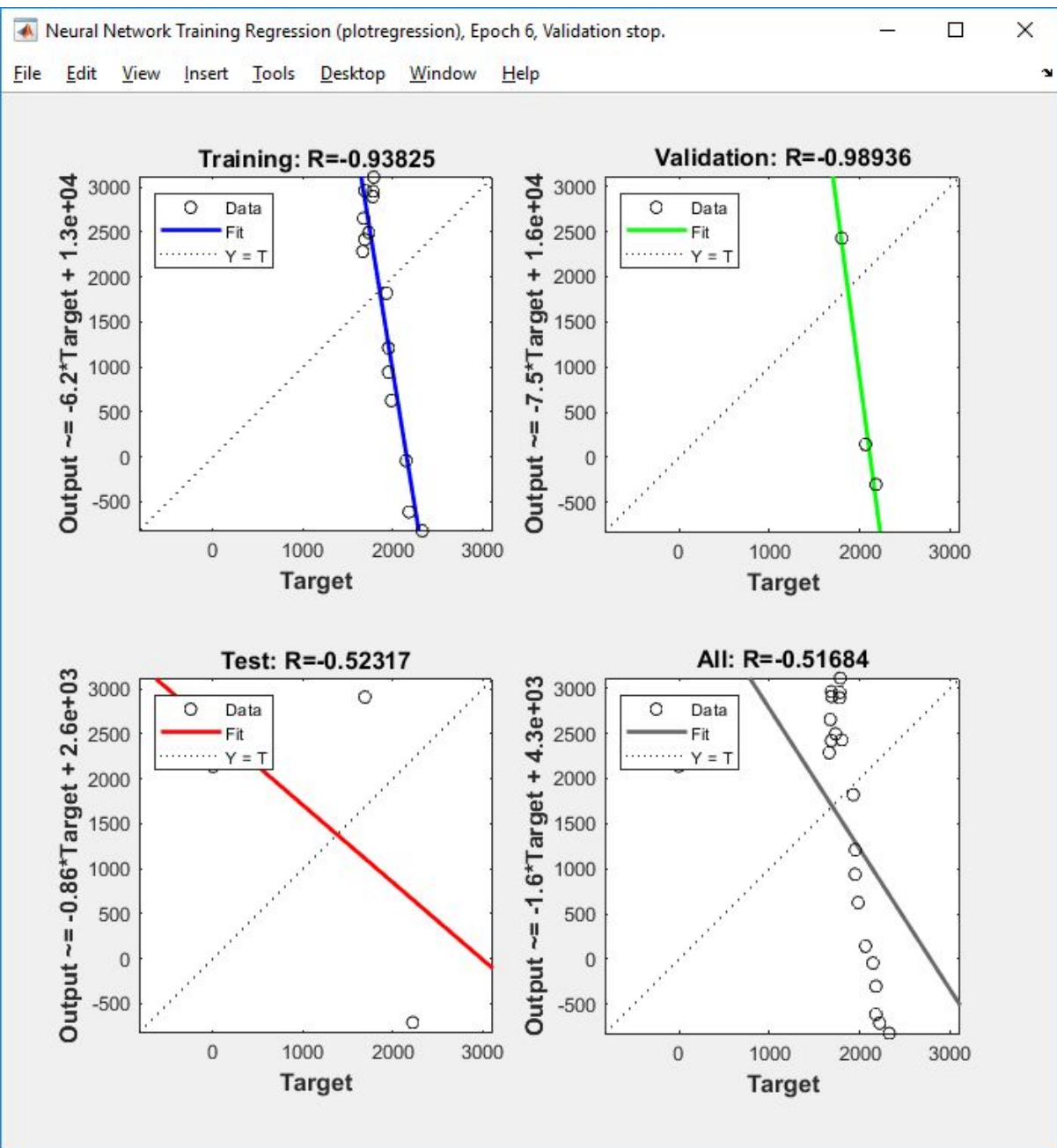
Nauka polega na doborze wag neuronów w taki sposób, by błąd sieci był mniejszy niż zadany błąd minimalny.

Algorytm ten charakteryzuje się tym, że najpierw obliczane są błędy w ostatniej warstwie. Błąd w dowolnej warstwie jest wyliczany na podstawie odpowiedniej funkcji błędów warstw poprzedzających i idzie to od ostatniej do pierwszej warstwy.

```
net.trainParam.lr = 0.1; net.trainParam.mc = 0.5;
```

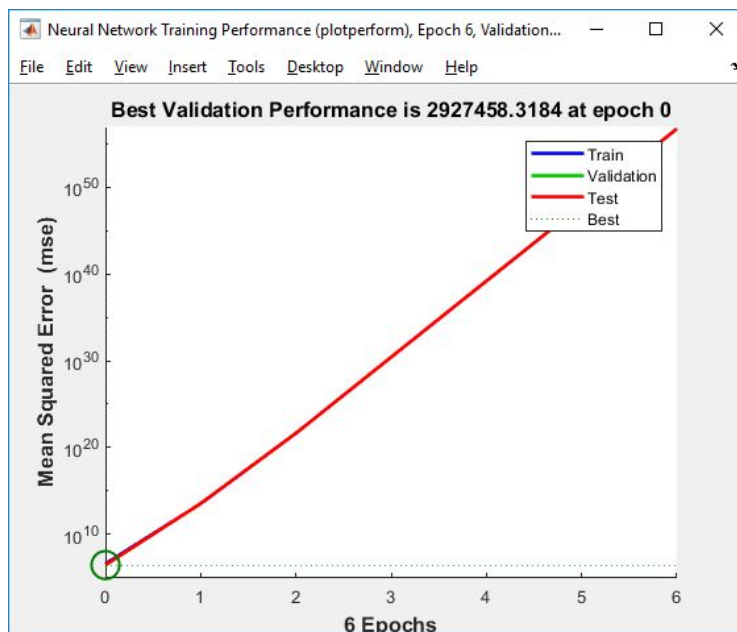
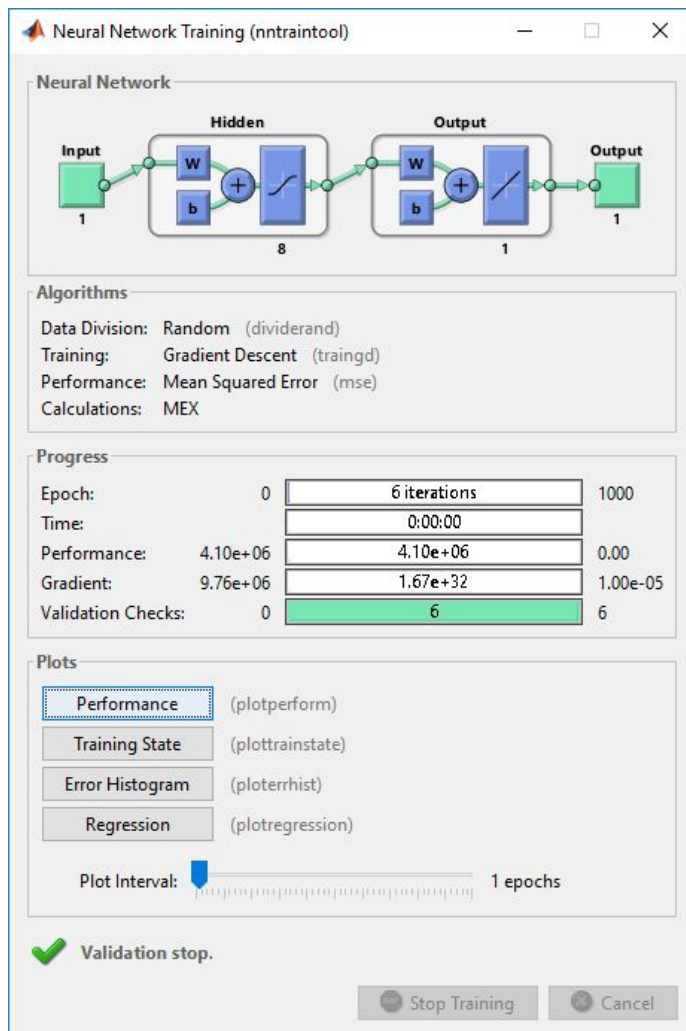


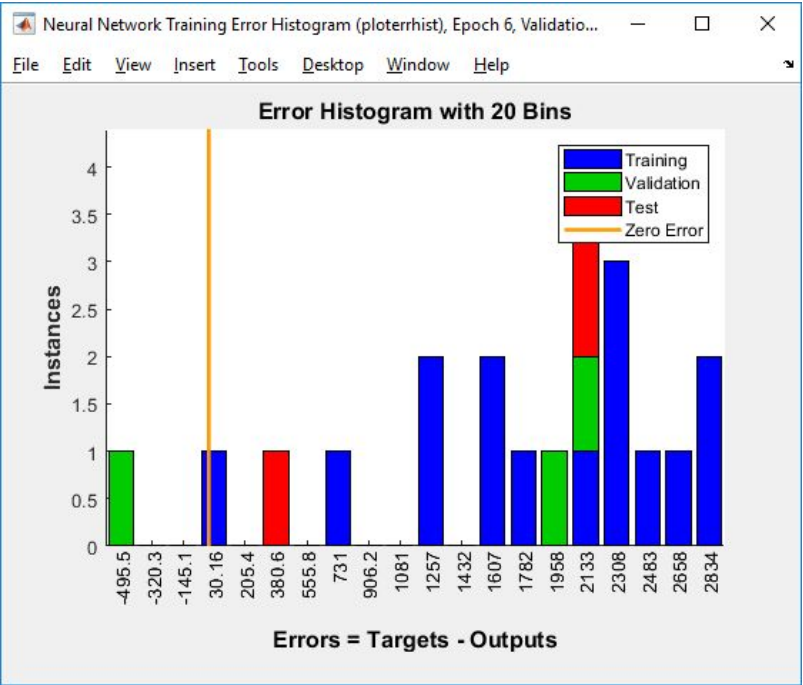
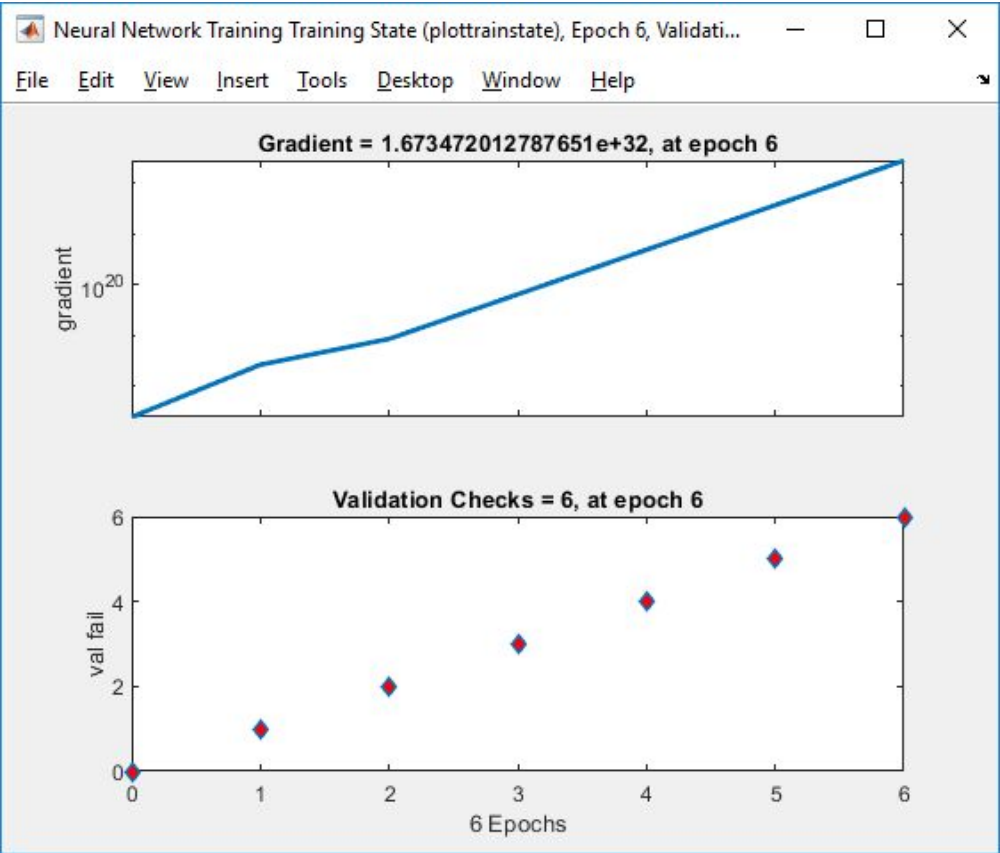


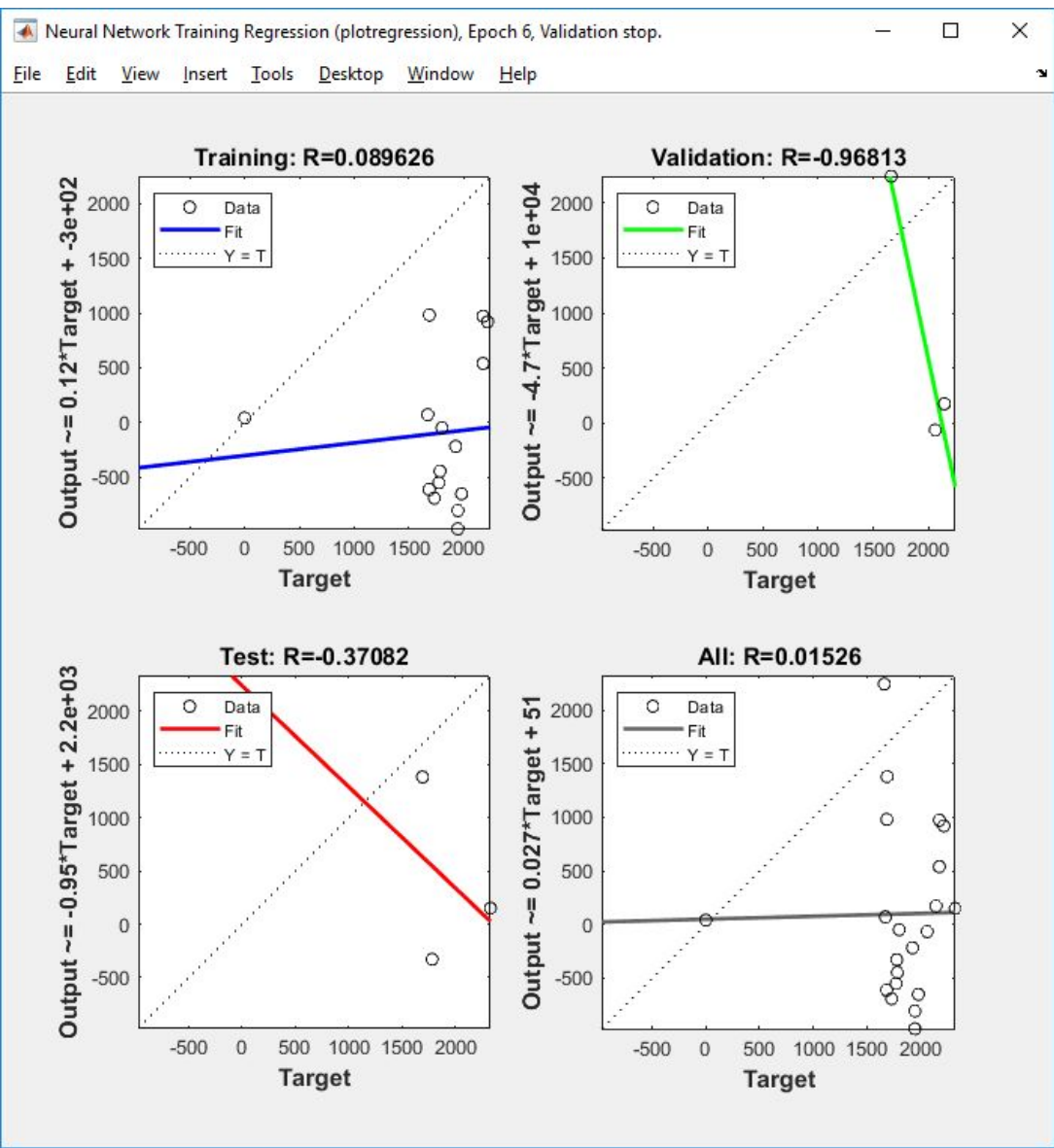


net.trainParam.lr = 0.001;

net.trainParam.mc = 0.0;

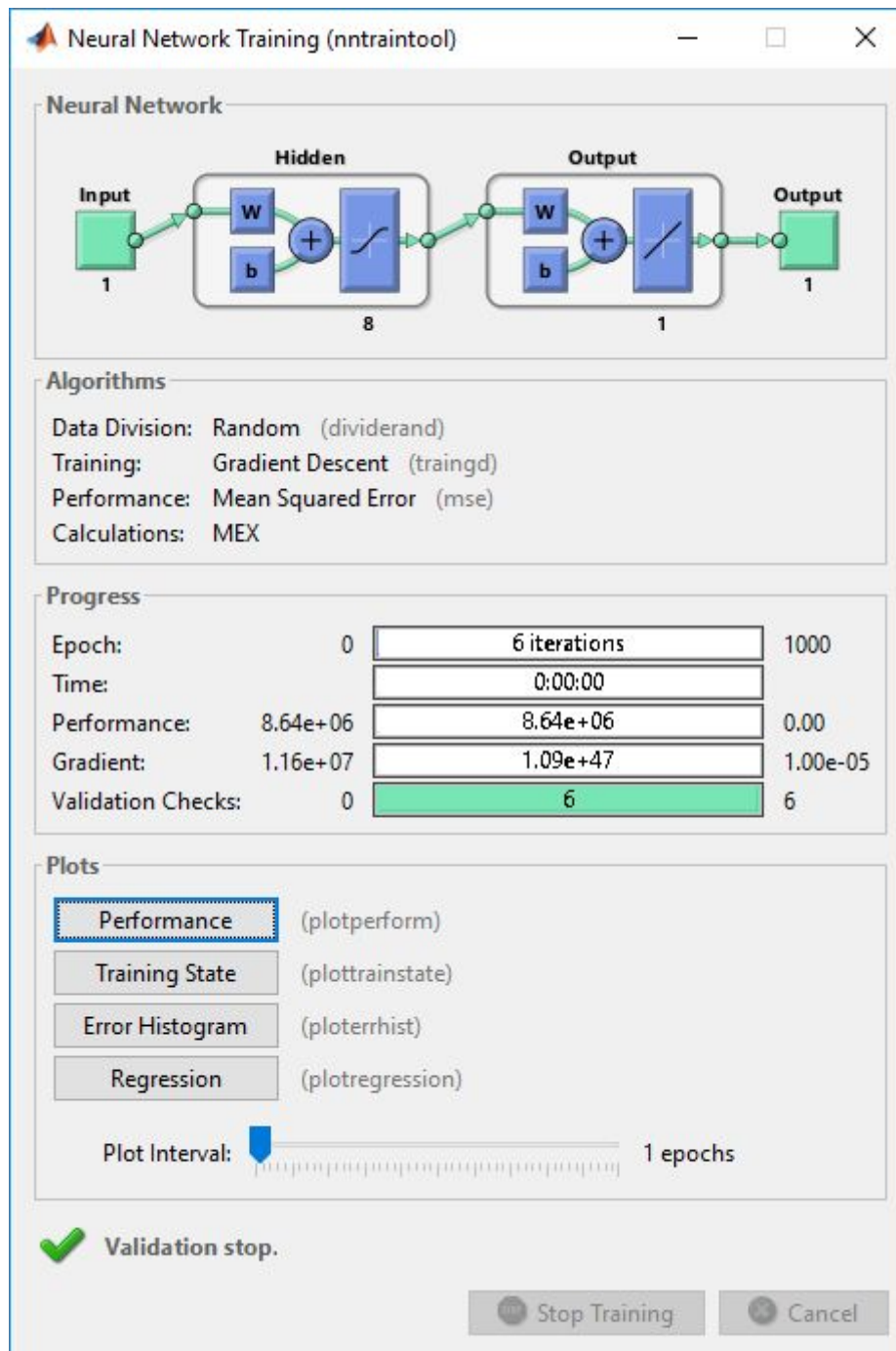




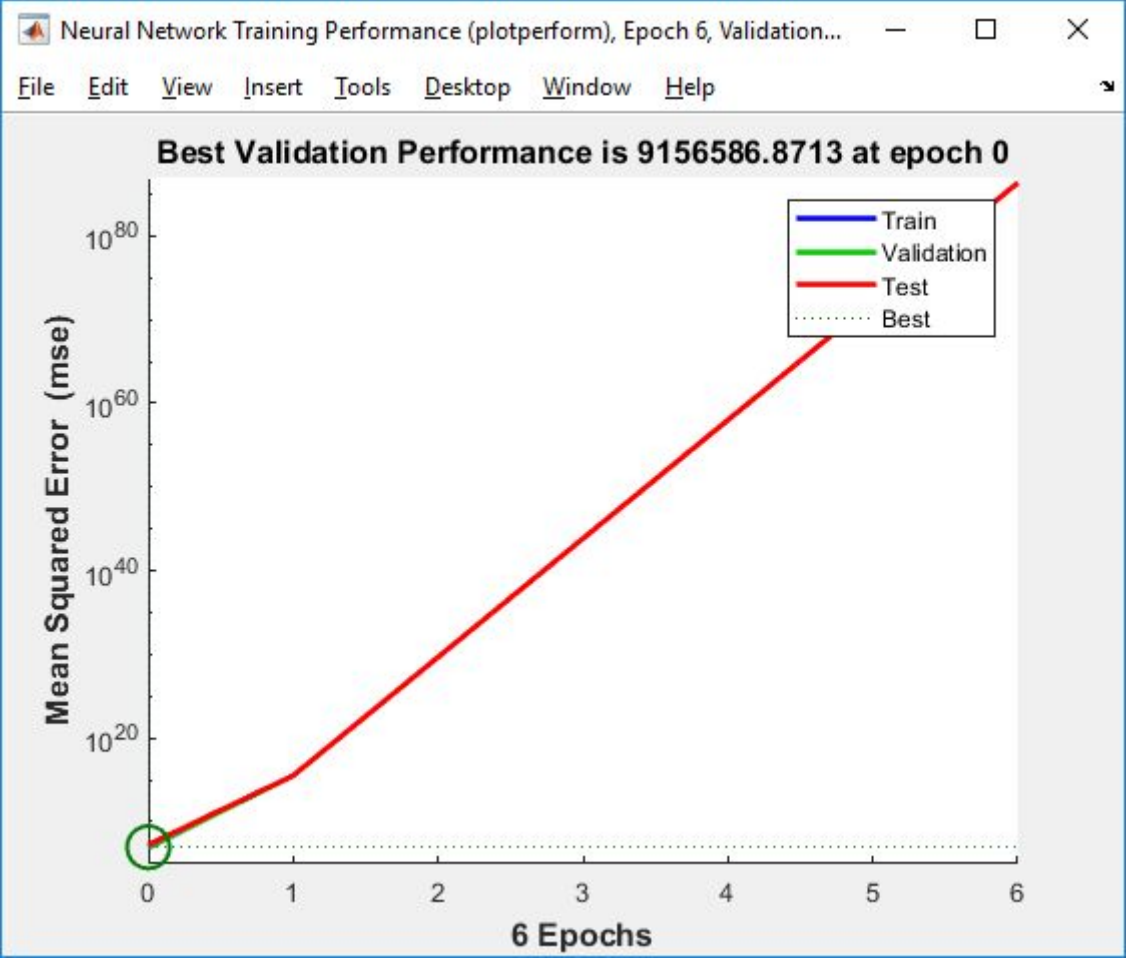


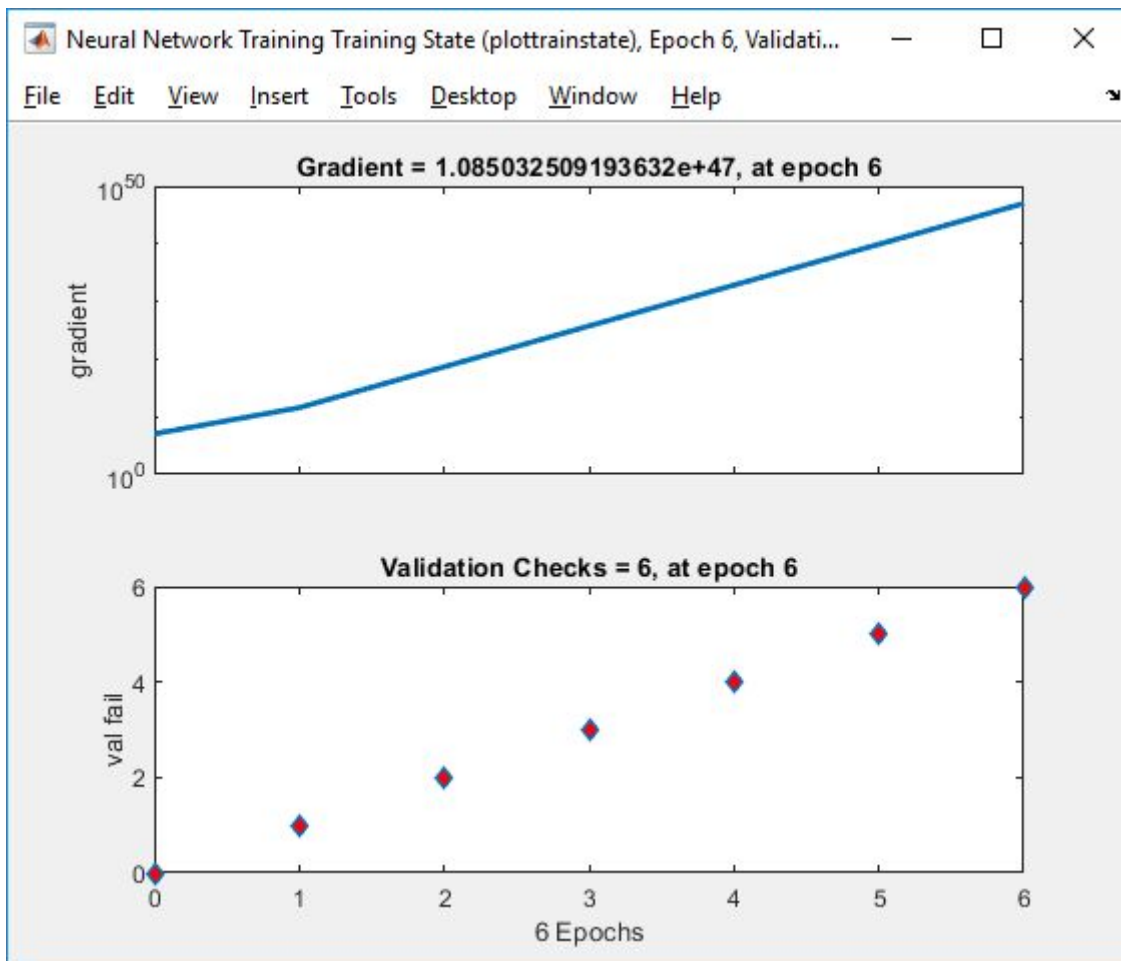


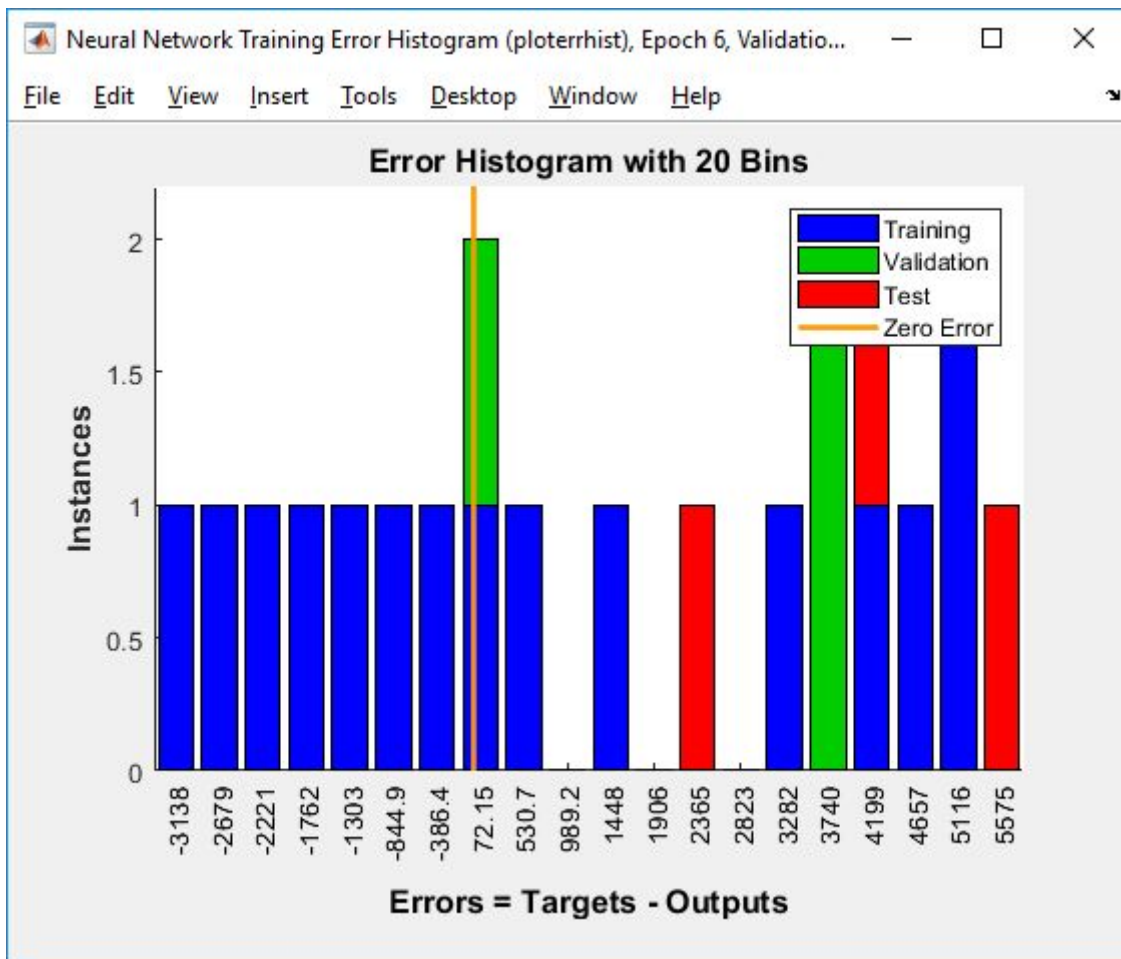
```
net.trainParam.lr = 0.5; net.trainParam.mc = 1.0;
```

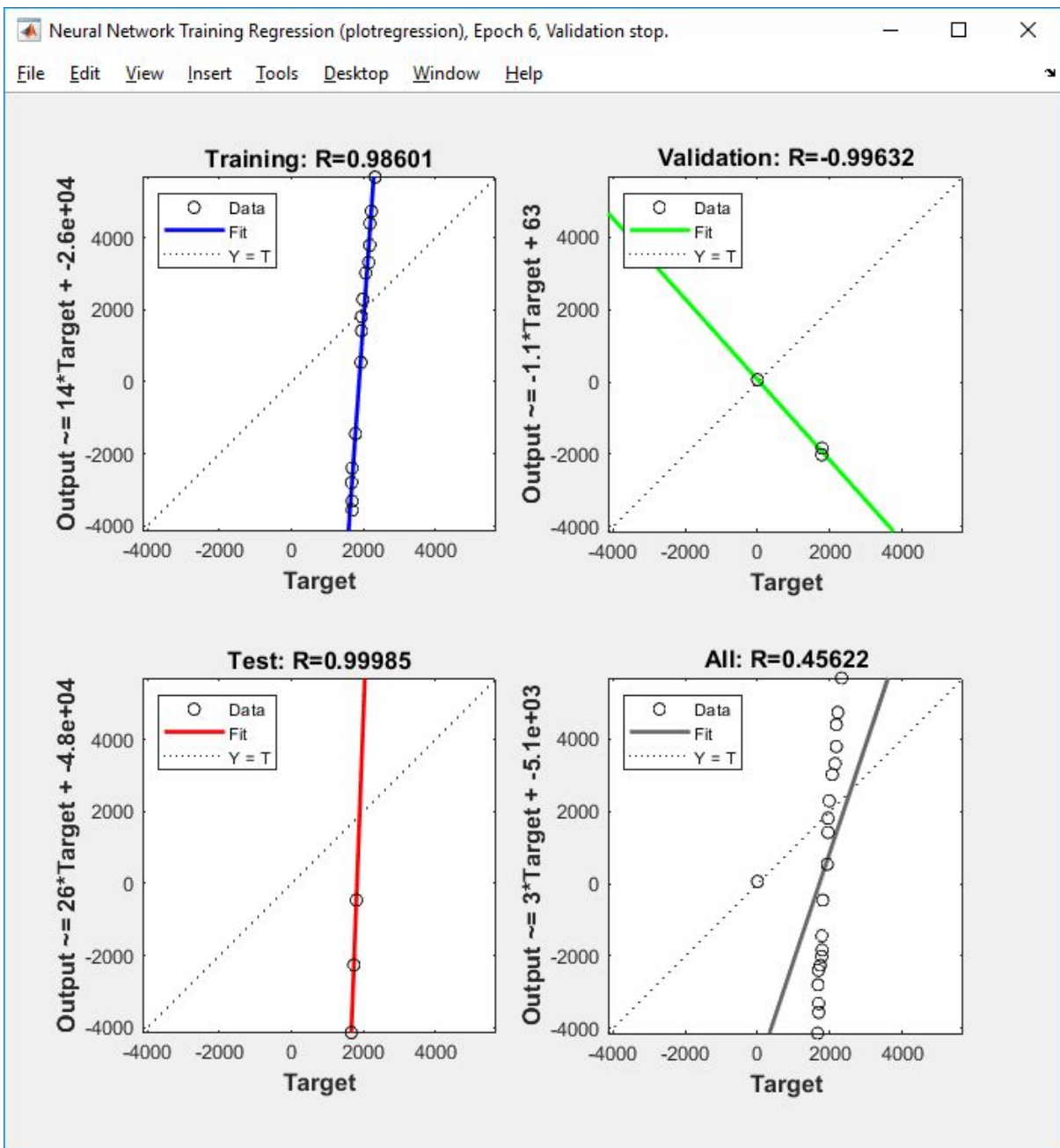




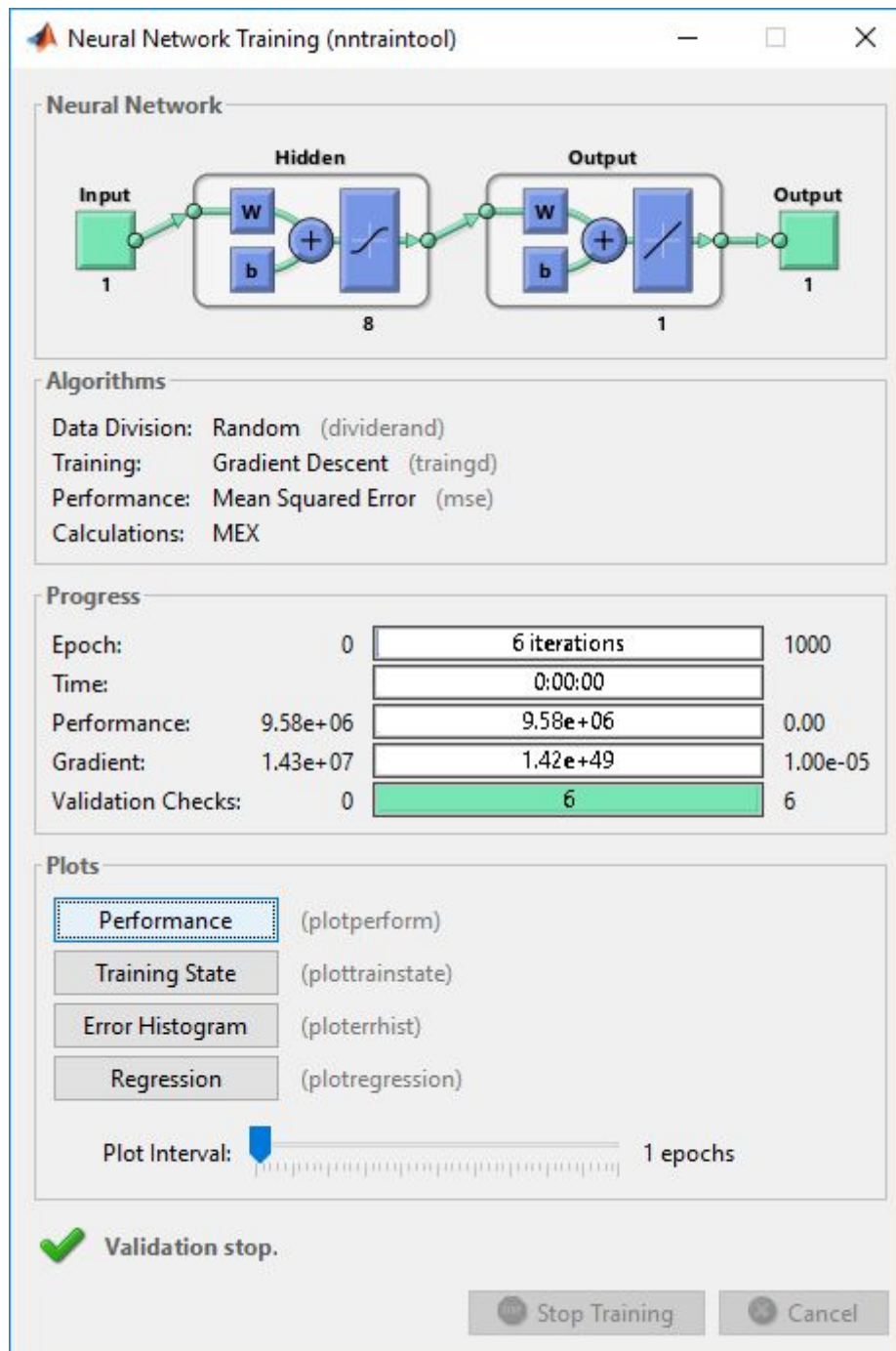


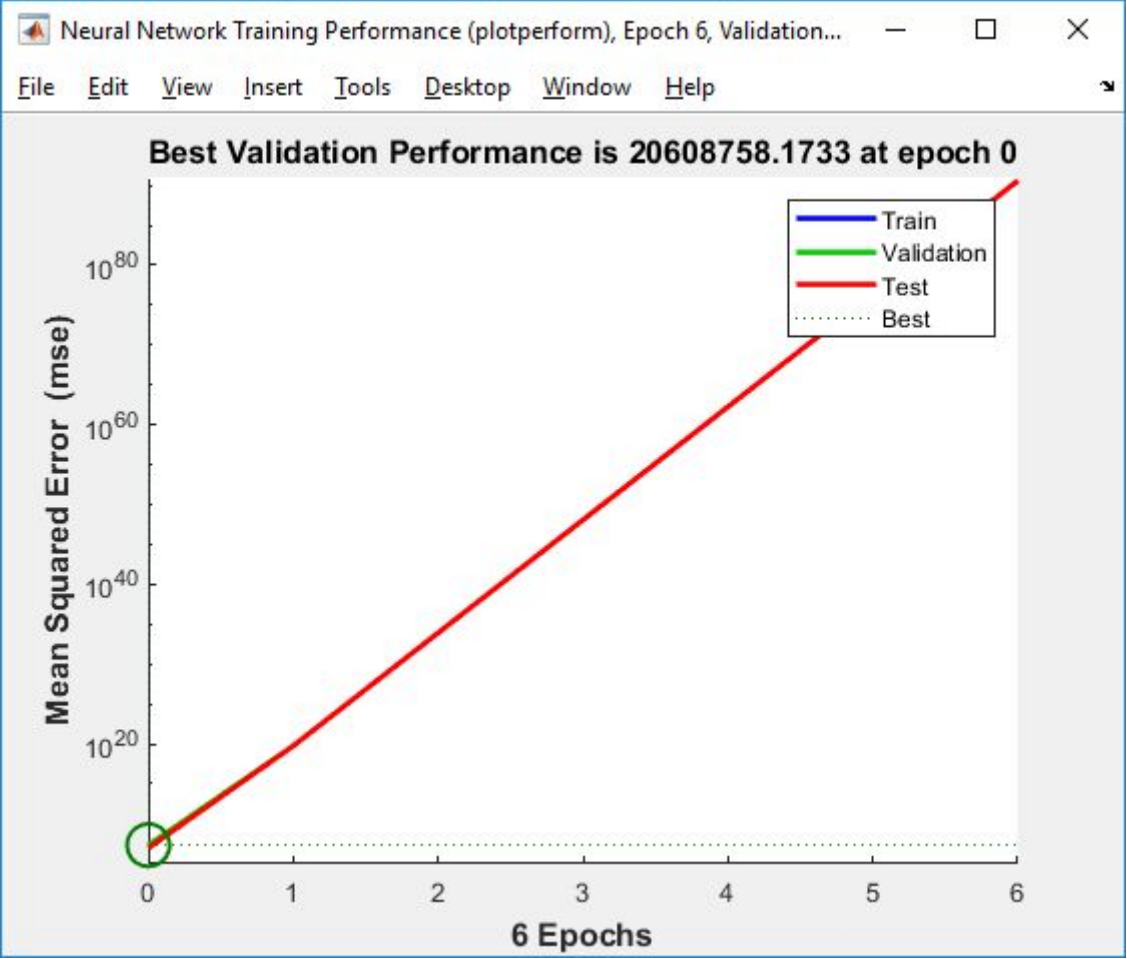


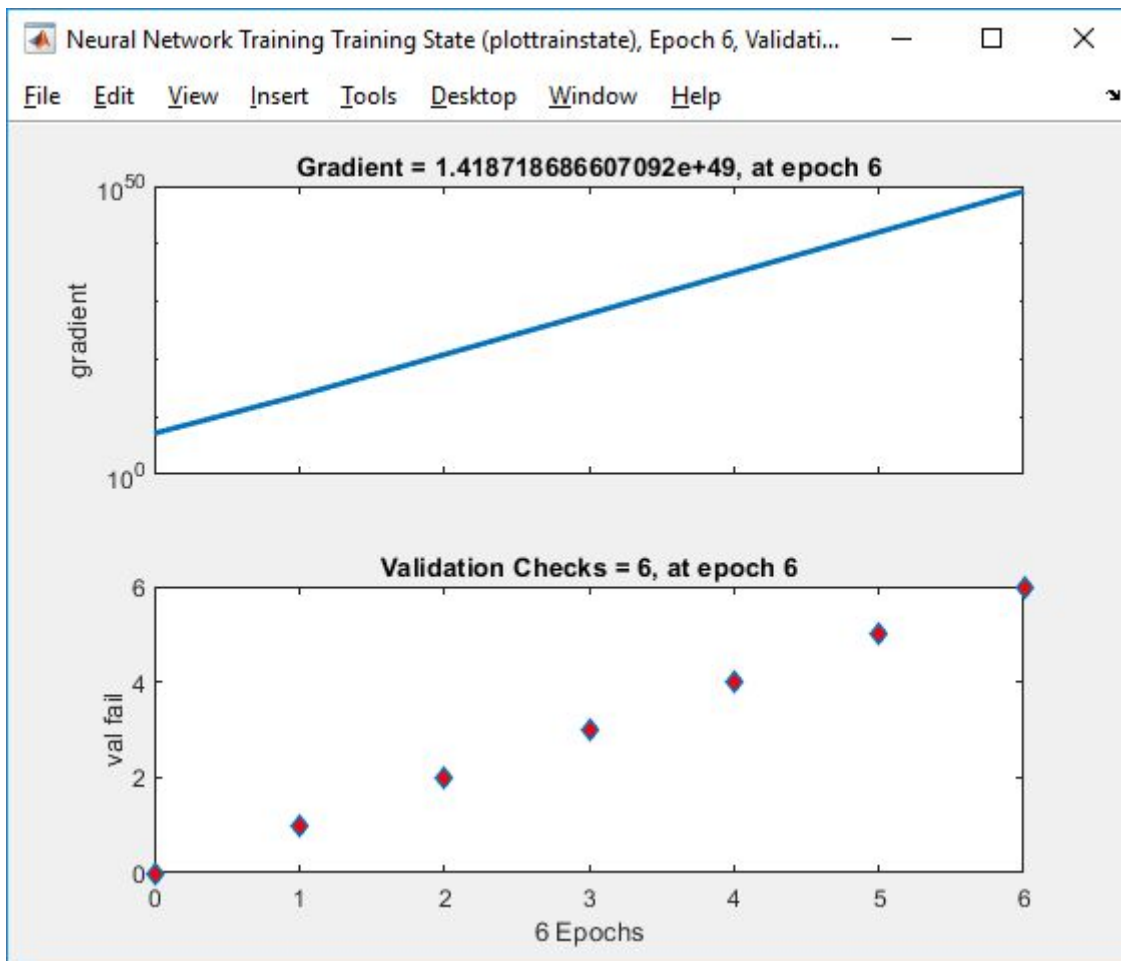




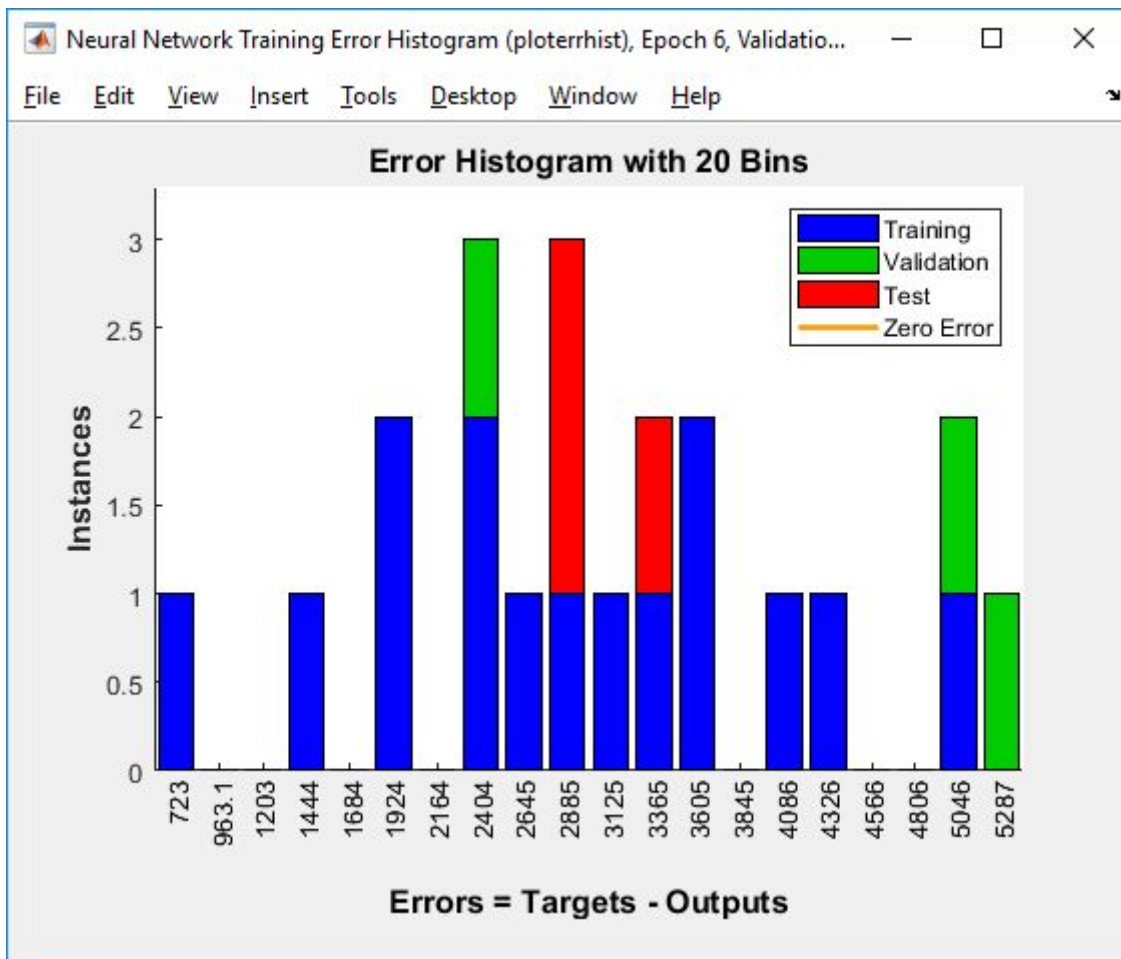
net.trainParam.lr = 0.5; net.trainParam.mc = 0.5;

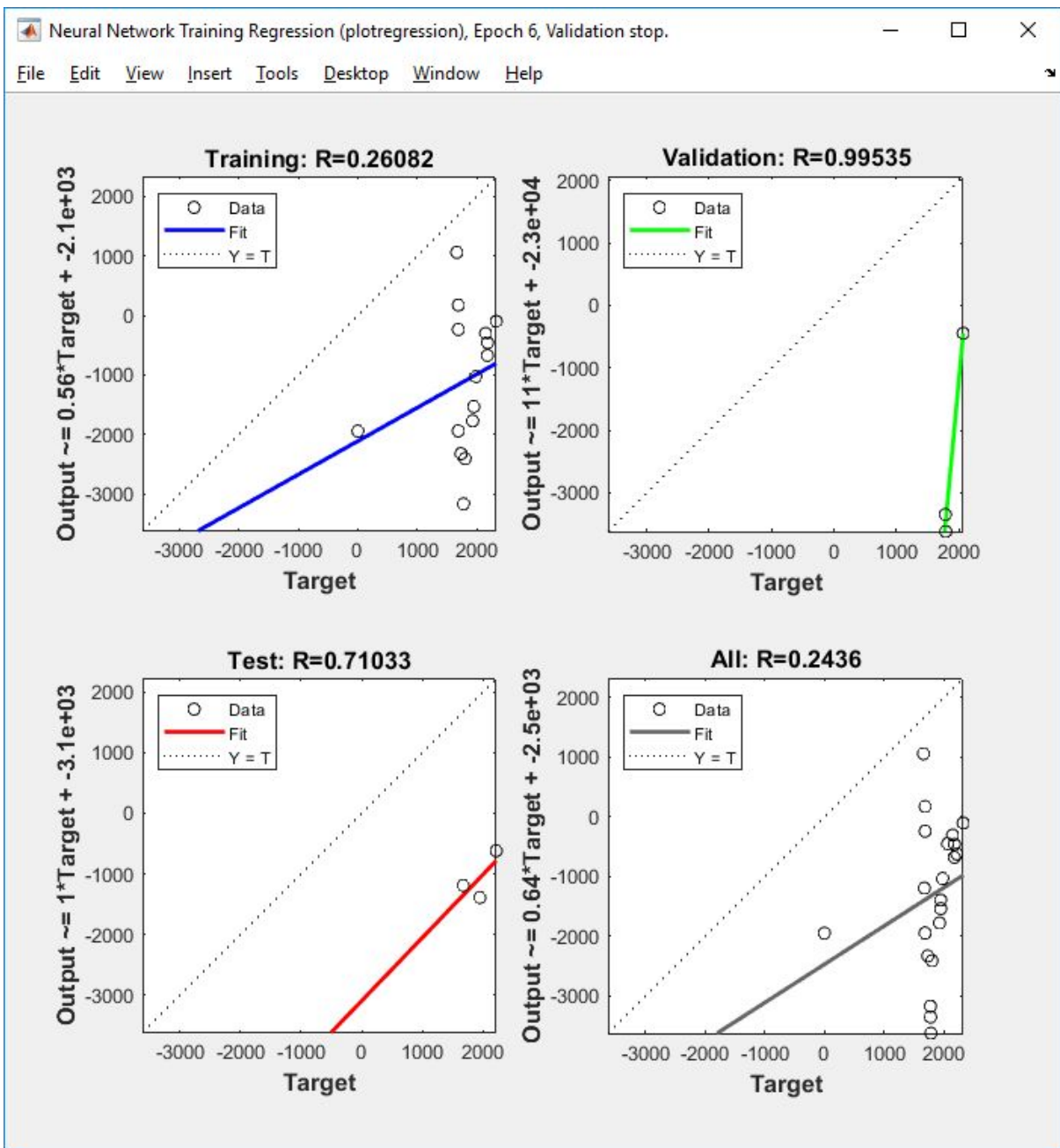




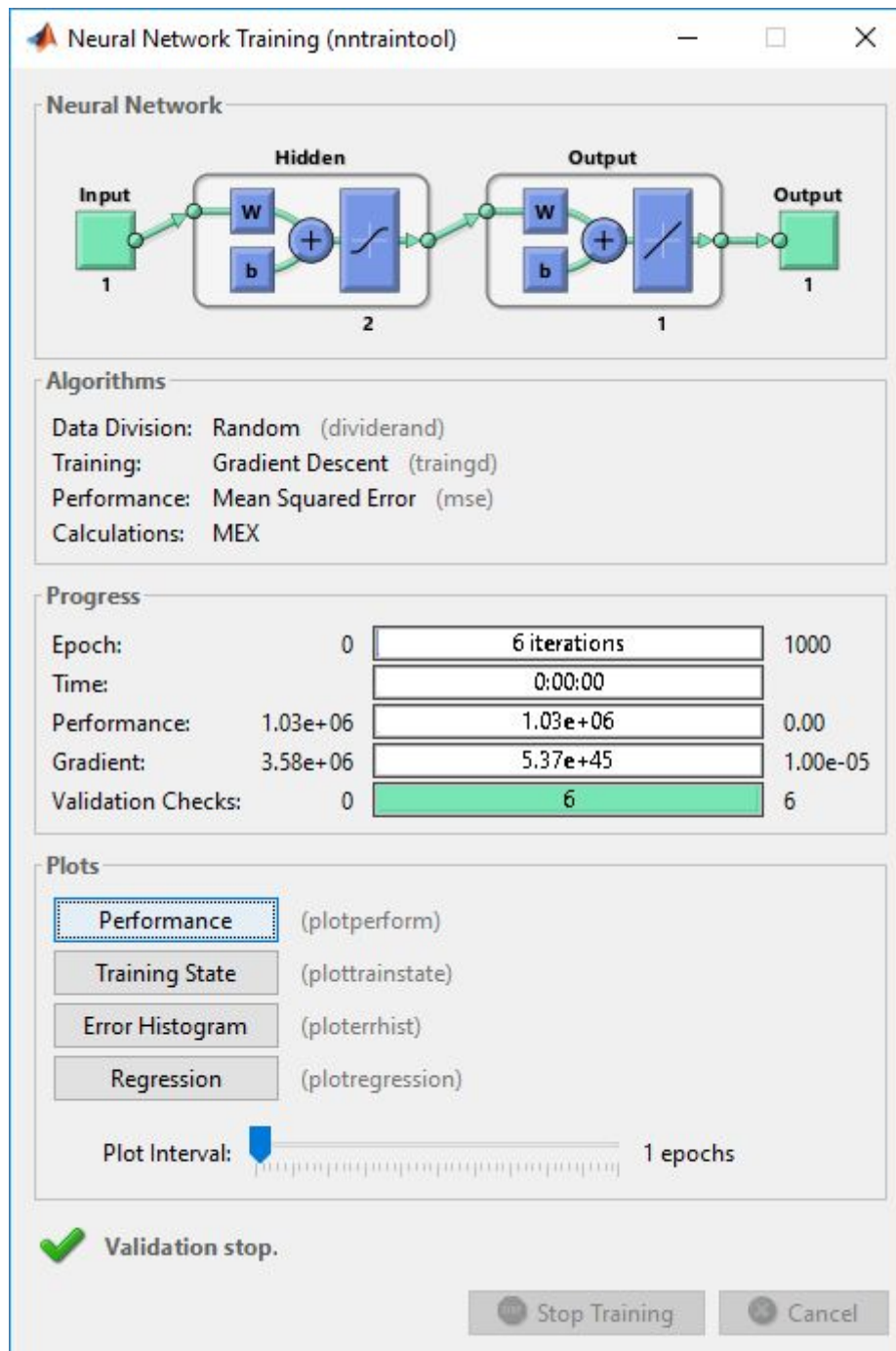


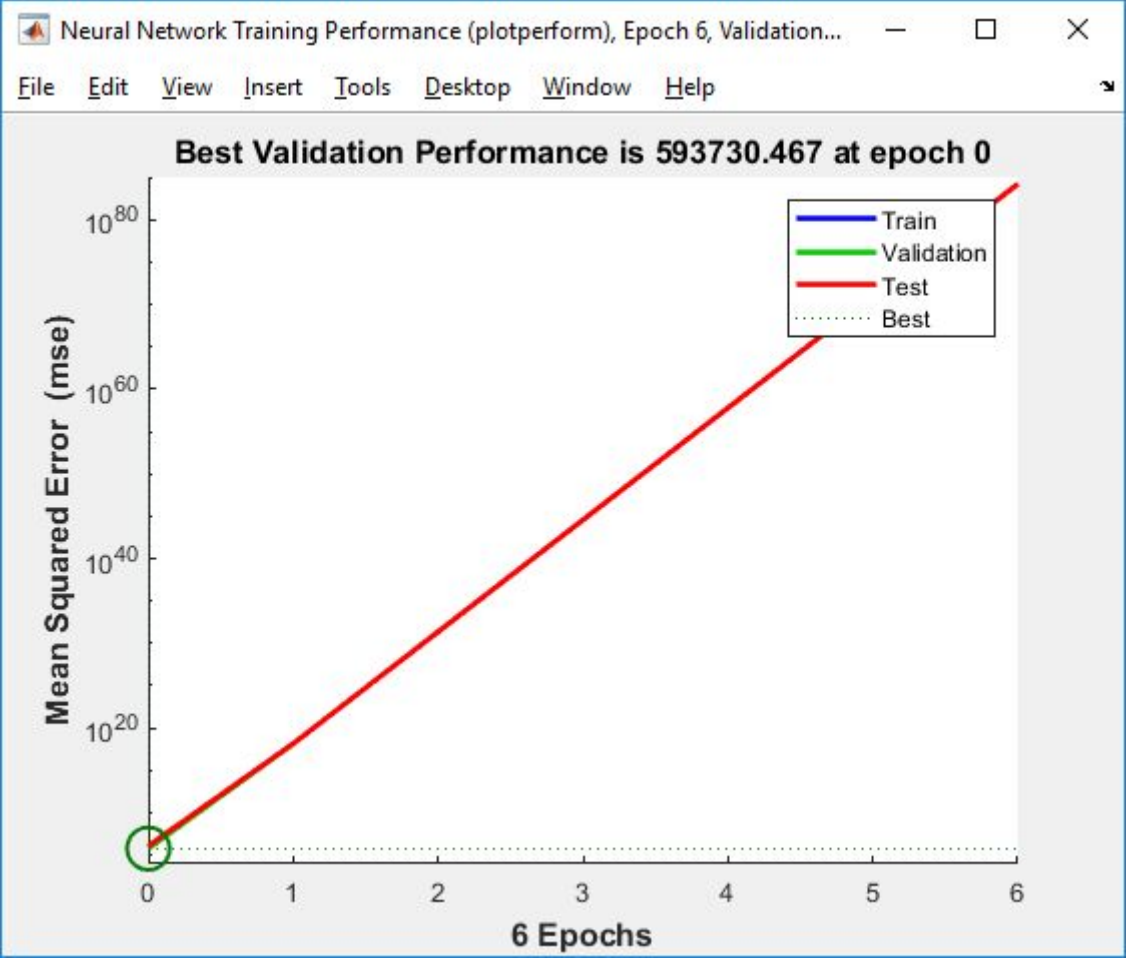


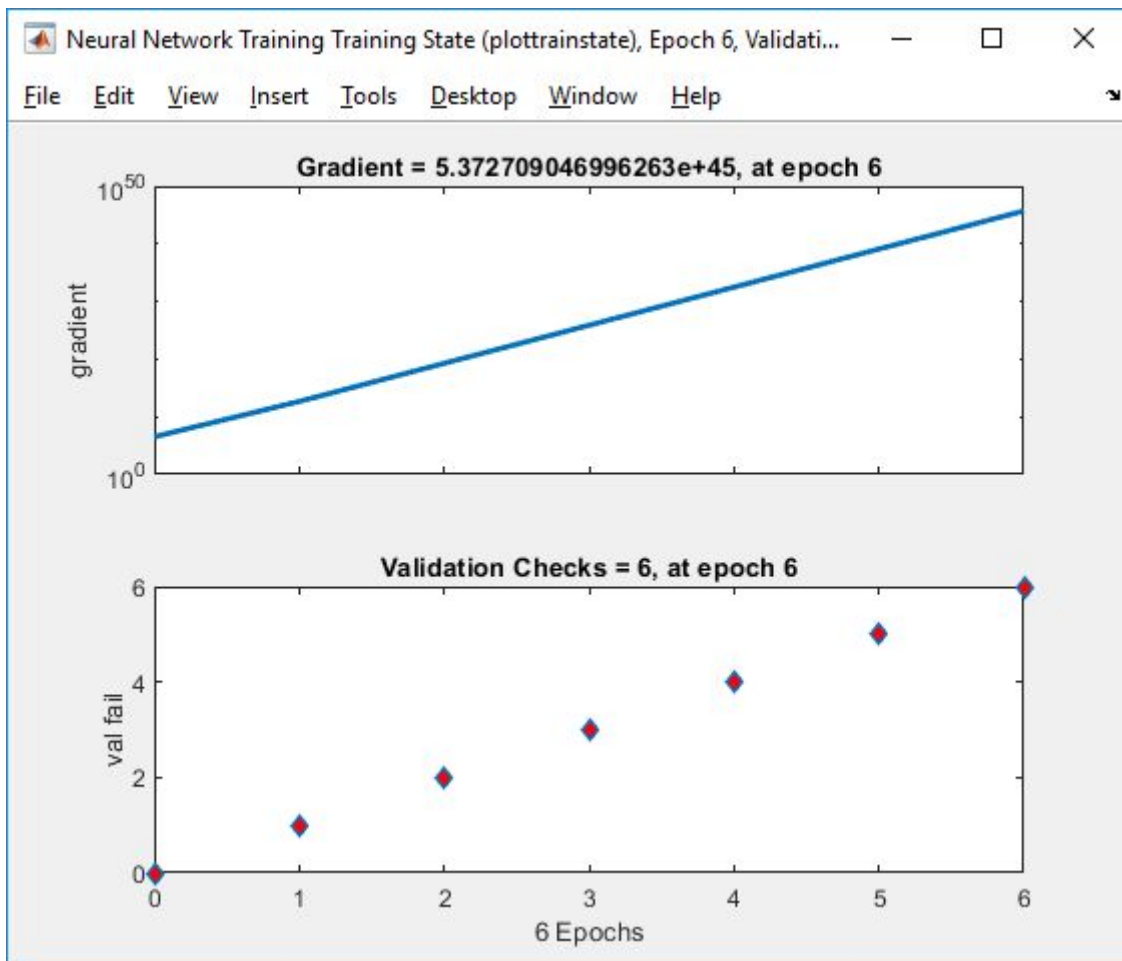


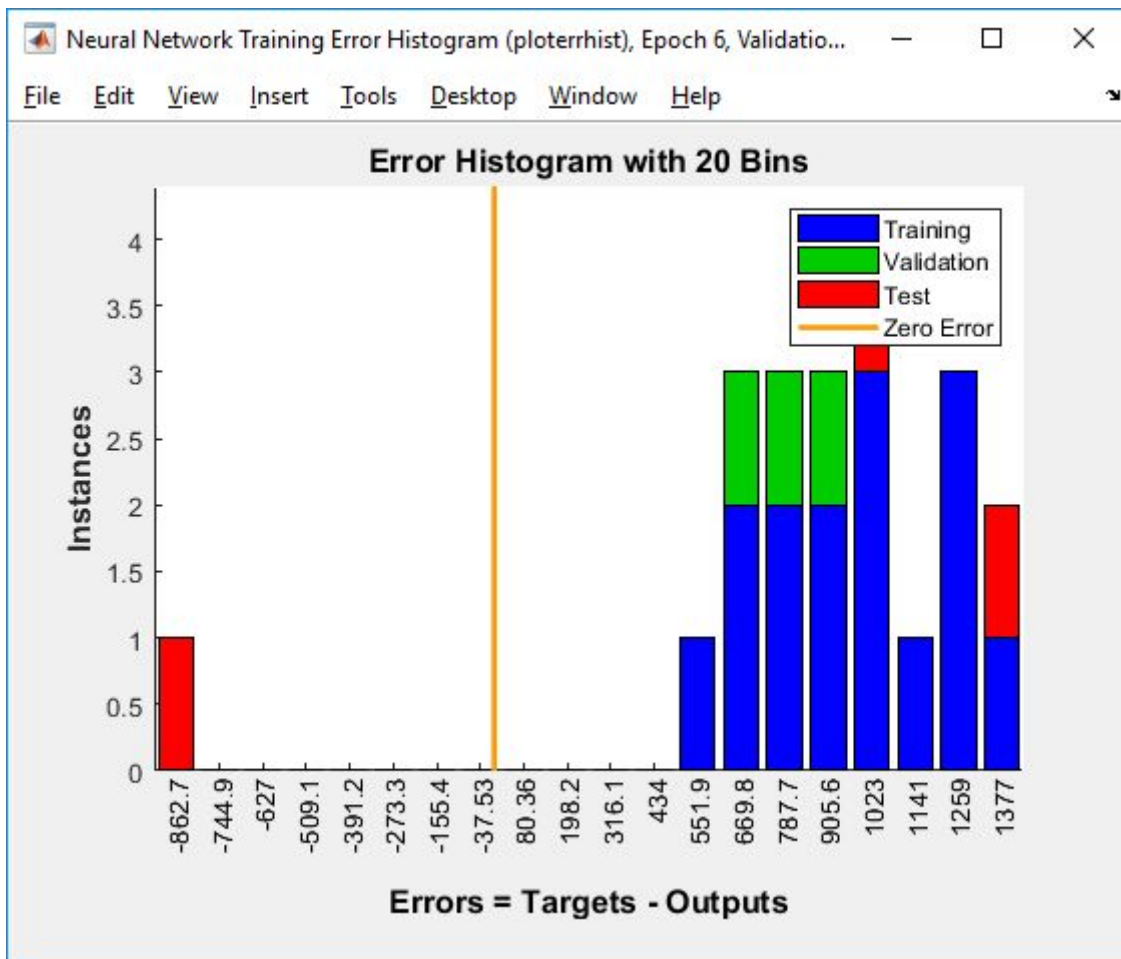


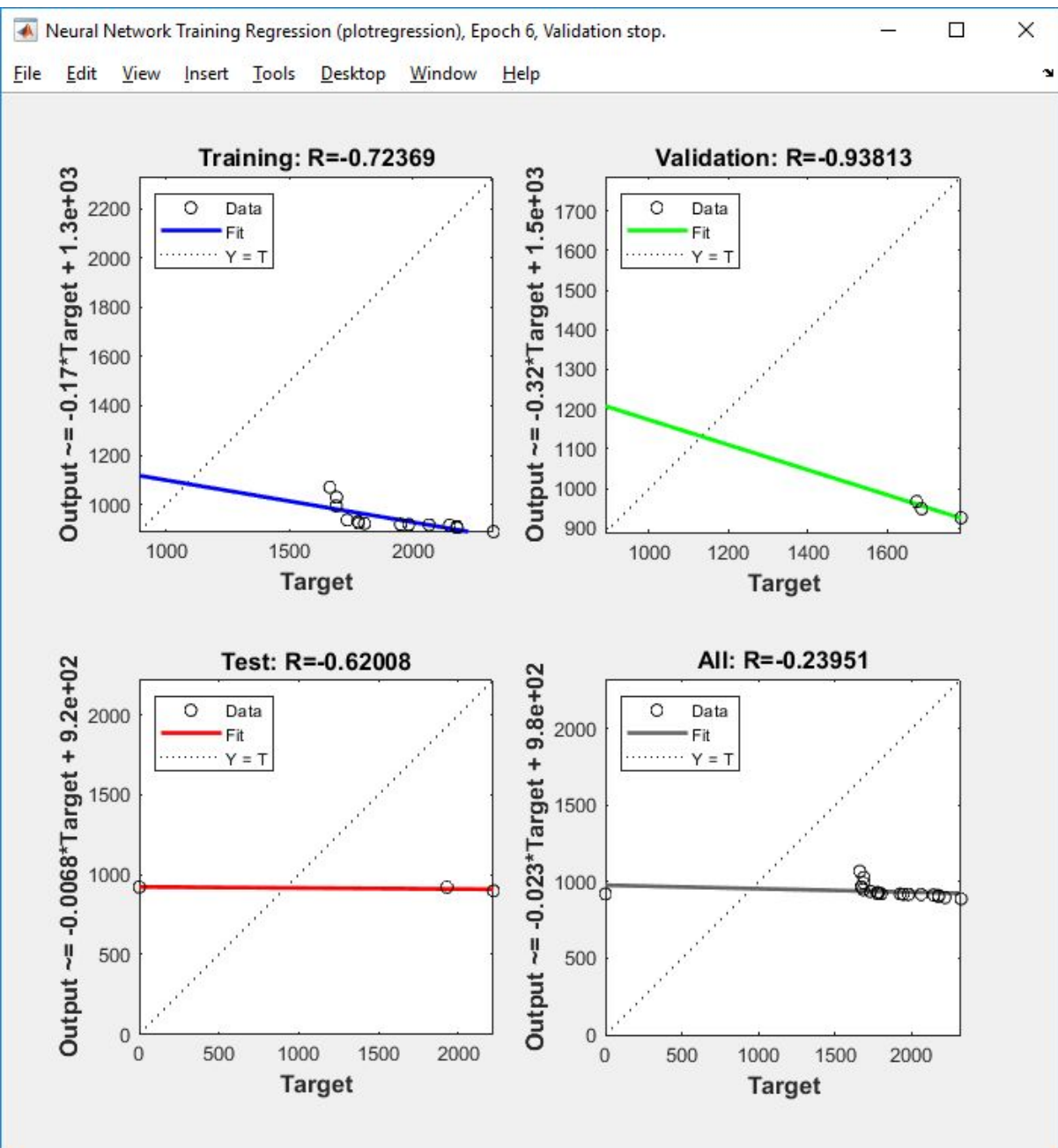
```
net.trainParam.lr = 0.5; net.trainParam.mc = 0.5;
```













## Wnioski

Najlepsze wyniki zostały uzyskane w momencie gdy współczynnik uczenia był różny od współczynnika bezwładności. Sieć wielowarstwowa bez algorytmu propagacji błędów działała lepiej niż sieci z wspomnianym algorytmem. Błędy, które na wejściu i w trakcie działania wydają się być pomijalne mogą na końcu spowodować duży rozrzut między tym, co zostało zasymulowane a tym, co powinna dać sieć w swoim wyniku. Sieci wielowarstwowe z algorytmem propagacji danych wykazują duże odsetki błędów. Im większy jest wsp. uczenia od bezwładności, tym częściej mogą zdarzać się błędy. Algorytmy z propagacją danych przyspieszają proces uczenia nawet o połowę epok w stosunku do sieci nie zawierających tego algorytmu. Gradient uczenia (z algorytmem propagacji błędów) posiada charakterystykę niemal że liniową, dzięki czemu można łatwiej przewidzieć, jakich współczynników do nauki sieci najlepiej użyć, by otrzymane rezultaty były zadowalające. Jeżeli nie korzystamy z algorytmu propagacji danych należy pamiętać, by były duże różnice między wsp. uczenia a bezwładnością.