

# Sprawozdanie lab 5

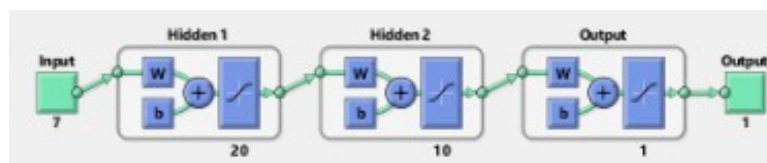
Metody inteligencji obliczeniowej – Informatyka Stosowana, WFIIS,  
Jakub Salamon, II rok

Celem zajęć laboratoryjnych nr 5 była predykcja szeregów czasowych z użyciem wielowarstwowych sieci neuronowych

Naszym zadaniem było skonstruowanie sieci neuronowej opartej na trzech warstwach, w tym dwóch ukrytych. Miała ona służyć do przewidywania notowań wartości indeksów giełdowych. Jako dane mieliśmy dostępne historyczne notowania od 19.11.2004 do 18.12.2018 roku jednego z azjatyckich indeksów.

Funkcją aktywacji dla każdej warstwy była funkcja tansig. Maksymalna liczba epok uczenia wynosiła 3000. Danymi wejściowymi były wartości zamknięcia giełdy. Były one przeskalowywane do przedziału  $[-0.5, 0.5]$  w celu uzyskania lepszych rezultatów.

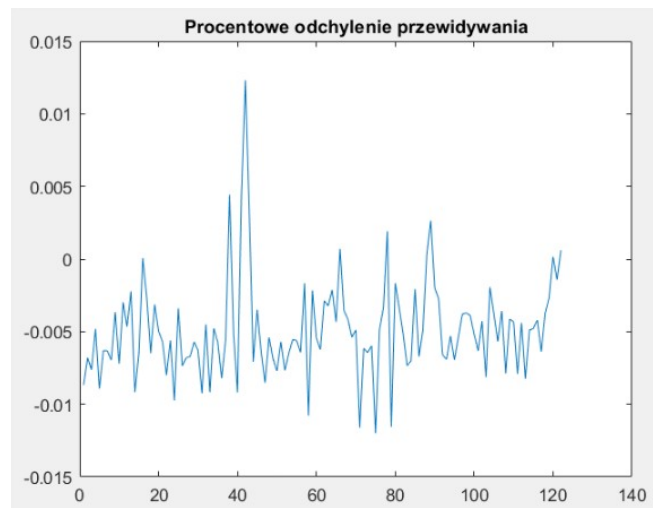
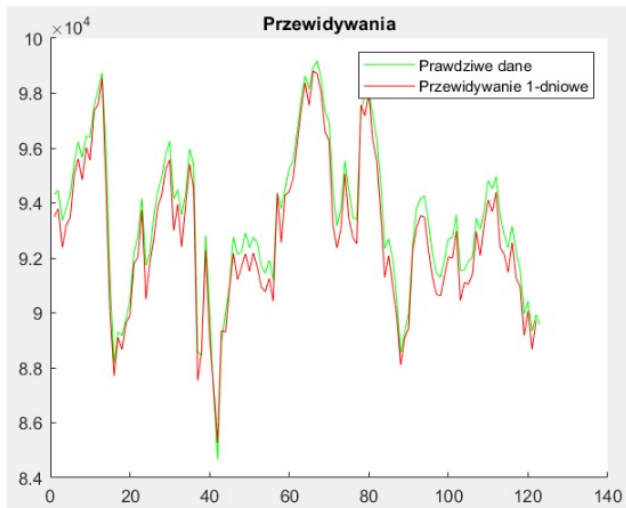
Schemat takiej sieci:



Przetestujemy działanie sieci dla różnej ilości wejść oraz różnych liczb neuronów w warstwach ukrytych.

Sprawdźmy dla jakiej liczby neuronów ukrytych uzyskamy najmniejszy błąd dla danych uczących oraz testowych. Danymi wejściowymi są wartości zamknięcia giełdy od początku 2017 roku do końca 2018 roku. Około półtora roku to dane uczące, a reszta testujące. Sieć przewiduje wartość akcji w następnym dniu na podstawie 5 poprzednich dni. Po wyliczeniu przewidywania sieć jest douczana na podstawie prawdziwej wartości w danym dniu. Proces powtarza się, aż do wyczerpania danych wejściowych. Przy określeniu jakości rozwiązania przyda się nam wykres procentowych odchyleń przewidzianych wartości akcji od wartości prawdziwych. Kolejnym wskaźnikiem jakości będzie współczynnik dobrze przewidzianych wzrostu i spadków cen akcji, który mówi nam czy w następnym dniu zarobimy czy stracimy.

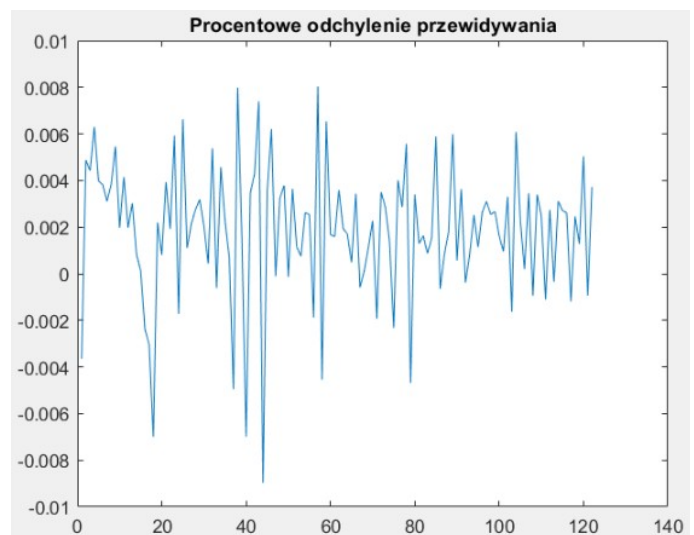
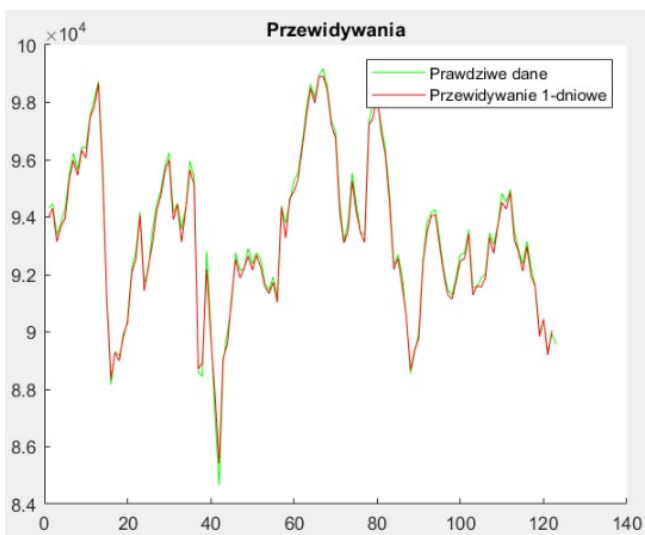
### 1) 1 neuron ukryty w I warstwie, 1 neuron ukryty w II warstwie



Błąd średniowarstwowy danych uczących 0.00026461  
Błąd średniowarstwowy dane testujące 309128.0574  
Średnie odchylenie przewidywań 0.0053989%  
Dobrze przewidziane wzrostu i spadki cen akcji 82.7869%

Taka sieć jest bardzo prosta, za prosta do takiego zadania. Już na wykresie przewidywań widzimy, że wartości akcji mało kiedy są do siebie zbliżone. Wykres odchylenia przewidywania rzadko pokazuje zbliżanie się wartości przewidywań do prawdziwej wartości akcji. Przewidywania nie oscylują wokół właściwej wartości.

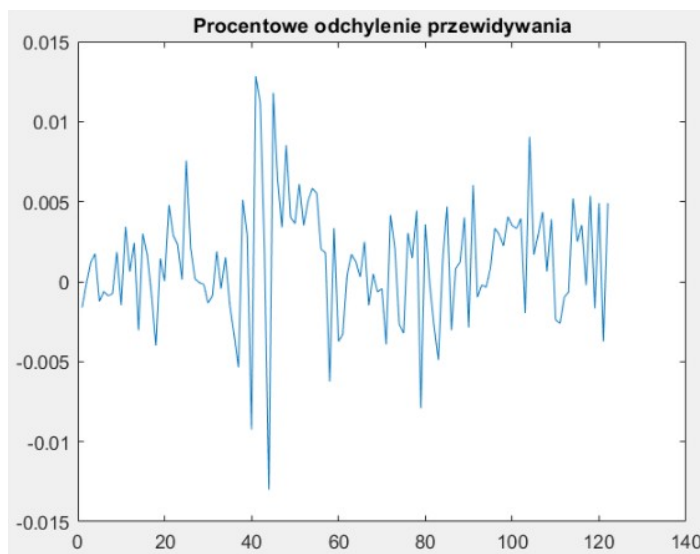
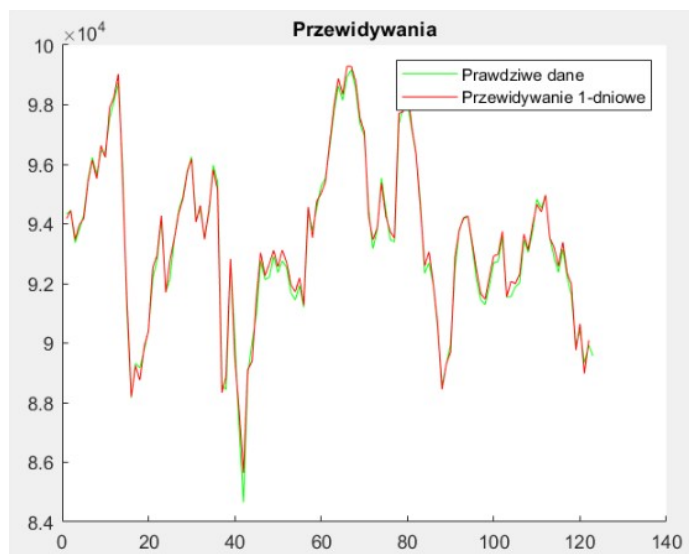
### 2) 1 neuron ukryty w I warstwie, 3 neurony ukryte w II warstwie



Błąd średniowarstwowy danych uczących 0.00026461  
Błąd średniowarstwowy dane testujące 104285.9117  
Średnie odchylenie przewidywań 0.0028828%  
Dobrze przewidziane wzrostu i spadki cen akcji 96.7213%

Błąd średniokwadratowy dla danych testujących jest prawie 3 razy mniejszy niż poprzednio. Odchylenie przewidywanej wartości oraz współczynnik dobrze przewidzianych możliwych zarobków lub strat również są lepsze.

### 3) 1 neuron ukryty w I warstwie, 7 neuronów ukrytych w II warstwie

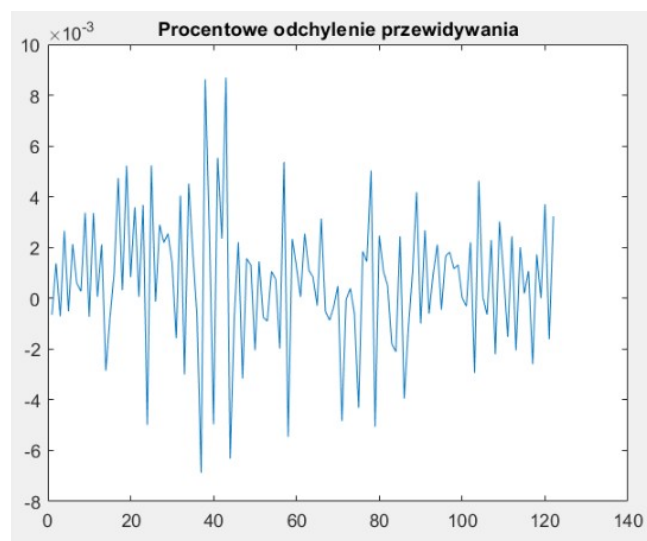
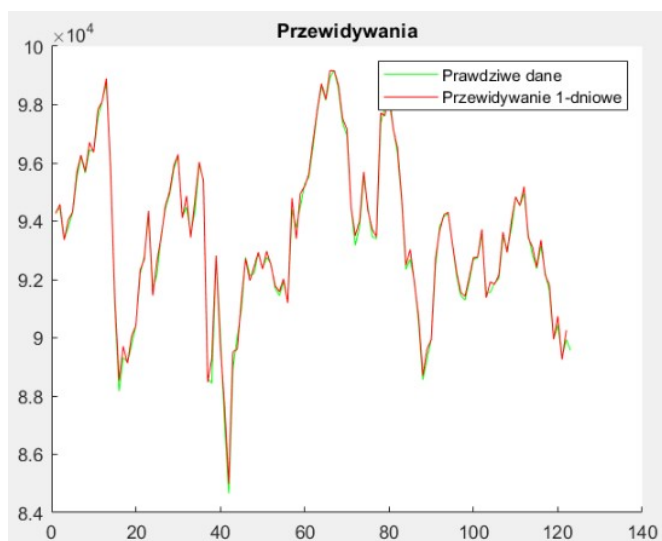


Błąd średniokwadratowy danych uczących  
Błąd średniokwadratowy dane testujące  
Średnie odchylenie przewidywań  
Dobrze przewidziane wzrostu i spadki cen akcji

0.00026461  
136394.3997  
0.0030747%  
93.4426%

Wyniki są gorsze niż poprzednio. Oscylacje odchylenia przewidywań są wokół prawdziwej wartości jednak są one większe niż poprzednio.

### 4) 1 neuron ukryty w I warstwie, 11 neuronów ukrytych w II warstwie

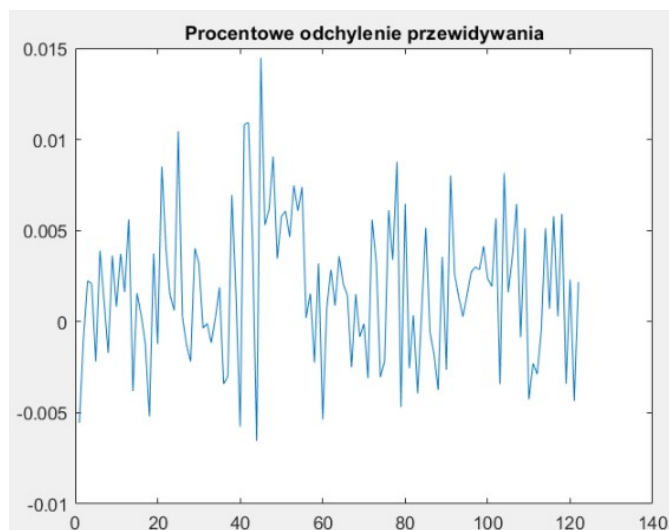
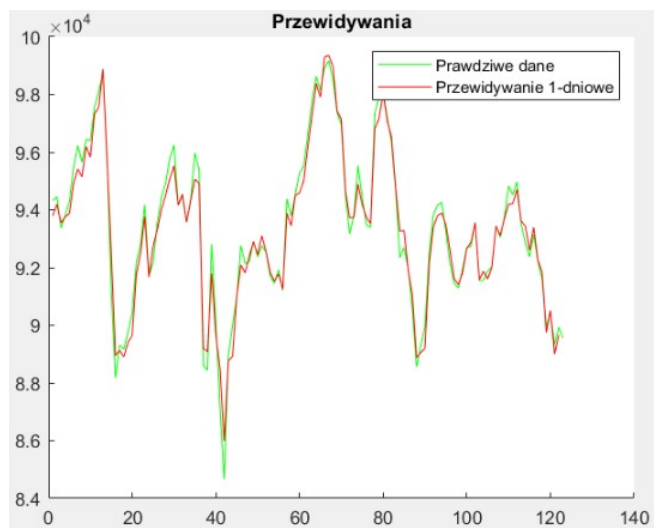


Błąd średniokwadratowy danych uczących  
Błąd średniokwadratowy dane testujące  
Średnie odchylenie przewidywań  
Dobrze przewidziane wzrostu i spadki cen akcji

0.00026461  
67170.1098  
0.0021678%  
98.3607%

Uzyskaliśmy najmniejszy jak do tej pory błąd średniokwadratowy dla danych testujących. Również średnie odchylenie przewidywanej kwoty jest mniejsze. Otrzymujemy prawie 100% współczynnik określający, czy zarobimy czy stracimy.

##### 5) 1 neuron ukryty w I warstwie, 17 neuronów ukrytych w II warstwie



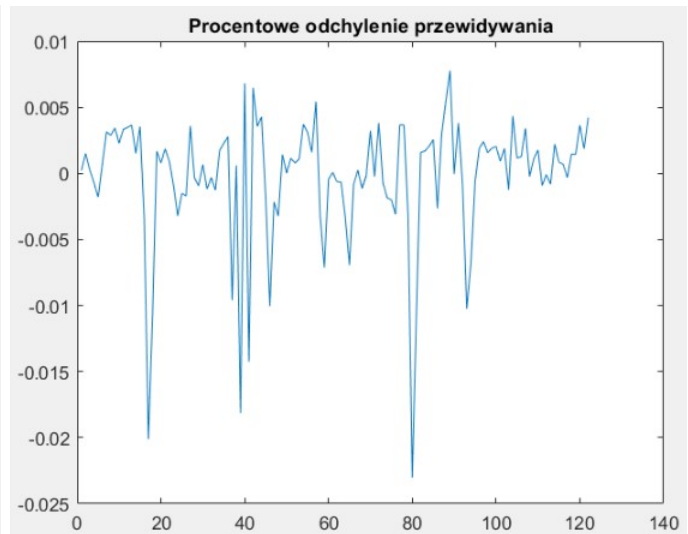
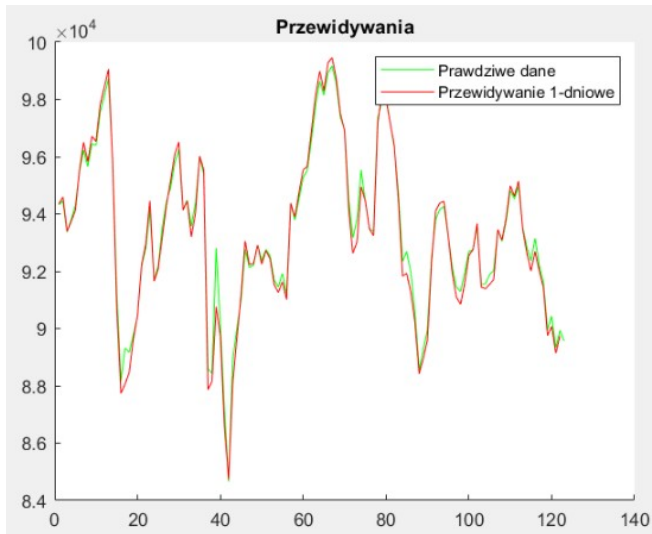
Błąd średniokwadratowy danych uczących	0.00026461
Błąd średniokwadratowy dane testujące	163805.4138
Średnie odchylenie przewidywań	0.0035026%
Dobrze przewidziane wzrostu i spadki cen akcji	94.2623%

Wydajność sieci spadła, błąd średniokwadratowy dla danych testujących bardzo wzrósł, odchylenie przewidywań kwoty również. Nadal otrzymujemy dość wysoki współczynnik przewidywań zarobku i strat, ale jest on niższy niż dla 11 neuronów ukrytych. Na podstawie wcześniejszych doświadczeń z sieciami neuronowymi możemy stwierdzić, że dalsze dodawanie liczby neuronów mija się z celem i tylko pogorszy wyniki, a czas oczekiwania na nie będzie niepotrzebnie długi.

Najlepszą spośród powyższych jest sieć oparta na 11 neuronach w II warstwie ukrytej. Warto także zauważyć, że błąd średniokwadratowy dla danych uczących był zawsze taki sam.

Sprawdźmy, jaki wpływ na sieć opartą na 11 neuronach w II warstwie ukrytej będzie miało zwiększenie ilości neuronów w I warstwie ukrytej.

### 1) 3 neurony ukryte w I warstwie, 11 neuronów ukrytych w II warstwie

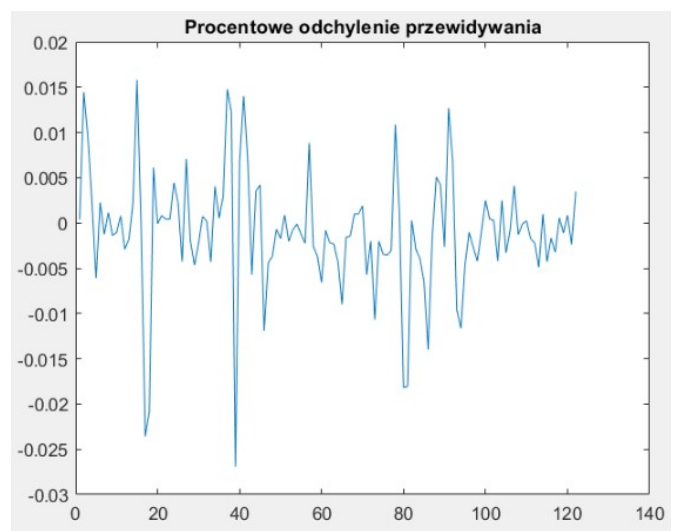
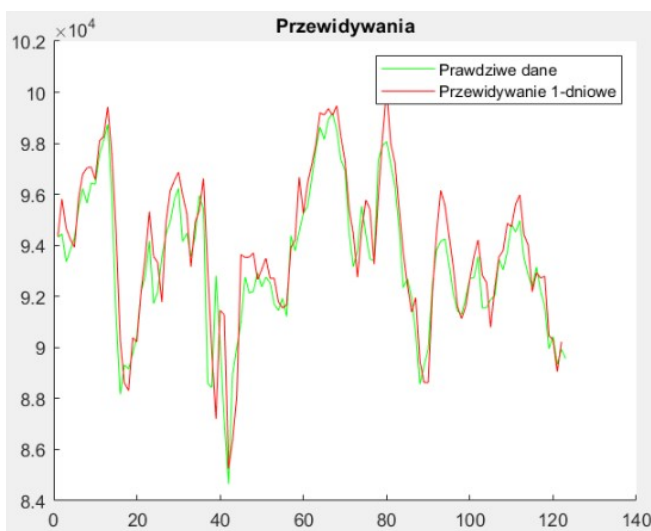


Błąd średniokwadratowy danych uczących  
Błąd średniokwadratowy dane testujące  
Średnie odchylenie przewidywań  
Dobrze przewidziane wzrostu i spadki cen akcji

0.00026461  
201026.8817  
0.0030483%  
93.4426%

Otrzymujemy dużo gorszą sieć niż w przypadku 1 neuronu w I warstwie ukrytej. Widzimy, że na wykresie odchylenia przewidywań oscylacje nie są wokół wartości prawdziwej.

### 2) 7 neuronów ukrytych w I warstwie, 11 neuronów ukrytych w II warstwie



Błąd średniokwadratowy danych uczących  
Błąd średniokwadratowy dane testujące  
Średnie odchylenie przewidywań  
Dobrze przewidziane wzrostu i spadki cen akcji

0.00026461  
394463.0943  
0.0045%  
92.623%

Wykres przewidywań wartości akcji w dużym stopniu odbiega od rzeczywistości. Błąd średniokwadratowy jest największy spośród wszystkich przetestowanych przypadków. Współczynnik mówiący nam o tym, czy zarobimy czy stracimy jest jednak wciąż dość dobry.

Widzimy, że wyniki są coraz gorsze, dlatego nie warto zwiększać już ilości neuronów w I warstwie ukrytej.

### Podsumowanie wyników

Liczba neuronów w warstwach ukrytych		Błąd średniokwadratowy danych testowych [ $\times 10^4$ ]
I	II	
1	1	30.91280574
1	3	10.42859117
1	7	13.63943997
1	11	6.71701098
1	17	16.38054138
3	11	20.10268817
7	11	39.44630943

Podsumowując najlepszą sieć uzyskaliśmy dla 1 neuronu w I warstwie ukrytej oraz 11 w II warstwie ukrytej. Błąd średniokwadratowy dla danych uczących był zawsze taki sam. Przypomnijmy, że wejściem dla sieci było 5 ostatnich dni. Sprawdźmy jak najlepsza sieć będzie zachowywać się dla innej liczby dni wejściowych.

Liczba wejść (dni na wejściu)	Błąd średniokwadratowy dane uczące	Błąd średniokwadratowy dane testowe
2	0.00012844	134723.0952
3	0.00017109	199095.2528
4	0.00021655	148835.2784
5	0.00026461	67824.0505
7	0.00036664	272520.9945
10	0.00051652	299450.2834
25	0.001297	491499.8754
70	0.0041844	181629884.454

Warto zauważyć, że dopiero zmiana liczby wejść na wpływ na błąd średniokwadratowy dla danych uczących. Widzimy, że wartość tego błędu nie ma bezpośredniego związku z wydajnością sieci. Najlepsze wyniki uzyskujemy dla 5 dni wejściowych. Błąd średniokwadratowy dla danych testowych jest wtedy najmniejszy, pomimo tego że taki błąd dla danych uczących nie należy do najmniejszych. W przypadku 70 wejść sieć działała bardzo źle, nawet współczynnik przewidywania zarobku lub straty wyniósł mniej niż 50%. Nie należy przesadzać z ilością wejść dla sieci. Optymalną ich ilość można wyznaczyć eksperymentalnie. Należy także zwrócić uwagę, że im większa liczba wejść oraz neuronów tym dłużej przebiega proces uczenia sieci.

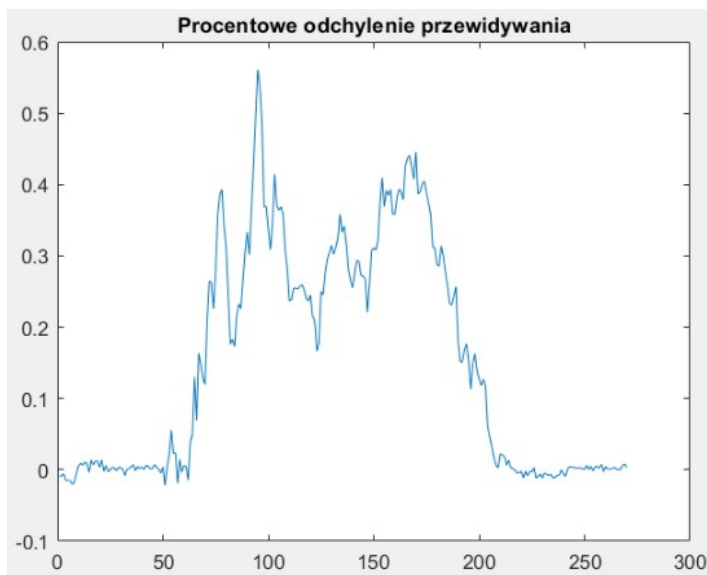
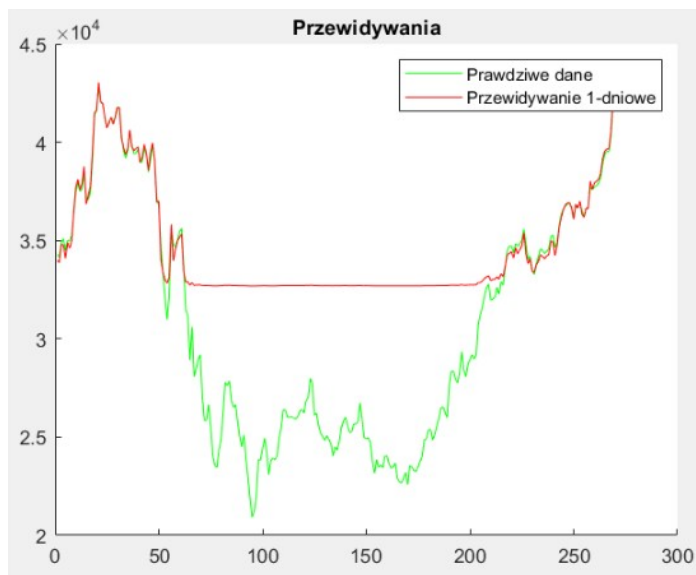


## Odporność sieci na kryzysy

Nasze dane zawierają wartości akcji od 2004 do końca 2018 roku. Skorzystamy z tego, że 15 września 2008 roku czwarty co do wielkości bank inwestycyjny Lehman Brothers ogłosił upadłość, po tym, jak amerykański bank centralny (Fed) nie udzielił mu oczekiwanej pomocy.

Przetestujemy odporność wybranej wcześniej najlepszej sieci na kryzysy. Będziemy uczyć się na podstawie danych od początku 2007 do połowy 2009 roku. Ponownie przywidujemy jeden dzień, zapisujemy przewidywanie i uczymy sieć dodając prawidłową wartość akcji w danym dniu.

Otrzymaliśmy następujące rezultaty:



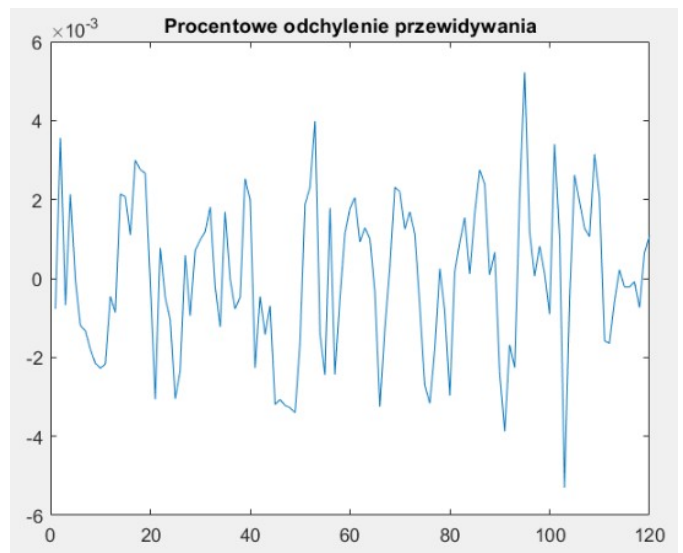
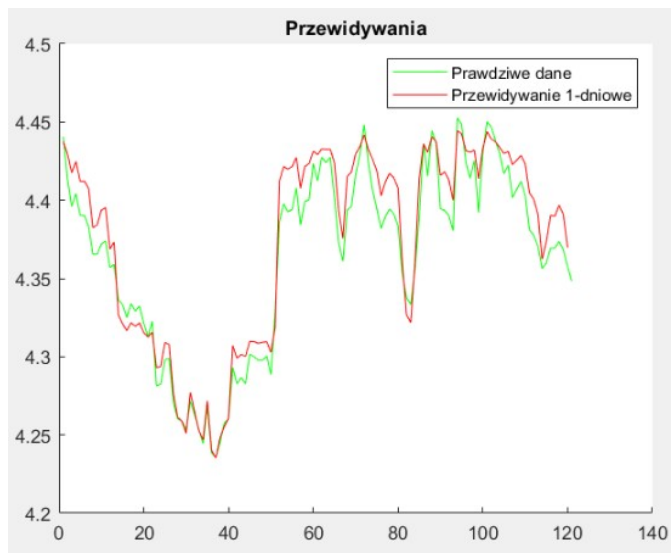
Błąd średniokwadratowy danych uczących	0.00073397
Błąd średniokwadratowy dane testujące	28724852.9513
Średnie odchylenie przewidywań	0.15375%
Dobrze przewidziane wzrosty i spadki cen akcji	71.1111%

Sieć nie była w stanie przewidzieć kryzysu. Widzimy, że sieć działa dobrze przed kryzysem i po nim. Jednak podczas niego nie była w stanie się dostosować, pomimo przewidywania tylko jednego dnia. Odchylenia przewidywanych wartości akcji są ogromne w porównaniu do tych przed i po kryzysie. Błąd średniokwadratowy dla danych testowych był również bardzo wysoki. Pomimo dobrego wskaźnika wzrostów i spadków wartości akcji, takiej sieci nie można uznać za dobrze działającą.

## Przewidywanie kursu Euro

Naszą sieć możemy wykorzystać również do innych celów jak np. przewidywanie kursów walut.

Wyniki przewidywań kursu euro. Dane uczące to przedział od grudnia 2004 roku do marca 2006. Testujące to kolejne 125 dni. Ponownie przywidujemy jeden dzień, zapisujemy przewidywanie i uczymy sieć dodając prawidłowy kurs w danym dniu.



Błąd średniokwadratowy danych uczących	3.4437e-05
Błąd średniokwadratowy dane testujące	7.3019e-05
Średnie odchylenie przewidywań	0.0016086%
Dobrze przewidziane wzrosty i spadki cen akcji	84.1667%

Sieć nie działa już tak dobrze jak dla przewidywania wartości akcji. Błędy średniokwadratowe są mniejsze, ale wynika to też z faktu, że wartość kursu euro jest mniejsza niż wartość wcześniej przewidywanych akcji. Odchylenie przewidywanych kursów do prawdziwych oscyluje wokół właściwych wartości.

### Interpretacja prezentowanych wyników z użyciem błędu relatywnego (względego)

Błąd ten wyraża się wzorem:

$$\delta_x = \frac{\Delta x}{x_0} = \frac{x - x_0}{x_0}$$

$x$  – wartość przewidziana  
 $x_0$  – wartość prawidłowa

Użycie takiego rodzaju błędu pozwala na dokładniejszą interpretację wyników, ponieważ błąd ten nie zależy od wielkości wartości danych tylko ich realnej różnicy wyrażonej najczęściej w procentach. Dzięki temu różnica przewidywań wartości może wynosić np. 100000zł co relatywnie może być małym błędem przy wartości akcji 98000000000zł. Błąd w takim przypadku wynosi 0.00010204081632653061224489795918367%. Jest on bardzo mały, pomimo że wydaje nam się, że 100000zł to duża różnica.