

Zadanie nr 1 - Aproksymacja

Inteligentna analiza danych

Przemysław Zdrzałik, 224466 Karol Domański 224285

07 kwietnia 2020 roku

1 Cel zadania

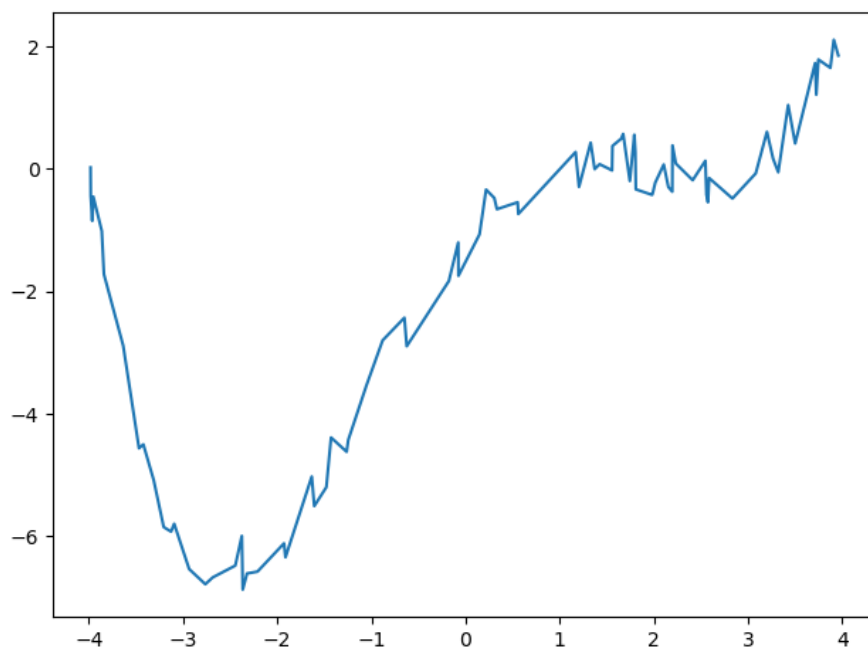
Celem zadania było wykorzystanie wielowarstwowego perceptronu (MLP) w celu aproksymacji funkcji.

2 Wstęp teoretyczny

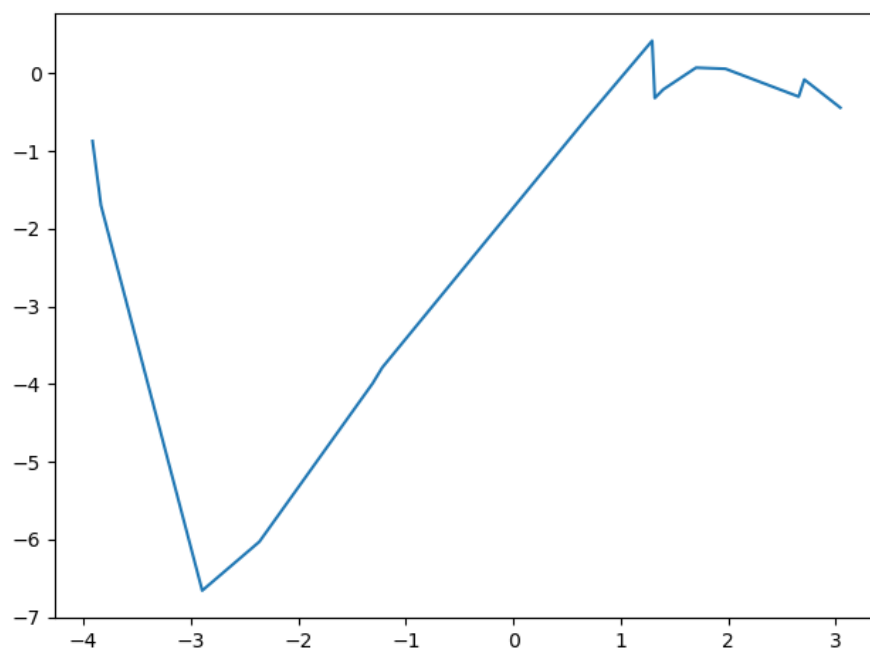
W zadaniu wykorzystujemy sieć neuronową o jednej warstwie ukrytej, jednym wejściu oraz jednym wyjściu.

Warstwa ukryta posiada sigmoidalne funkcje aktywacji, a warstwa wyjściowa identycznościową funkcję aktywacji. Sieć uczylismy korzystając z dwu różnych zbiorów treningowych, a naukę testowaliśmy na trzecim zbiorze testowym. Jakość aproksymacji określaliśmy za pomocą błędu średniokwadratowego.

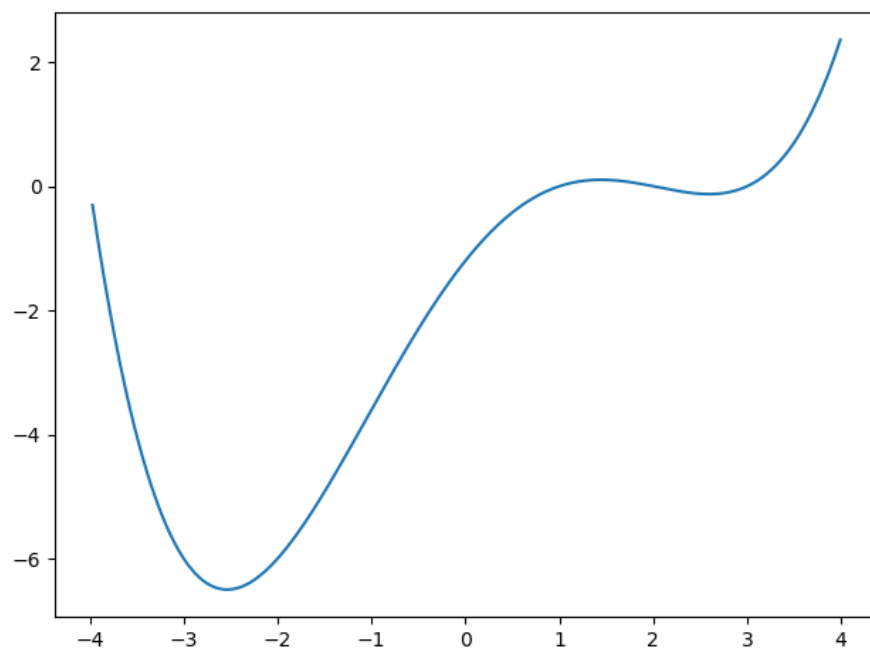
Oto przebiegi funkcji odpowiednio dla pliku treningowego 1, pliku treningowego 2 i pliku testowego:



Rysunek 1: Wykres aproksymowanej funkcji na podstawie plików treningowych 1



Rysunek 2: Wykres aproksymowanej funkcji na podstawie plików treningowych 2



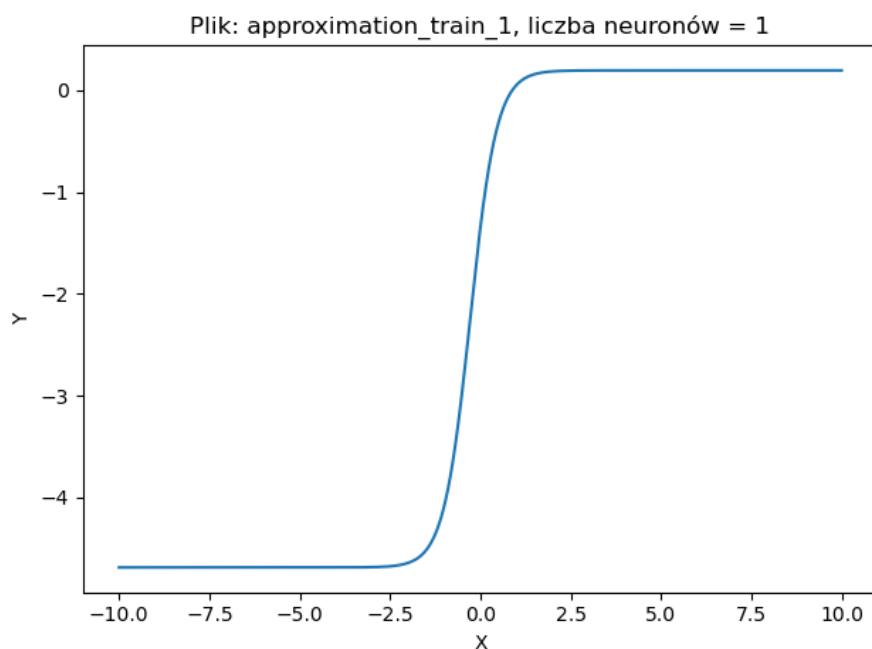
Rysunek 3: Wykres aproksymowanej funkcji na podstawie plików testowych

3 Eksperymenty i wyniki

Wszystkie eksperymenty wykonujemy ustawiając `numpy.random.seed()` na wartość 0.

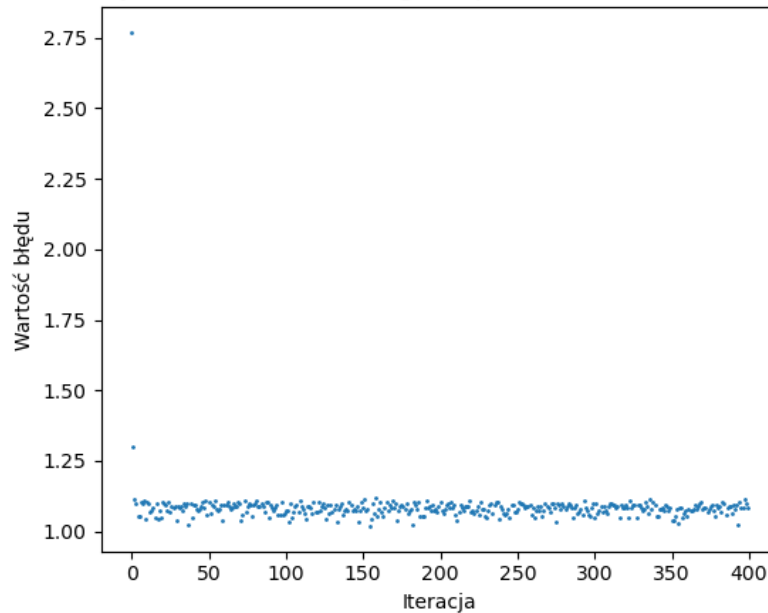
3.1 Eksperyment nr 1

W pierwszym eksperymencie postanowiliśmy przetestować możliwości aproksymacji funkcji dla 1 neuronu w warstwie ukrytej.



Rysunek 4: Wykres apoksymowanej funkcji po 400 epokach

Zmiana Błędu Średniokwadratowego, wsp. uczenia = 0.1 momentum = 0.1



Rysunek 5: Wykres zmiany błędu dla 400 epok

Jak widać na rysunku Rys. 5 błąd cały czas pozostawał taki sam.

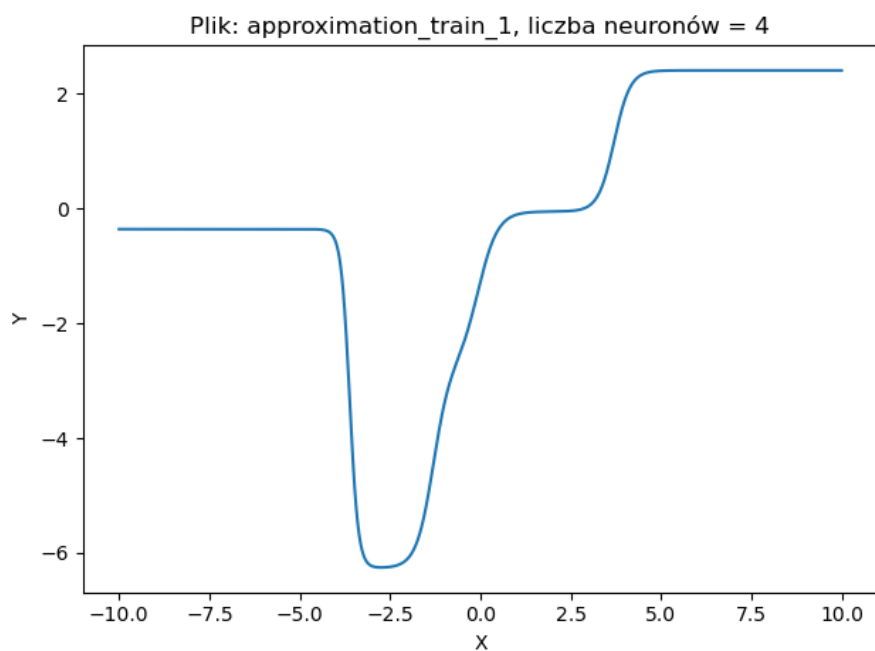
Testując sieć na zbiorze testowym otrzymaliśmy błąd równy 0.536, który, chociaż duży okazuje się dwa razy mniejszy niż błędy uzyskiwane na zbiorze testowym.

Eksperyment ten powtórzyliśmy na drugim zbiorze treningowym. Wyniki okazały się bardzo zbliżone, błąd pozostawał taki sam podczas nauki a testując sieć na zbiorze testowym uzyskaliśmy duży błąd równy 0.840. W próbie zachowania zwięzłości sprawozdania nie umieszczamy wyników.

3.2 Eksperyment nr 2

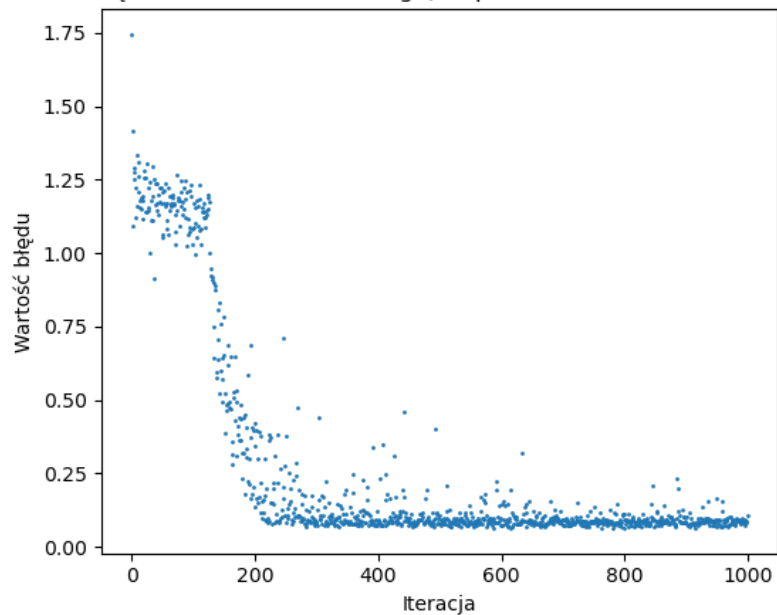
W tym eksperymencie powtarzamy warunki z eksperymentu nr. 1, zwiększając jednak liczbę epok do tysiąca, tym razem dla 4 neuronów.

Wyniki eksperymentów dla 2 i 3 neuronów były bardzo zbliżone do wyników poprzedniego eksperymentu, dlatego pomijamy je w sprawozdaniu.



Rysunek 6: Wykres apoksymowanej funkcji po 1000 epokach

Zmiana Błędu Średniokwadratowego, wsp. uczenia = 0.1 momentum = 0.1



Rysunek 7: Wykres zmiany błędu dla 1000 epok

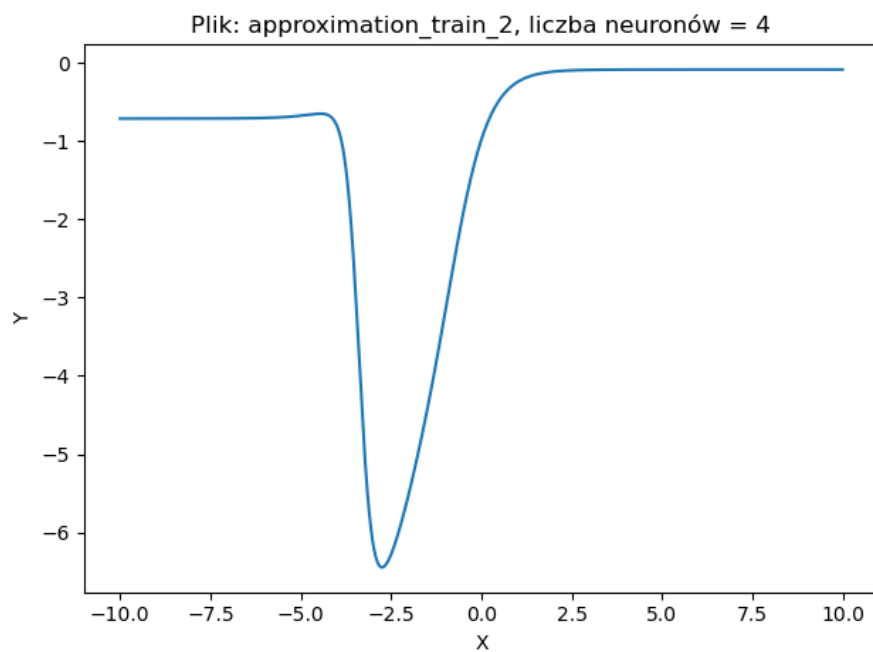
Jak widać na rysunku Rys. 7 w tym eksperymencie nastąpiła 'nauka' sieci neuronowej.

Błąd na zbiorze testowym systematycznie malał aż do około 300 epoki, od którego momentu pozostawał w przybliżeniu nie zmieniony. Błąd ten był równy w przybliżeniu 0.13.

Testując sieć na zbiorze testowym otrzymaliśmy błąd równy 0.029, jest on dużo mniejszy niż ten otrzymywany dla ilości neuronów od 1 do 3.

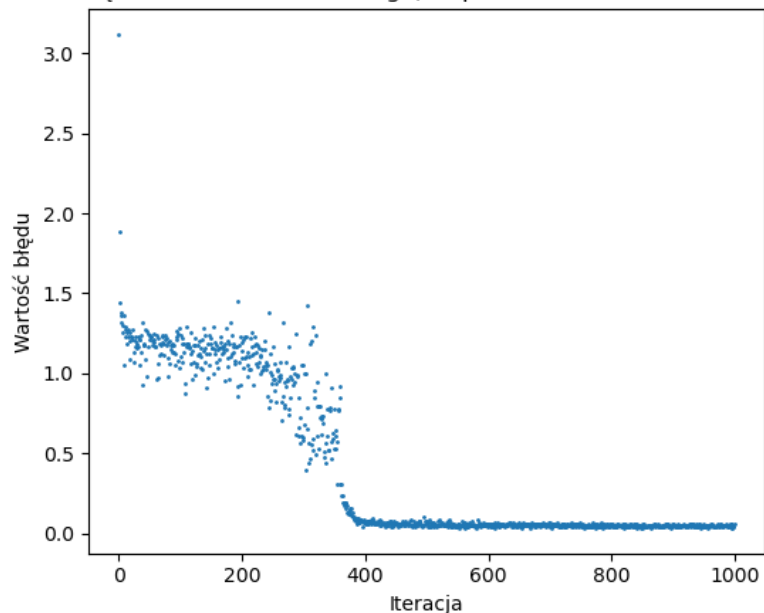
Analizując wykres aproksymowanej funkcji na rysunku Rys. 6 można zauważyć dużą różnicę względem poprzedniego eksperymentu.

Eksperyment ten powtórzyliśmy na drugim zbiorze treningowym.



Rysunek 8: Wykres apoksymowanej funkcji po 1000 epokach

Zmiana Błędu Średniokwadratowego, wsp. uczenia = 0.1 momentum = 0.1



Rysunek 9: Wykres zmiany błędu dla 1000 epok

Widać zauważalną różnicę w nauce neuronu.

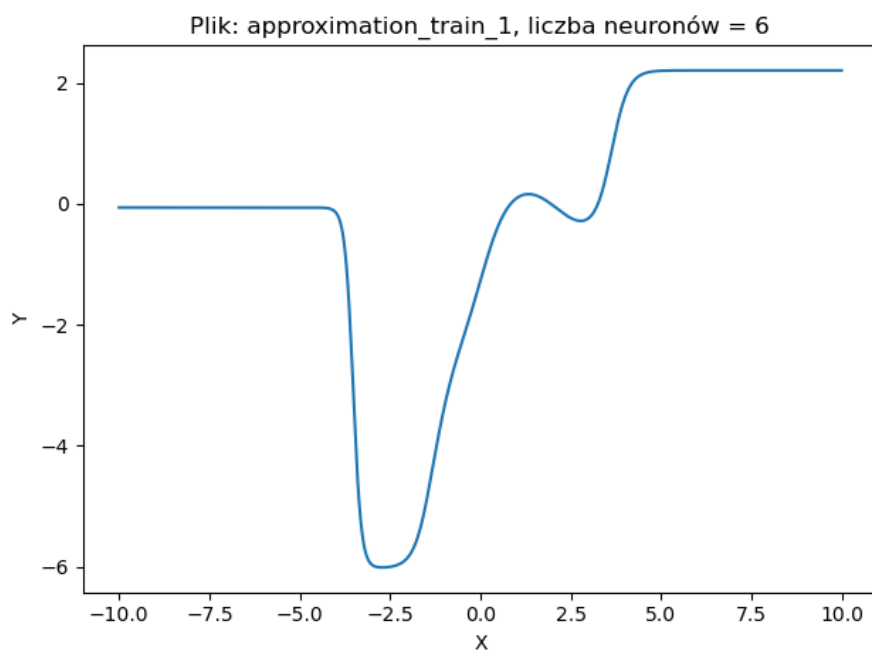
Otrzymany wykres funkcji na rysunku Rys. 8 jest zauważalnie mniej dokładny niż dla zbioru treningowego 1 na rysunku Rys. 6.

Różnice widać też na wykresie zmian błędu średniokwadratowego. Dla drugiego zbioru testowego jak widać na rysunku Rys. 9 zauważamy że po 'nauczeniu' się sieci około epoki 400 błąd pozostawał dużo bardziej stały niż dla pierwszego zbioru testowego, na rysunku Rys. 7.

Testując sieć na zbiorze testowym otrzymaliśmy błąd równy 0.150, a więc około 5 razy większy błąd niż dla zbioru treningowego 1. Możemy na podstawie tego przypuszczać że nauka sieci dla zbioru treningowego 1 postępuje lepiej.

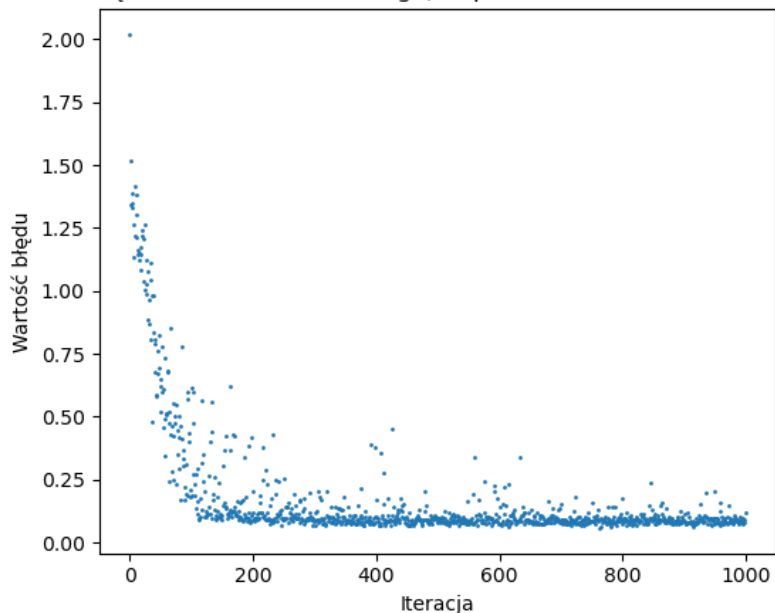
3.3 Eksperyment nr 3

W tym eksperymencie powtarzamy warunki z poprzedniego eksperymentu, z wykorzystaniem 6 neuronów w warstwie ukrytej.



Rysunek 10: Wykres apoksymowanej funkcji po 1000 epokach

Zmiana Błędu Średniokwadratowego, wsp. uczenia = 0.1 momentum = 0.1



Rysunek 11: Wykres zmiany błędu dla 1000 epok

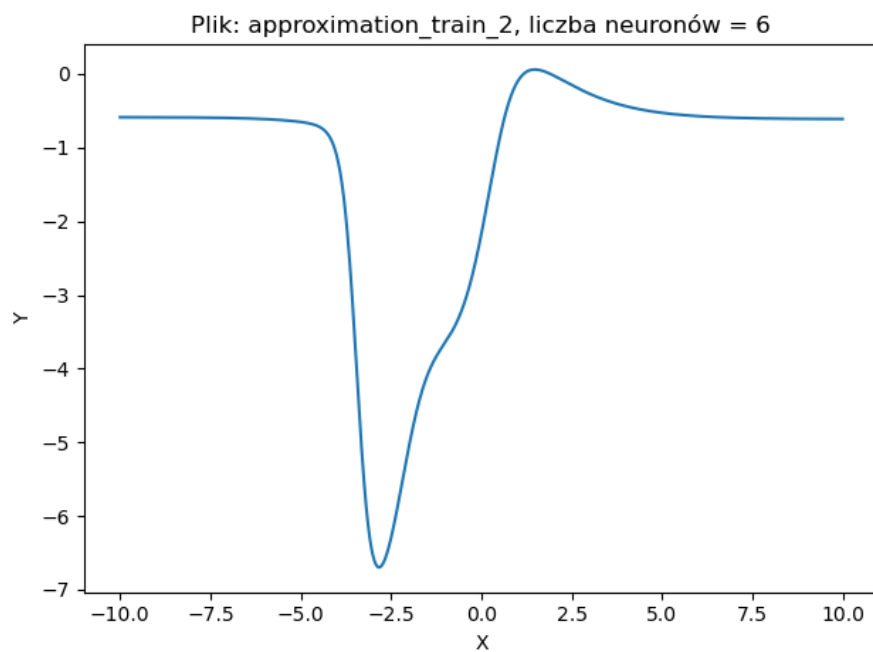
W tym eksperymencie jesteśmy w stanie zauważyć zmianę w szkicowanym wykresie funkcji w okolicach $x = 2.5$, jak widać na rysunku Rys. 10.

Wykres zmiany błędu średniokwadratowego jest bardzo podobny do poprzedniego eksperymentu, dlatego jego komentarz pomijamy.

Mimo tego, że dokładność aproksymacji wydaje się być większa niż w poprzednim eksperymencie testując sieć na zbiorze testowym otrzymaliśmy błąd równy 0.072, więc zauważalnie większy, bo ponad trzy krotnie, niż w eksperymencie dla 4 neuronów.

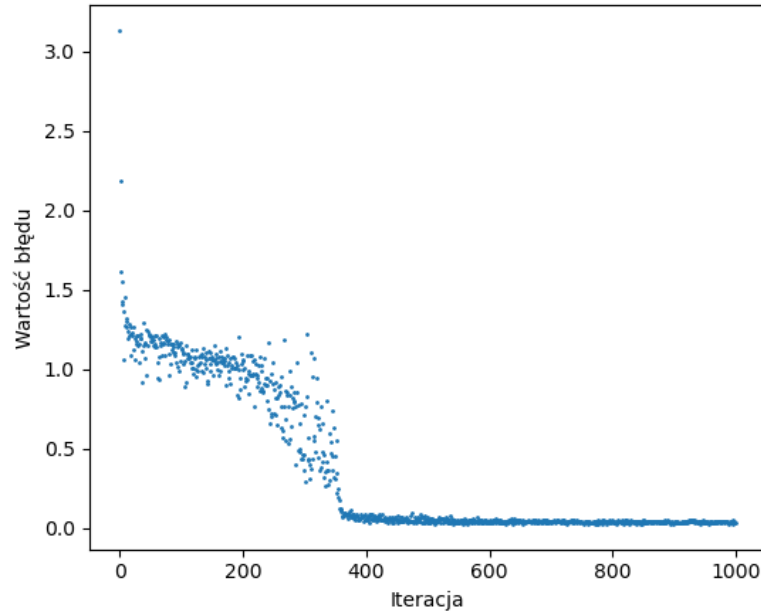
Analizując wykres aproksymowanej funkcji na rysunku Rys. 6 można zauważyć dużą różnicę względem poprzedniego eksperymentu.

Eksperyment ten powtórzyliśmy na drugim zbiorze treningowym.



Rysunek 12: Wykres apoksymowanej funkcji po 1000 epokach

Zmiana Błędu Średniokwadratowego, wsp. uczenia = 0.1 momentum = 0.1



Rysunek 13: Wykres zmiany błędu dla 1000 epok

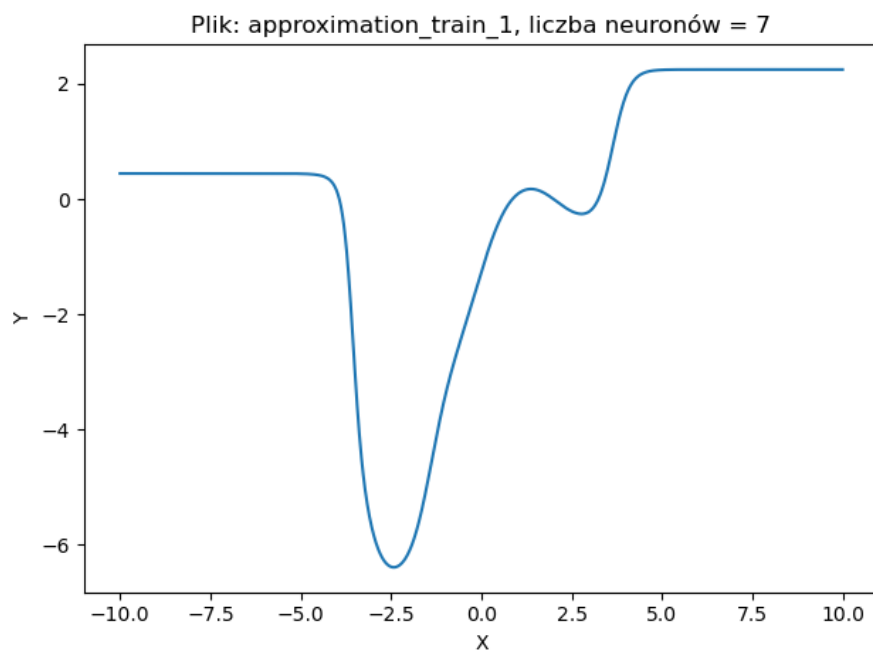
Wykres zmiany błędu średniokwadratowego jest bardzo podobny do poprzedniego eksperymentu, dlatego jego komentarz pomijamy.

Ponownie, otrzymany wykres funkcji na rysunku Rys. 12 jest zauważalnie mniej dokładny niż dla zbioru treningowego 1 na rysunku Rys. 10.

Testując sieć na zbiorze testowym otrzymaliśmy błąd równy 0.228, a więc podobnie jak dla zbioru treningowego 1 wraz z wzrostem ilości neuronów wzrósł błąd.

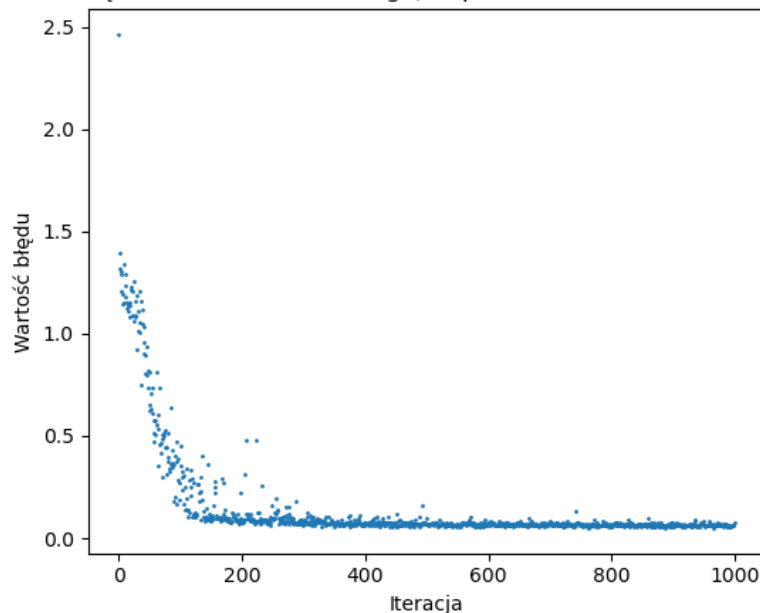
3.4 Eksperyment nr 4

Ponawiamy poprzednią konfigurację dla 7 neuronów.



Rysunek 14: Wykres apoksymowanej funkcji po 1000 epokach

Zmiana Błędu Średniokwadratowego, wsp. uczenia = 0.1 momentum = 0.1



Rysunek 15: Wykres zmiany błędu dla 1000 epok

Zauważamy dużą zmianę w wykresie błędu średniokwadratowego, jak wiadać na rysunku Rys. 15 dla 7 neuronów po osiągnięciu momentu kiedy błąd średniokwadratowy dla kolejnych epok się ustabilizował, w odróżnieniu od poprzedniej próby, rysunek Rys. 11, nie pojawiają się żadne epoki dla których zauważylibyśmy duże zmiany błędu średniokwadratowego. Sam błąd również ustabilizował się na bardzo niskich wartościach, bo bliskich 0.07.

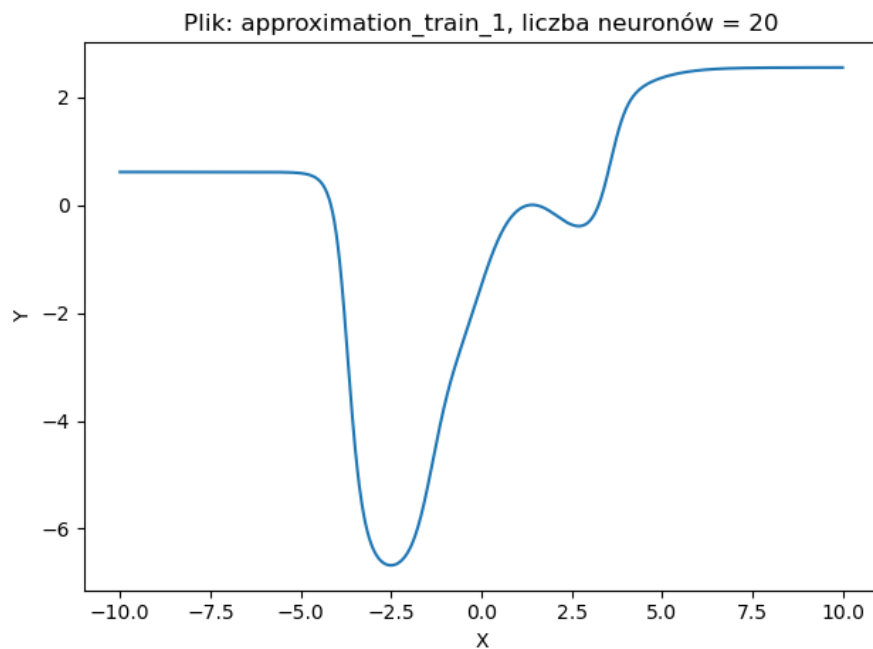
W wykresie funkcji brak zauważalnych zmian względem poprzedniego eksperymentu.

Testując sieć na zbiorze testowym otrzymaliśmy błąd równy 0.075, a więc podobnie jak w poprzednim eksperymencie.

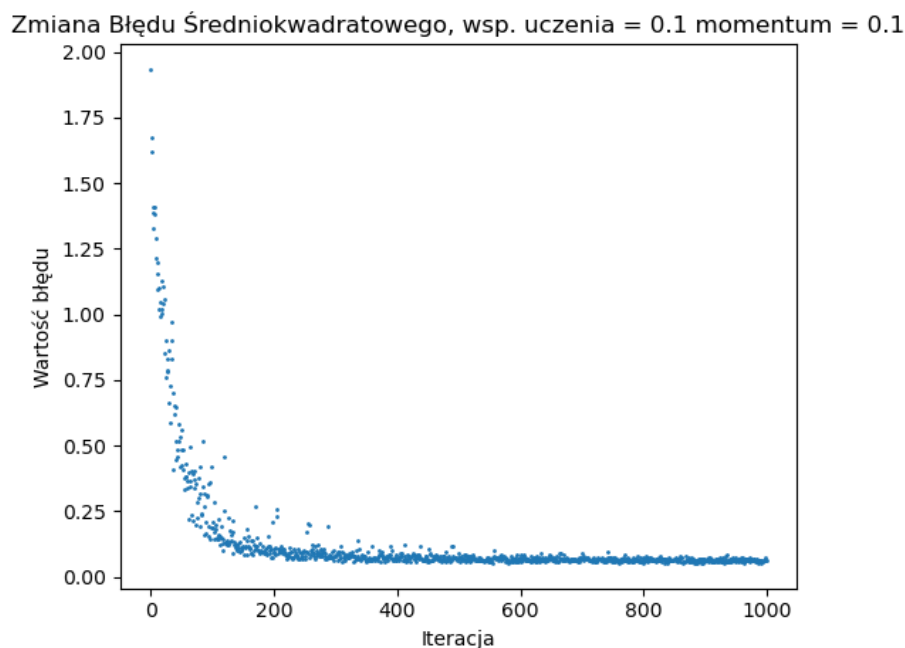
Dla drugiego zbioru testowego nie zaobserwowaliśmy zauważalnych różnic względem poprzedniego eksperymentu.

3.5 Eksperyment nr 5

Znowu powtarzamy te same warunki jak w poprzednim eksperymencie. Eksperyment wykonujemy dla 20 neuronów, nie zauważyliśmy widocznych zmian dla mniejszych ilości neuronów, więc tamte eksperymenty pomijamy.



Rysunek 16: Wykres apoksymowanej funkcji po 1000 epokach



Rysunek 17: Wykres zmiany błędu dla 1000 epok

Wyniki są bardzo zbliżone do poprzednich eksperymentów, z wyjątkiem błędu otrzymanego testując aproksymację na zbiorze testowym. Błąd ten okazał się być równy zaledwie 0.033, a więc był on najmniejszy poza błędem otrzymanym dla 4 neuronów. Błąd ten maleł powoli dla każdego dodanego neuronu, jednak była to jedyna zauważalna różnica pomiędzy kolejnymi dodanymi neuronami więc te eksperymenty pomijamy.

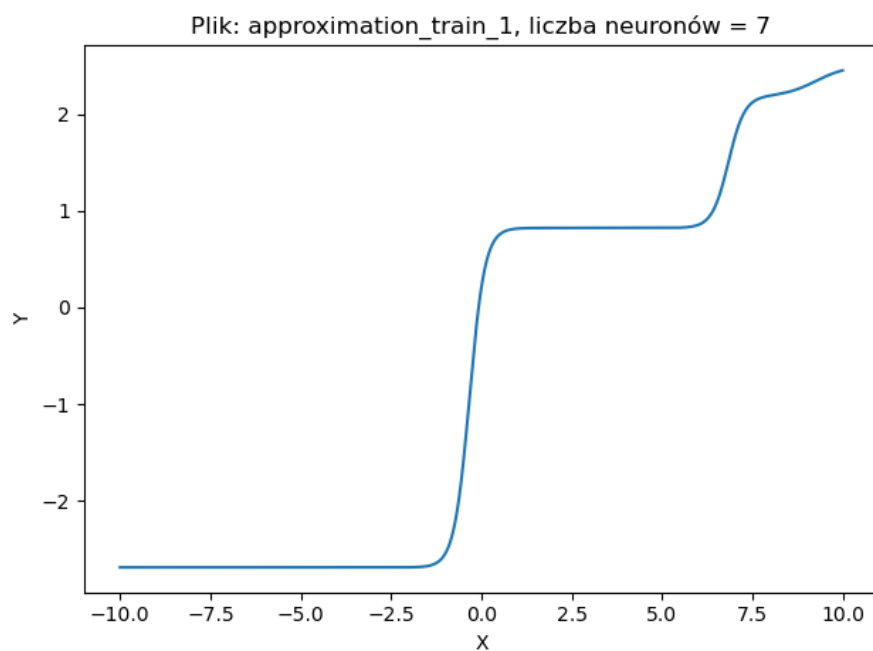
Otrzymane rezultaty zachęciły nas do sprawdzenia eksperymentu dla 4 neuronów przy innym 'seedzie' losowego generatora. Przy seedzie = 1 otrzymaliśmy błąd około 0.1 na zbiorze testowym, więc według nas to, że błąd dla 4 neuronów okazał się tak mały było przypadkiem, a sam błąd maleje, choć powoli, dla coraz większych ilości neuronów. Dla porównania błąd dla seedu = 1 dla 20 neuronów okazał się równy 0.017.

Dla pliku treningowego nr 2 nie zauważyliśmy dużych różnic względem poprzedniego eksperymentu więc te wyniki pomijamy.

3.6 Eksperyment nr 6

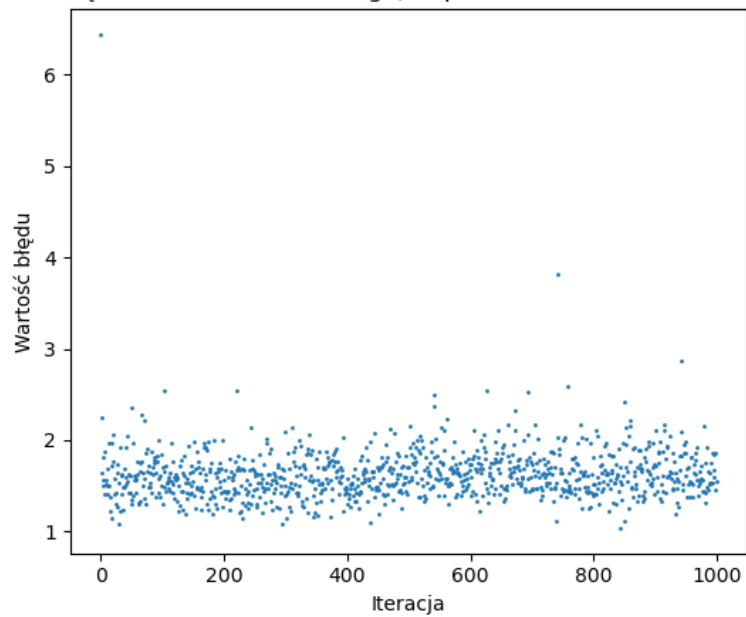
W tym eksperymencie analizujemy wpływ współczynnika uczenia i momentum na wyniki.

Ostatnią zauważalną zmianę w przebiegu nauki zaobserwowaliśmy dla 7 neuronów, więc dla takiej ilości przeprowadzamy te eksperymenty.



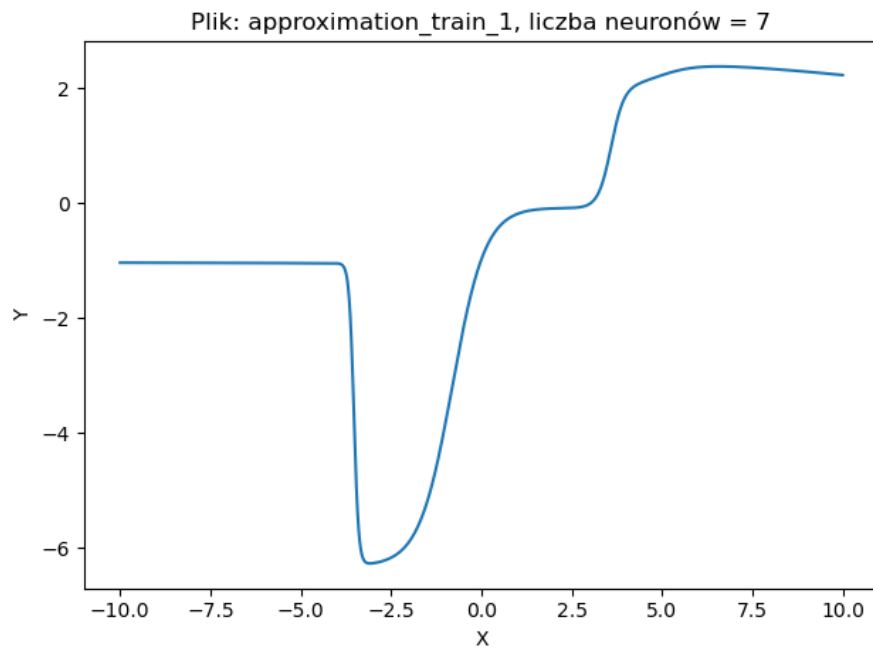
Rysunek 18: Wykres apoksymowanej funkcji po 1000 epokach

Zmiana Błędu Średniokwadratowego, wsp. uczenia = 0.1 momentum = 0.8



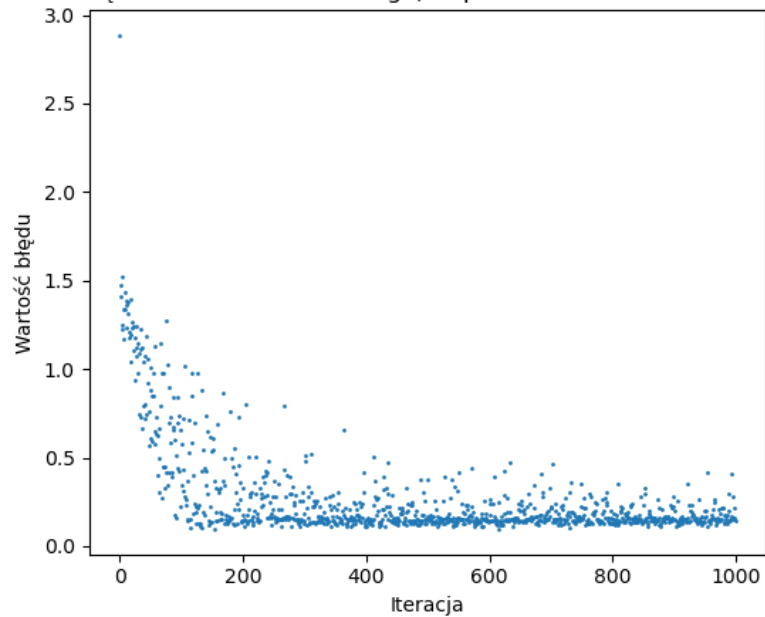
Rysunek 19: Wykres zmiany błędu dla 1000 epok

Jak widać na rysunku Rys. 19 dla $\text{momentum} = 0.8$ proces nauki sieci nie udał się. Otrzymany błąd dla zbioru testowego jest równy 1.6.



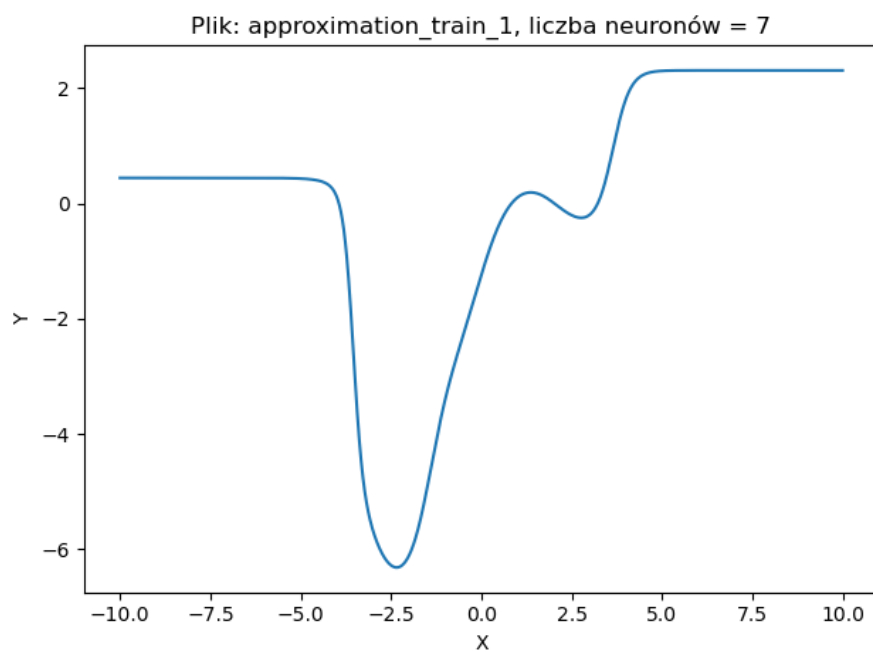
Rysunek 20: Wykres apoksymowanej funkcji po 1000 epokach

Zmiana Błędu Średniokwadratowego, wsp. uczenia = 0.1 momentum = 0.5

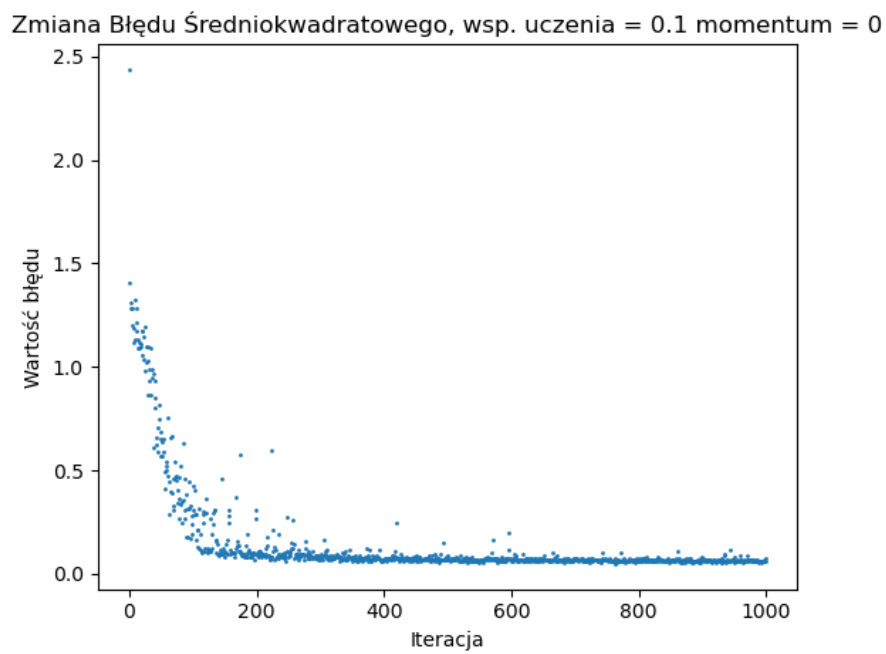


Rysunek 21: Wykres zmiany błędu dla 1000 epok

Możemy zauważyć na rysunku Rys. 21, że dla $\text{momentum} = 0.5$ nauka już przebiega, chociaż zauważalnie gorzej niż przy $\text{momentum} = 0.1$. Mimo tego, że błąd podczas nauki okazywał się 'skoczny', to błąd na zbiorze testowym $= 0.06$. Sprawdzenie dla $\text{seedu} = 1$ pokazało, że był to przypadek, tam błąd $= 0.123$

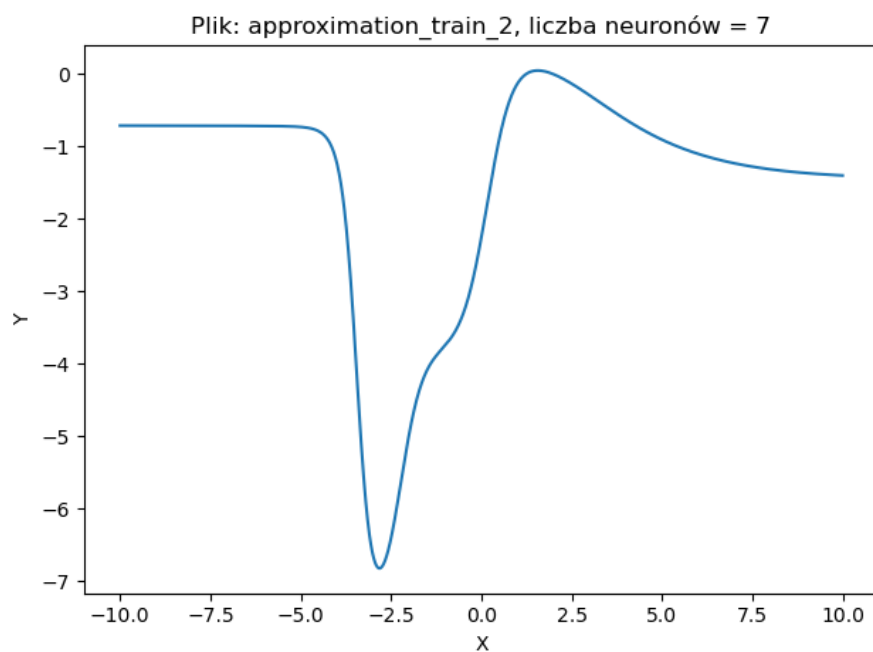


Rysunek 22: Wykres apoksymowanej funkcji po 1000 epokach



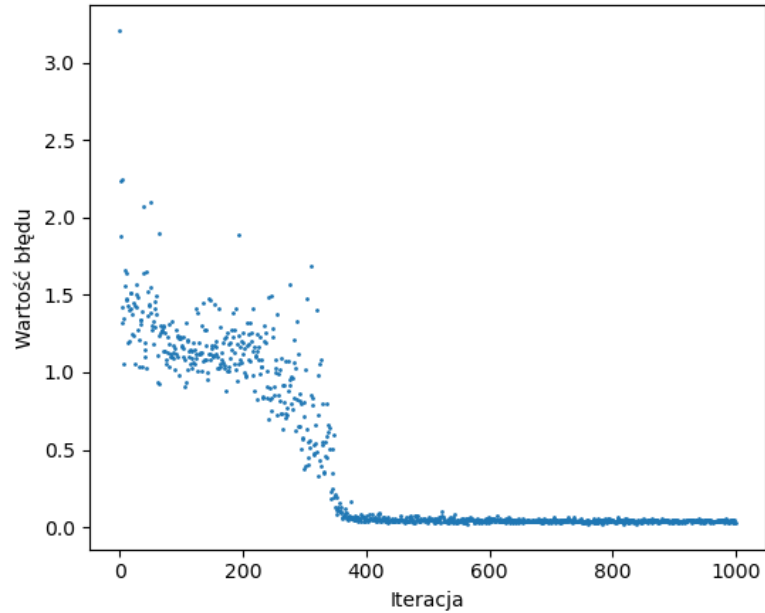
Rysunek 23: Wykres zmiany błędu dla 1000 epok

Wyeliminowanie momentum wydaje się dawać najlepsze rezultaty. Błąd na zbiorze testowym = 0.076, a wykres błędu na zbiorze treningowym wygląda 'dobrze'.



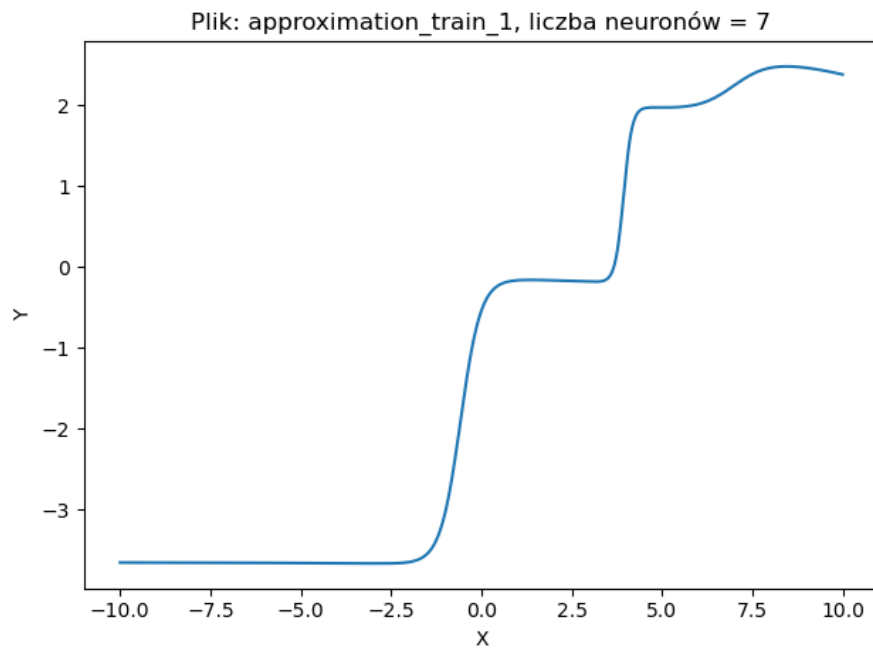
Rysunek 24: Wykres apoksymowanej funkcji po 1000 epokach

Zmiana Błędu Średniokwadratowego, wsp. uczenia = 0.1 momentum = 0.5



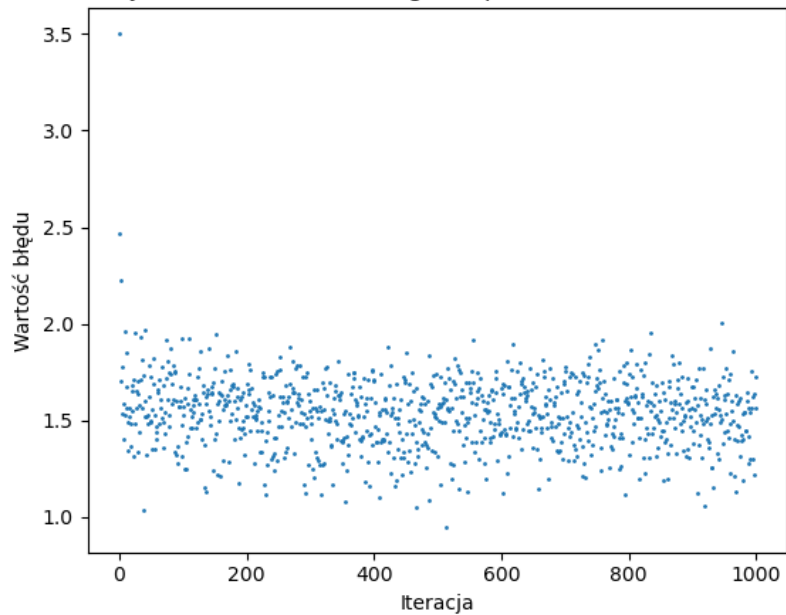
Rysunek 25: Wykres zmiany błędu dla 1000 epok

Inaczej sytuacja wygląda dla drugiego zbioru treningowego. Tutaj, jak widać na Rysunku Rys. 25 wykres zmiany błędu dla drugiego zestawu treningowego wygląda 'dobrze', również wykres funkcji zgadza się z oczekiwanym a błąd na zbiorze testowym, równy 0.27 jest zbliżony do tego otrzymanego dla $\text{momentum} = 0.1$.



Rysunek 26: Wykres apoksymowanej funkcji po 1000 epokach

Zmiana Błędu Średniokwadratowego, wsp. uczenia = 0.4 momentum = 0



Rysunek 27: Wykres zmiany błędu dla 1000 epok

Zauważalnie, wykres na rysunku Rys. 26 wygląda 'źle', po wykresie zmiany błędu widzimy, że nauka nie nastąpiła. Błąd na zbiorze testowym = 0.81.

4 Wnioski

Zauważamy, że przeprowadzona przez nas aproksymacja wykorzystująca sieć neuronową jest poprawna. Zauważamy, że błąd średnio kwadratowy spada 'szybko', aż osiągnie okolice swojego minimum około 200-400 epoki, kiedy to stabilizuje się i zaczyna bardzo powoli maleć. Wtedy też można uznać że sieć jest nauczona.

W najlepszych przypadkach otrzymujemy przybliżenia wartości funkcji dla testowanego zbioru o średnim błędzie bliskim 0.02.

Zauważamy również, że niskie wartości momentum i współczynnika uczenia oraz duża liczba neuronów daje nam najlepsze rezultaty.

Zauważamy, że do odpowiedniego odwzorowania tej funkcji potrzebujemy przynajmniej 7 neuronów, a nauka na większym, pierwszym, zbiorze pozwala lepiej aproksymować wartości funkcji oryginalnej jak zadanej w zbiorze testowym, nauka na zbiorze treningowym 2 daje nam wykres funkcji, który mniej pokrywa się z tym oryginalnym niż zbiór treningowy 1.

5 Bibliografia

- [1] Wzory związane z perceptronem wielowarstwowym
- [2] Bartłomiej Stasiak Sztuczne sieci neuronowe
- [3] Bartłomiej Stasiak Jednokierunkowe wielowarstwowe sieci neuronowe