

Sieci Neuronowe – laboratorium

Ćwiczenie nr. 2

Kajetan Pynka, 254495

Spis treści

| | |
|---|----|
| Wstęp teoretyczny..... | 3 |
| Wpływ liczby neuronów..... | 4 |
| Liczba neuronów - 10..... | 4 |
| Liczba neuronów - 20..... | 5 |
| Liczba neuronów - 30..... | 6 |
| Liczba neuronów - 50..... | 7 |
| Wnioski..... | 8 |
| Wpływ współczynnika uczenia..... | 9 |
| Współczynnik uczenia - $1e-4$ | 9 |
| Współczynnik uczenia $5e-5$ | 10 |
| Współczynnik uczenia $1e-5$ | 11 |
| Współczynnik uczenia - $1e-6$ | 12 |
| Wnioski..... | 13 |
| Wpływ wielkości paczki (batcha)..... | 14 |
| Wielkość paczki - 6..... | 14 |
| Wielkość paczki - 60..... | 15 |
| Wielkość paczki - 600..... | 16 |
| Wielkość paczki - 6000..... | 17 |
| Wnioski..... | 18 |
| Wpływ inicjalizacji wag - odchylenie standardowe..... | 19 |
| Odchylenie standardowe - 0.1..... | 19 |
| Odchylenie standardowe - 0.01..... | 20 |
| Odchylenie standardowe - 0.001..... | 21 |
| Odchylenie standardowe - 0.0001..... | 22 |
| Wnioski..... | 23 |
| Wpływ funkcji aktywacji..... | 24 |
| Funkcja aktywacji - sigmoid..... | 24 |
| Funkcja aktywacji - tangens hiperboliczny..... | 25 |
| Funkcja aktywacji - RELU..... | 26 |
| Wnioski..... | 27 |

Wstęp teoretyczny

W ramach ćwiczenia zaimplementowałem sieć MLP. Wszelkie algorytmy zostały napisane od podstaw, natomiast do przeprowadzenia obliczeń macierzowych wykorzystałem bibliotekę 'numpy'. W ramach badań analizowałem z osobna następujących pięć czynników: liczbę neuronów warstwy ukrytej, wartość współczynnika uczenia, wielkość paczki (batcha), początkowe wartości wag oraz rodzaj funkcji aktywacji warstwy ukrytej. Każdy eksperyment był oceniany ze względu na liczbę epok potrzebnych do wyuczenia modelu (model uznawany był za wyuczony gdy różnica między wartością funkcji kosztu dla epoki poprzedniej i aktualnej wynosiła poniżej 30). Dodatkowo, dla każdej epoki zapisywany był aktualny błąd na zbiorze testowym (w skali [0;1] gdzie 0.1 oznacza, że model miał 10% błędnych predykcji = 90% poprawnych predykcji). Każdy eksperyment został powtórzony 10 razy (z takimi samymi wartościami hiperparametrów). Macierz pomyłek jest średnią z tych 10 uruchomień, dla każdego uruchomienia obliczana dla zbioru testowego PO zakończeniu nauki.

Legenda do wykresów:

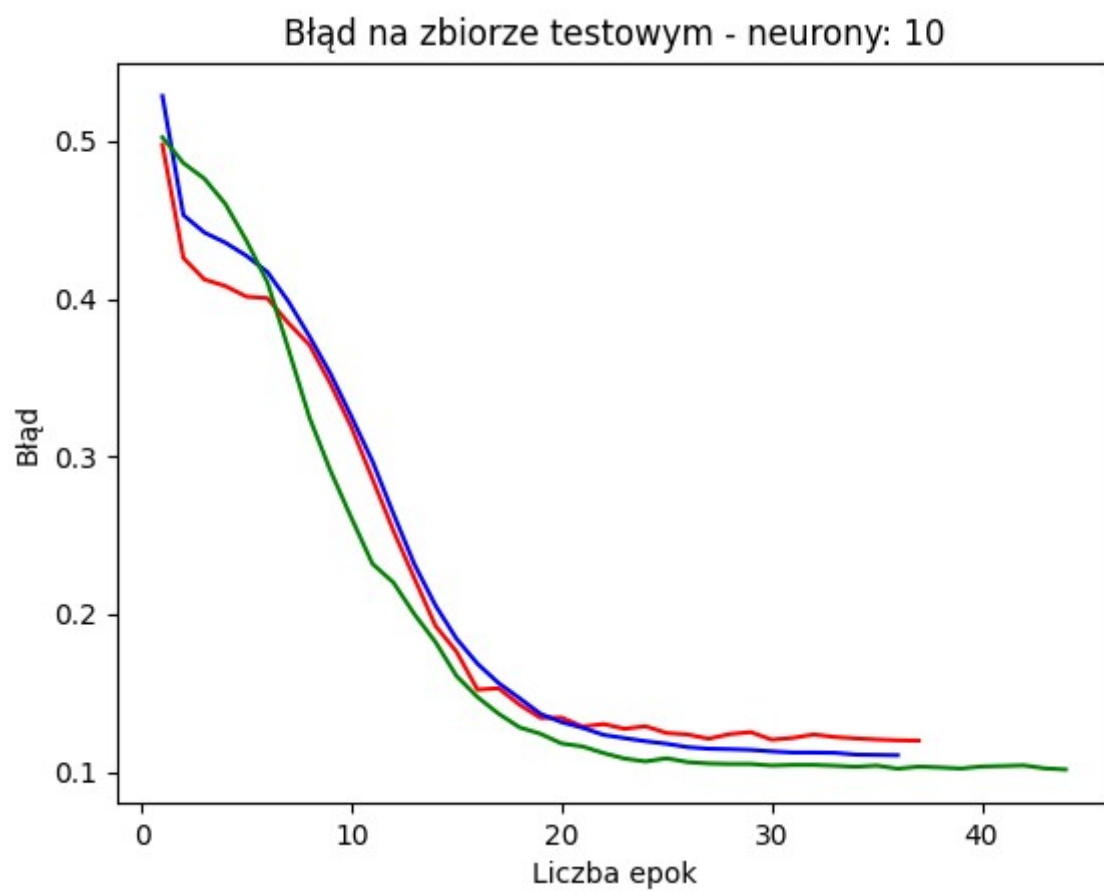
- Kolor **czzerwony** oznacza przebieg eksperymentu o najgorszym finalnym rezultacie (najwyższym błędzie)
- Kolor **niebieski** oznacza średnią ze wszystkich 10 powtórzeń eksperymentu
- Kolor **zielony** oznacza przebieg eksperymentu o najlepszym finalnym rezultacie (najniższym błędzie)

Dla każdego eksperymentu przyjęte zostały następujące wartości hiperparametrów (tam, gdzie dany parametr nie był oczywiście przedmiotem badania):

- Liczba neuronów warstwy ukrytej – 20
- Rozmiar zbioru treningowego – 60000
- Rozmiar zbioru testowego – 10000
- Rozmiar paczki (batcha) – 60
- Dopuszczalna wartość błędu między funkcją kosztu – 30
- Współczynnik uczenia – $1e-5$ (0.00001)
- Early stopping – aktywny, przy czym jeśli przez 3 kolejne epoki i tak następuje regres to proces nauki był przerywany
- Tylko jedna warstwa ukryta
- Wagi inicjalizowane za pomocą rozkładu normalnego gdzie
 - o Mediana / wartość oczekiwana – 0
 - o Odchylenie standardowe – 0.001

Wpływ liczby neuronów

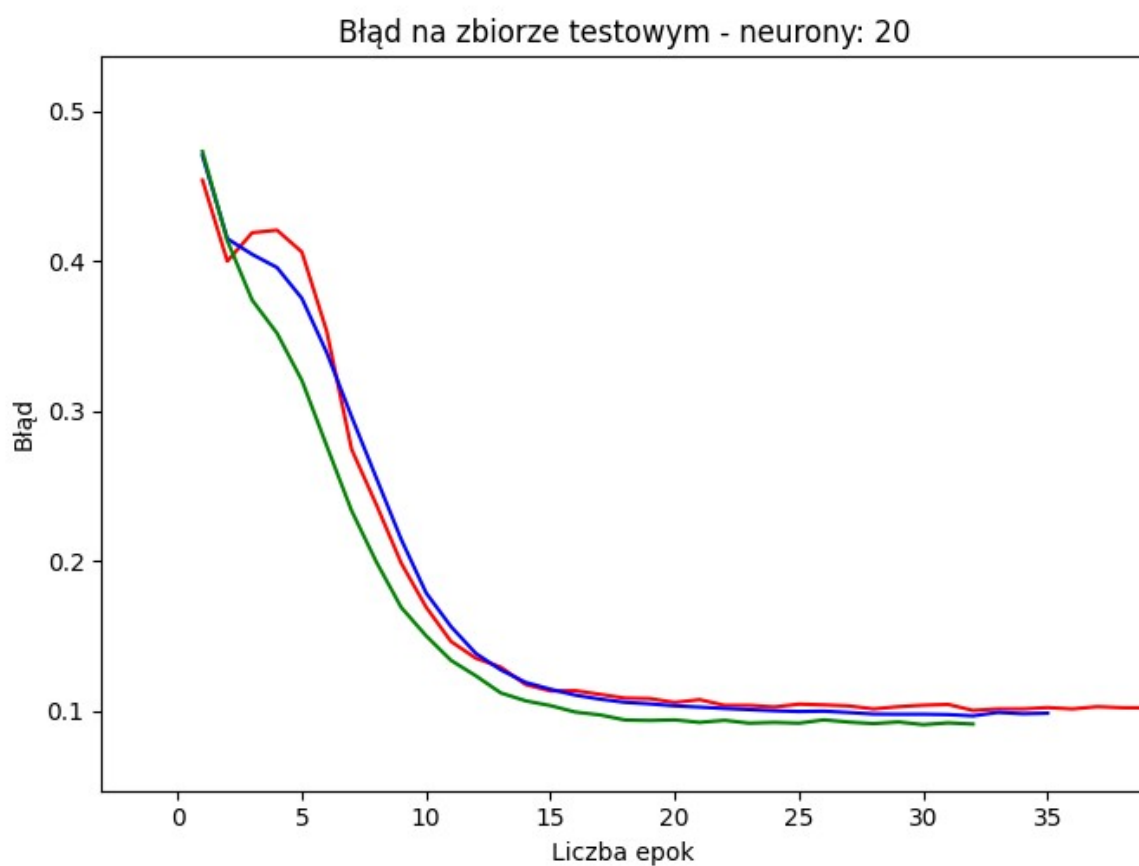
Liczba neuronów - 10



Macierz pomyłek:

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|--------|---------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0 | 943.00 | 0.10 | 17.80 | 5.30 | 1.80 | 27.10 | 25.30 | 5.10 | 11.20 | 9.30 |
| 1 | 0.10 | 1108.10 | 8.30 | 1.60 | 3.60 | 1.80 | 3.50 | 21.30 | 12.80 | 5.20 |
| 2 | 3.20 | 3.60 | 863.20 | 27.00 | 4.20 | 9.00 | 8.50 | 16.10 | 11.10 | 1.00 |
| 3 | 4.00 | 5.40 | 35.50 | 886.60 | 0.00 | 66.00 | 0.60 | 6.60 | 26.10 | 10.30 |
| 4 | 1.70 | 0.60 | 18.30 | 0.90 | 909.10 | 14.50 | 14.30 | 10.20 | 19.00 | 40.10 |
| 5 | 9.40 | 0.40 | 4.20 | 36.60 | 0.10 | 675.10 | 14.90 | 0.20 | 33.20 | 10.30 |
| 6 | 8.90 | 3.50 | 15.70 | 1.50 | 14.50 | 16.20 | 882.50 | 0.20 | 12.50 | 2.30 |
| 7 | 5.10 | 1.70 | 15.90 | 15.90 | 1.50 | 16.10 | 0.60 | 918.50 | 16.40 | 27.10 |
| 8 | 4.60 | 11.40 | 46.20 | 24.50 | 4.60 | 54.90 | 7.60 | 3.50 | 813.90 | 11.00 |
| 9 | -0.00 | 0.20 | 6.90 | 10.10 | 42.60 | 11.30 | 0.20 | 46.30 | 17.80 | 892.40 |

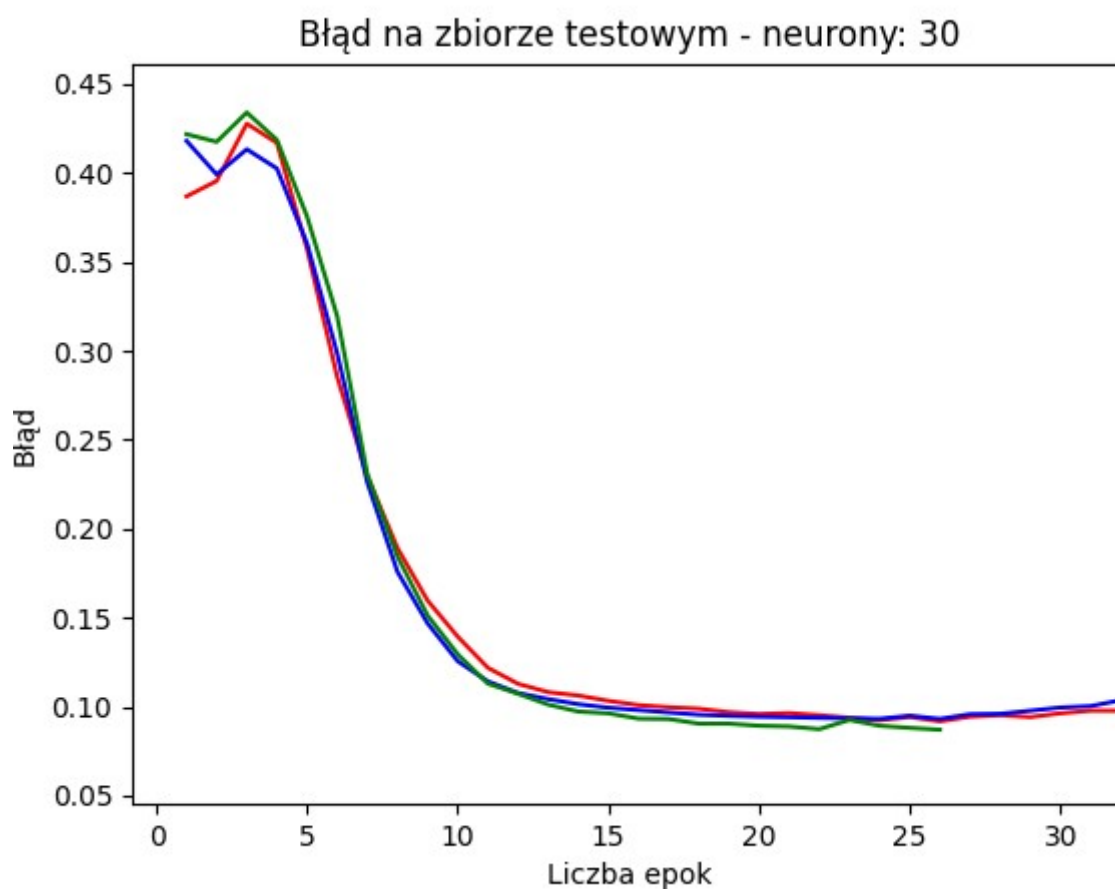
Liczba neuronów - 20



Macierz pomyłek:

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|--------|---------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0 | 952.10 | 0.00 | 16.30 | 8.00 | 4.10 | 27.30 | 21.10 | 2.90 | 7.70 | 8.70 |
| 1 | 0.00 | 1112.40 | 7.10 | 2.40 | 3.40 | 4.20 | 4.20 | 17.50 | 11.40 | 6.90 |
| 2 | 1.10 | 2.20 | 874.50 | 21.60 | 3.90 | 7.80 | 7.30 | 16.10 | 5.50 | 0.60 |
| 3 | 2.40 | 4.60 | 22.90 | 882.80 | 0.00 | 48.00 | 0.80 | 6.80 | 17.60 | 11.80 |
| 4 | 0.30 | 0.30 | 15.40 | 0.30 | 905.90 | 10.40 | 13.60 | 9.60 | 12.60 | 29.80 |
| 5 | 6.30 | 0.20 | 3.00 | 41.30 | 0.20 | 700.40 | 14.10 | 0.10 | 27.20 | 11.60 |
| 6 | 9.20 | 3.40 | 18.60 | 2.50 | 18.90 | 19.70 | 884.50 | 0.10 | 10.40 | 1.40 |
| 7 | 2.30 | 1.50 | 14.40 | 13.80 | 1.50 | 13.10 | 1.00 | 928.90 | 13.00 | 22.90 |
| 8 | 6.30 | 10.10 | 55.90 | 28.10 | 4.40 | 52.40 | 11.30 | 7.10 | 858.80 | 12.20 |
| 9 | 0.00 | 0.30 | 3.90 | 9.20 | 39.70 | 8.70 | 0.10 | 38.90 | 9.80 | 903.10 |

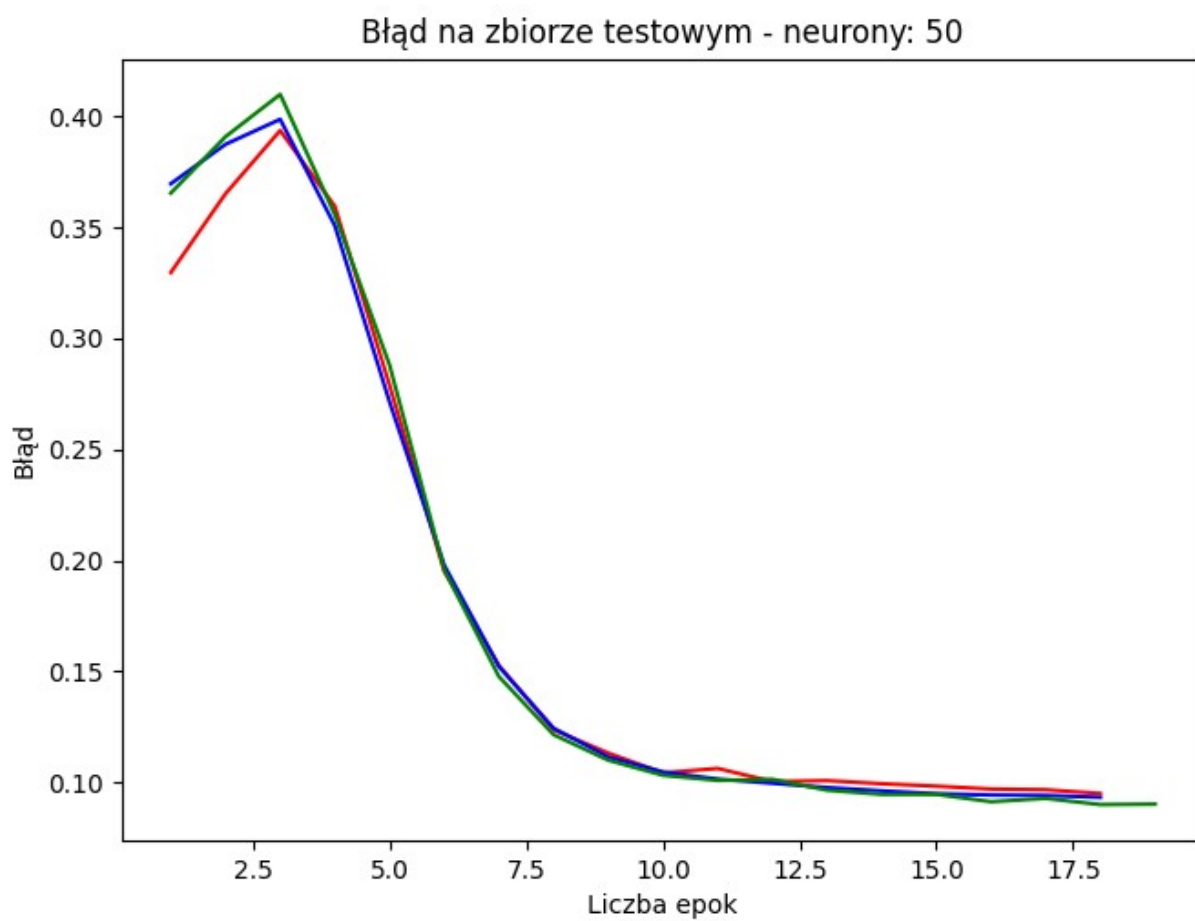
Liczba neuronów - 30



Macierz pomyłek:

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|--------|---------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0 | 944.10 | 0.10 | 13.00 | 6.00 | 4.00 | 23.10 | 19.00 | 2.70 | 6.00 | 7.50 |
| 1 | 0.20 | 1107.00 | 9.40 | 1.60 | 3.80 | 3.30 | 5.30 | 15.30 | 11.30 | 5.30 |
| 2 | 3.30 | 2.40 | 858.80 | 23.80 | 5.30 | 14.30 | 19.00 | 20.20 | 11.50 | 3.60 |
| 3 | 2.30 | 4.80 | 25.00 | 895.50 | 0.80 | 42.80 | 0.60 | 6.80 | 20.60 | 11.90 |
| 4 | 0.60 | 1.00 | 15.90 | 0.40 | 898.80 | 13.10 | 23.40 | 9.80 | 14.10 | 31.80 |
| 5 | 7.80 | 0.60 | 6.90 | 31.60 | 0.50 | 713.60 | 23.60 | 0.90 | 24.90 | 13.90 |
| 6 | 10.80 | 3.20 | 26.30 | 3.20 | 20.00 | 20.30 | 846.60 | 1.50 | 8.80 | 2.60 |
| 7 | 3.50 | 1.80 | 20.80 | 12.70 | 2.40 | 12.50 | 0.80 | 915.80 | 12.20 | 15.90 |
| 8 | 7.40 | 13.60 | 50.20 | 25.00 | 4.90 | 41.20 | 19.00 | 6.10 | 851.20 | 14.50 |
| 9 | 0.00 | 0.50 | 5.70 | 10.20 | 41.50 | 7.80 | 0.70 | 48.90 | 13.40 | 902.00 |

Liczba neuronów - 50



Macierz pomyłek:

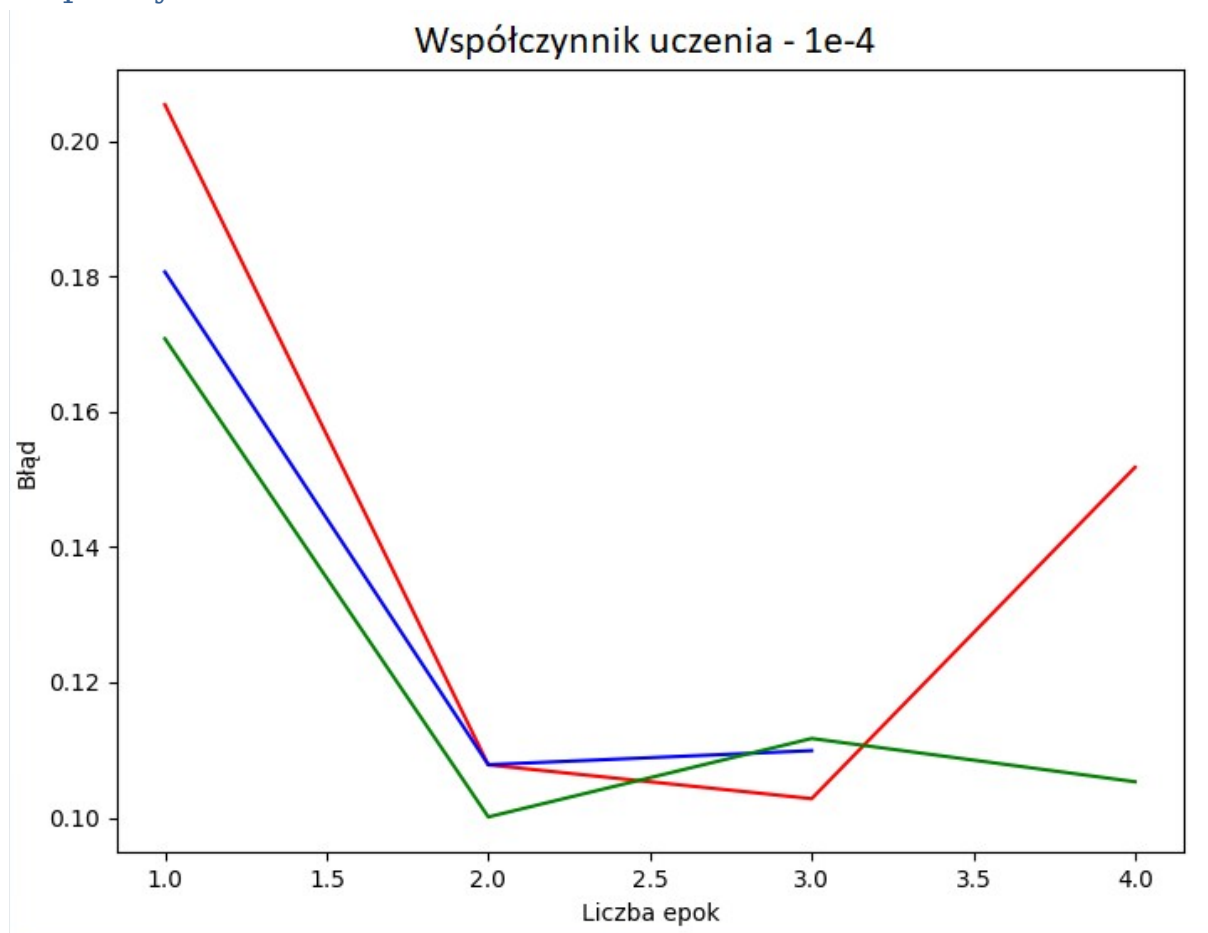
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|--------|---------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0 | 951.00 | 0.00 | 9.40 | 3.80 | 2.30 | 21.60 | 16.70 | 1.50 | 4.20 | 6.80 |
| 1 | 0.00 | 1104.00 | 7.60 | 0.10 | 4.10 | 3.70 | 3.10 | 13.70 | 7.10 | 5.70 |
| 2 | 1.50 | 1.60 | 873.20 | 17.70 | 4.70 | 8.20 | 8.80 | 18.80 | 4.40 | 1.20 |
| 3 | 1.40 | 6.60 | 25.00 | 900.30 | 0.40 | 39.00 | 0.30 | 8.00 | 16.40 | 12.10 |
| 4 | 0.00 | 0.90 | 16.20 | 0.00 | 908.10 | 10.40 | 14.80 | 8.70 | 9.80 | 28.30 |
| 5 | 11.50 | 0.30 | 7.80 | 39.40 | 0.10 | 732.50 | 20.50 | 0.50 | 27.90 | 13.60 |
| 6 | 7.00 | 3.60 | 15.90 | 2.00 | 14.60 | 15.80 | 879.20 | 0.00 | 5.70 | 0.10 |
| 7 | 1.60 | 1.90 | 13.50 | 12.50 | 1.30 | 9.50 | 0.90 | 928.70 | 9.90 | 15.20 |
| 8 | 6.00 | 16.10 | 58.80 | 26.00 | 4.80 | 44.10 | 13.70 | 8.10 | 882.10 | 15.20 |
| 9 | 0.00 | 0.00 | 4.60 | 8.20 | 41.60 | 7.20 | 0.00 | 40.00 | 6.50 | 910.80 |

Wnioski

Ogólnie można powiedzieć, że im więcej neuronów w warstwie ukrytej tym model szybciej się uczył (błąd osiągał rząd 0.1 na przestrzeni mniejszej liczby epok). Dodatkowo modele z większą liczbą neuronów osiągały lepsze wyniki (choć nieznacznie). Gdyby problem rozszerzyć o dodatkowe warstwy ukryte z pewnością doprowadziłoby to do uzyskania lepszych wyników (większej dokładności).

Wpływ współczynnika uczenia

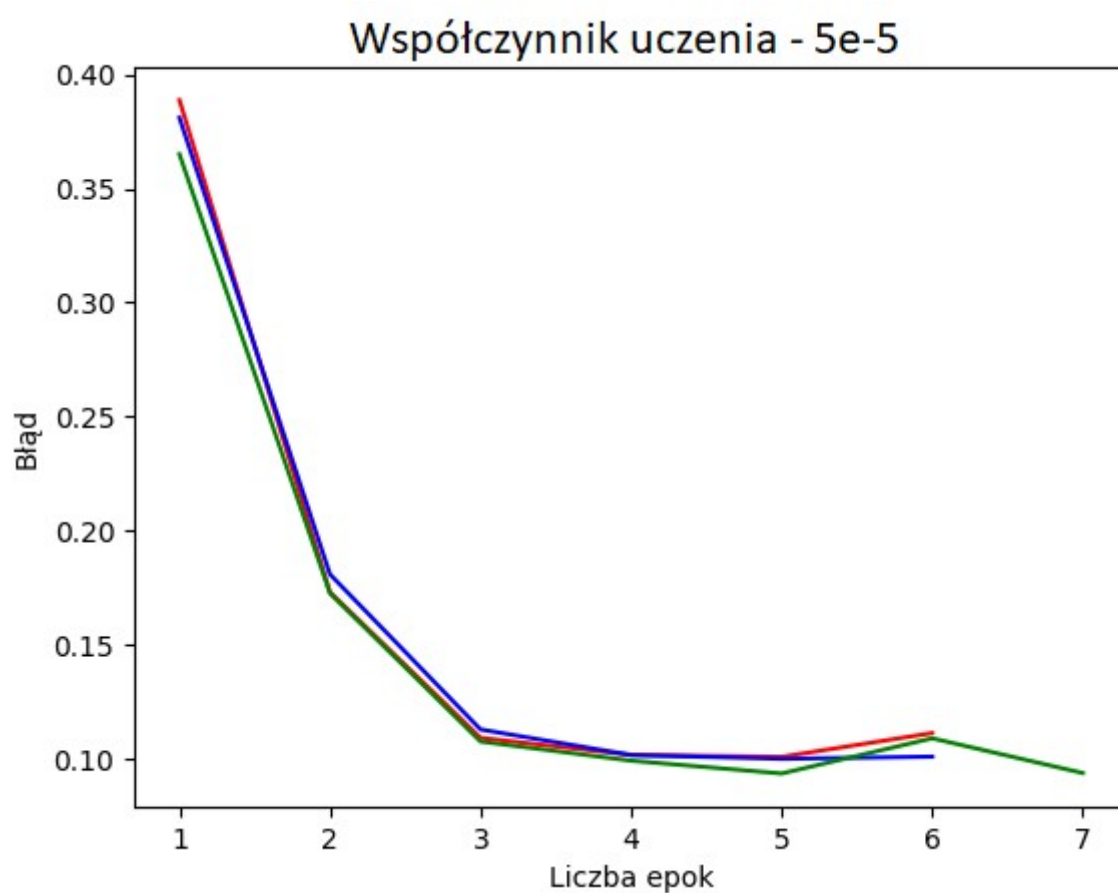
Współczynnik uczenia - $1e-4$



Macierz pomyłek:

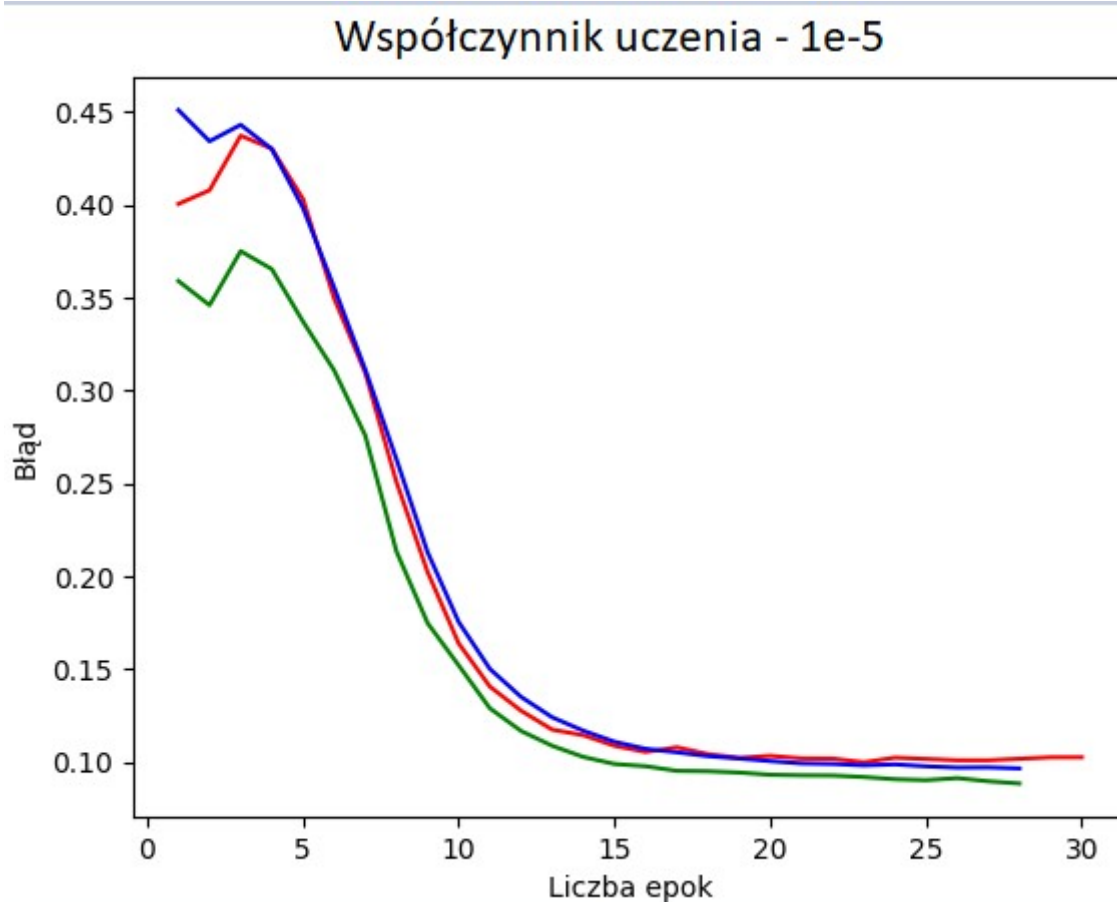
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|--------|---------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0 | 938.10 | 0.60 | 32.00 | 33.30 | 13.90 | 71.00 | 38.50 | 12.10 | 21.80 | 20.40 |
| 1 | 0.10 | 1105.20 | 10.60 | 1.60 | 3.80 | 6.00 | 4.00 | 19.50 | 10.80 | 6.70 |
| 2 | 6.20 | 6.30 | 877.50 | 46.20 | 8.80 | 16.40 | 25.40 | 24.20 | 18.70 | 4.30 |
| 3 | 2.30 | 4.10 | 21.20 | 838.70 | 0.40 | 36.80 | 0.20 | 6.80 | 15.30 | 8.10 |
| 4 | 1.60 | 0.40 | 8.40 | 1.10 | 866.80 | 16.60 | 16.80 | 13.60 | 11.80 | 52.80 |
| 5 | 3.40 | 0.60 | 1.00 | 30.10 | 0.70 | 640.10 | 9.80 | 0.20 | 21.40 | 14.10 |
| 6 | 16.80 | 3.60 | 23.30 | 8.00 | 40.10 | 34.50 | 853.60 | 2.70 | 16.60 | 9.60 |
| 7 | 6.20 | 2.10 | 15.40 | 15.10 | 4.50 | 16.30 | 0.80 | 908.20 | 10.70 | 37.20 |
| 8 | 5.20 | 11.50 | 40.60 | 26.50 | 8.80 | 47.00 | 8.90 | 4.70 | 839.30 | 17.90 |
| 9 | 0.10 | 0.60 | 2.00 | 9.40 | 34.20 | 7.30 | 0.00 | 36.00 | 7.60 | 668.93 |

Współczynnik uczenia 5e-5



Macierz pomyłek:

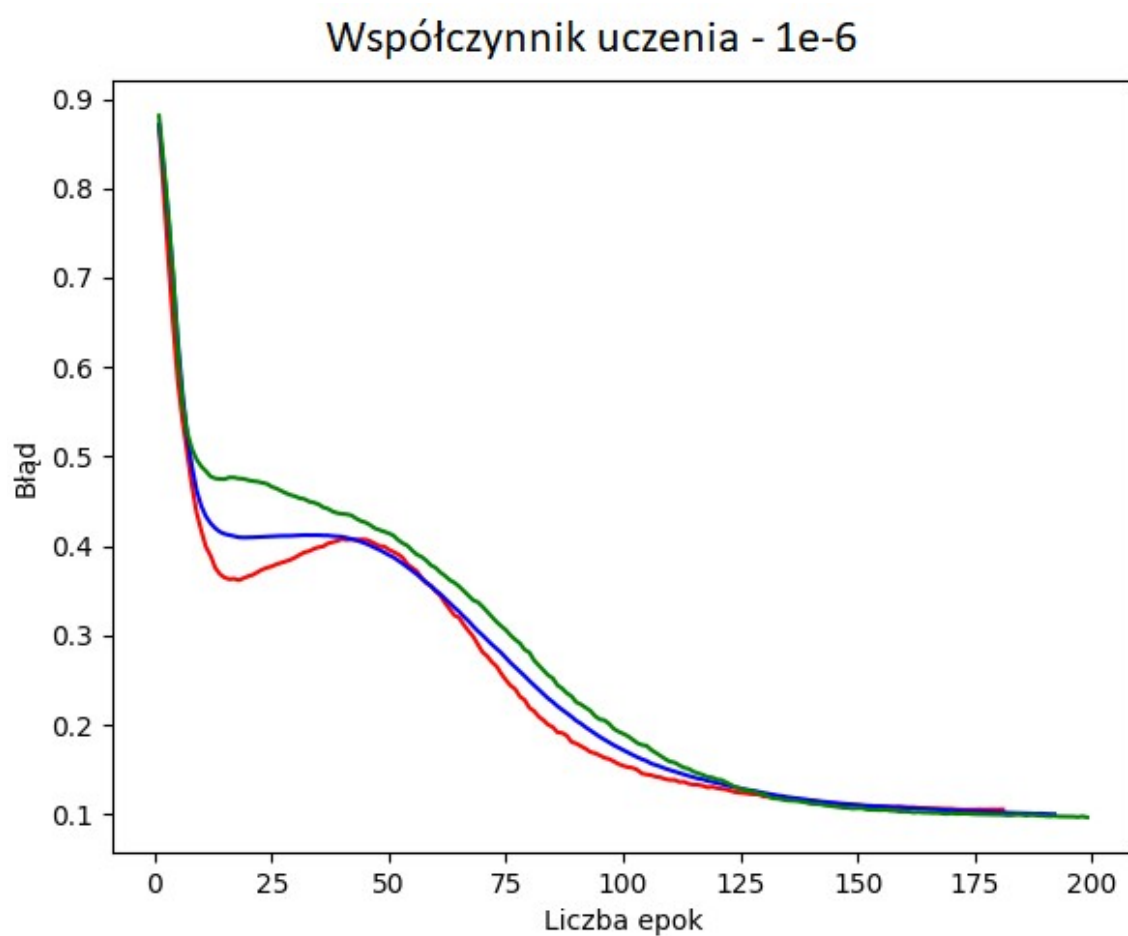
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|--------|---------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0 | 955.50 | 0.10 | 24.90 | 16.80 | 8.20 | 51.40 | 22.70 | 6.70 | 16.00 | 14.70 |
| 1 | 0.00 | 1109.70 | 7.90 | 1.70 | 3.20 | 5.60 | 3.90 | 17.20 | 11.90 | 5.60 |
| 2 | 1.10 | 2.70 | 876.40 | 25.00 | 1.80 | 7.70 | 9.10 | 18.80 | 12.00 | 1.40 |
| 3 | 1.40 | 4.00 | 17.40 | 884.50 | 0.10 | 50.20 | 0.30 | 5.60 | 21.30 | 11.30 |
| 4 | 0.40 | 0.80 | 14.60 | 0.80 | 911.90 | 18.40 | 15.20 | 11.10 | 16.80 | 48.30 |
| 5 | 5.50 | 0.90 | 3.10 | 28.70 | 0.80 | 678.40 | 14.20 | 0.10 | 22.40 | 13.90 |
| 6 | 8.40 | 3.70 | 25.20 | 6.40 | 21.10 | 27.10 | 885.90 | 0.70 | 11.80 | 1.70 |
| 7 | 2.80 | 1.70 | 17.40 | 16.60 | 0.60 | 13.80 | 0.40 | 924.80 | 11.20 | 27.60 |
| 8 | 4.90 | 11.30 | 40.90 | 22.60 | 4.70 | 31.60 | 6.30 | 4.30 | 839.80 | 13.40 |
| 9 | 0.00 | 0.10 | 4.20 | 6.90 | 29.60 | 7.80 | 0.00 | 38.70 | 10.80 | 871.10 |



Macierz pomyłek:

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|--------|---------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0 | 947.80 | 0.00 | 14.70 | 6.00 | 1.40 | 23.60 | 16.20 | 2.20 | 5.10 | 6.20 |
| 1 | 0.00 | 1109.90 | 5.00 | 1.10 | 3.10 | 3.30 | 3.90 | 16.30 | 11.80 | 7.00 |
| 2 | 2.20 | 1.80 | 883.30 | 22.80 | 4.40 | 6.90 | 7.90 | 19.30 | 6.60 | 1.20 |
| 3 | 2.50 | 5.50 | 27.00 | 891.10 | 0.10 | 52.80 | 0.00 | 5.40 | 17.50 | 9.90 |
| 4 | 0.20 | 0.50 | 16.70 | 0.20 | 910.80 | 11.10 | 14.80 | 9.30 | 12.30 | 32.00 |
| 5 | 8.30 | 0.30 | 2.70 | 38.20 | 0.20 | 708.80 | 16.40 | 0.30 | 29.00 | 14.20 |
| 6 | 9.60 | 3.20 | 16.70 | 2.50 | 18.50 | 16.20 | 890.20 | 0.50 | 8.20 | 1.00 |
| 7 | 2.60 | 1.60 | 12.20 | 13.40 | 0.80 | 11.30 | 0.70 | 925.90 | 11.60 | 21.70 |
| 8 | 6.60 | 12.00 | 49.20 | 25.60 | 4.40 | 49.20 | 7.90 | 5.70 | 863.80 | 14.10 |
| 9 | 0.20 | 0.20 | 4.50 | 9.10 | 38.30 | 8.80 | 0.00 | 43.10 | 8.10 | 901.70 |

Współczynnik uczenia – $1e-6$



Macierz pomyłek:

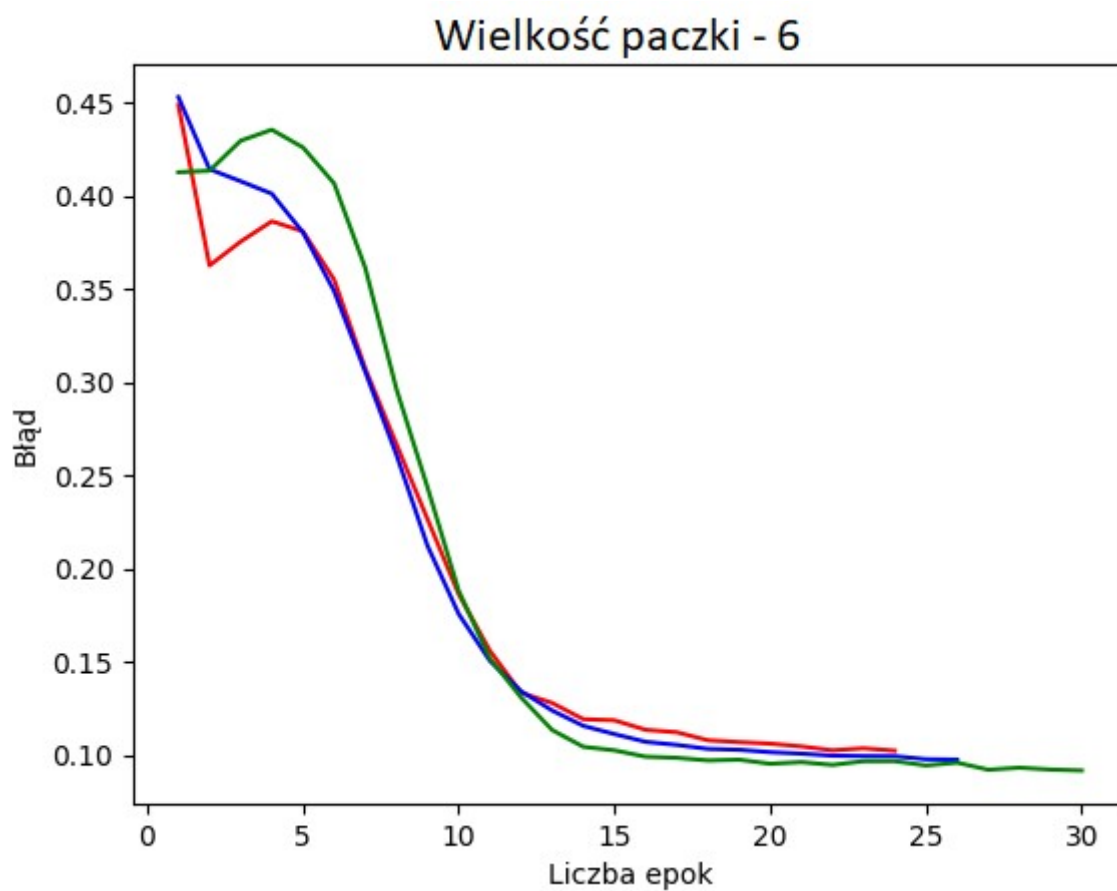
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|--------|---------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0 | 952.60 | 0.00 | 14.30 | 2.90 | 1.90 | 20.90 | 20.40 | 1.60 | 5.50 | 6.10 |
| 1 | 0.00 | 1115.20 | 8.80 | 3.70 | 4.40 | 2.10 | 3.60 | 23.30 | 16.70 | 7.00 |
| 2 | 2.80 | 2.00 | 876.40 | 25.60 | 3.70 | 9.50 | 7.70 | 19.20 | 9.70 | 1.10 |
| 3 | 3.60 | 4.60 | 22.00 | 895.40 | 0.10 | 56.90 | 0.30 | 4.50 | 18.50 | 10.30 |
| 4 | 0.60 | 0.70 | 20.40 | 0.30 | 905.30 | 14.40 | 14.70 | 9.00 | 15.30 | 32.30 |
| 5 | 6.40 | 0.00 | 1.90 | 30.60 | 0.20 | 689.90 | 13.70 | 0.20 | 26.60 | 12.70 |
| 6 | 7.20 | 3.10 | 14.10 | 0.90 | 15.00 | 14.70 | 891.20 | 0.20 | 13.60 | 1.80 |
| 7 | 2.20 | 1.80 | 14.10 | 17.10 | 1.10 | 14.70 | 0.40 | 924.40 | 14.20 | 20.20 |
| 8 | 4.60 | 7.60 | 54.40 | 24.10 | 3.30 | 58.60 | 6.00 | 3.20 | 839.10 | 10.40 |
| 9 | 0.00 | 0.00 | 5.60 | 9.40 | 47.00 | 10.30 | 0.00 | 42.40 | 14.80 | 907.10 |

Wnioski

Powyższe wykresy bardzo dobrze ilustrują kwestię odpowiedniego dobrania współczynnika uczenia. Za mały współczynnik ($1e-4$) skutkował wybuchającym gradientem a wagi zbyt dynamicznie się zmieniały. Dla współczynnika $5e-5$ widać już pewien proces nauki natomiast przeskoki błędu dalej są dość „ostre”. Dla współczynnika $1e-5$ widać ładny, a przy tym stopniowy spadek błędu. Gdy współczynnik był stosunkowo mały ($1e-6$) to model musiał „wydostać” się w niektórych momentach z optimum lokalnych a sam proces nauki zdecydowanie się wydłużył.

Wpływ wielkości paczki (batcha)

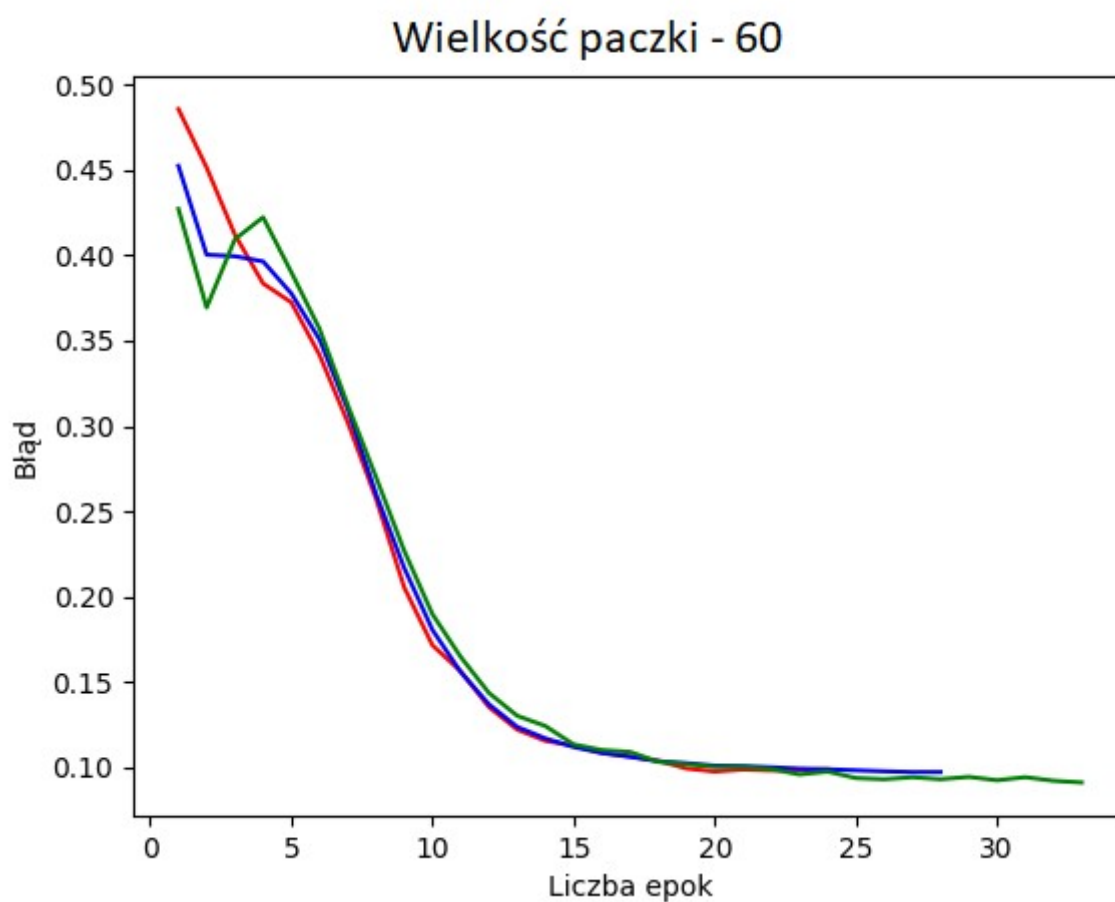
Wielkość paczki - 6



Macierz pomyłek:

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|--------|---------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0 | 953.20 | 0.00 | 13.90 | 3.70 | 1.90 | 22.90 | 19.30 | 2.60 | 5.70 | 6.30 |
| 1 | 0.00 | 1108.90 | 6.10 | 1.80 | 3.50 | 2.80 | 4.30 | 17.00 | 9.20 | 5.70 |
| 2 | 2.50 | 2.00 | 880.50 | 23.50 | 4.00 | 10.70 | 8.60 | 18.40 | 7.30 | 1.70 |
| 3 | 2.40 | 4.90 | 21.20 | 885.20 | 0.10 | 47.40 | 0.40 | 5.00 | 14.00 | 9.20 |
| 4 | 0.20 | 0.60 | 16.20 | 0.00 | 902.70 | 13.00 | 15.10 | 9.80 | 12.60 | 29.90 |
| 5 | 5.80 | 0.40 | 3.40 | 41.30 | 0.10 | 702.90 | 16.80 | 0.40 | 19.50 | 13.60 |
| 6 | 7.00 | 3.10 | 17.60 | 1.50 | 17.10 | 17.70 | 883.00 | 0.20 | 7.40 | 0.80 |
| 7 | 2.90 | 1.90 | 13.40 | 13.50 | 1.20 | 13.00 | 0.30 | 927.50 | 13.10 | 21.80 |
| 8 | 6.00 | 13.10 | 54.80 | 29.70 | 5.90 | 51.40 | 10.20 | 5.60 | 874.00 | 16.30 |
| 9 | 0.00 | 0.10 | 4.90 | 9.80 | 45.50 | 10.20 | 0.00 | 41.50 | 11.20 | 903.70 |

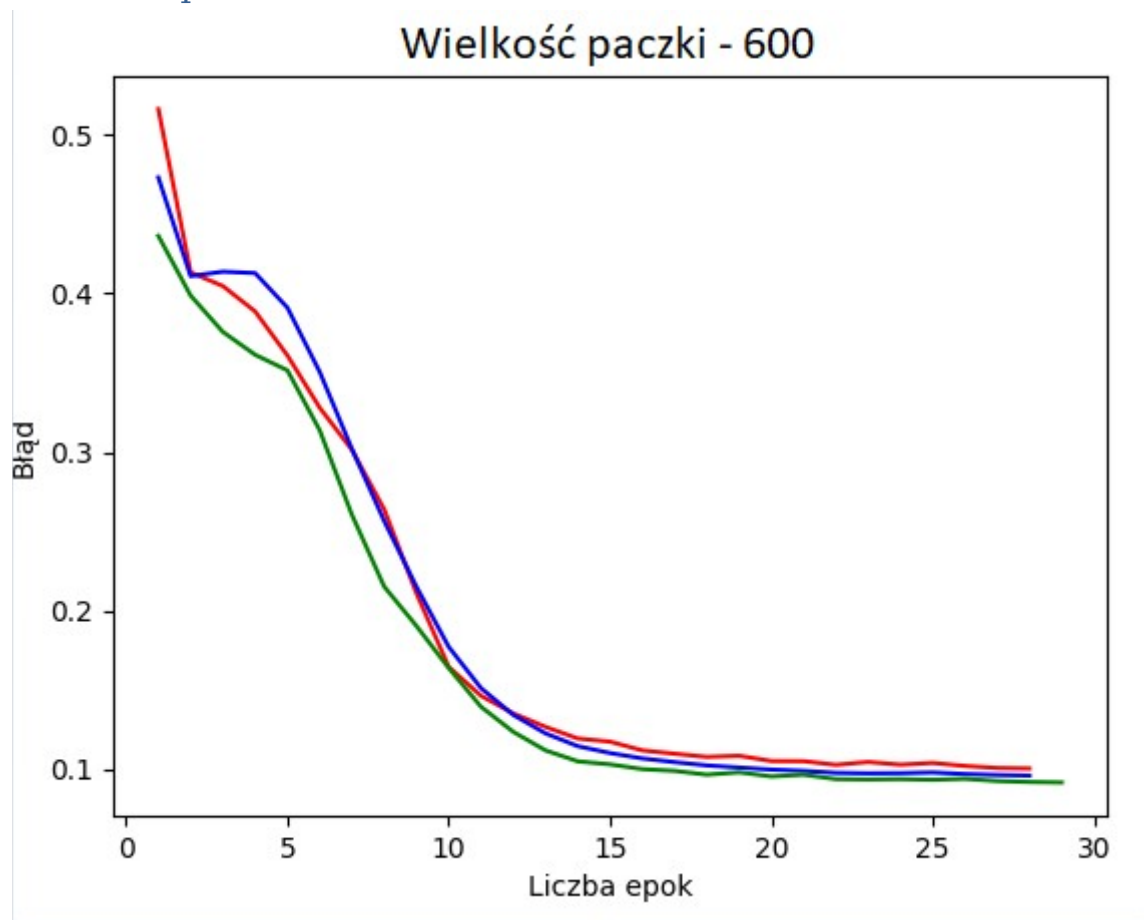
Wielkość paczki – 60



Macierz pomyłek:

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0 | 953.80 | 0.00 | 12.30 | 0.00 | 2.50 | 24.70 | 0.00 | 2.30 | 0.00 | 0.00 |
| 1 | 0.00 | 722.76 | 6.50 | 0.00 | 0.00 | 0.00 | 4.00 | 0.00 | 0.00 | 6.40 |
| 2 | 0.00 | 0.00 | 118.03 | 0.00 | 0.00 | 0.00 | 0.00 | 19.10 | 0.00 | 0.00 |
| 3 | 0.00 | 0.00 | 0.00 | 893.90 | 0.00 | 0.00 | 0.00 | 0.00 | 18.20 | 0.00 |
| 4 | 0.00 | 0.00 | 0.00 | 0.00 | 909.80 | 0.00 | 0.00 | 0.00 | 0.00 | 32.10 |
| 5 | 0.00 | 0.00 | 4.00 | 0.00 | 0.00 | 707.50 | 16.70 | 0.00 | 0.00 | 0.00 |
| 6 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 14.90 | 881.80 | 0.00 | 0.00 | 0.00 |
| 7 | 0.00 | 1.50 | 13.40 | 0.00 | 0.00 | 0.00 | 0.00 | 537.63 | 0.00 | 19.90 |
| 8 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 49.90 | 0.00 | 0.00 | 867.00 | 0.00 |
| 9 | 0.00 | 0.00 | 0.00 | 0.00 | 41.60 | 8.90 | 0.00 | 40.90 | 12.20 | 114.12 |

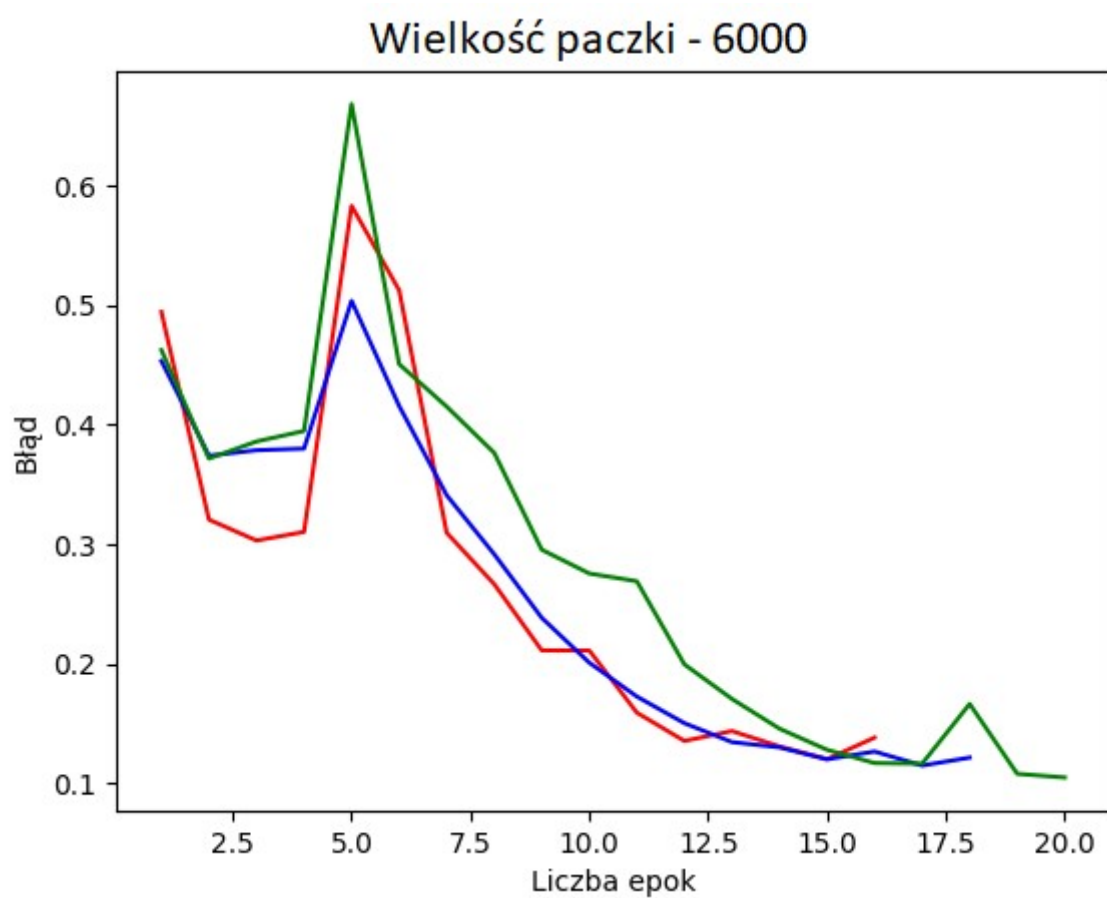
Wielkość paczki – 600



Macierz pomyłek:

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|--------|---------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0 | 952.20 | 0.00 | 13.10 | 5.00 | 3.00 | 24.30 | 16.80 | 2.30 | 5.10 | 7.30 |
| 1 | 0.00 | 1111.10 | 7.20 | 1.10 | 2.70 | 3.10 | 3.70 | 16.70 | 10.40 | 6.60 |
| 2 | 1.90 | 1.90 | 878.00 | 23.50 | 4.30 | 7.40 | 7.20 | 19.50 | 6.90 | 1.10 |
| 3 | 1.50 | 4.40 | 20.40 | 882.90 | 0.00 | 44.70 | 0.20 | 6.60 | 15.50 | 10.40 |
| 4 | 0.40 | 1.00 | 16.60 | 0.70 | 915.80 | 13.40 | 14.70 | 9.30 | 16.50 | 35.30 |
| 5 | 7.70 | 0.30 | 5.00 | 42.20 | 0.30 | 715.70 | 16.60 | 0.20 | 23.10 | 12.10 |
| 6 | 8.20 | 3.70 | 16.40 | 2.70 | 13.90 | 12.70 | 889.40 | 0.40 | 9.20 | 0.50 |
| 7 | 1.60 | 1.40 | 12.60 | 14.70 | 1.10 | 12.00 | 0.60 | 918.60 | 10.00 | 18.20 |
| 8 | 6.50 | 11.10 | 56.70 | 27.30 | 4.70 | 49.00 | 8.80 | 6.90 | 868.10 | 14.80 |
| 9 | 0.00 | 0.10 | 6.00 | 9.90 | 36.20 | 9.70 | 0.00 | 47.50 | 9.20 | 902.70 |

Wielkość paczki – 6000



Macierz pomyłek:

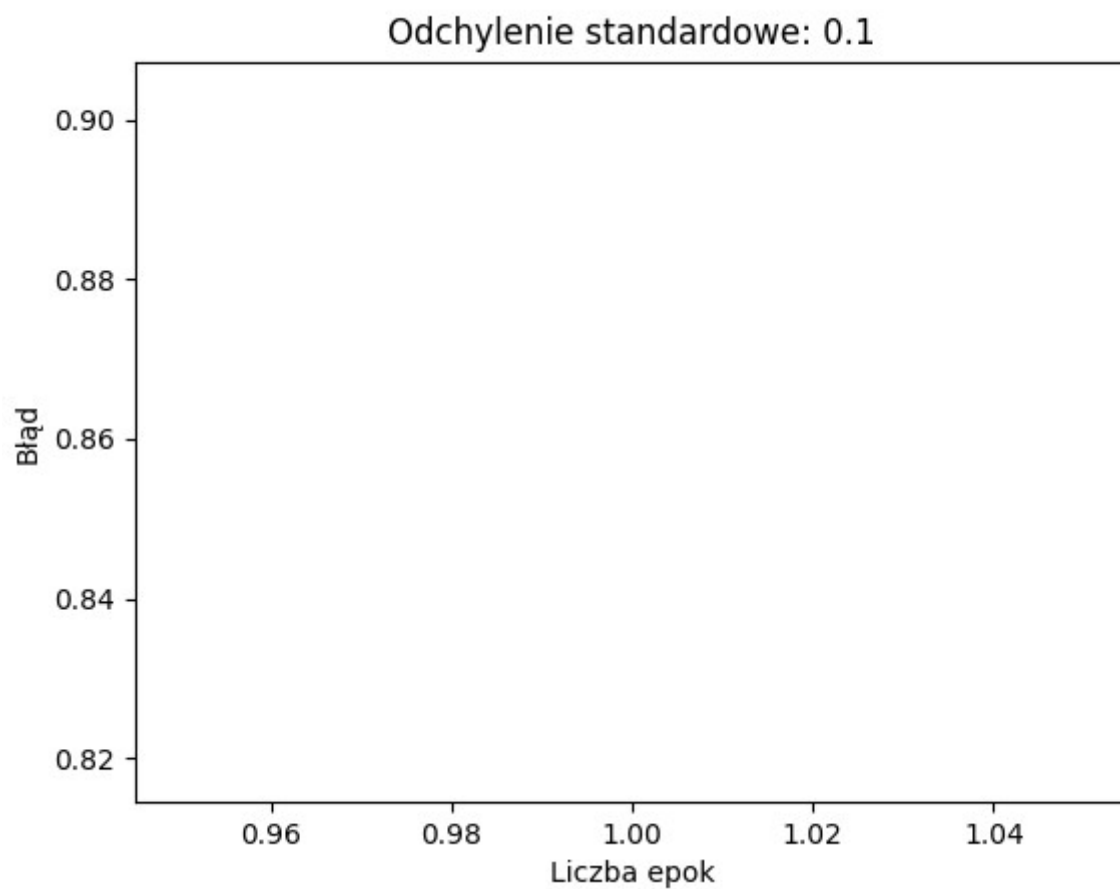
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|--------|---------|--------|--------|---------|--------|--------|--------|--------|--------|
| 0 | 696.77 | 0.00 | 18.20 | 5.70 | 0.00 | 22.50 | 0.00 | 3.50 | 7.40 | 0.00 |
| 1 | 0.00 | 1094.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 12.60 | 4.60 |
| 2 | 1.80 | 2.80 | 931.73 | 23.80 | 0.00 | 0.00 | 0.00 | 0.00 | 9.80 | 0.00 |
| 3 | 0.00 | 0.00 | 34.70 | 865.20 | 0.00 | 49.20 | 0.00 | 0.00 | 20.50 | 0.00 |
| 4 | 0.00 | 1.00 | 0.00 | 0.60 | 1532.52 | 19.80 | 0.00 | 0.00 | 17.10 | 0.00 |
| 5 | 18.00 | 0.50 | -0.00 | 0.00 | 0.10 | 702.40 | 18.30 | 0.00 | 47.10 | 0.00 |
| 6 | 0.00 | 0.00 | 19.00 | 0.00 | 14.70 | 20.50 | 877.70 | 0.20 | 16.30 | 1.10 |
| 7 | 1.70 | 0.00 | 0.00 | 0.00 | 1.40 | 12.70 | 0.70 | 873.99 | 14.30 | 24.60 |
| 8 | 4.70 | 23.50 | 61.70 | 29.00 | 5.60 | 42.50 | 9.60 | 8.30 | 808.40 | 9.00 |
| 9 | 0.00 | 0.10 | 6.10 | 9.70 | 67.40 | 11.30 | 0.00 | 57.90 | 20.50 | 743.70 |

Wnioski

Zasadniczo im większa wielkość paczki tym rzadziej następuje korekcja wag (przy czym korekcja ta powinna być lepiej zorientowana względem całości zbioru treningowego). Im mniejsza wielkość paczki natomiast tym częściej następuje korekcja. Przy obu podejściach model jest w stanie się wyuczyć (wiąże się to natomiast z różnym przebiegiem nauki, dla dużych paczek postępy są bardziej skokowe i czasem następuje regres).

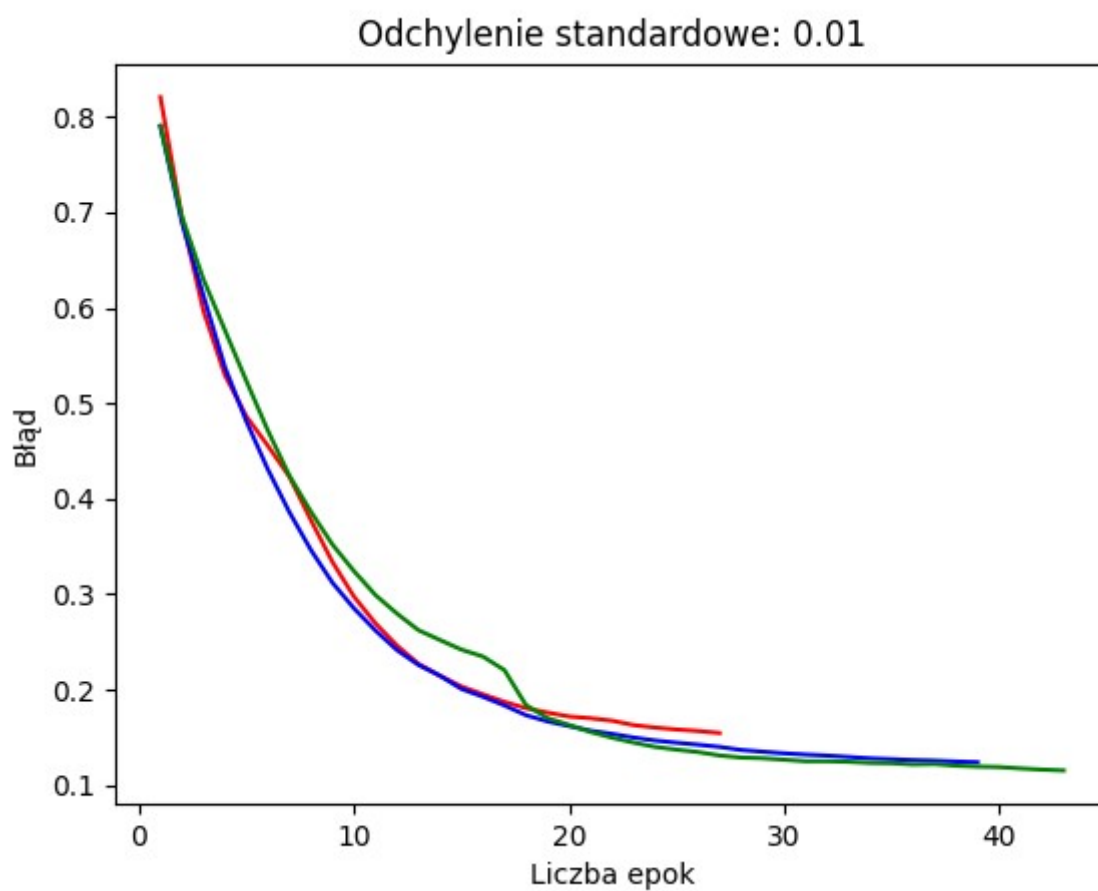
Wpływ inicjalizacji wag – odchylenie standardowe

Odchylenie standardowe – 0.1



W ogóle się nie wyuczył i osiągnął skuteczność około 10%. Prawdopodobnie doszło do sytuacji wystrzelającego gradientu.

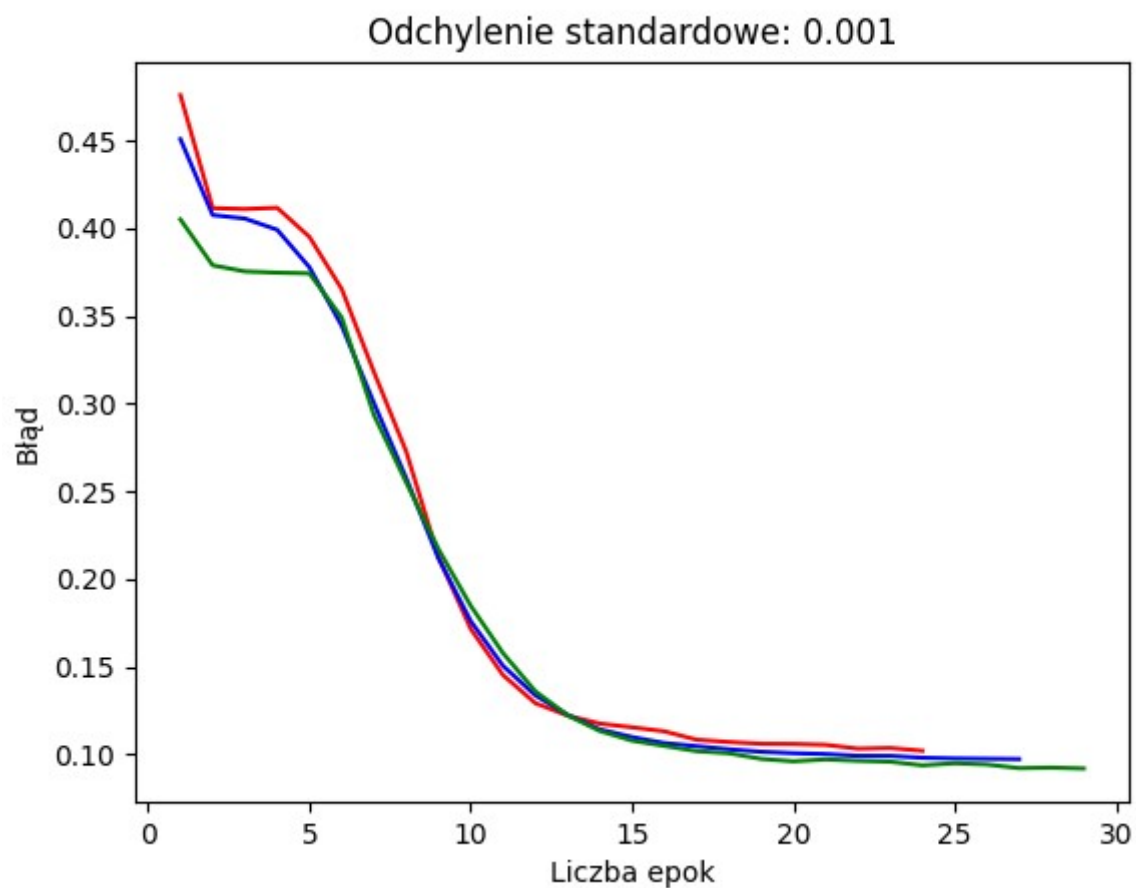
Odchylenie standardowe – 0.01



Macierz pomyłek:

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|--------|---------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0 | 939.50 | 0.00 | 16.20 | 8.80 | 3.00 | 28.10 | 20.70 | 4.00 | 8.20 | 9.30 |
| 1 | 0.40 | 1089.80 | 8.60 | 2.10 | 3.00 | 6.80 | 3.40 | 15.90 | 8.60 | 4.50 |
| 2 | 4.80 | 3.40 | 856.60 | 24.30 | 6.30 | 12.00 | 11.90 | 24.30 | 16.00 | 9.00 |
| 3 | 2.70 | 8.70 | 24.50 | 862.90 | 1.10 | 49.60 | 1.60 | 8.90 | 35.00 | 9.60 |
| 4 | 0.90 | 0.90 | 17.00 | 1.30 | 883.10 | 17.10 | 15.80 | 11.80 | 18.90 | 55.90 |
| 5 | 14.80 | 2.30 | 4.60 | 48.20 | 0.90 | 666.10 | 20.10 | 1.20 | 47.00 | 17.20 |
| 6 | 8.90 | 4.30 | 25.60 | 5.20 | 20.60 | 26.40 | 874.20 | 1.50 | 16.60 | 2.90 |
| 7 | 2.20 | 1.30 | 19.00 | 14.90 | 1.60 | 13.20 | 0.60 | 903.90 | 13.70 | 29.20 |
| 8 | 5.50 | 24.20 | 52.90 | 29.70 | 8.60 | 61.30 | 9.20 | 9.20 | 787.30 | 15.90 |
| 9 | 0.30 | 0.10 | 7.00 | 12.60 | 53.80 | 11.40 | 0.50 | 47.30 | 22.70 | 855.50 |

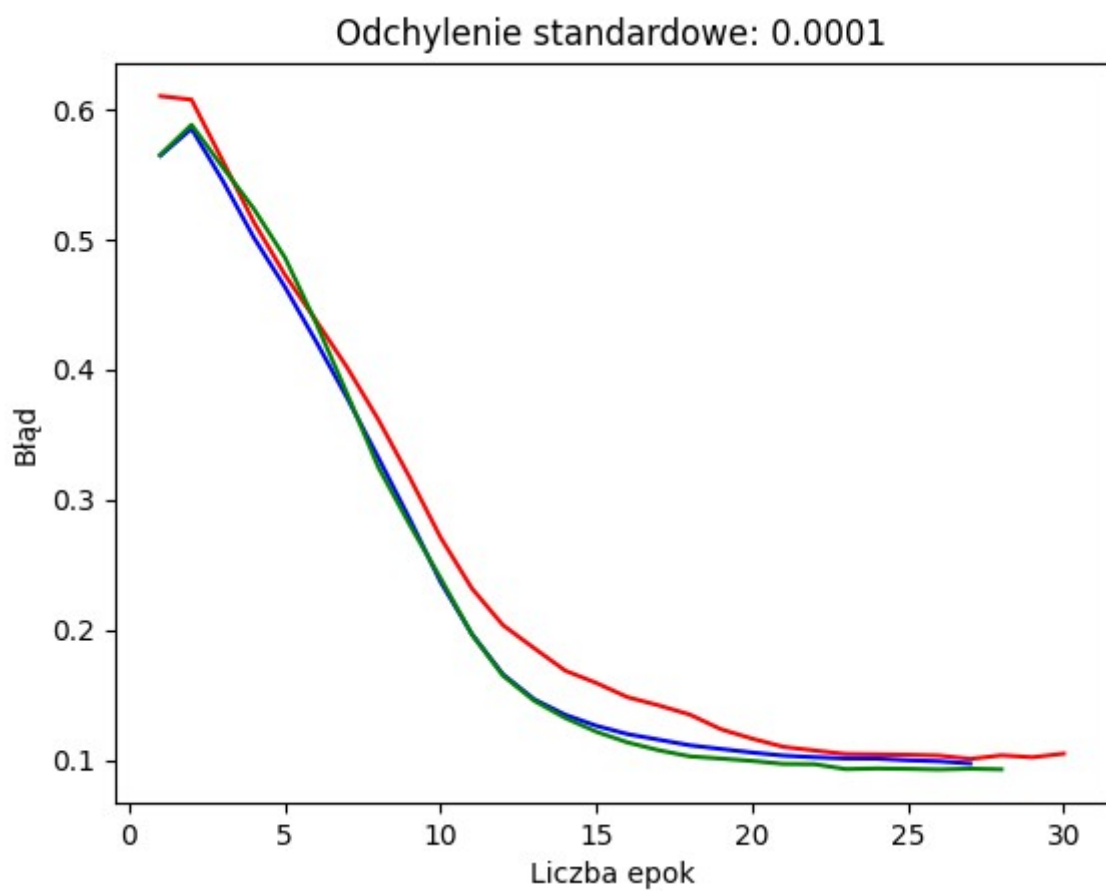
Odchylenie standardowe – 0.001



Macierz pomyłek:

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|--------|---------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0 | 953.70 | 0.00 | 14.10 | 5.60 | 1.60 | 25.00 | 17.00 | 2.00 | 5.50 | 6.60 |
| 1 | 0.00 | 1109.40 | 5.70 | 1.90 | 3.50 | 3.40 | 3.90 | 17.40 | 11.10 | 5.80 |
| 2 | 1.80 | 1.60 | 880.40 | 22.80 | 4.60 | 8.70 | 7.20 | 18.60 | 8.20 | 1.20 |
| 3 | 1.70 | 4.80 | 22.00 | 884.70 | 0.00 | 43.30 | 0.10 | 5.80 | 14.70 | 8.10 |
| 4 | 0.20 | 0.70 | 16.40 | 0.70 | 908.80 | 15.70 | 14.50 | 9.30 | 14.70 | 34.10 |
| 5 | 5.80 | 0.10 | 4.40 | 41.70 | 0.00 | 706.60 | 16.90 | 0.40 | 20.90 | 14.30 |
| 6 | 8.70 | 3.80 | 17.30 | 1.70 | 15.80 | 17.00 | 888.10 | 0.20 | 9.20 | 0.60 |
| 7 | 2.00 | 2.10 | 13.90 | 14.70 | 1.70 | 13.00 | 1.20 | 928.20 | 12.60 | 20.30 |
| 8 | 6.10 | 12.20 | 52.10 | 26.80 | 4.50 | 49.10 | 9.10 | 4.10 | 865.20 | 14.50 |
| 9 | 0.00 | 0.30 | 5.70 | 9.40 | 41.50 | 10.20 | 0.00 | 42.00 | 11.90 | 903.50 |

Odchylenie standardowe – 0.0001



Macierz pomyłek:

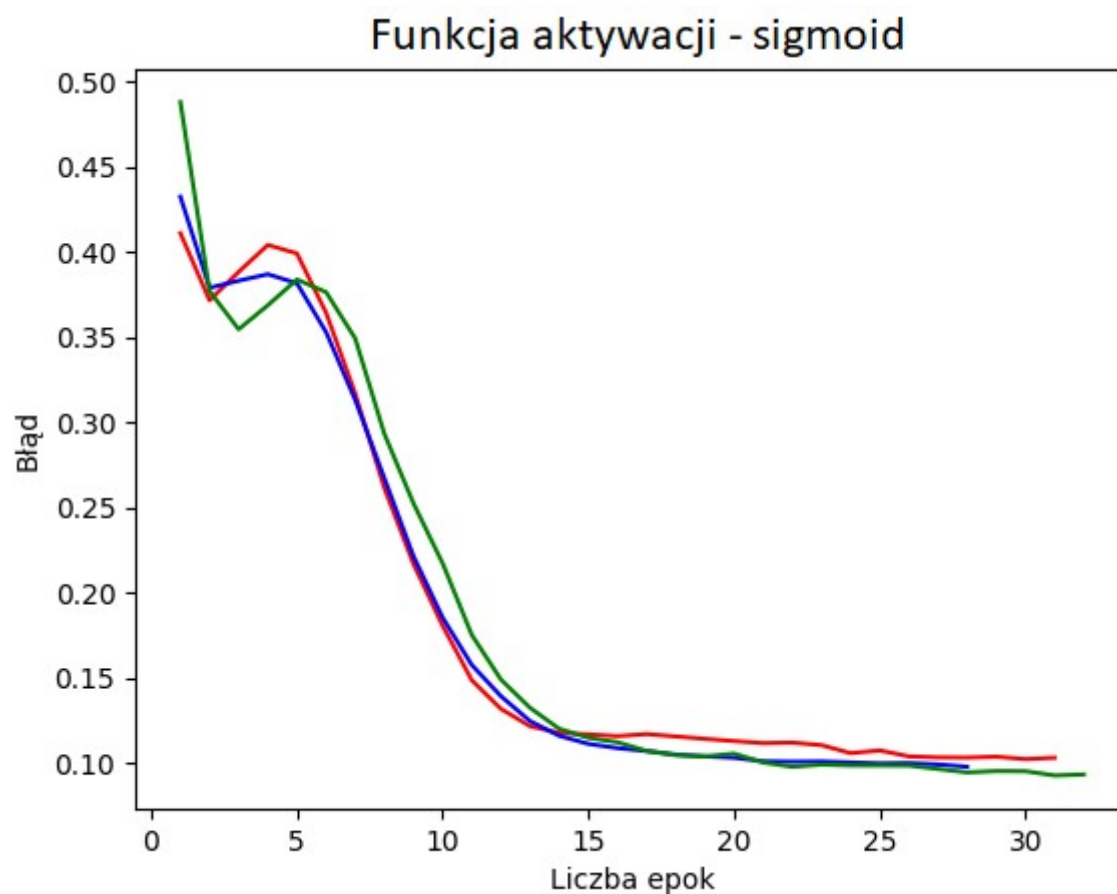
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|--------|---------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0 | 948.90 | 0.00 | 13.70 | 4.60 | 2.30 | 22.00 | 20.00 | 3.50 | 5.10 | 7.20 |
| 1 | 0.00 | 1110.40 | 5.20 | 4.10 | 4.30 | 3.20 | 3.60 | 22.10 | 12.10 | 8.00 |
| 2 | 1.70 | 2.40 | 865.70 | 23.50 | 4.00 | 6.40 | 9.70 | 15.50 | 7.50 | 0.30 |
| 3 | 4.00 | 6.90 | 26.50 | 893.80 | 0.10 | 50.30 | 0.40 | 2.70 | 16.90 | 10.60 |
| 4 | 0.20 | 0.50 | 21.70 | 0.10 | 896.30 | 11.60 | 18.80 | 10.40 | 10.80 | 23.70 |
| 5 | 6.60 | 0.60 | 2.70 | 30.20 | 0.00 | 705.60 | 11.70 | 0.30 | 21.30 | 10.30 |
| 6 | 8.40 | 2.90 | 14.50 | 1.50 | 16.20 | 15.20 | 883.10 | 0.20 | 6.20 | 0.70 |
| 7 | 3.40 | 1.60 | 14.30 | 14.30 | 1.70 | 14.40 | 0.40 | 925.90 | 16.40 | 23.20 |
| 8 | 6.70 | 9.30 | 61.90 | 29.30 | 5.50 | 55.20 | 10.30 | 4.90 | 868.80 | 15.60 |
| 9 | 0.10 | 0.40 | 5.80 | 8.60 | 51.60 | 8.10 | 0.00 | 42.50 | 8.90 | 909.40 |

Wnioski

Poprawna inicjalizacja wag określa czy model będzie w stanie się w ogóle wyuczyć oraz z jakiego „pułapu” rozpocznie naukę. Najlepszy startowy błąd został osiągnięty dla trzeciego eksperymentu (0.001), czyli pomiędzy dwoma pozostałymi eksperymentami.

Wpływ funkcji aktywacji

Funkcja aktywacji – sigmoid

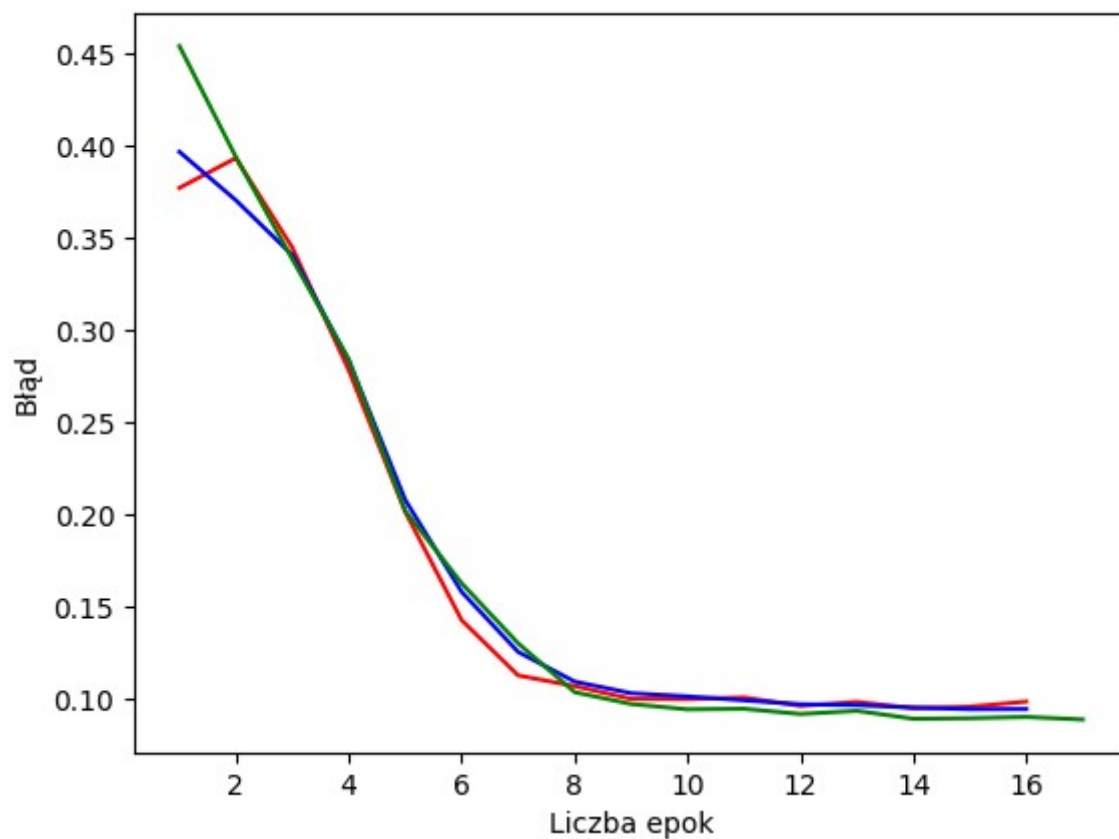


Macierz pomyłek:

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|--------|---------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0 | 951.30 | 0.00 | 12.70 | 4.30 | 2.50 | 19.90 | 18.50 | 2.00 | 4.70 | 6.30 |
| 1 | 0.10 | 1108.80 | 9.00 | 2.40 | 4.00 | 4.20 | 4.60 | 18.30 | 10.70 | 6.60 |
| 2 | 1.70 | 1.80 | 868.40 | 21.20 | 4.50 | 8.70 | 7.90 | 20.70 | 6.90 | 1.40 |
| 3 | 2.50 | 6.30 | 26.60 | 891.30 | 0.10 | 47.40 | 0.30 | 7.40 | 19.70 | 9.80 |
| 4 | 0.10 | 0.60 | 17.30 | 0.10 | 908.20 | 8.10 | 15.10 | 10.40 | 11.60 | 29.20 |
| 5 | 8.60 | 0.60 | 5.70 | 41.90 | 0.10 | 716.30 | 18.70 | 0.30 | 29.40 | 12.90 |
| 6 | 7.70 | 4.10 | 15.40 | 1.90 | 17.50 | 15.70 | 880.10 | 0.50 | 8.90 | 1.20 |
| 7 | 2.60 | 1.60 | 11.20 | 10.50 | 1.20 | 11.60 | 0.50 | 921.00 | 9.90 | 22.10 |
| 8 | 5.40 | 11.10 | 61.30 | 26.30 | 3.40 | 49.40 | 12.30 | 5.80 | 861.10 | 14.70 |
| 9 | 0.00 | 0.10 | 4.40 | 10.10 | 40.50 | 10.70 | 0.00 | 41.60 | 11.10 | 904.80 |

Funkcja aktywacji – tangens hiperboliczny

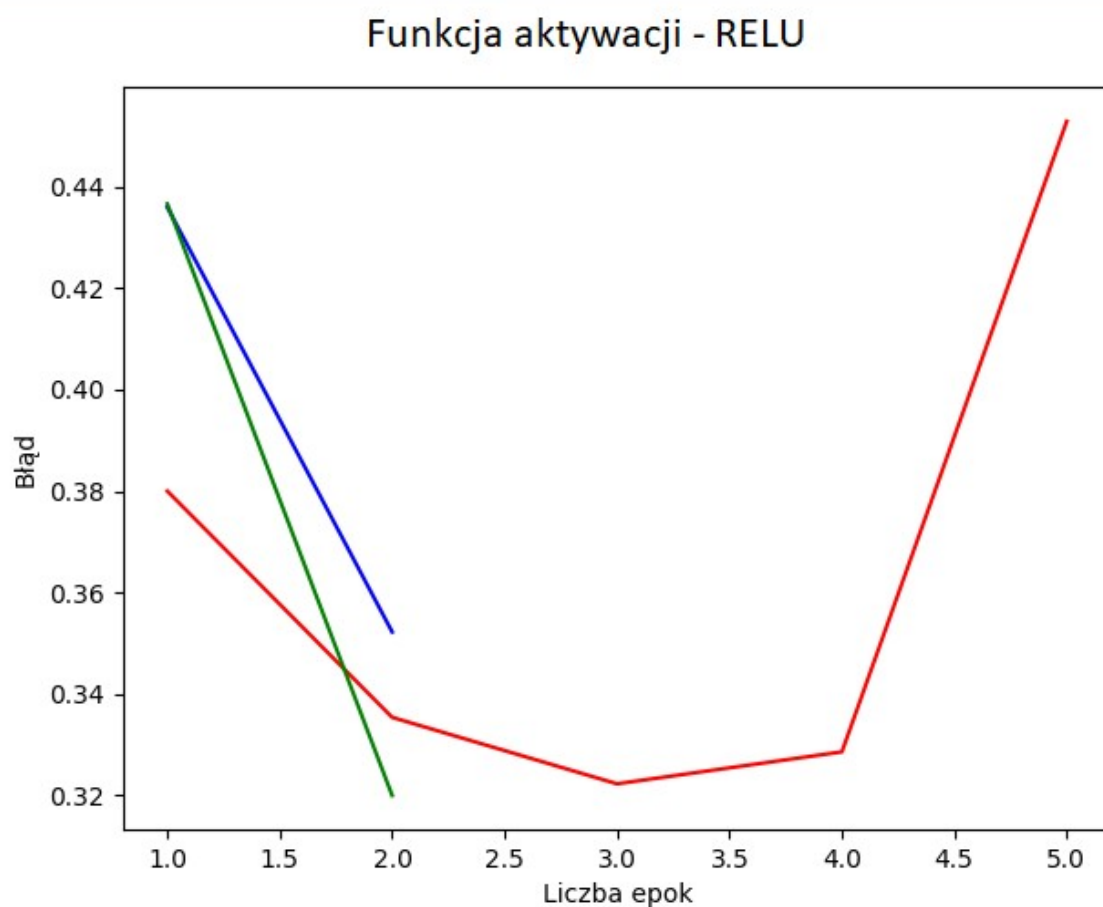
Funkcja aktywacji - htan



Macierz pomyłek:

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|--------|---------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0 | 955.10 | 0.10 | 16.50 | 7.80 | 2.40 | 19.00 | 17.30 | 3.70 | 10.40 | 9.10 |
| 1 | 0.00 | 1111.50 | 8.50 | 1.20 | 2.70 | 4.40 | 3.80 | 15.30 | 12.40 | 4.90 |
| 2 | 2.10 | 2.70 | 888.10 | 24.50 | 3.70 | 8.10 | 8.30 | 18.70 | 6.30 | 1.10 |
| 3 | 1.50 | 4.10 | 25.40 | 888.70 | 0.30 | 38.70 | 0.40 | 6.30 | 16.50 | 8.80 |
| 4 | 0.00 | 0.60 | 12.30 | 0.50 | 911.70 | 13.70 | 10.00 | 10.00 | 14.20 | 38.00 |
| 5 | 6.20 | 0.40 | 4.00 | 37.70 | 0.10 | 726.80 | 17.30 | 0.40 | 29.80 | 14.00 |
| 6 | 8.50 | 3.30 | 13.60 | 2.60 | 15.90 | 13.50 | 893.60 | 0.50 | 11.60 | 1.90 |
| 7 | 2.80 | 1.50 | 11.40 | 12.50 | 1.20 | 13.00 | 1.20 | 928.40 | 9.30 | 22.90 |
| 8 | 3.80 | 10.50 | 45.30 | 24.70 | 4.80 | 44.70 | 6.10 | 3.60 | 852.80 | 10.40 |
| 9 | 0.00 | 0.30 | 6.90 | 9.80 | 39.20 | 10.10 | 0.00 | 41.10 | 10.70 | 897.90 |

Funkcja aktywacji - RELU



Macierz pomyłek:

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|--------|-------|--------|--------|--------|--------|--------|-------|--------|--------|
| 0 | 520.80 | 0.00 | 8.80 | 10.40 | 0.10 | 102.20 | 4.10 | 0.70 | 34.80 | 1.90 |
| 1 | 0.40 | 0.00 | 39.90 | 8.10 | 7.00 | 23.70 | 7.20 | 22.00 | 108.60 | 14.80 |
| 2 | 0.00 | 17.50 | 791.50 | 47.40 | 1.80 | 43.20 | 20.20 | 27.20 | 126.50 | 5.20 |
| 3 | 0.00 | 11.50 | 46.60 | 850.20 | 3.60 | 151.10 | 2.70 | 22.60 | 156.60 | 0.00 |
| 4 | 43.80 | 0.70 | 42.50 | 5.10 | 813.50 | 68.50 | 62.00 | 49.50 | 101.10 | 0.00 |
| 5 | 68.60 | 12.00 | 9.00 | 18.10 | 0.00 | 170.60 | 3.50 | 0.70 | 0.00 | 0.00 |
| 6 | 136.20 | 5.10 | 45.30 | 11.70 | 42.90 | 131.40 | 848.80 | 3.80 | 0.00 | 9.70 |
| 7 | 34.60 | 1.00 | 20.90 | 24.80 | 2.80 | 76.00 | 0.10 | 0.00 | 57.70 | 78.30 |
| 8 | 18.20 | 6.30 | 4.40 | 9.00 | 0.10 | 59.30 | 1.90 | 0.00 | 40.50 | 0.10 |
| 9 | 69.90 | 2.20 | 23.10 | 25.20 | 110.20 | 66.00 | 0.00 | 60.40 | 131.00 | 555.90 |

Wnioski

Funkcja sigmoidalna jak i tangensa hiperbolicznego działają podobnie dobrze (przy czym nieznacznie lepsze rezultaty osiągnął tangens hiperboliczny). W przypadku RELU określiłem próg odcięcia na poziomie wartości 5 (żeby zapobiec wystrzelającemu gradientowi). Jak widać, pomimo ustalenia progu i tak doszło do rozbieżności wag co z resztą dobrze ilustruje problem związany z wykorzystaniem RELU.