WSI – Ćwiczenie 5

Zadanie wykonywane w parach

Proszę zaimplementować perceptron dwuwarstwowy i nauczyć go reprezentować funkcję $J: [-5,5] \rightarrow R$, daną wzorem: $J(x) = \sin(x*\operatorname{sqrt}(p[0]+1)) + \cos(x*\operatorname{sqrt}(p[1]+1))$, gdzie p[0] i p[1] to najmłodsze cyfry numerów indeksów wykonawców.

W sprawozdaniu powinny znaleźć się wykresy funkcji aproksymowanej i jej aproksymacji. Powinny również znaleźć się wskaźniki jakości aproksymacji. Jak liczba neuronów w warstwie ukrytej wpływa na jakość aproksymacji?

Przygotowałem dla Państwa <u>kod</u>, który powinien ułatwić wykonanie zadania a jednocześnie utrudnić bezmyślne użycie kodu z Internetu. Program ten należy dokończyć, a nie zastąpić całkowicie innym kodem. Nie można używać kodu z Internetu, czy bardziej ogólnie, kodu, którego nie jest się autorem.

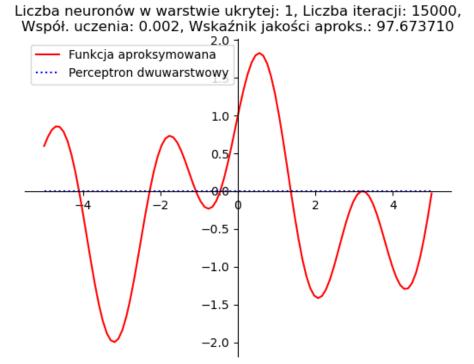
Wyniki

W ramach zadania zaimplementowano perceptron dwuwarstwowy, który nauczono reprezentować funkcję podaną wzorem:

$$J(x) = \sin(x * \operatorname{sqrt}(6)) + \cos(x * \operatorname{sqrt}(1)) = \sin(x * \operatorname{sqrt}(6)) + \cos(x)$$

Następnie badano jak liczba neuronów w warstwie ukrytej i (opcjonalnie) wartości pozostałych parametrów wpływają na jakość aproksymacji.

Test 1: 1 neuron w warstwie ukrytej, 15 000 iteracji, 0,002 współczynnik uczenia.

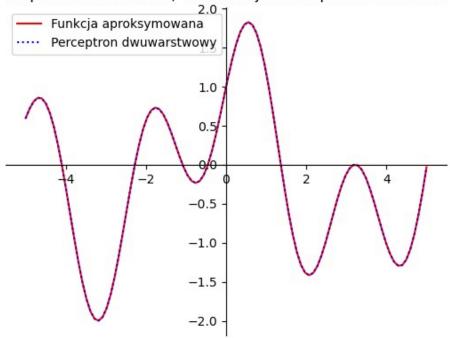


Wskaźnik jakości aproksymacji: 97,67370980554897

Zakładając tę samą liczbę iteracji (15 000) i współczynnik uczenia (0,002) sprawdzano, począwszy od liczby 1, to jak zwiększanie liczby neuronów w warstwie ukrytej wpływa na jakość aproksymacji. W pierwszym teście użyto tylko jednego neuronu i okazało się, że niemożliwe było uzyskanie przyzwoitego rozwiązania w przyjętej liczbie iteracji. Strata wyniosła ponad 97.

Test 2: 2 neurony w warstwie ukrytej, 15 000 iteracji, 0,002 współczynnik uczenia.

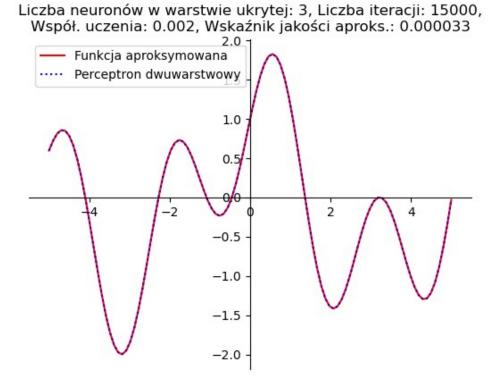
Liczba neuronów w warstwie ukrytej: 2, Liczba iteracji: 15000, Współ. uczenia: 0.002, Wskaźnik jakości aproks.: 0.000017



Wskaźnik jakości aproksymacji: 1,714753*10^(-5)

W drugim teście udało się natomiast uzyskać już o wiele lepsze wyniki. Wykres wyuczonej funkcji aproksymującej pokrywa się niemal w całości z wykresem funkcji aproksymowanej a wyliczona wartość funkcji straty jest bardzo niska (0,000017).

Test 3: 3 neurony w warstwie ukrytej, 15 000 iteracji, 0,002 współczynnik uczenia.

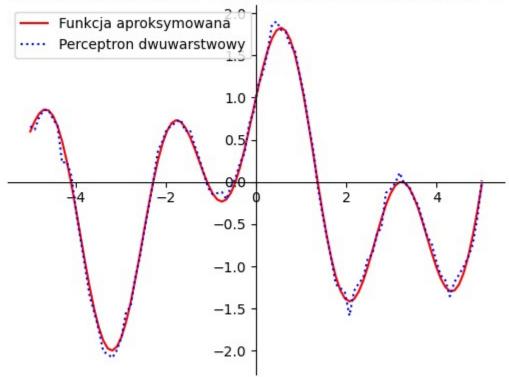


Wskaźnik jakości aproksymacji: 3,324402*10^(-5)

Dodanie jeszcze kolejnego neuronu, do warstwy ukrytej trenowanej sieci, pogorszyło z kolei uzyskane wyniki prawie dwukrotnie (0,000033). Jednakże, była to niewielka różnica ze względu na spory rząd wielkości wartości po przecinku (10^-5). Perceptron nadal bardzo dobrze reprezentował funkcję aproksymowaną a wykresy obu funkcji pokrywały się ze sobą.

Test 4: 10 neuronów w warstwie ukrytej, 15 000 iteracji, 0,002 współczynnik uczenia.

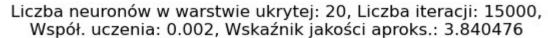


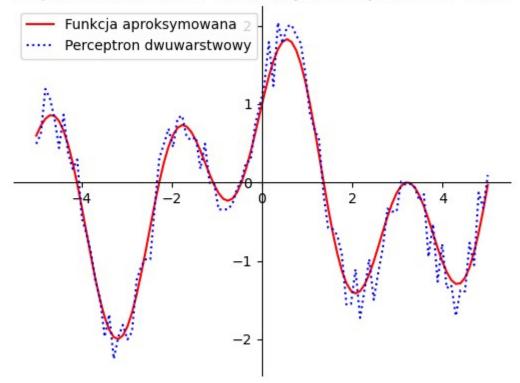


Wskaźnik jakości aproksymacji: 0,48434015305689737

Zachowując ten sam współczynnik uczenia i liczbę przeprowadzanych iteracji, w teście czwartym postanowiono sprawdzić jak wpłynie, na jakość aproksymacji modelu, użycie jeszcze większej liczby neuronów dla warstwy ukrytej. Ze względu na niewielkie zmiany spróbowano od razu wykorzystać 10 neuronów. Zgodnie z podejrzeniami, jakość aproksymacji znacząco zmalała, bo aż z 0,000033 do 0,48.

Test 5: 20 neuronów w warstwie ukrytej, 15 000 iteracji, 0,002 współczynnik uczenia.





Wskaźnik jakości aproksymacji: 3,840476236726488

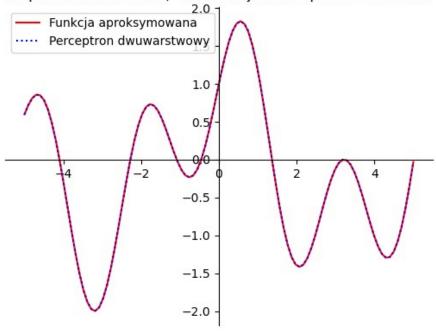
W teście piątym, nadal zachowując ten sam współczynnik uczenia i liczbę iteracji, sprawdzono jak zmieni się wartość wskaźnika jakości aproksymacji funkcji przez użycie dwukrotnie większej liczby neuronów (20) w warstwie ukrytej sieci. Po raz ponowny spełniły się przypuszczenia i jakość aproksymacji uległa znacznemu pogorszeniu z 0,48 do 3,84.

Testy dodatkowe (wpływ zmiany liczby epok i współczynnika uczenia)

Dla perceptronu dwuwarstwowego z warstwą ukrytą zawierającą 2 neurony, postanowiono sprawdzić dodatkowo jak szybko był w stanie on nauczyć się aproksymować zadaną funkcję. Otrzymane rezultaty przedstawiają wykresy zamieszczone na następnych stronach.

Test 6: 2 neurony w warstwie ukrytej, 12 500 iteracji, 0,002 współczynnik uczenia.

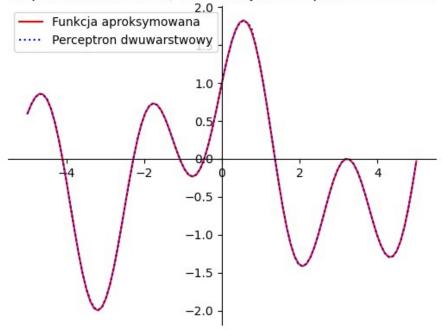
Liczba neuronów w warstwie ukrytej: 2, Liczba iteracji: 12500, Współ. uczenia: 0.002, Wskaźnik jakości aproks.: 0.000221



Wskaźnik jakości aproksymacji: 0,0002210693541896076

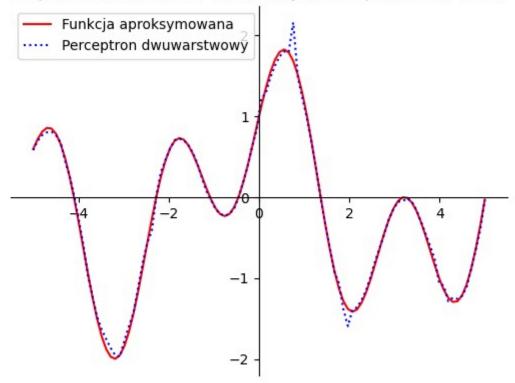
Test 7: 2 neurony w warstwie ukrytej, 10 000 iteracji, 0,002 współczynnik uczenia.

Liczba neuronów w warstwie ukrytej: 2, Liczba iteracji: 10000, Współ. uczenia: 0.002, Wskaźnik jakości aproks.: 0.002881



Test 8: 2 neurony w warstwie ukrytej, 5 000 iteracji, 0,002 współczynnik uczenia.

Liczba neuronów w warstwie ukrytej: 2, Liczba iteracji: 5000, Współ. uczenia: 0.002, Wskaźnik jakości aproks.: 0.443267

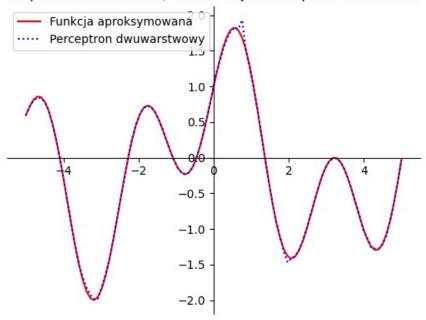


Wskaźnik jakości aproksymacji: 0,44326683191180716

Dla sieci z dwu-neuronową warstwą ukrytą można było wytrenować już po 10 000 iteracjach całkiem przyzwoity aproksymator. Błąd w tym przypadku wynosił około 0,003.

Test 9: (zmniejszenie) 2 neurony w warstwie ukrytej, 15 000 iteracji, 0,001 współcz. uczenia.

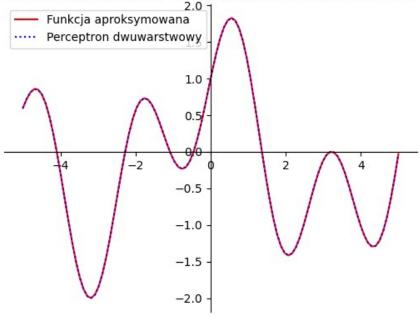
Liczba neuronów w warstwie ukrytej: 2, Liczba iteracji: 15000, Współ. uczenia: 0.001, Wskaźnik jakości aproks.: 0.088534



Wskaźnik jakości aproksymacji: 0,08853364057725138

Test 10: (zwiększenie) 2 neurony w warstwie ukrytej, 15 000 iteracji, 0,003 współczynnik uczenia.

Liczba neuronów w warstwie ukrytej: 2, Liczba iteracji: 15000, Współ. uczenia: 0.003, Wskaźnik jakości aproks.: 0.000000



Wskaźnik jakości aproksymacji: 9,544799 * 10^(-9)

Większy współczynnik uczenia może jako rezultat zwiększyć jakość aproksymacji funkcji i vice versa, mniejszy współczynnik może tę jakość pogorszyć.

Wnioski

Pytanie: Jak liczba neuronów w warstwie ukrytej wpływa na jakość aproksymacji?

Zbyt mała liczba neuronów w warstwie ukrytej powoduje, że model nie jest w stanie wystarczająco nauczyć się aproksymować zadaną funkcję. Zwiększenie ilości neuronów może polepszyć jakość aproksymacji, jednakże trzeba uważać i starać się znaleźć ich optymalną liczbę, ponieważ może dojść do przeuczenia modelu a otrzymywane błędy będą coraz większe.