

Sprawozdanie WSI ćwiczenie 5.

Treść zadania:

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

Czyli naszą funkcją jest $J(x) = \sin(x \cdot \sqrt{1+1}) + \cos(x \cdot \sqrt{7+1})$

Liczba próbek wejściowych to 300.

Wskaźnik jakości aproksymacji, którego używaliśmy do porównywania otrzymanych wyników to **średni błąd kwadratowy (ang. MSE)** oraz **błąd maksymalny**.

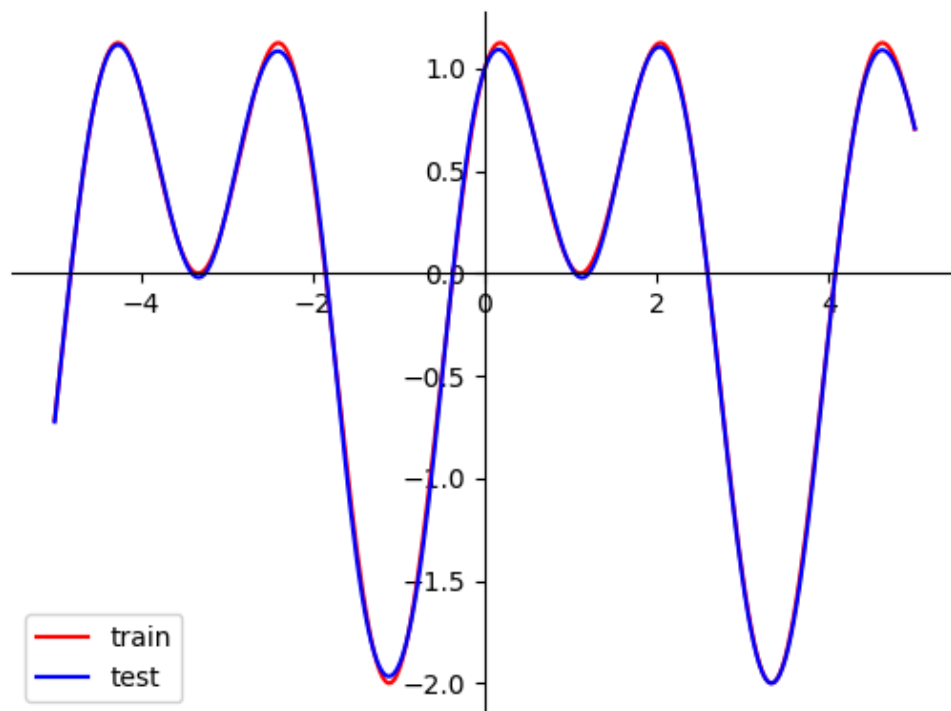
Parametry dla najlepszej aproksymacji funkcji i jej wykres:

Liczba iteracji: 15000

liczba neuronów: 13

współczynnik uczenia: 0,1

rozmiar mini zbioru: 10



MSE: 0,00042

Błąd maksymalny: 0,051

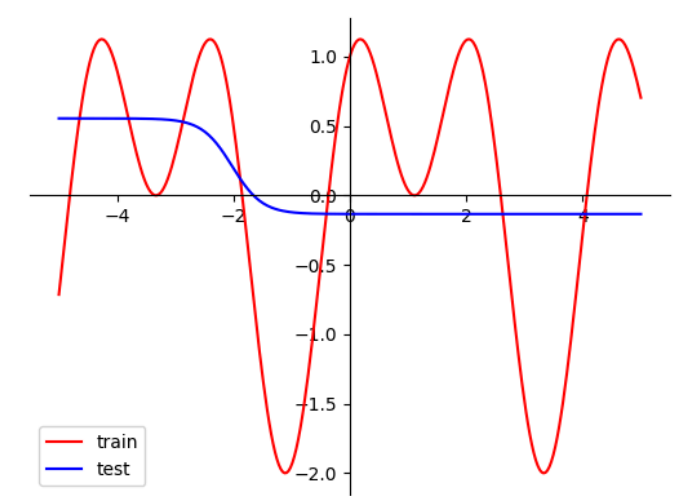
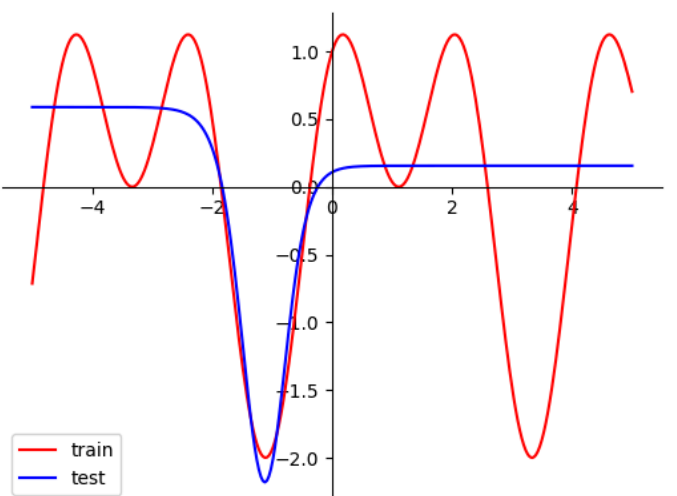
Jak liczba neuronów wpływa na jakość aproksymacji?

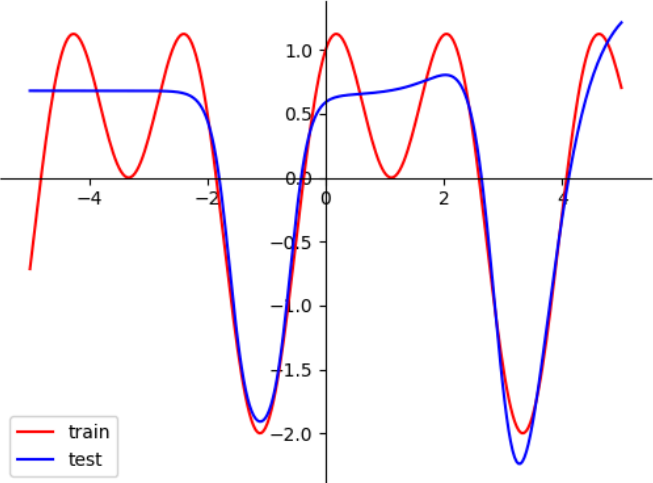
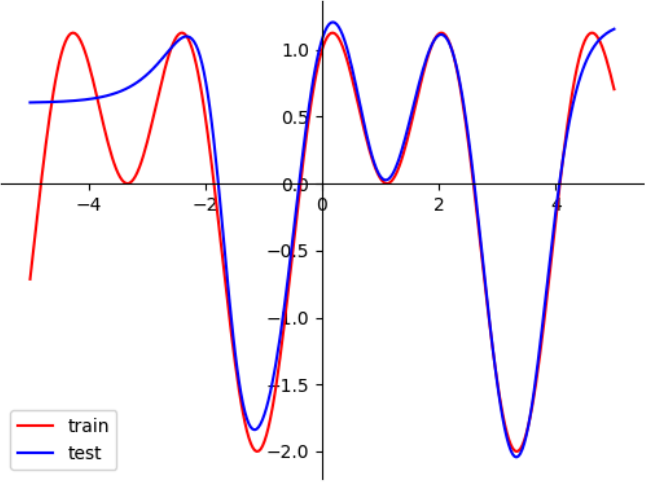
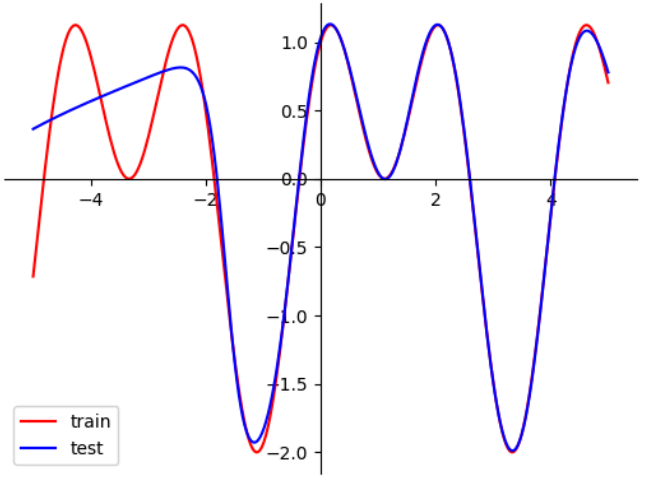
Dla każdego pomiaru parametry to:

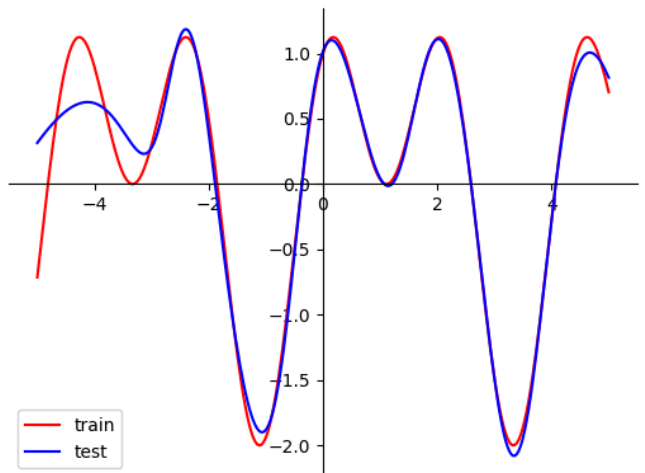
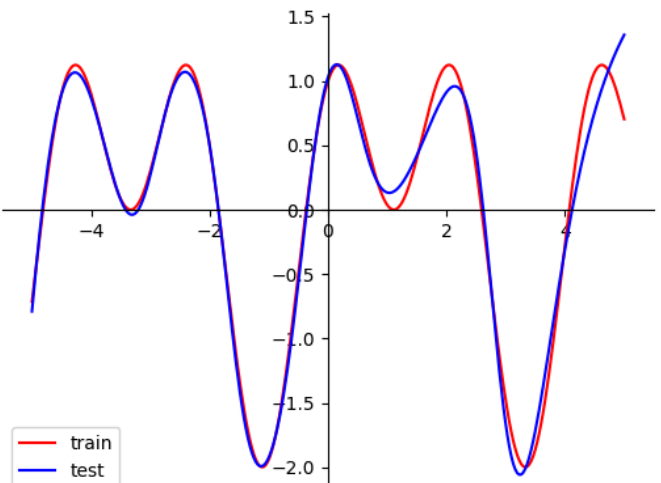
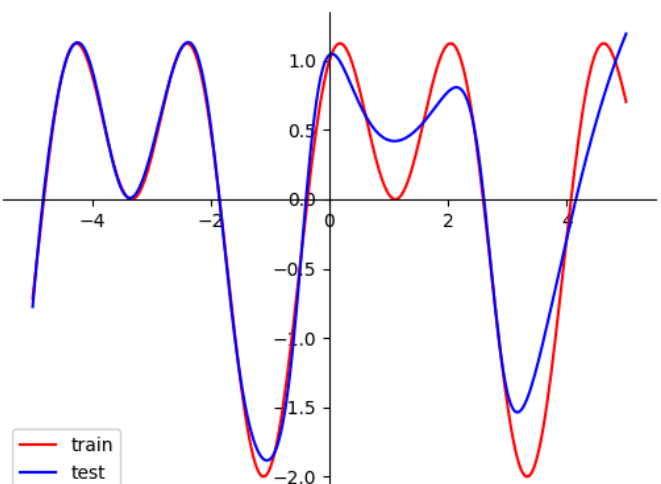
liczba iteracji: 5000

współczynnik uczenia: 0,1

rozmiar mini zbioru: 10

Liczba neuronów	MSE	Błąd maksymalny	Wykres
1	0,871	1,934	 <p>The plot shows a red 'train' curve that oscillates between approximately -2.0 and 1.0. The blue 'test' curve is a nearly flat horizontal line at y ≈ 0.5, failing to capture the underlying pattern of the training data. The x-axis ranges from -5 to 5, and the y-axis ranges from -2.0 to 1.0.</p>
2	0,585	2,167	 <p>The plot shows a red 'train' curve identical to the one above. The blue 'test' curve is a smooth, S-shaped curve that follows the general trend of the training data, indicating a better approximation. The axes and legend are the same as in the first plot.</p>

4	0,121	1,391	
8	0,073	1,312	
13	0,055	1,072	

30	0,032	1,029	
50	0,015	0,652	
80	0,058	0,653	

Wnioski:

Liczba neuronów w warstwie ukrytej ma znaczący wpływ na wyniki aproksymacji. Początkowo dla małej liczby neuronów dodanie niewielkiej ilości nowych neuronów znacznie zwiększa jakość aproksymacji. Lecz później dodawanie kolejnych neuronów niekoniecznie musi oznaczać poprawę aproksymacji, gdyż dla 50 neuronów uzyskaliśmy lepsze wyniki niż dla 80 neuronów. Także należy uważać przy wyborze liczby neuronów w warstwie ukrytej, aby nie występowało przeuczenie sieci przy zbyt dużej liczbie neuronów.

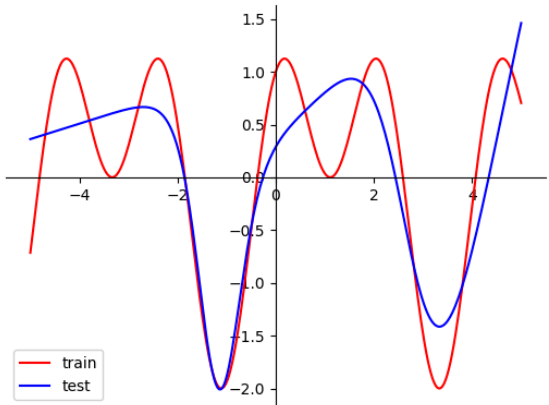
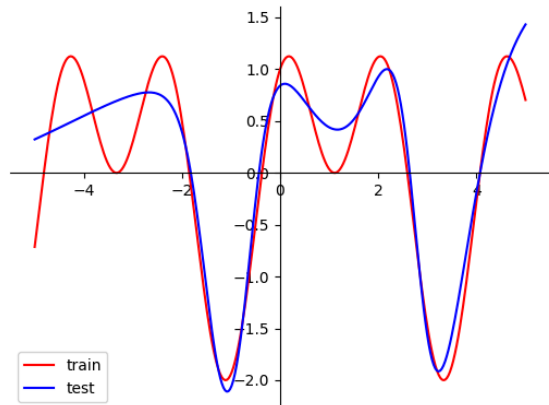
Wpływ liczby iteracji na wyniki aproksymacji

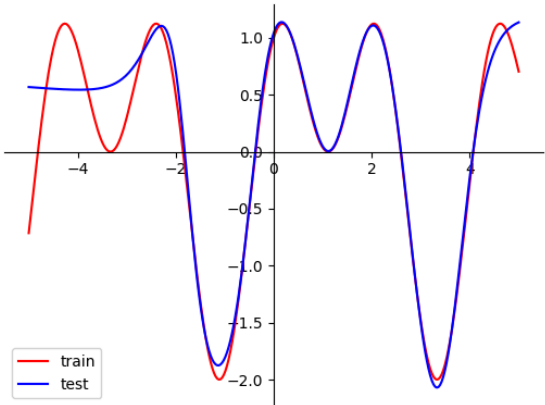
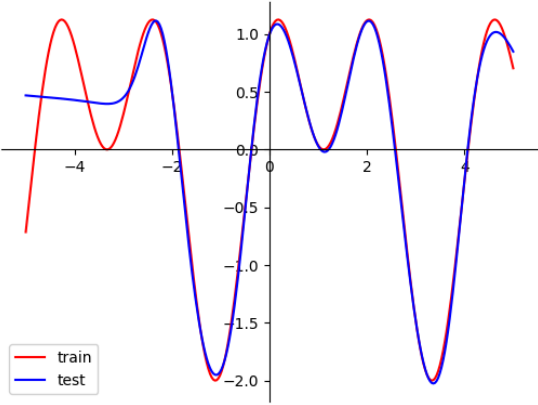
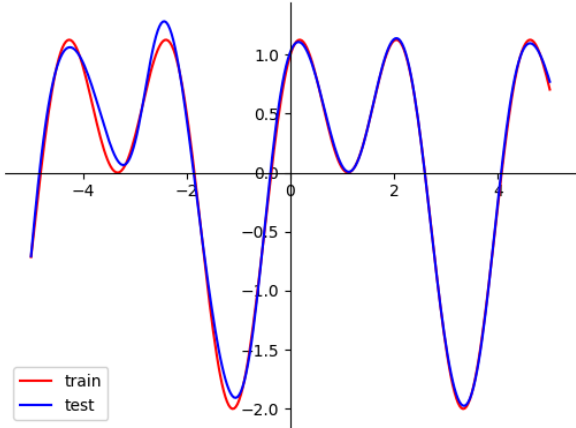
Dla każdego pomiaru parametry to:

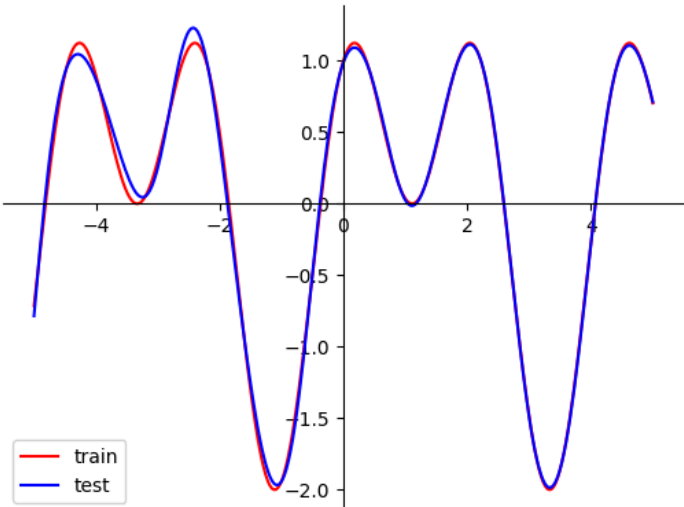
liczba neuronów: 10

współczynnik uczenia: 0,1

rozmiar mini zbioru: 10

Liczba iteracji	MSE	Błąd maksymalny	Wykres
500	0,207	1,076	
1000	0,093	1,034	

5000	0,054	1,214	
10000	0,047	1,138	
20000	0,0045	0,197	

40000	0,0043	0,148	
-------	--------	-------	--

Wnioski:

Czym większa liczba iteracji, tym dostajemy lepsze rezultaty. Jednak po przekroczeniu pewnego momentu rezultaty nie poprawiają się znacząco lub pozostają podobne i utrzymują się na pewnym poziomie.

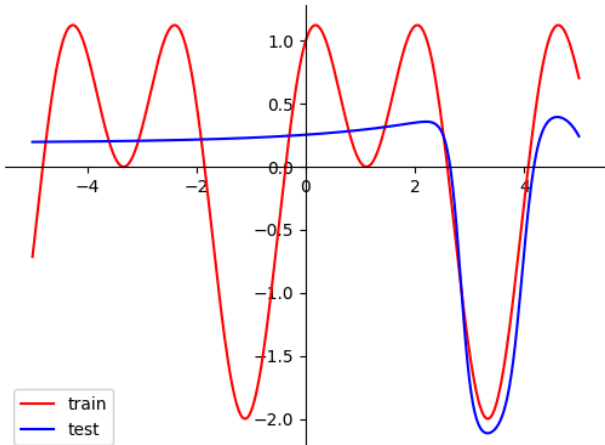
Wpływ współczynnik uczenia się na aproksymację funkcji

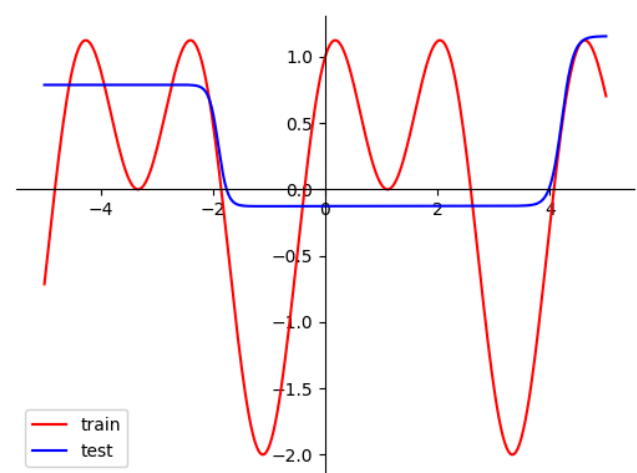
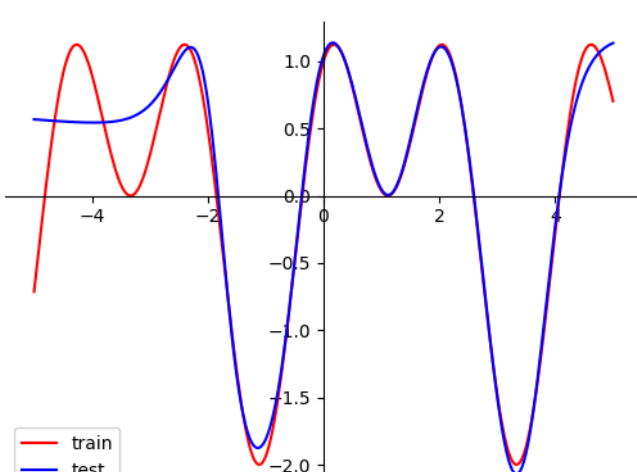
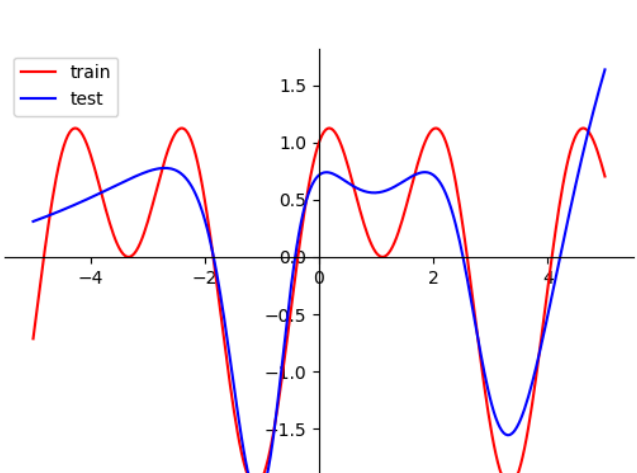
Dla każdego pomiaru parametry to:

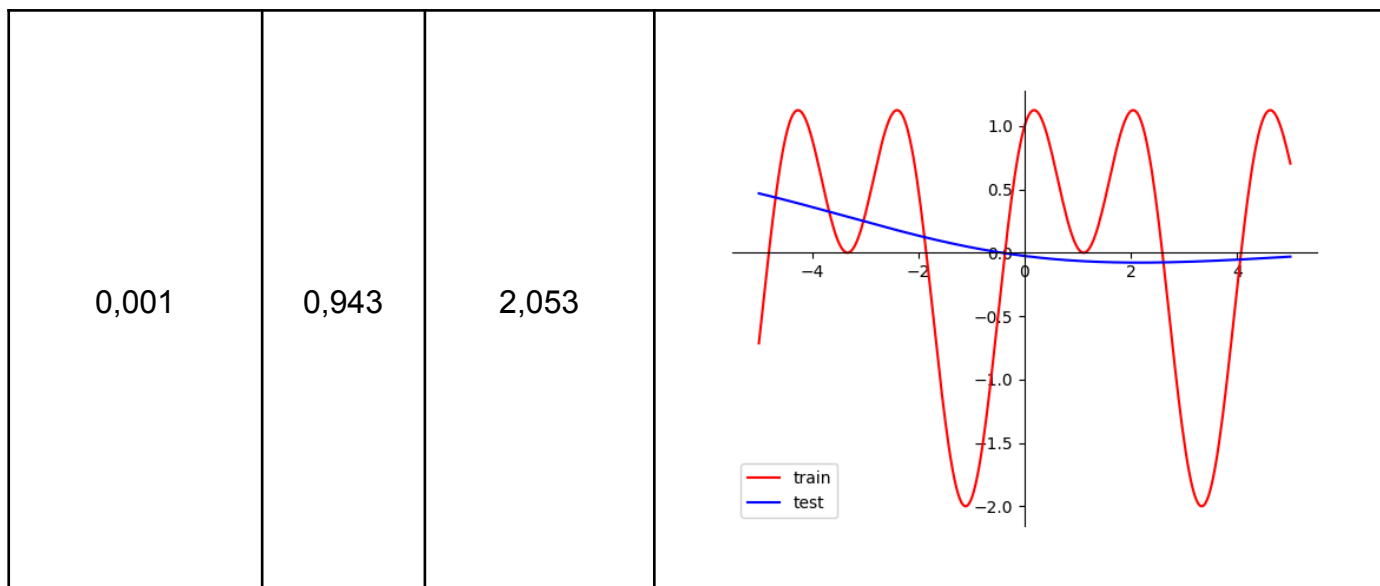
liczba neuronów: 10

liczba iteracji: 5000

rozmiar mini zbioru: 10

Współczynnik uczenia	MSE	Błąd maksymalny	Wykres
0,75	0,593	2,228	

0,5	0,733	1,872	 <p>A plot showing 'train' data (red line) and 'test' data (blue line). The x-axis ranges from -5 to 5, and the y-axis ranges from -2.0 to 1.0. The 'train' data is a periodic function with peaks at approximately 1.0 and troughs at approximately -2.0. The 'test' data is a constant function at y=0.5.</p>
0,1	0,054	1,214	 <p>A plot showing 'train' data (red line) and 'test' data (blue line). The x-axis ranges from -5 to 5, and the y-axis ranges from -2.0 to 1.0. The 'train' data is a periodic function. The 'test' data is a smooth curve that follows the general shape of the 'train' data but with some deviations.</p>
0,01	0,135	1,146	 <p>A plot showing 'train' data (red line) and 'test' data (blue line). The x-axis ranges from -5 to 5, and the y-axis ranges from -2.0 to 1.5. The 'train' data is a periodic function. The 'test' data is a smooth curve that follows the general shape of the 'train' data but with some deviations.</p>



Wnioski:

Zbyt duży współczynnik uczenia się powoduje, że sieć zbyt szybko chce znaleźć optymalne parametry do minimalizowania funkcji straty. W efekcie optymalizator może "przeskakiwać" nad minimum lub nawet doprowadzić do destabilizacji procesu uczenia, co skutkuje brakiem zbieżności lub oscylacjami wartości funkcji straty. Z drugiej strony zbyt mały współczynnik uczenia się powoduje że kroki do znalezienia odpowiednich wag są zbyt małe i znalezienie dobrej aproksymacji wymagałoby znacznie większej liczby iteracji. Co więcej, istnieje ryzyko, że model utknie w minimum lokalnym, nie osiągając globalnego minimum.

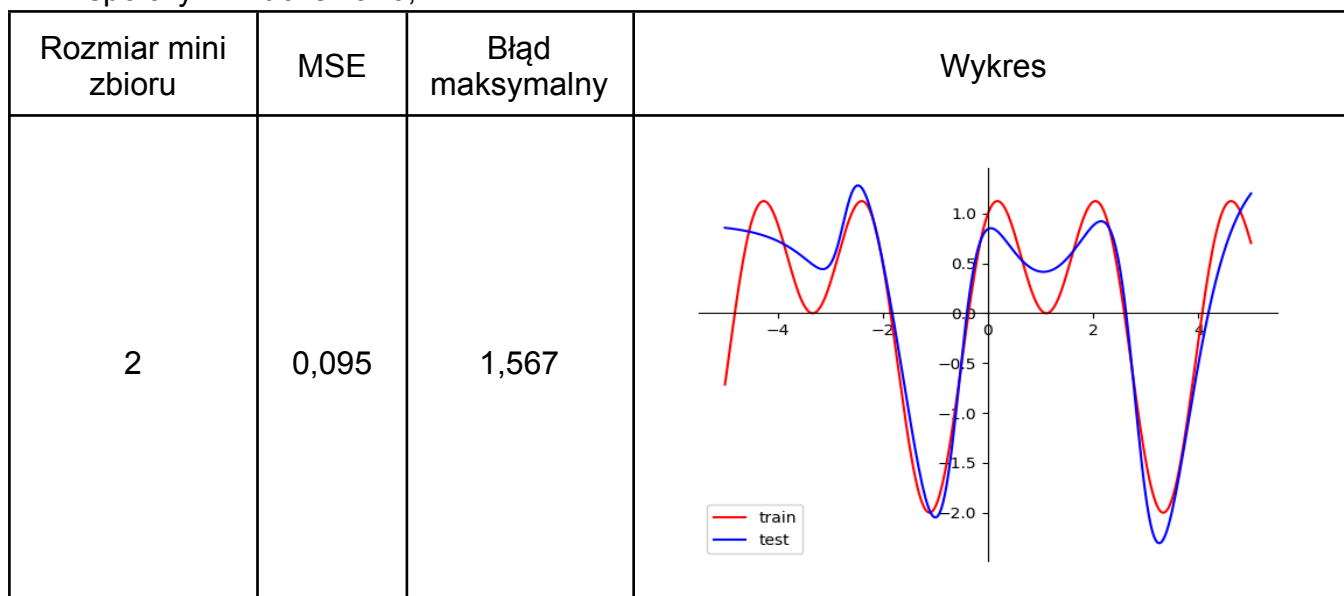
Wpływ rozmiaru mini zbioru na aproksymację funkcji

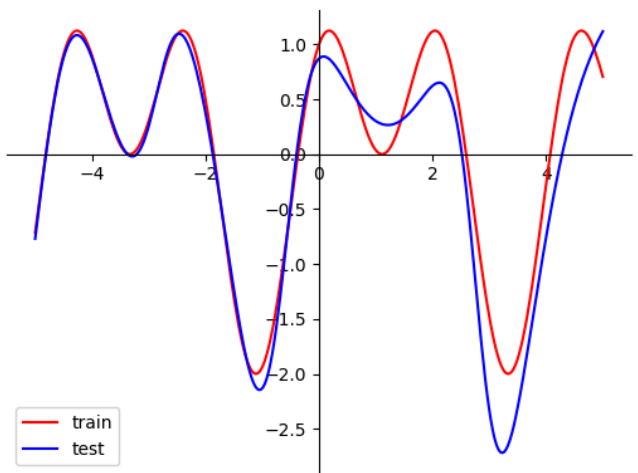
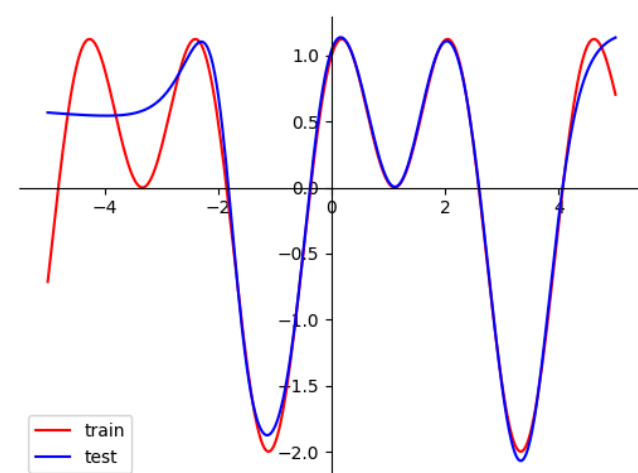
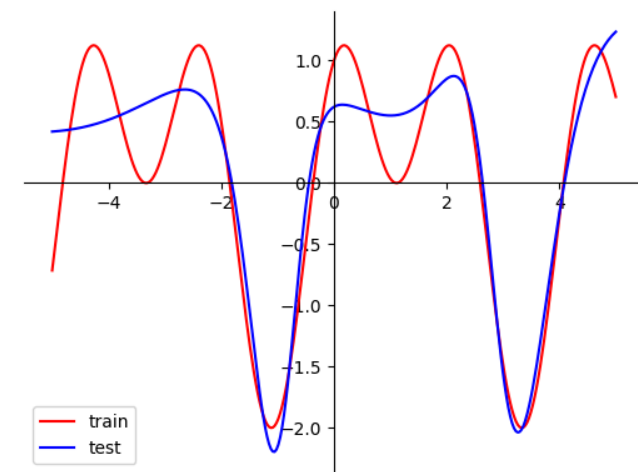
Dla każdego pomiaru parametry to:

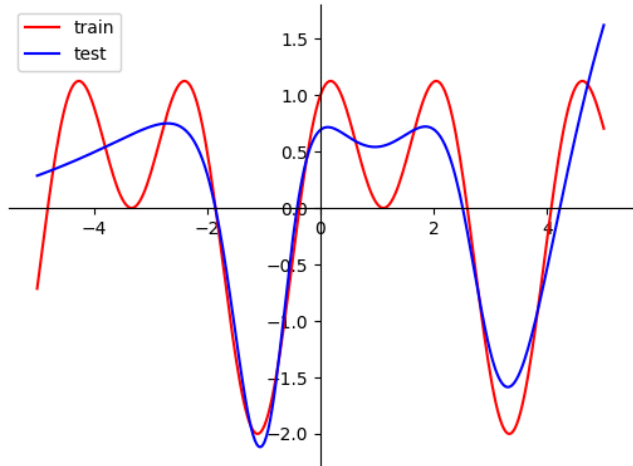
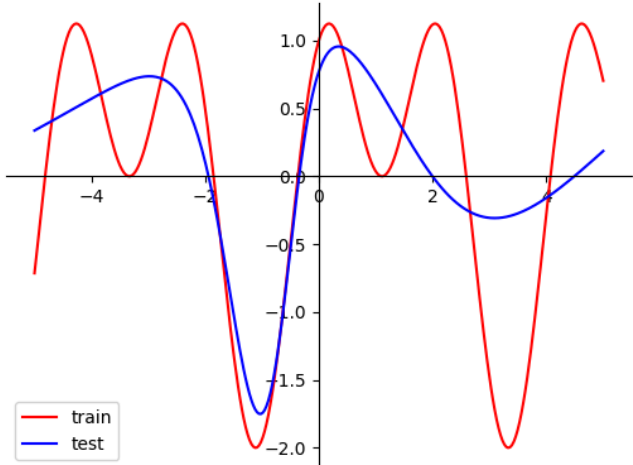
liczba neuronów: 10

liczba iteracji: 5000

współczynnik uczenia: 0,1



5	0,093	0,865	
10	0,054	1,214	
50	0,105	1,131	

100	0,136	0,993	
200	0,445	1,702	

Wnioski:

Wybór rozmiaru mini zbioru przy trenowaniu sieci neuronowej powinien zależeć od wartości innych parametrów. Na przykład mniejszy mini zbiór będzie wymagał większej liczby iteracji, aby odpowiednio aproksymować funkcję. Za to duży mini zbiór może prowadzić do przeuczenia i tym samym przeciętnej aproksymacji funkcji, co widać po uzyskanych rezultatach.