

PODSTAWY SIECI NEURONOWYCH

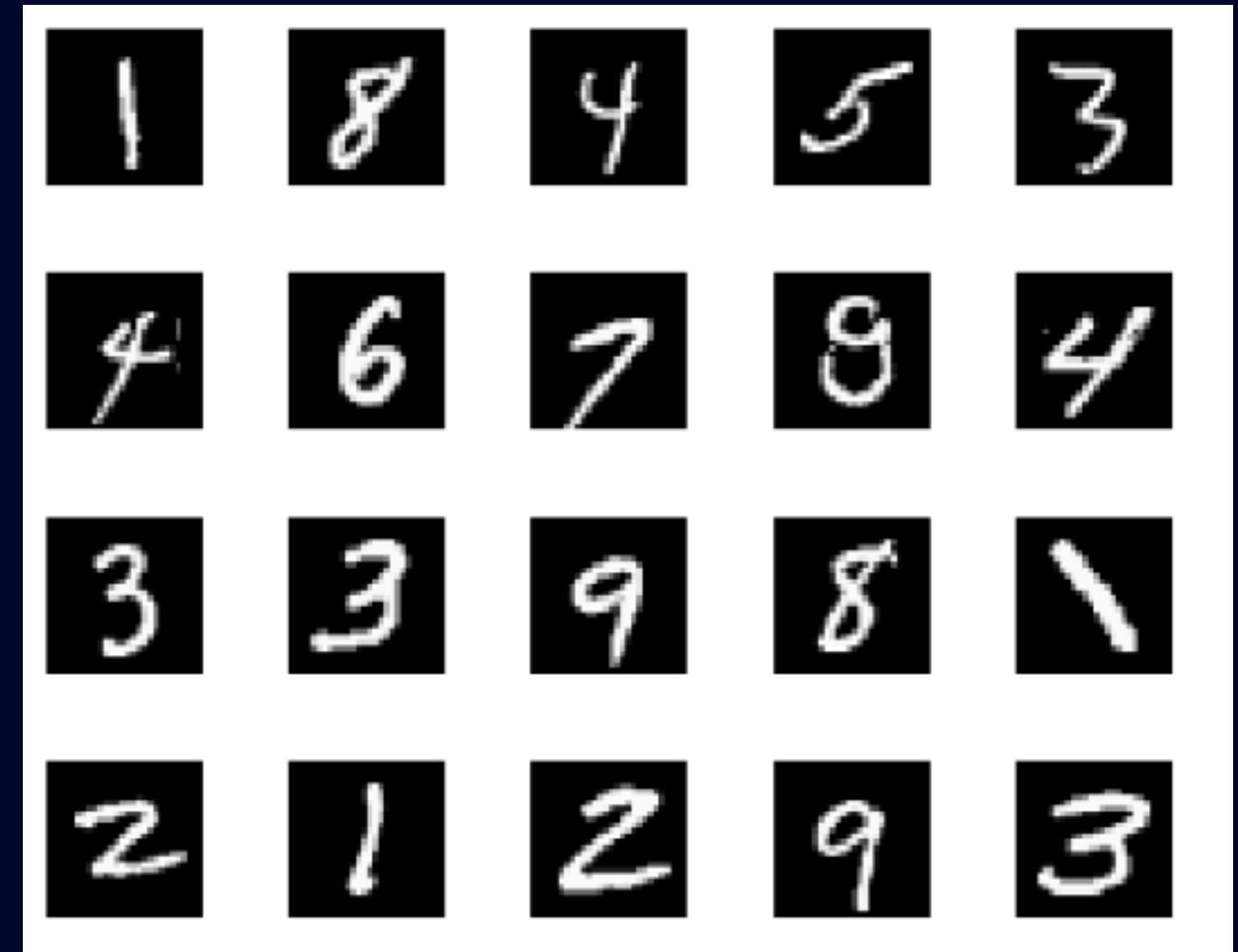
PROJEKT -
WARIANT 1

OSKAR KRUPSKI 272511
IGOR FRYSIAK 272548

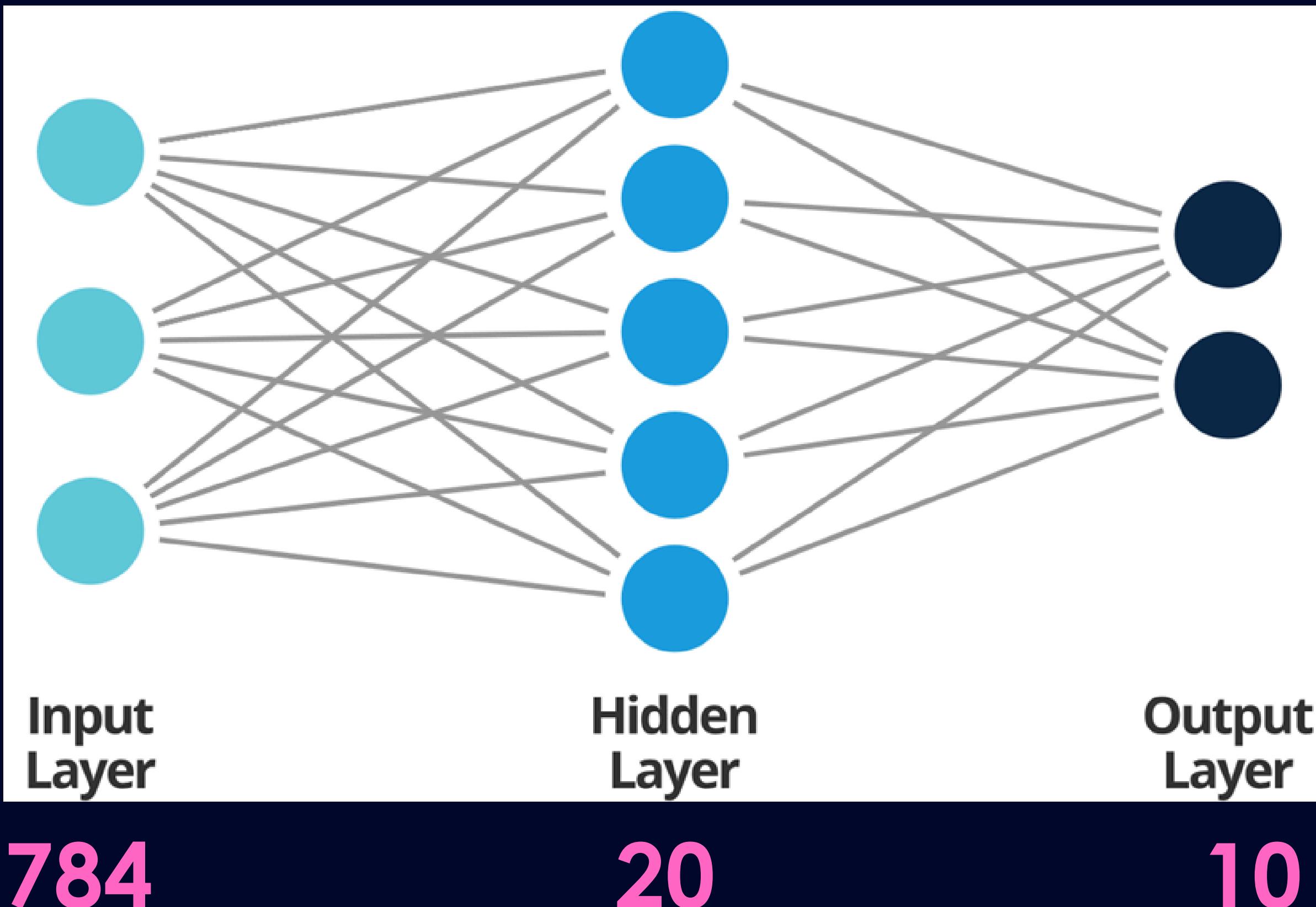
Zadanie 1 - Klasyfikacja obrazów

Zadanie polegało na zaprojektowaniu

i implementacji sztucznej sieci neuronowej do rozwiązania prostego problemu klasyfikacji obrazów dla zestawu danych z bazy MNIST.

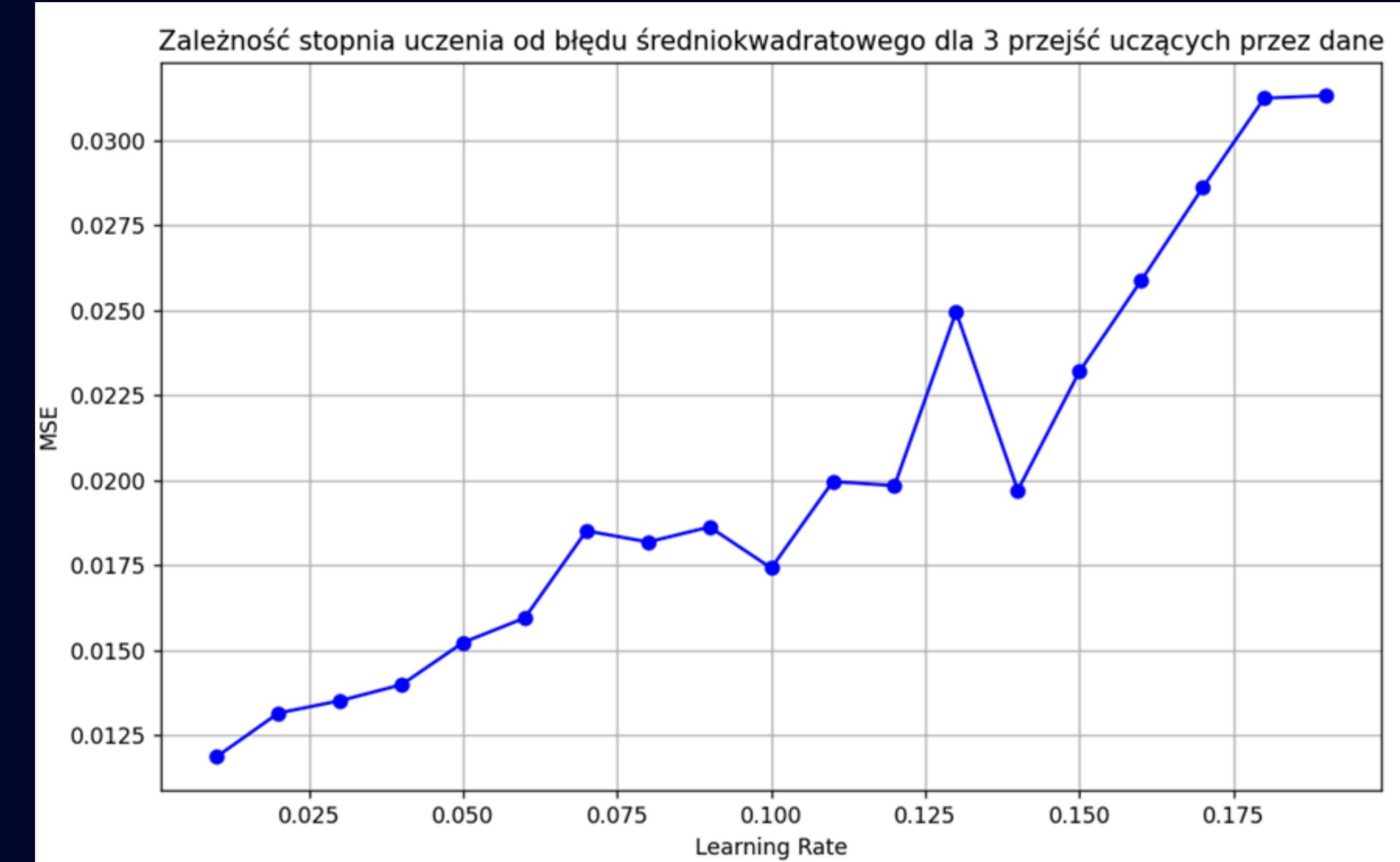


Model poglądowy sieci neuronowej



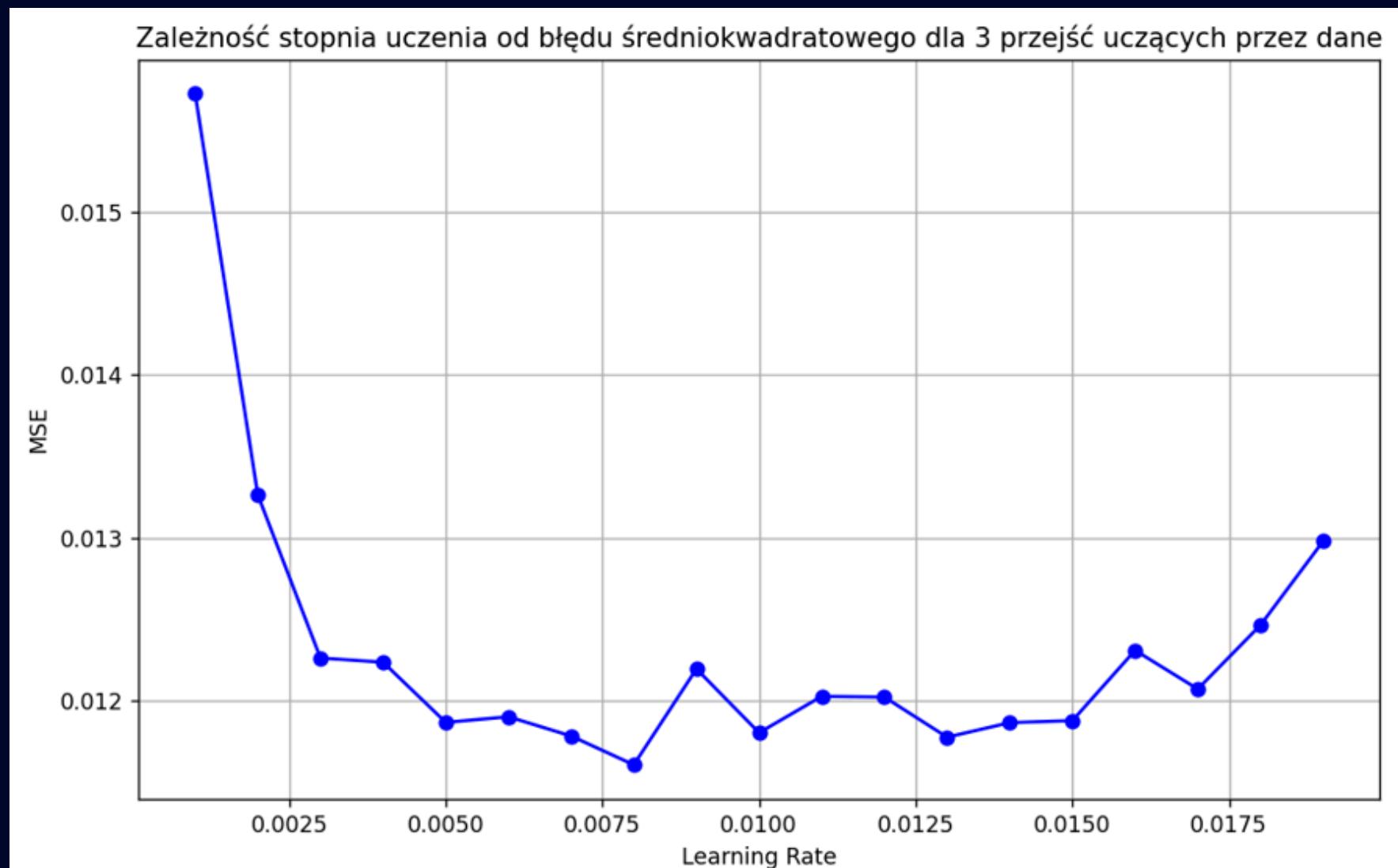
Wyniki

Skuteczność naszej sieci i poszczególnych funkcji aktywacji testowaliśmy dla różnych parametrów, np. różnych współczynników uczenia. Jakość tych metod oceniliśmy licząc błąg średniokwadratowy.

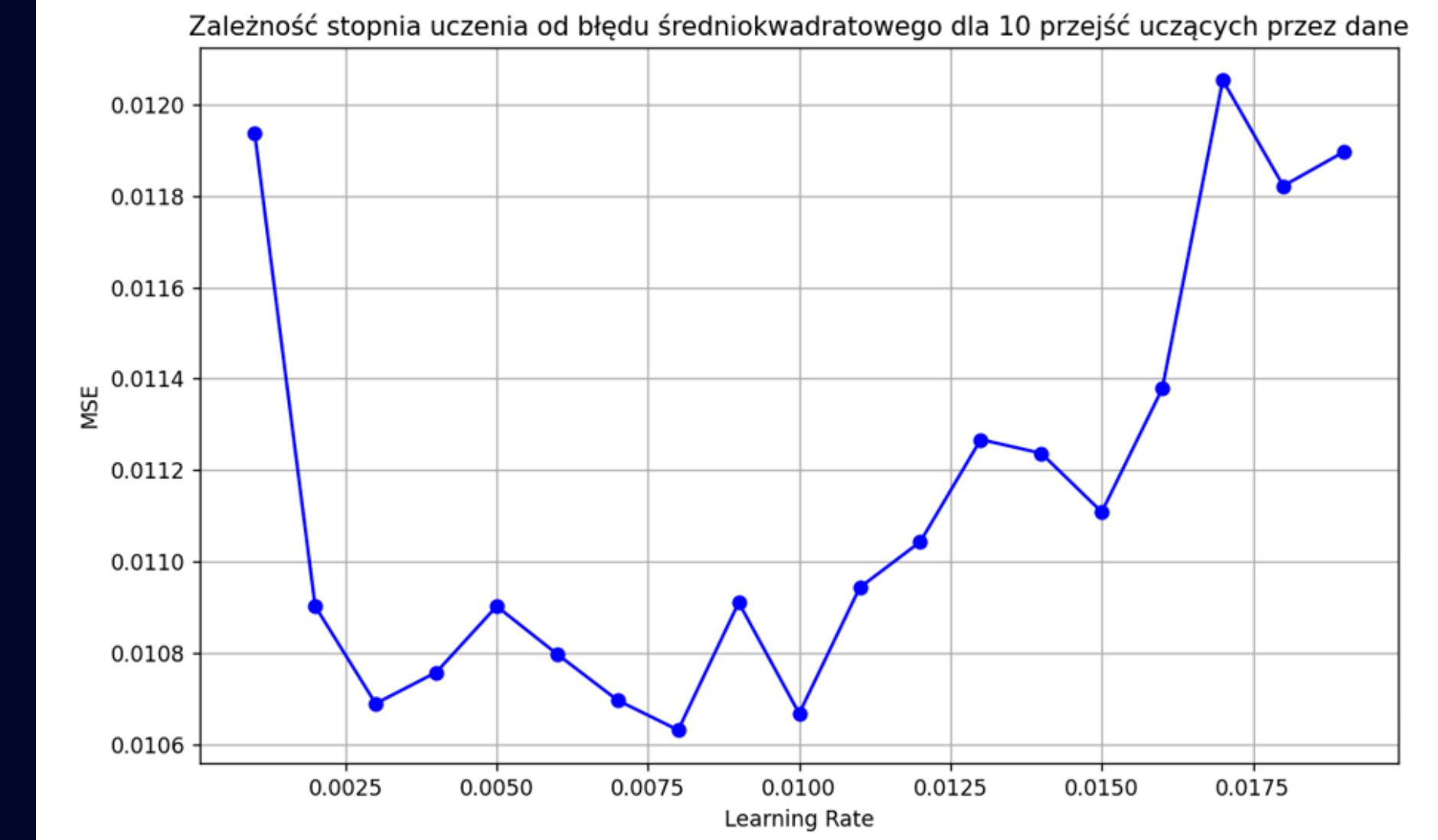


Funkcja aktywacji sigmoid,
współczynnik nauczania = 0.01,
3 epoki.

Wyniki

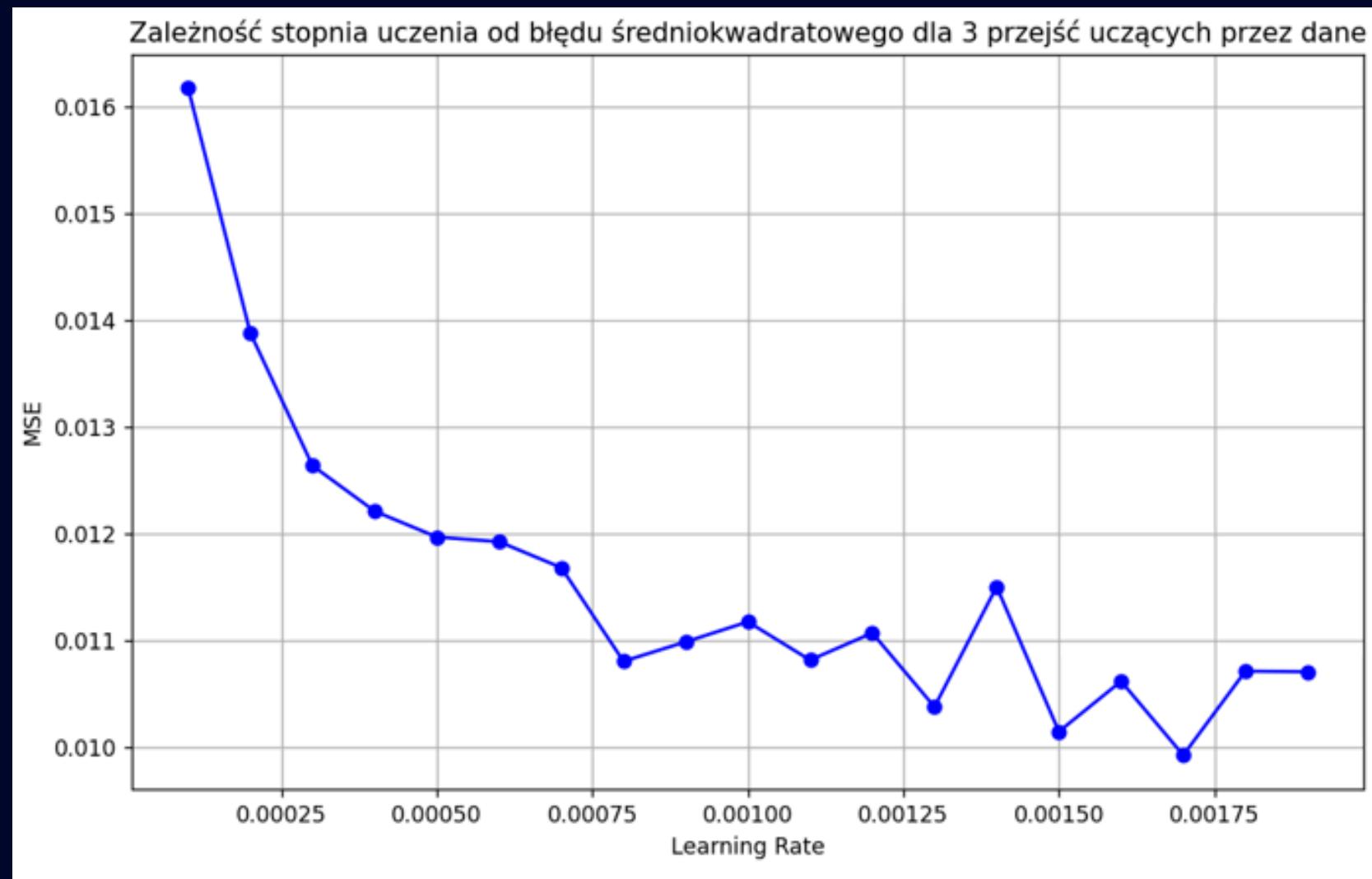


Funkcja aktywacji sigmoid,
współczynnik nauczania = 0.001,
3 epoki.

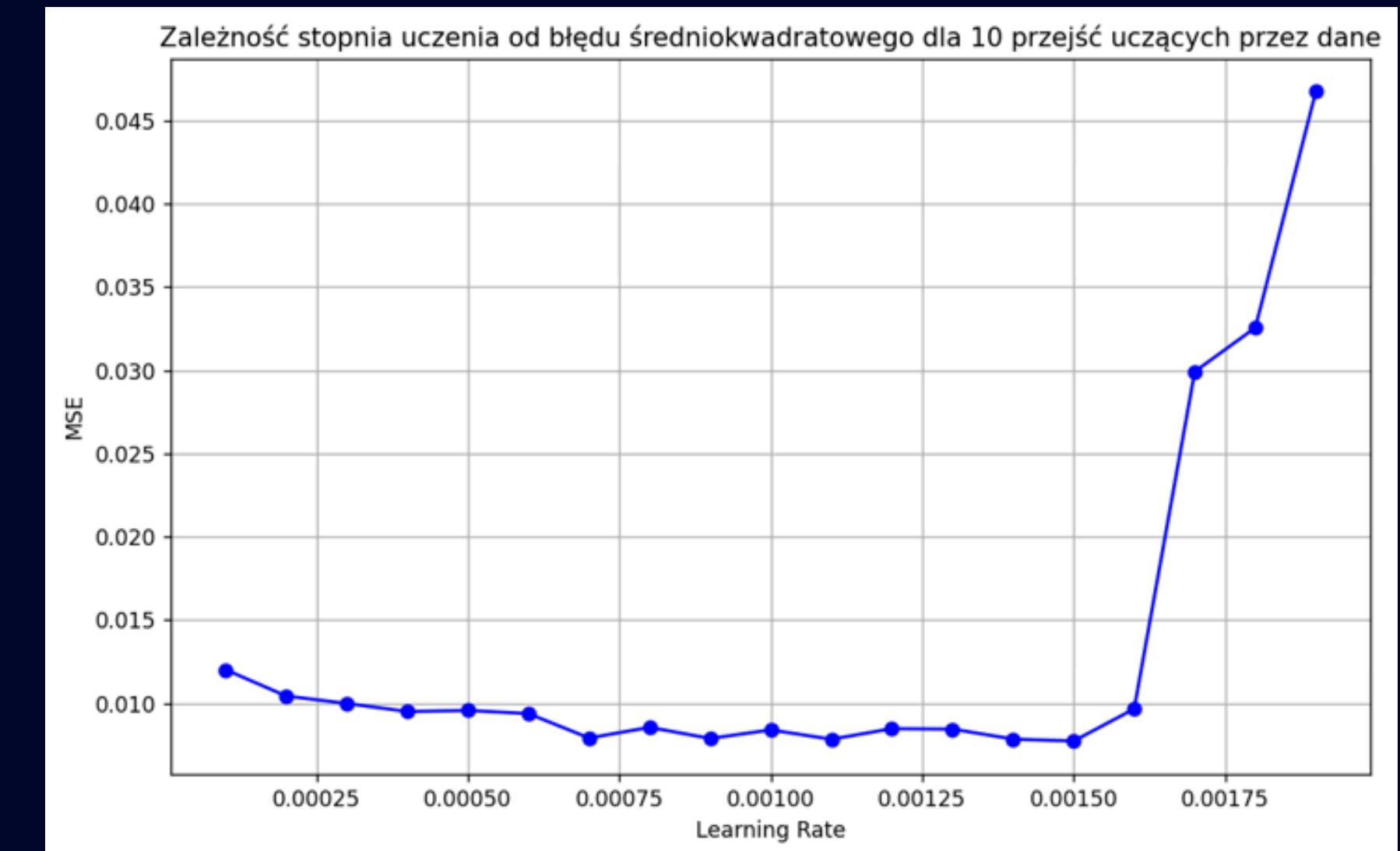


Funkcja aktywacji sigmoid,
współczynnik nauczania = 0.001,
10 epok.

Wyniki



Funkcja aktywacji ReLU,
współczynnik nauczania = 0.0001,
3 epoki.



Funkcja aktywacji ReLU,
współczynnik nauczania = 0.0001,
10 epok.

Wnioski

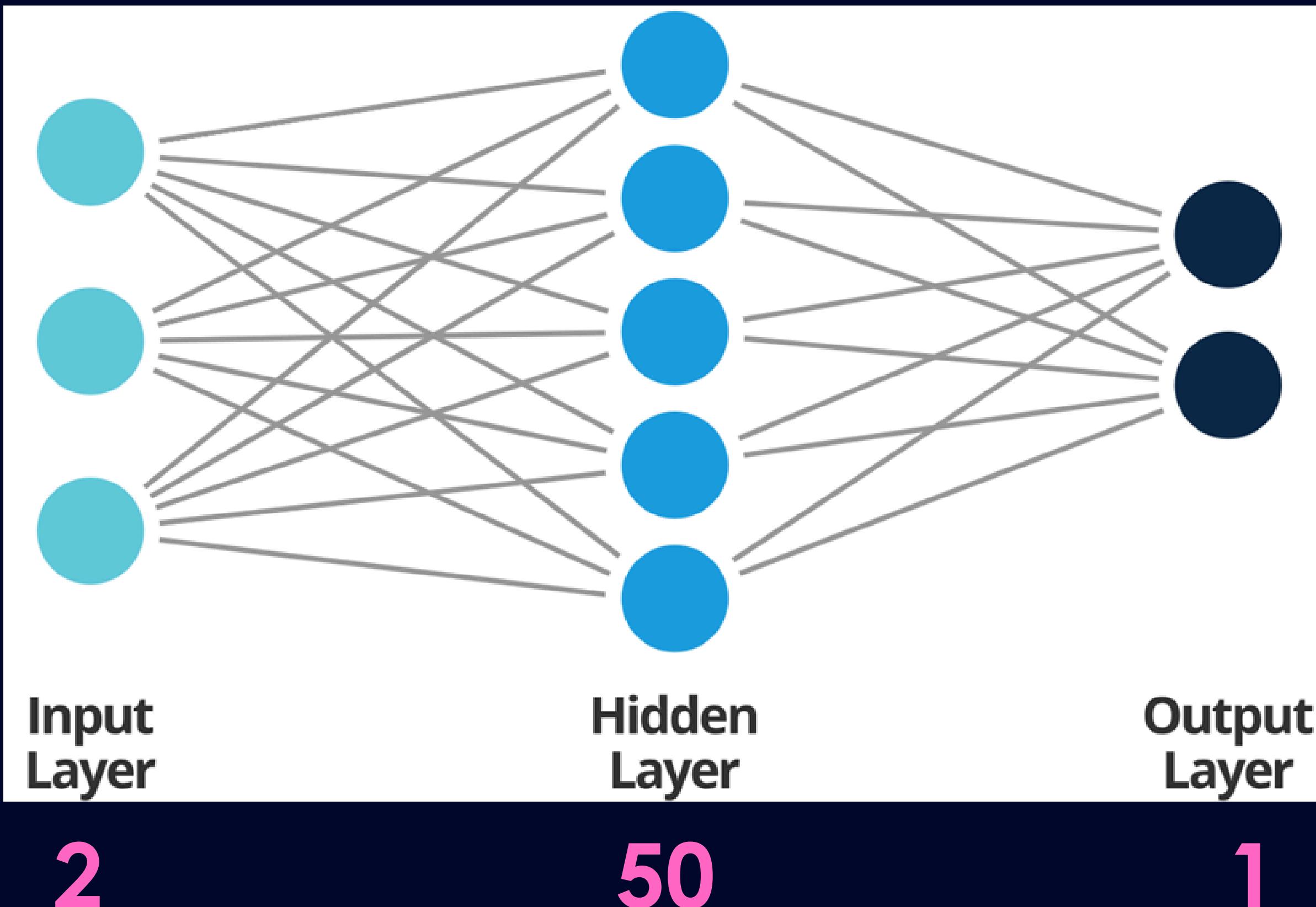
- Funkcja sigmoidalna może być przydatna przy wyższych współczynnikach uczenia (np. 0.01), jednak przy niższych wartościach (np. 0.0001) funkcje aktywacji ReLU i Leaky ReLU sprawują się znacznie lepiej.
- Najlepsze wyniki osiągnięto dla Leaky ReLU - 95.5% (współczynnik nauczania 0.0001, 10 epok).
- Im wyższa liczba epok, tym lepsze dokładności. Oznacza to, że sieć potrzebuje odpowiedniej ilości czasu, aby dobrze się nauczyć.
- Wybór funkcji aktywacji powinien być powiązany z odpowiednim współczynnikiem nauczania, gdyż różne funkcje działają lepiej lub gorzej dla różnych wartości - nie ma uniwersalnie najlepszej funkcji.

Zadanie 2 - Aproksymacja funkcji dwuwymiarowej

Zadanie drugie polegało na aproksymacji dwuwymiarowej funkcji Ackley'a na podstawie zadanego zestawu losowych punktów i wartości funkcji.

$$f_A(x_1, x_2) = -20 \exp [-0.2 \sqrt{0.5(x_1^2 + x_2^2)}] - \exp [0.5(\cos(2\pi x_1) + \cos(2\pi x_2))] + \exp [1] + 20$$

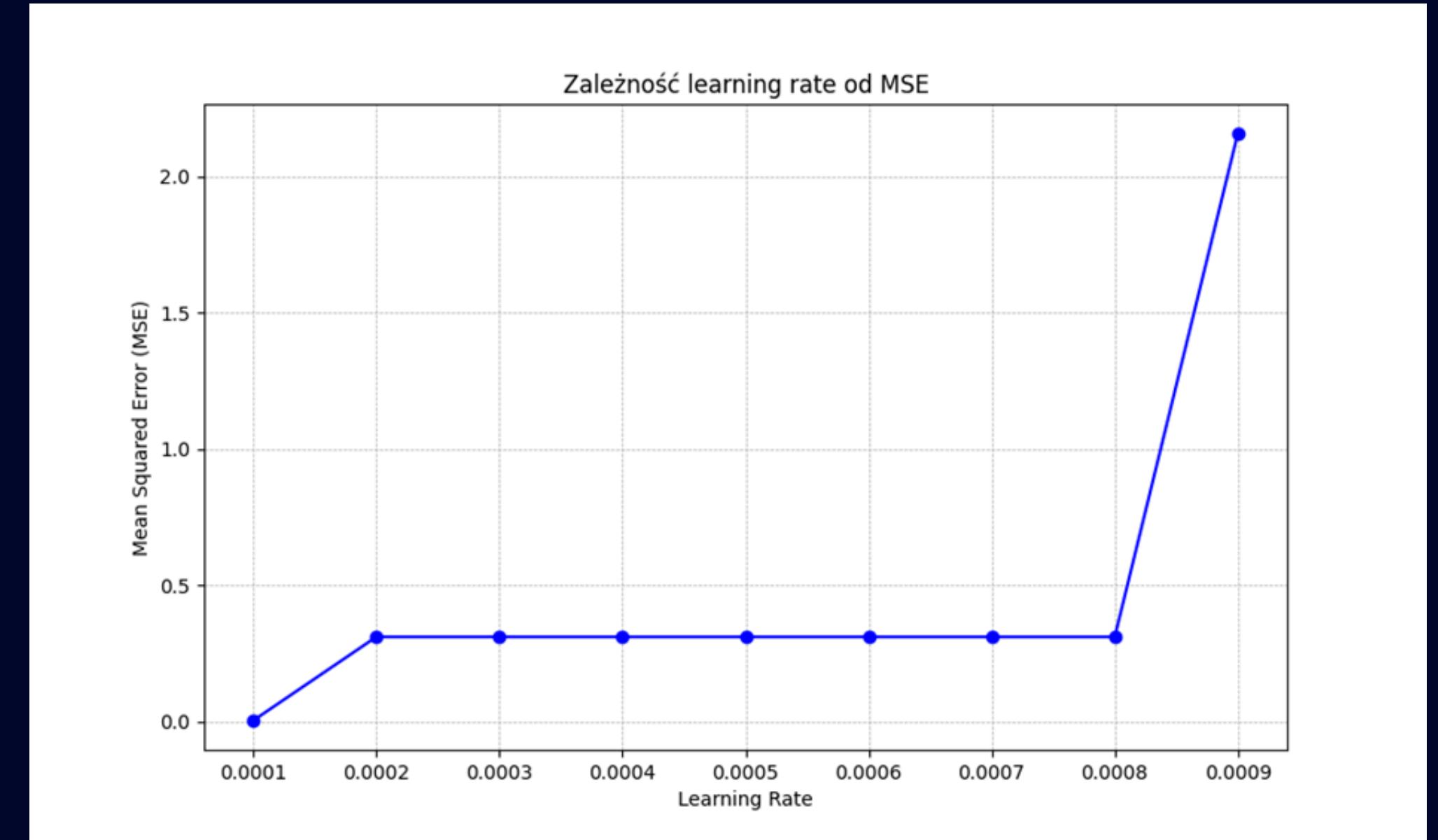
Model poglądowy sieci neuronowej



Wyniki

W tym zadaniu również zbadaliśmy zależność MSE od współczynnika uczenia dla różnych funkcji aktywacji, np. dla ReLU błęd średniokwadratowy rośnie wzraz ze wzrostem współczynnika uczenia.

Dodatkowo wprowadziliśmy 2 nowe funkcje aktywacji: Elu oraz Swish.

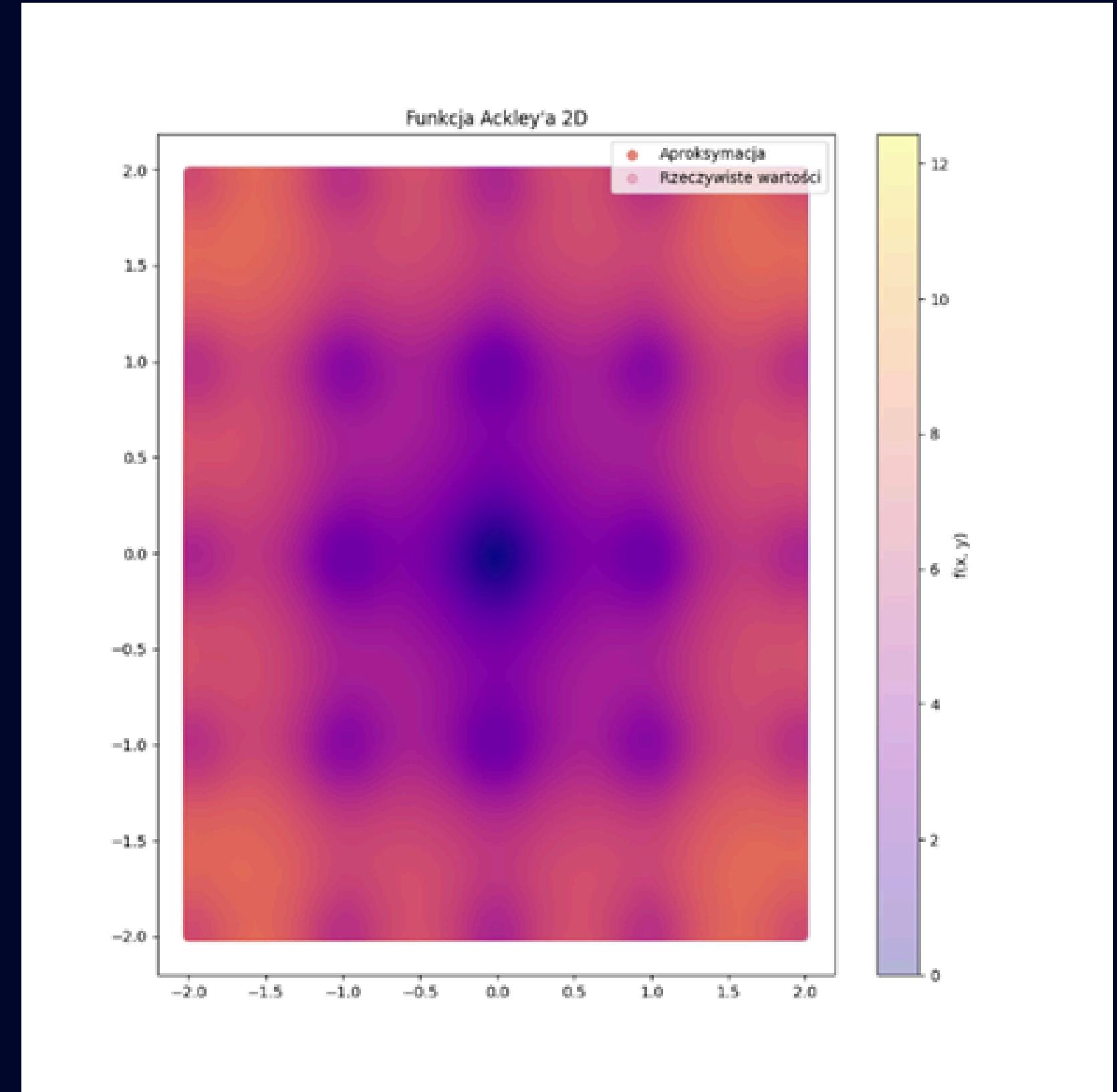


Wykres zależności stopnia uczenia od MSE

Wyniki

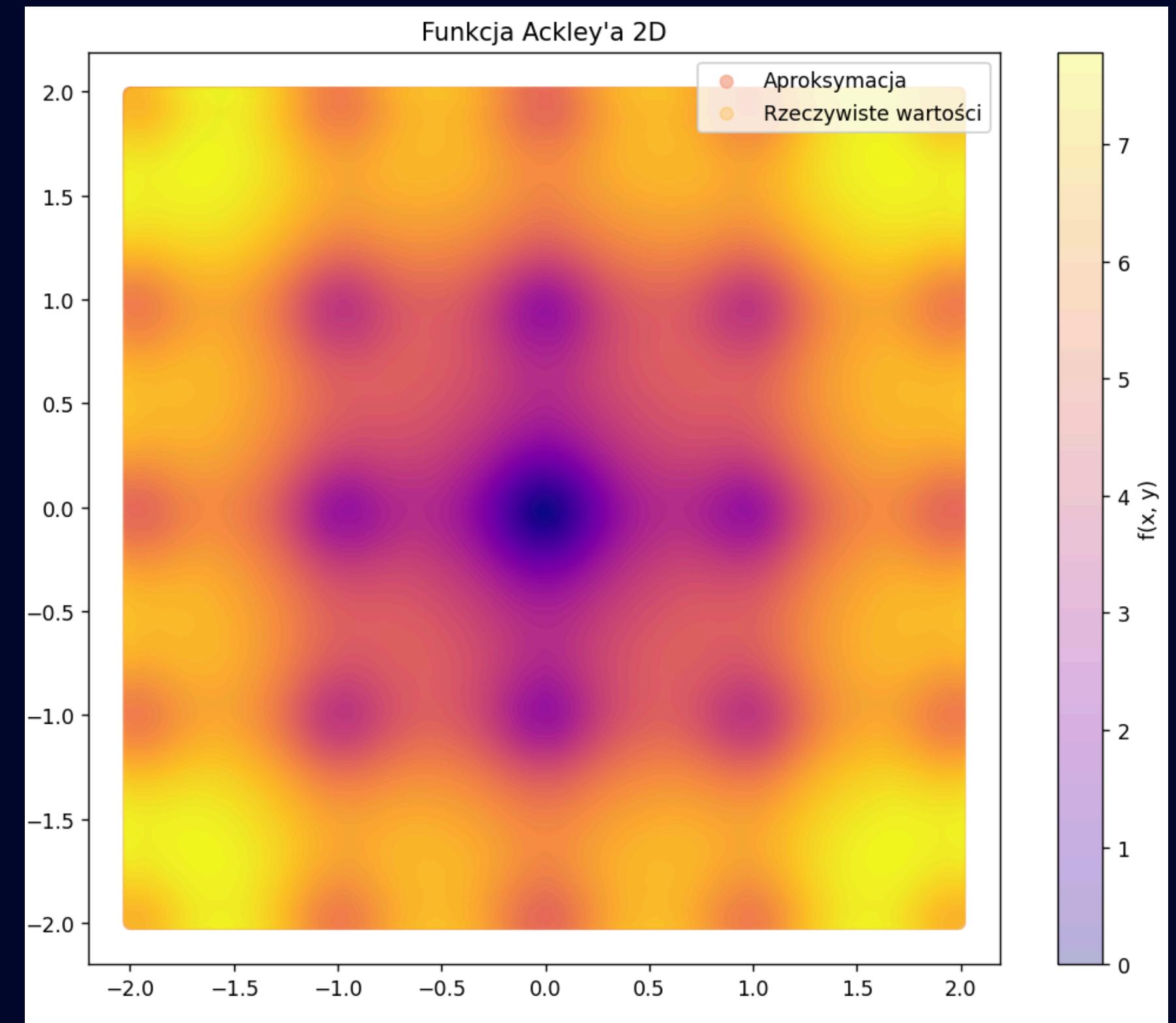
Przeprowadziliśmy również badania aproksymacji w zależności ilości epok.

Na przykład wykres 2D funkcji Ackley'a dla funkcji aktywacji Sigmoid przy 5 epokach
(learning rate = 0.0001):

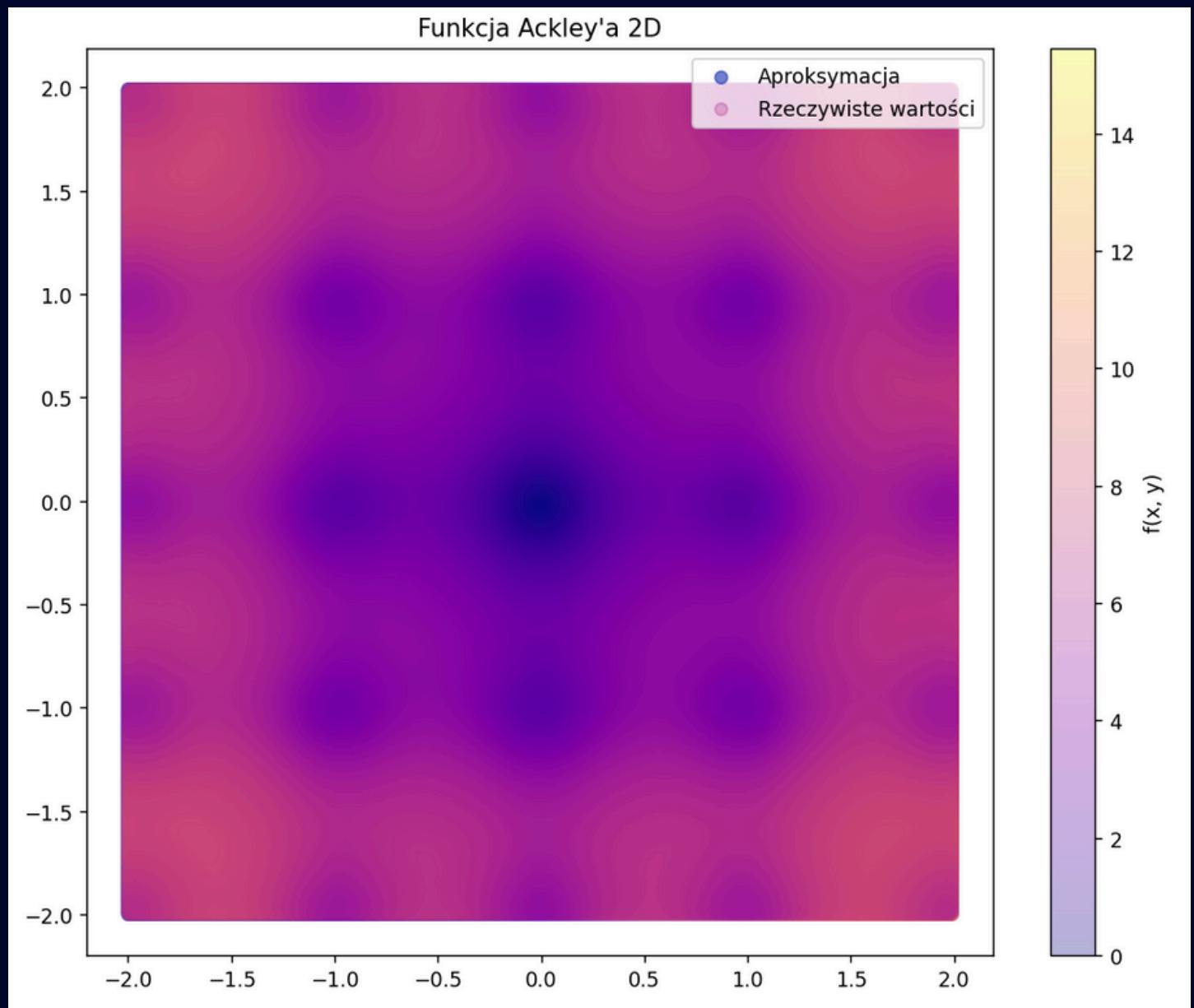


Wyniki

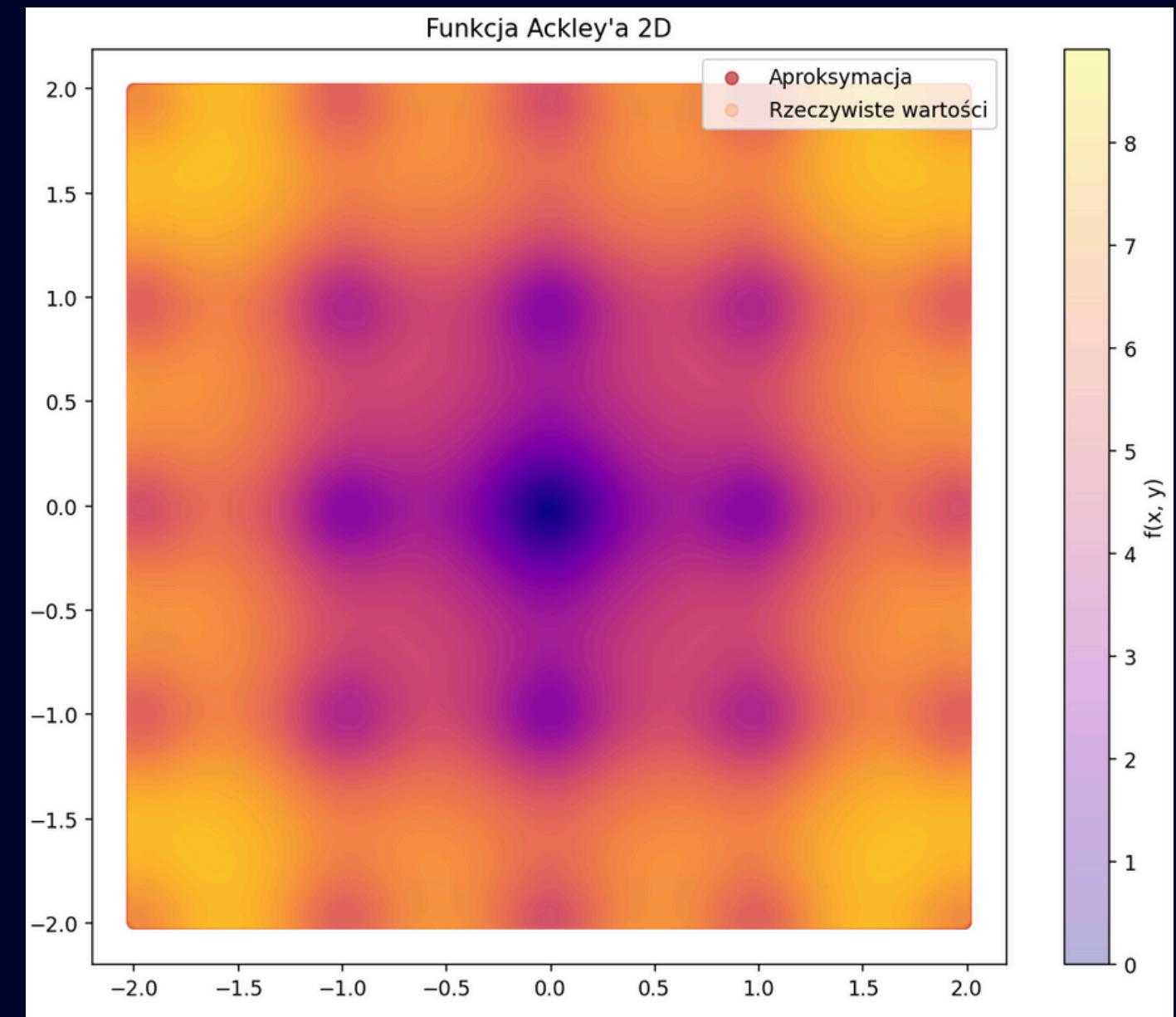
Z kolei wykres 2D funkcji Ackley'a dla tej samej funkcji aktywacji, jednak większej ilości epok (500):



Wyniki



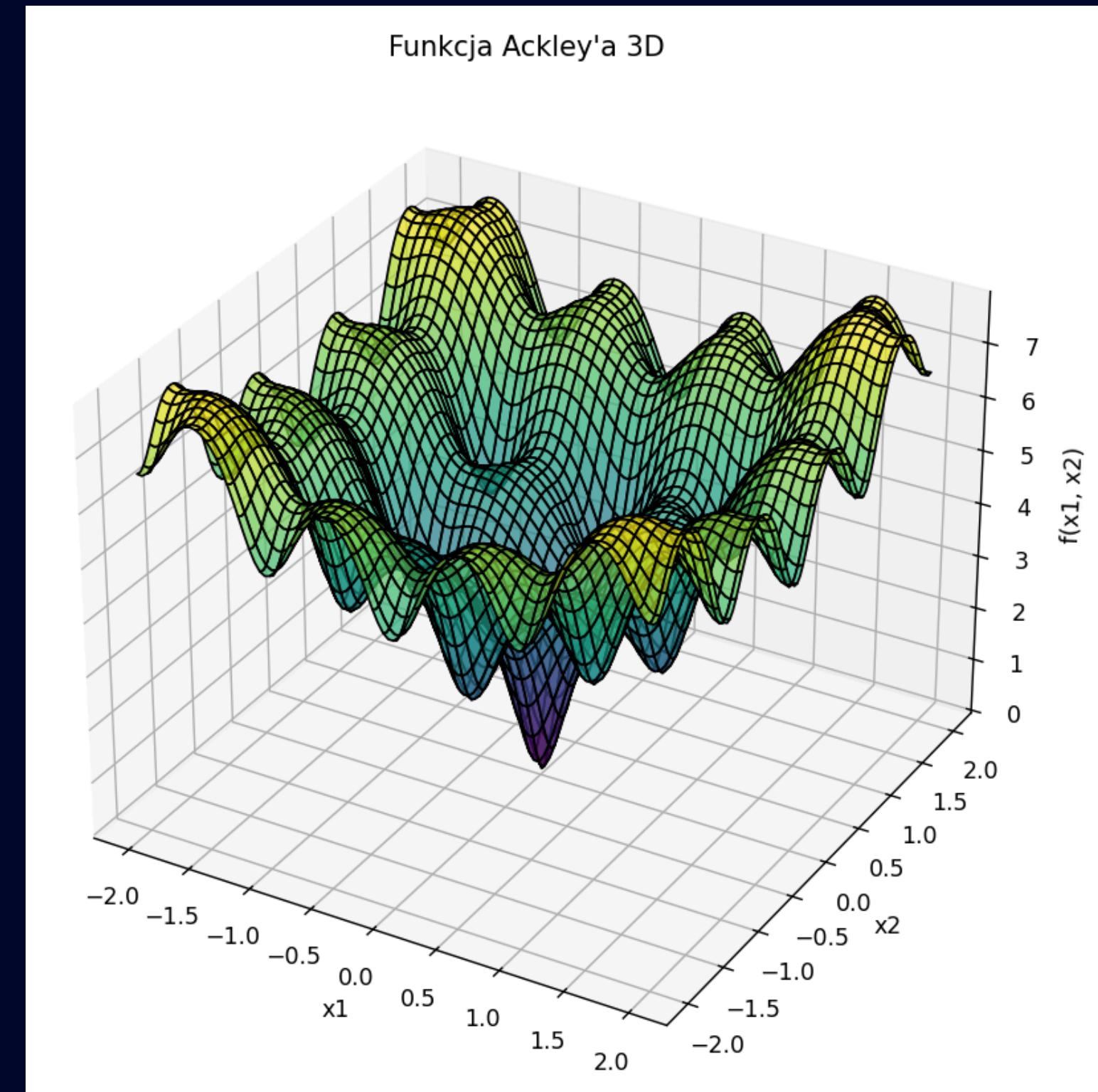
Funkcja aktywacji Elu,
współczynnik nauczania = 0.0001,
3 epoki,
20 neuronów w warstwie ukrytej.



Funkcja aktywacji Elu,
współczynnik nauczania = 0.0001,
3000 epok,
20 neuronów w warstwie ukrytej.

Wyniki

Wykres 3D funkcji Ackley'a:



Wnioski

- Badając wyniki dla mniejszych ilości epok, zaobserwowaliśmy, że sieć uczy się najszybciej dla funkcji aktywacji Sigmoid,
- a następnie dla Elu, Leaky ReLU, Swish, a najwolniej dla ReLU.
- Jeśli natomiast zwrócimy uwagę na dokładność sieci przy większej ilości epok (np. 3000), wartości MSE w kolejności od najmniejszej (najlepszej) do najgorszej są takie: ReLU > Leaky ReLU > Sigmoid > Swish > Elu.
- Dla niektórych funkcji aktywacji zbyt duża ilość neuronów w warstwie ukrytej (50+) prowadziła do niestabilności w predykcji, gdzie MSE zachowywało się nieprzewidywalnie, co ostatecznie oznaczało, że sieć nie jest nauczona.

**DZIĘKUJEMY ZA
UWAGĘ**

