

Równania różniczkowe - spectral bias

Laboratorium 10

Jakub Ciszewski, Wiktor Smaga

20 czerwca 2024

1 Opis zadania

Rozwiązano równanie różniczkowe zwyczajne

$$\frac{du(x)}{dx} = \cos(\omega x) \text{ dla } x \in \Omega \quad (1)$$

gdzie: $x, \omega, u \in \mathbb{R}$,

x to położenie,

Ω to dziedzina, na której rozwiązujemy równanie, $\Omega = \{x \mid -2\pi \leq x \leq 2\pi\}$,

u to funkcja, której postaci szukamy.

Z warunkiem początkowym:

$$u(0) = 0 \quad (2)$$

Analityczne rozwiązanie równania:

$$u(x) = \frac{1}{\omega} \sin(\omega x) \quad (3)$$

Koszt rezydualny:

$$\mathcal{L}_r(\theta) = \frac{1}{N} \sum_{i=1}^N \left\| \frac{d\hat{u}(x)}{dx} - \cos(\omega x_i) \right\|^2 \quad (4)$$

N - liczba kroków kolokacyjnych

Koszt związany z warunkiem początkowym:

$$\mathcal{L}_{IC}(\theta) = \|\hat{u}(0) - 0\|^2 \quad (5)$$

Funkcja kosztu:

$$\mathcal{L}(\theta) = \mathcal{L}_r(\theta) + \mathcal{L}_{IC}(\theta) \quad (6)$$

Przy użyciu sieci neuronowych PINN. Do wykonania zadania stworzono 8 modeli z poszczególnymi parametrami, które dalej będą nazywane jak niżej:

- Model **M1** - $\omega = 1, 2$ warstwy ukryte, 16 neuronów w każdej warstwie, 200 punktów treningowych, 1000 punktów testowych

- Model **M2** - $\omega = 1$, 4 warstwy ukryte, 64 neuronów w każdej warstwie, 200 punktów treningowych, 1000 punktów testowych
- Model **M3** - $\omega = 1$, 5 warstwy ukryte, 128 neuronów w każdej warstwie, 200 punktów treningowych, 1000 punktów testowych
- Model **M4** - $\omega = 15$, 2 warstwy ukryte, 16 neuronów w każdej warstwie, 3000 punktów treningowych, 5000 punktów testowych
- Model **M5** - $\omega = 15$, 4 warstwy ukryte, 64 neuronów w każdej warstwie, 3000 punktów treningowych, 5000 punktów testowych
- Model **M6** - $\omega = 15$, 5 warstwy ukryte, 128 neuronów w każdej warstwie, 3000 punktów treningowych, 5000 punktów testowych
- Model **M7** - $\omega = 15$, 2 warstwy ukryte, 16 neuronów w każdej warstwie, 3000 punktów treningowych, 5000 punktów testowych. W tym modelu zmieniono sposób obliczania kosztu rezydualnego. Przyjęto, że:

$$\hat{u}(x; \theta) = \tanh(\omega x) \cdot NN(x; \theta) \quad (7)$$

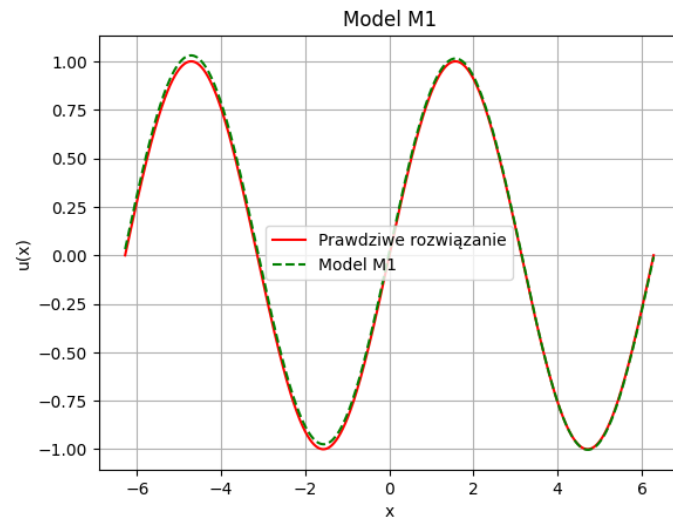
- Model **M8** - $\omega = 15$, 2 warstwy ukryte, 16 neuronów w każdej warstwie, 3000 punktów treningowych, 5000 punktów testowych. W tym modelu pierwsza warstwa ukryta jest inicjalizowana cechami Fouriera:

$$\gamma(x) = [\sin(2^0 \pi x), \cos(2^0 \pi x), \dots, \sin(2^{L-1} \pi x), \cos(2^{L-1} \pi x)] \quad (8)$$

$L = \lfloor \frac{K}{2} \rfloor$, K - ilość neuronów w warstwie

2 Wyniki

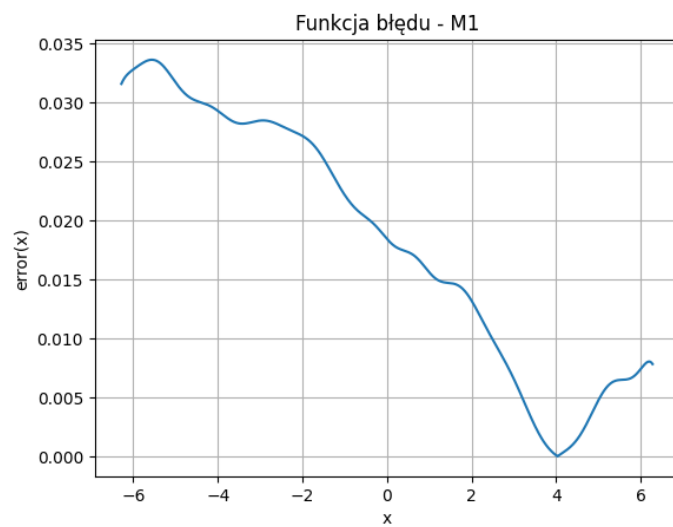
2.1 Model M1



Wykres 1: Rozwiązanie równania przez model M1

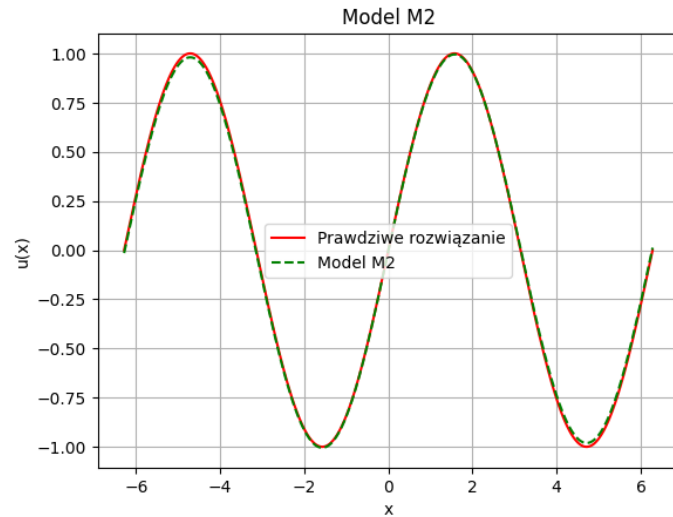


Wykres 2: Funkcja kosztu w zależności od epoki - M1

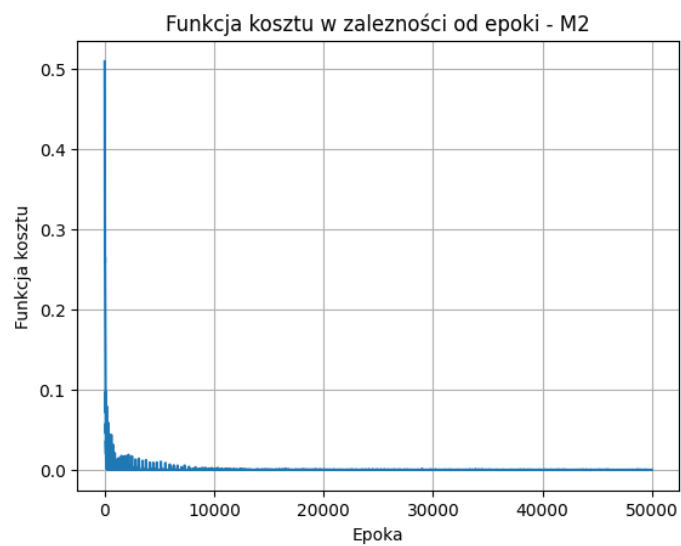


Wykres 3: Funkcja błędu modelu M1

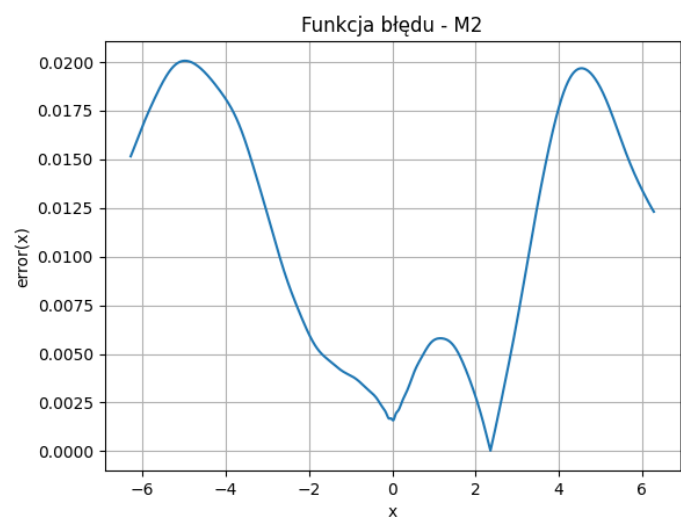
2.2 Model M2



Wykres 4: Rozwiązanie równania przez model M2

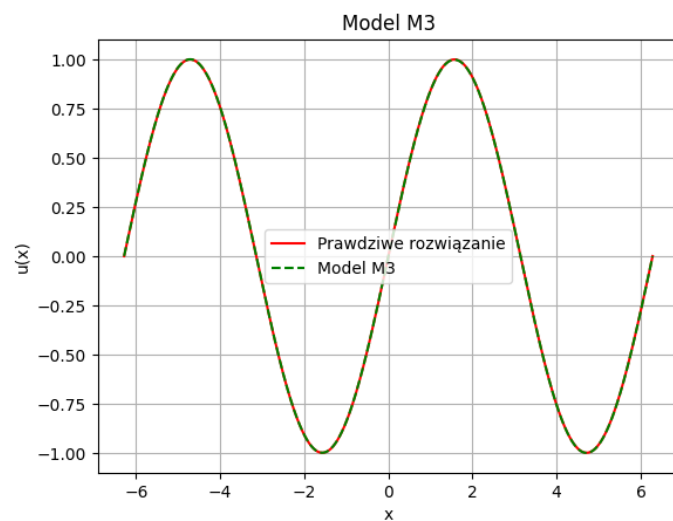


Wykres 5: Funkcja kosztu w zależności od epoki - M2

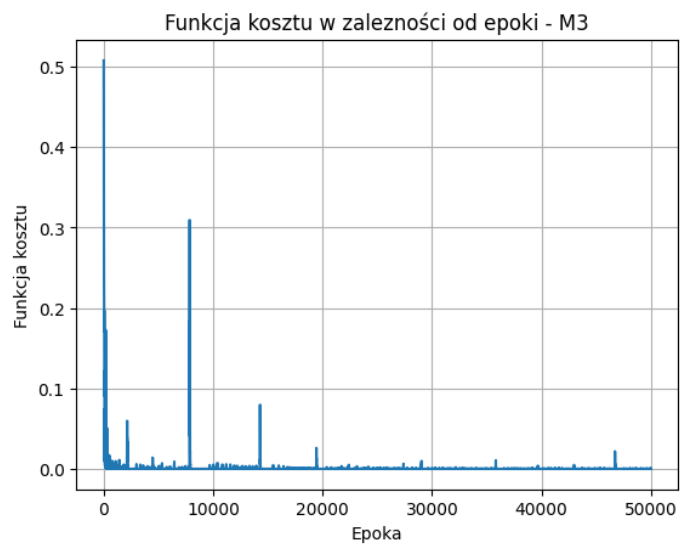


Wykres 6: Funkcja błędu modelu M2

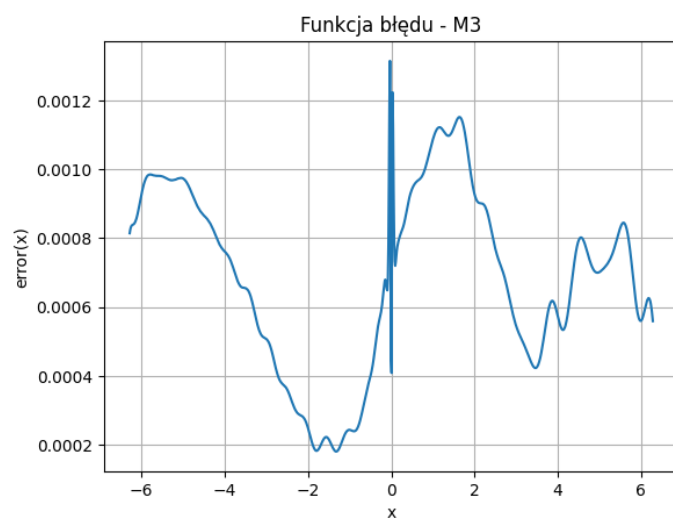
2.3 Model M3



Wykres 7: Rozwiązanie równania przez model M3

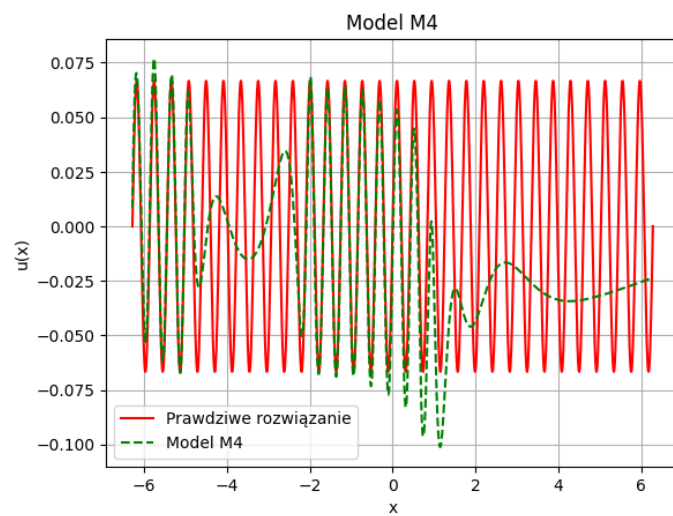


Wykres 8: Funkcja kosztu w zależności od epoki - M3

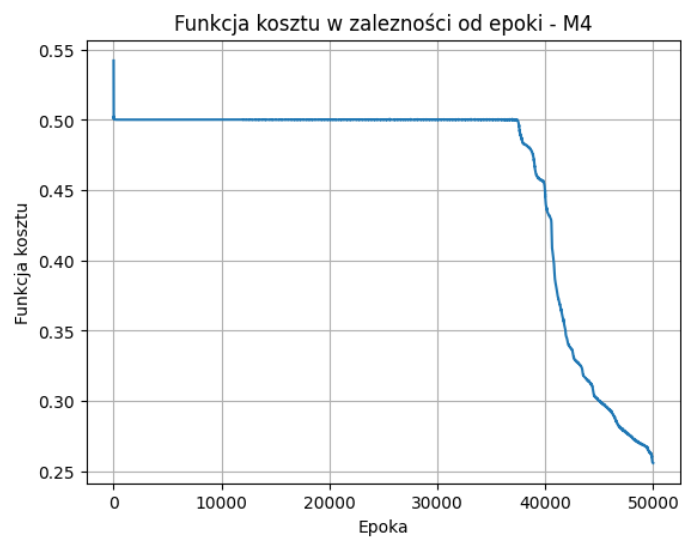


Wykres 9: Funkcja błędu modelu M3

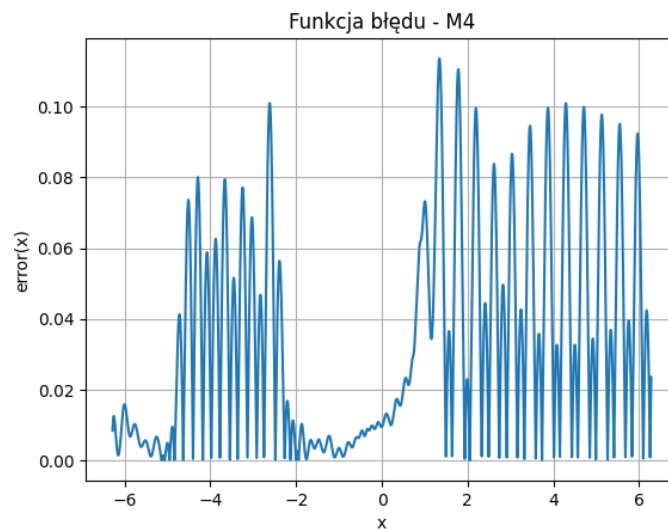
2.4 Model M4



Wykres 10: Rozwiązanie równania przez model M4

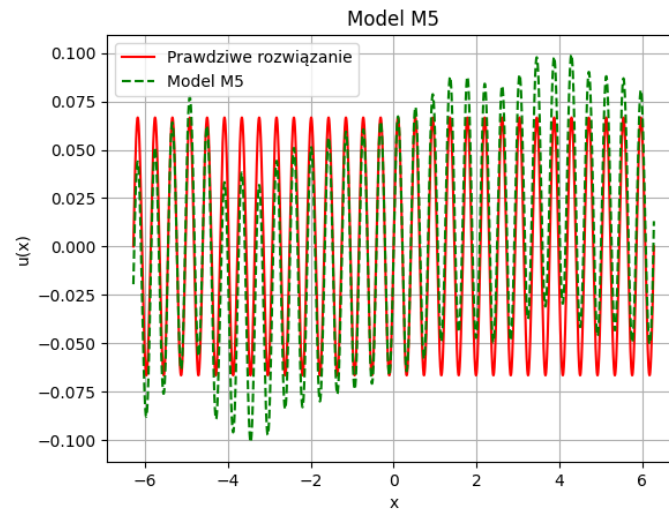


Wykres 11: Funkcja kosztu w zależności od epoki - M4

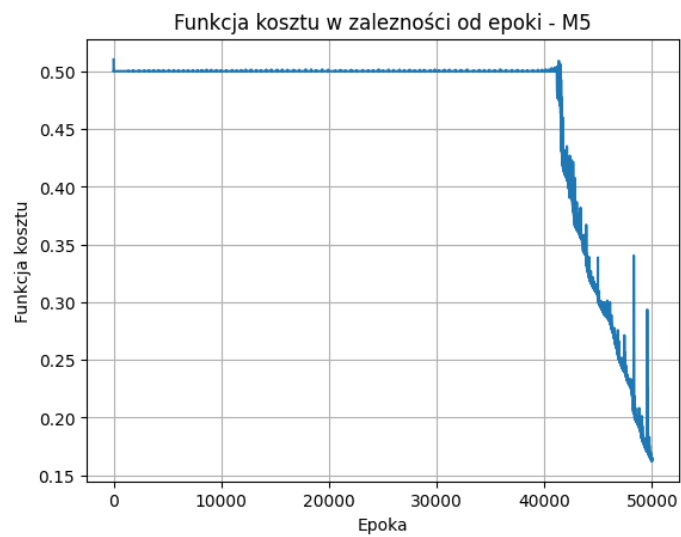


Wykres 12: Funkcja błędu modelu M4

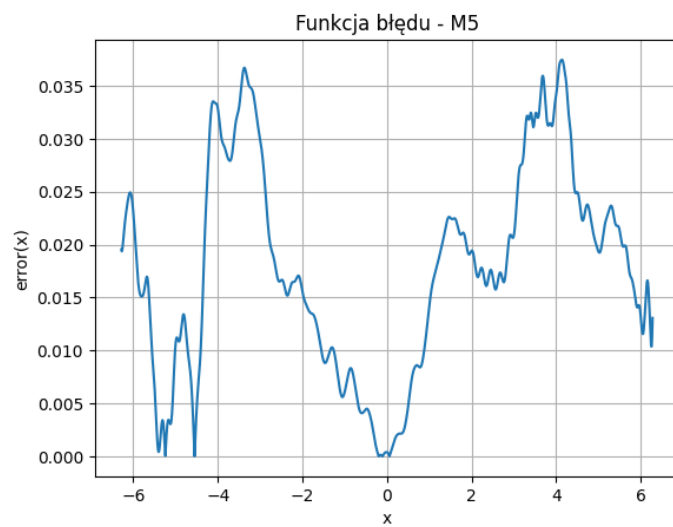
2.5 Model M5



Wykres 13: Rozwiązanie równania przez model M5

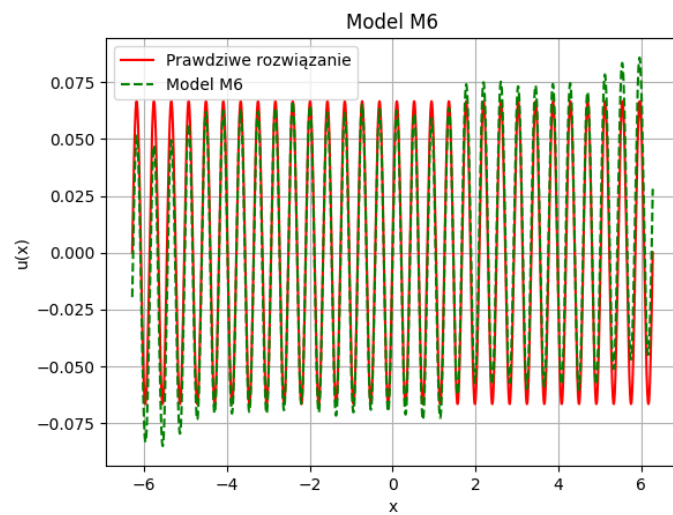


Wykres 14: Funkcja kosztu w zależności od epoki - M5



Wykres 15: Funkcja błędu modelu M5

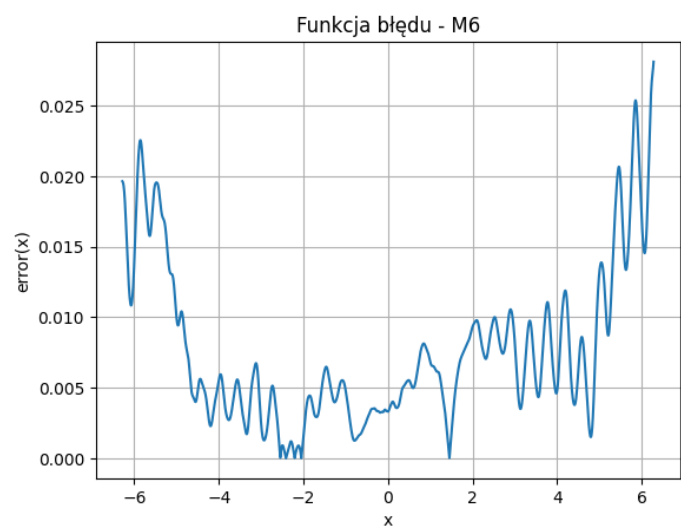
2.6 Model M6



Wykres 16: Rozwiązanie równania przez model M6

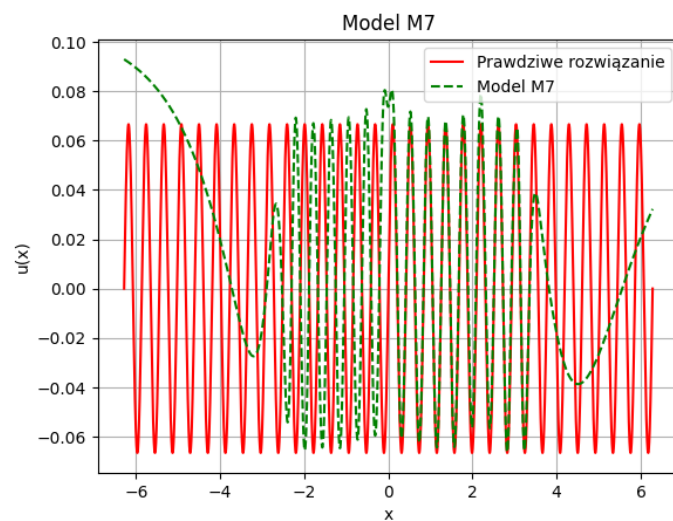


Wykres 17: Funkcja kosztu w zależności od epoki - M6

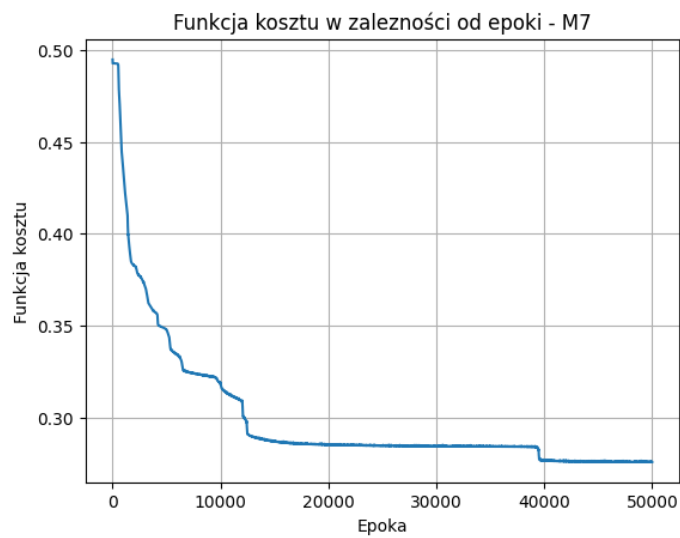


Wykres 18: Funkcja błędu modelu M6

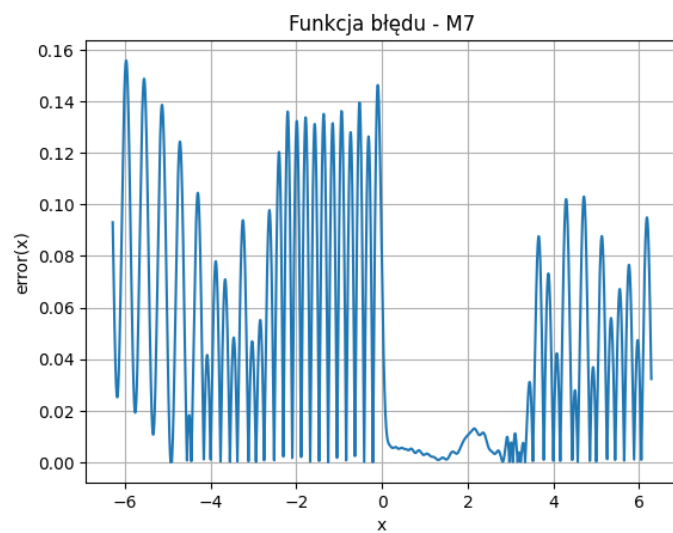
2.7 Model M7



Wykres 19: Rozwiązanie równania przez model M7

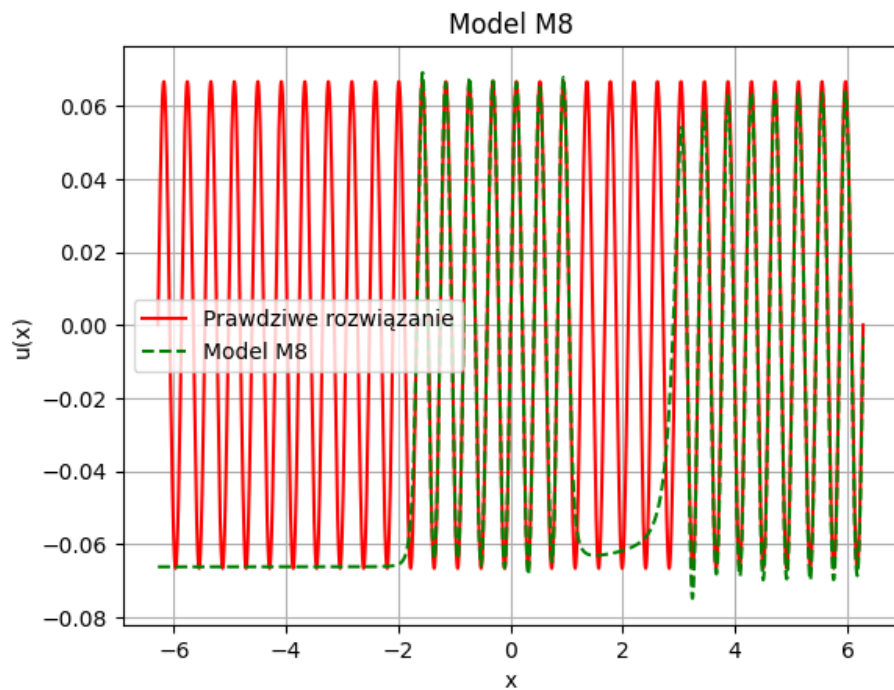


Wykres 20: Funkcja kosztu w zależności od epoki - M7

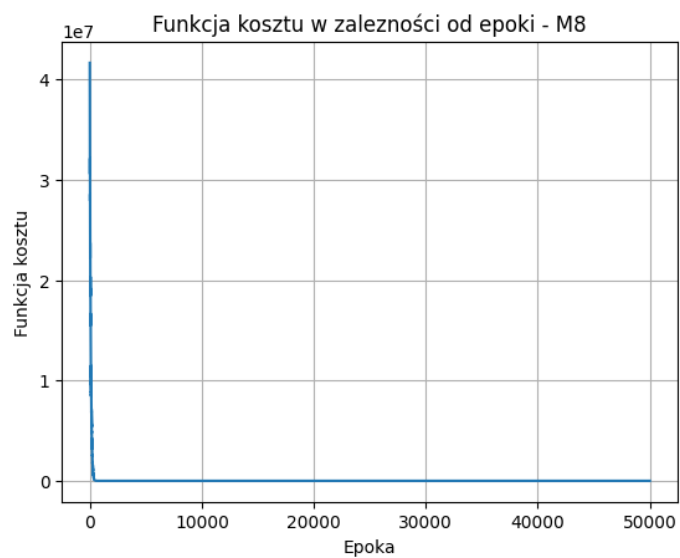


Wykres 21: Funkcja błędu modelu M7

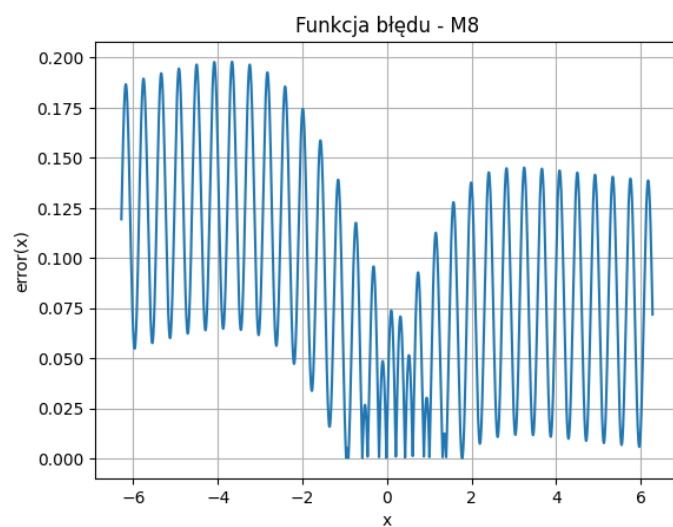
2.8 Model M8



Wykres 22: Rozwiązanie równania przez model M8



Wykres 23: Funkcja kosztu w zależności od epoki - M8



Wykres 24: Funkcja błędów modelu M8

3 Wnioski

- Proste modele takie jak model M1 dobrze radzą sobie z prostymi równaniami, w których częstotliwość zmian jest mała ($\omega = 1$). Natomiast nie radzą

sobie w trudniejszych równaniach, w których częstotliwość gwałtownych zmian jest wysoka ($\omega = 15$).

- Jednym ze sposobów zwiększenia dokładności modelu jest zwiększenie liczby warstw oraz neuronów - niestety niesie to za sobą znaczne wydłużenie czasu uczenia modelu.
- Dobranie odpowiedniej funkcji kosztu może znacznie polepszyć dokładność modelu. Można to zaobserwować porównując wyniki modelu M4 oraz M7.
- Warstwa ucząca z cechami Fouriera pozwala nam na zwiększenie dokładności modelu w przypadkach gdy rozwiązaniem równania jest funkcja, która bardzo gwałtownie i często się zmienia.