
Optymalizacja uczenia Sieci Neuronowe 2020

Jakub Ciszek
238035

Spis treści

1	Opis badań	3
1.1	Plan eksperymentów	3
1.2	Charakterystyka zbiorów danych	3
2	Eksperymenty	4
2.1	Wpływ optymalizatorów na przebieg procesu uczenia	4
2.2	Wpływ inicjalizacji wag na przebieg procesu uczenia	12
3	Wnioski	18

Cały kod wykorzystany w zadaniu znajduje się pod adresem: <https://github.com/Greenpp/sieci-neuronowe-pwr-2020>

1 Opis badań

1.1 Plan eksperymentów

Wszystkie eksperymenty zostały przeprowadzone 10 razy. Losowość przy inicjalizacji wag oraz generacji danych nie została narzucona żadnym ziarnem. Podczas badań przyjęto górną granicę 5 epok, po przekroczeniu której, uczenie zostawało przerywane. Ze względu na charakter zadania (klasyfikacja) na ostatniej warstwie użyto funkcji Softmax, a za funkcję straty przyjęto Entropię krzyżową. Warstwa ukryta składała się z 512 neuronów, a początkowy współczynnik uczenia wynosił 0.01. Z powodów wydajnościowych testowanie modelu przeprowadzano co każde 32 paczki, z których każda składała się z 32 przykładów.

Zgodnie z instrukcją zostały przeprowadzone następujące badania:

- Wpływ optymalizatorów na przebieg procesu uczenia
- Wpływ inicjalizacji wag na przebieg procesu uczenia

Podczas wizualizacji funkcji straty pominięto pierwsze 10 pomiarów dla lepszej czytelności.

1.2 Charakterystyka zbiorów danych

Danymi użytymi w zadaniu jest zbiór ręcznie pisanych cyfr 0 – 9 - MNIST. Na zbiór składa się 70,000 obrazów wielkości 28x28 pikseli, co po przekształceniu odpowiadało 784 elementom wektorowi wejściowemu i 10 klasom na wyjściu. Użyta w zadaniu wersja została podzielona na 3 zbiory:

- Uczący - 50,000 przykładów.
- Walidujący - 10,000 przykładów.
- Testowy - 10,000 przykładów.

W trakcie eksperymentów wykorzystano jedynie zbiory uczący i testowy.

2 Eksperymenty

2.1 Wpływ optymalizatorów na przebieg procesu uczenia

Założenia

Tabela 1: Stałe dla eksperymentu 1

Parametr	Wartość
Inicjalizacja wag	-0.1 – 0.1

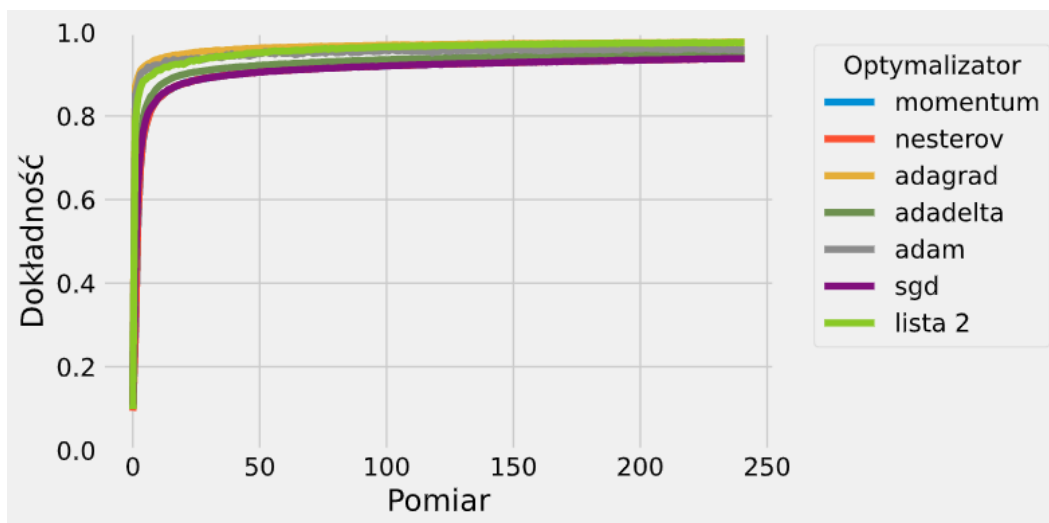
Zmienną w tym eksperymencie był użyty optymalizator uczenia. Użyto metod ze zbioru {SGD, Momentum, Nesterov, AdaGrad, AdaDelta, Adam}

Przebieg

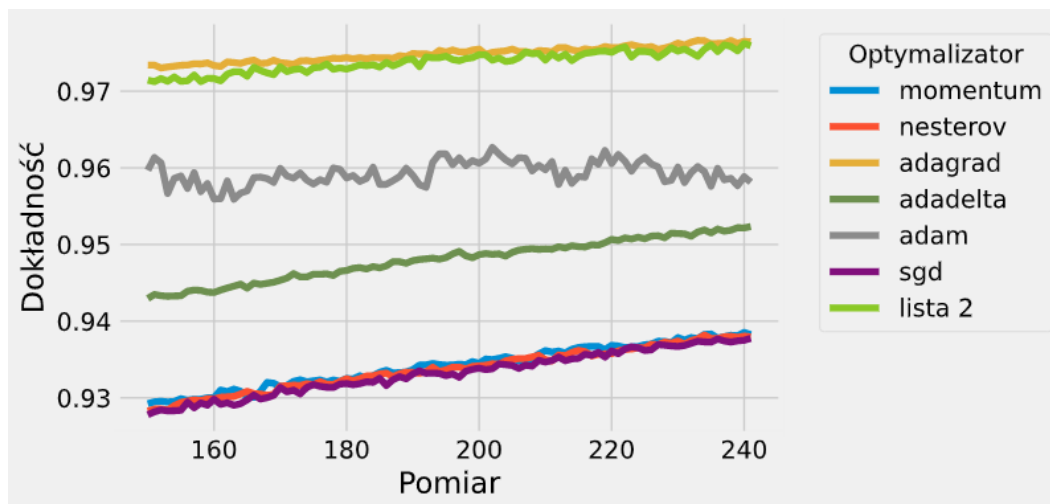
Podczas eksperymentu model został zainicjalizowany 10 razy dla każdej z badanych wartości oraz wyuczony, uzyskane wyniki zostały zapisane w postaci pliku .plk do dalszej analizy. Badania wykonano dla funkcji aktywacji Sigmoid oraz ReLU.

Wyniki

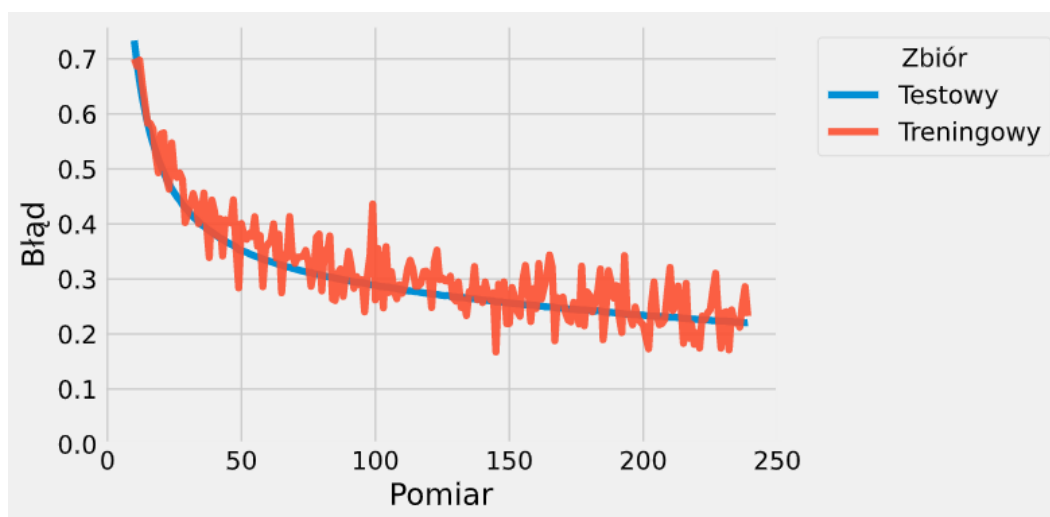
Wykres 1: Dokładność modelu w zależności od użytego optymalizatora dla funkcji ReLU



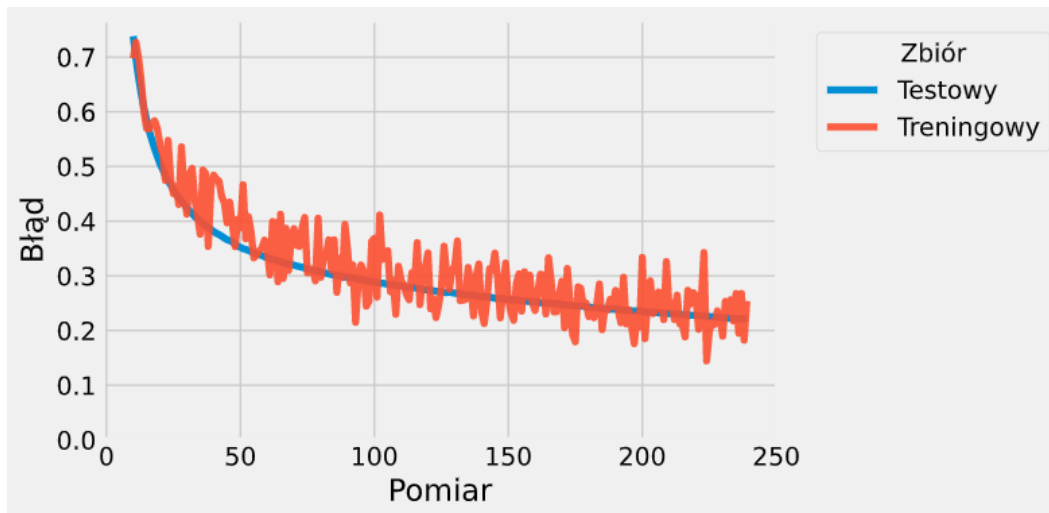
Wykres 2: Dokładność modelu w końcowym etapie uczenia w zależności od użytego optymalizatora dla funkcji ReLU



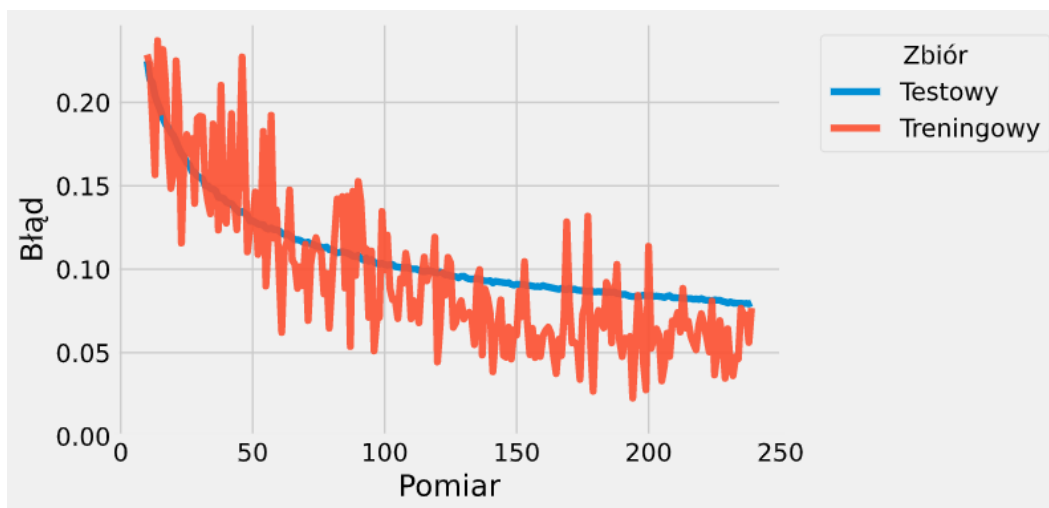
Wykres 3: Zachowanie funkcji błęd dla funkcji ReLU i optymalizatora Momentum



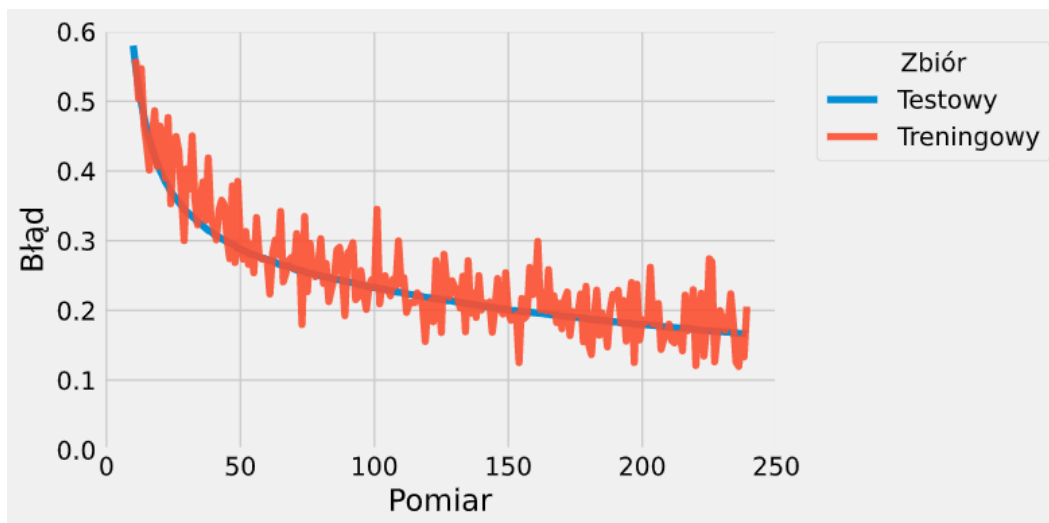
Wykres 4: Zachowanie funkcji błędu dla funkcji ReLU i optymalizatora Nesterov



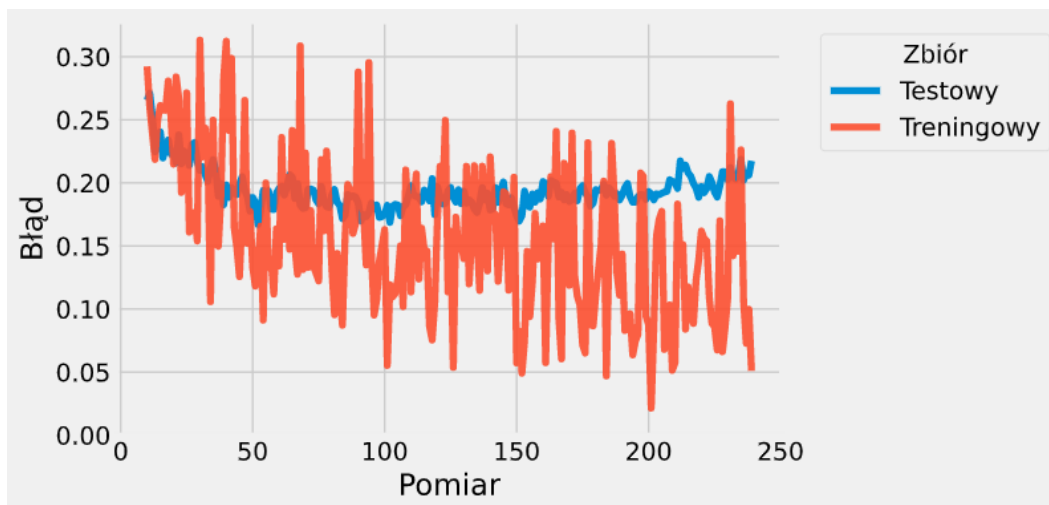
Wykres 5: Zachowanie funkcji błędów dla funkcji ReLU i optymalizatora AdaGrad



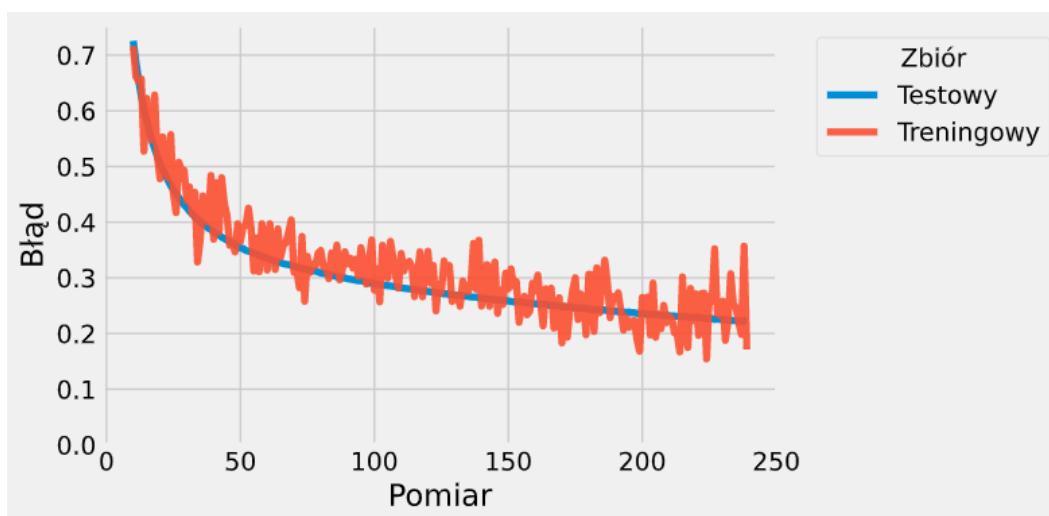
Wykres 6: Zachowanie funkcji błędów dla funkcji ReLU i optymalizatora AdaDelta



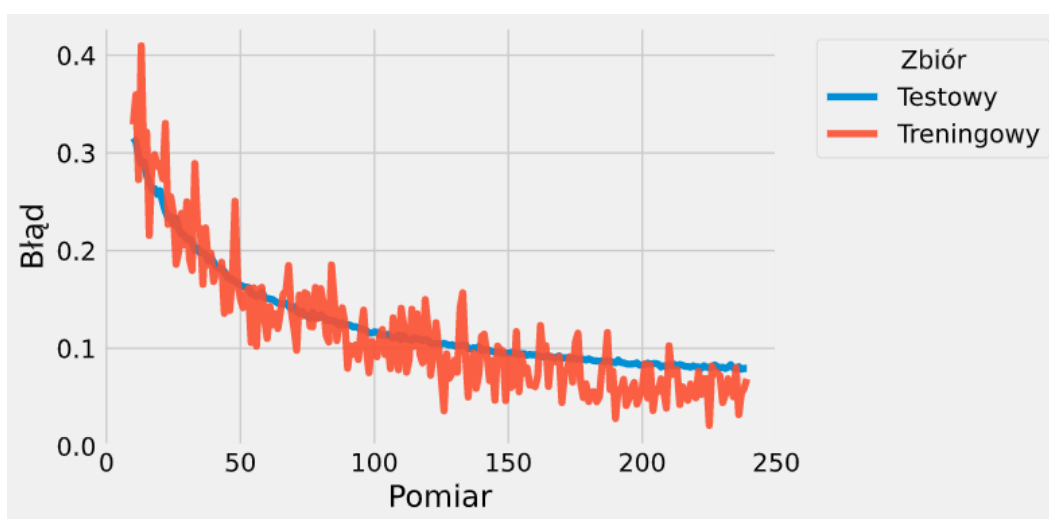
Wykres 7: Zachowanie funkcji błędu dla funkcji ReLU i optymalizatora Adam



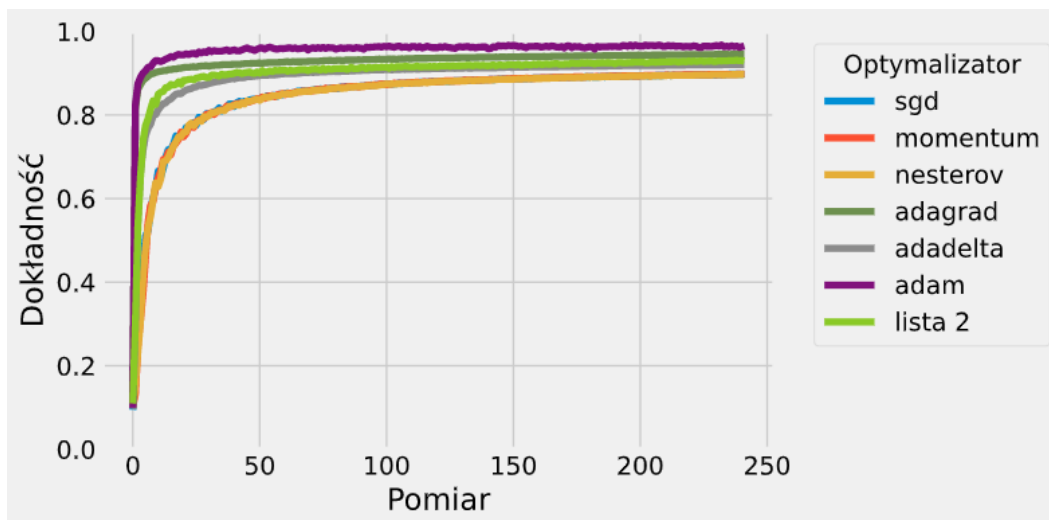
Wykres 8: Zachowanie funkcji błędu dla funkcji ReLU i optymalizatora SGD



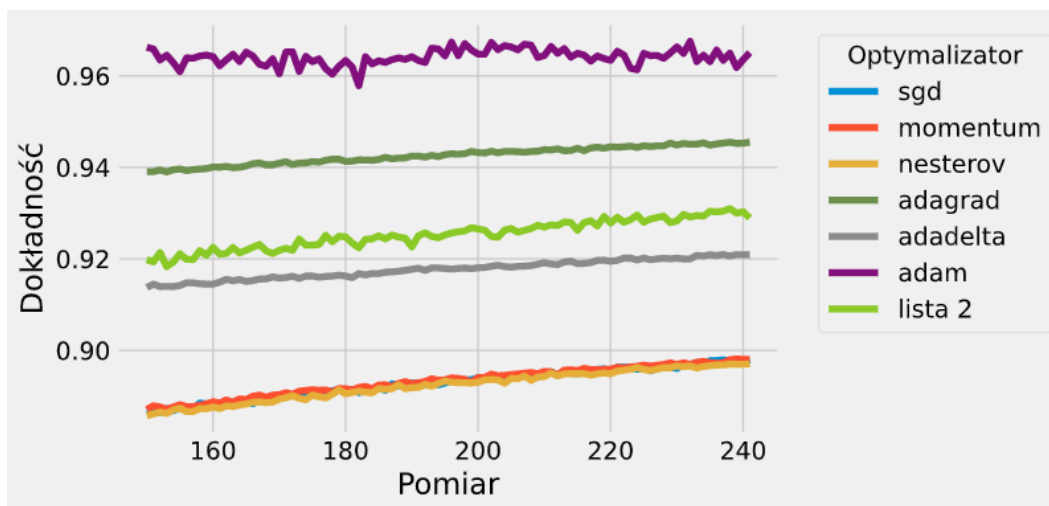
Wykres 9: Zachowanie funkcji błędu dla funkcji ReLU z listy 2



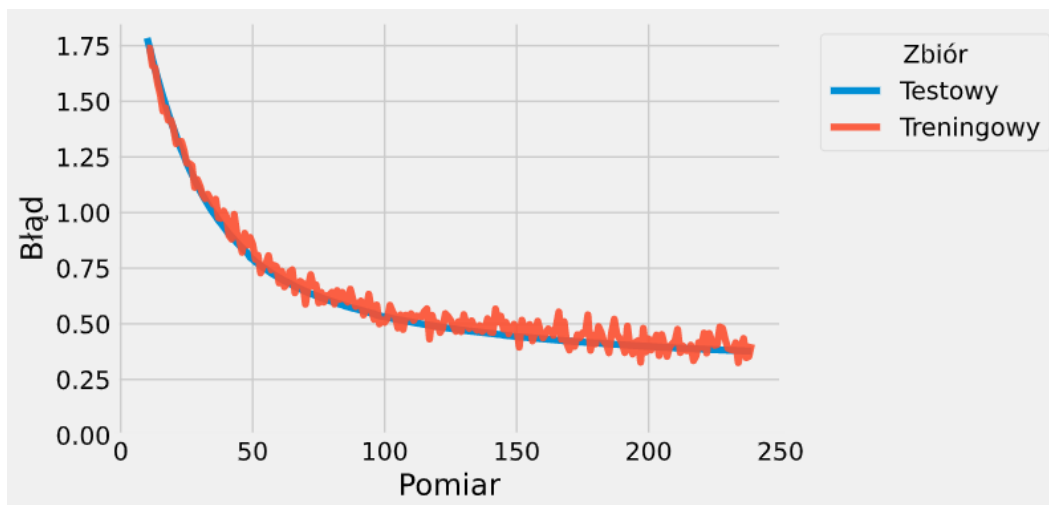
Wykres 10: Dokładność modelu w zależności od użytego optymalizatora dla funkcji Sigmoid



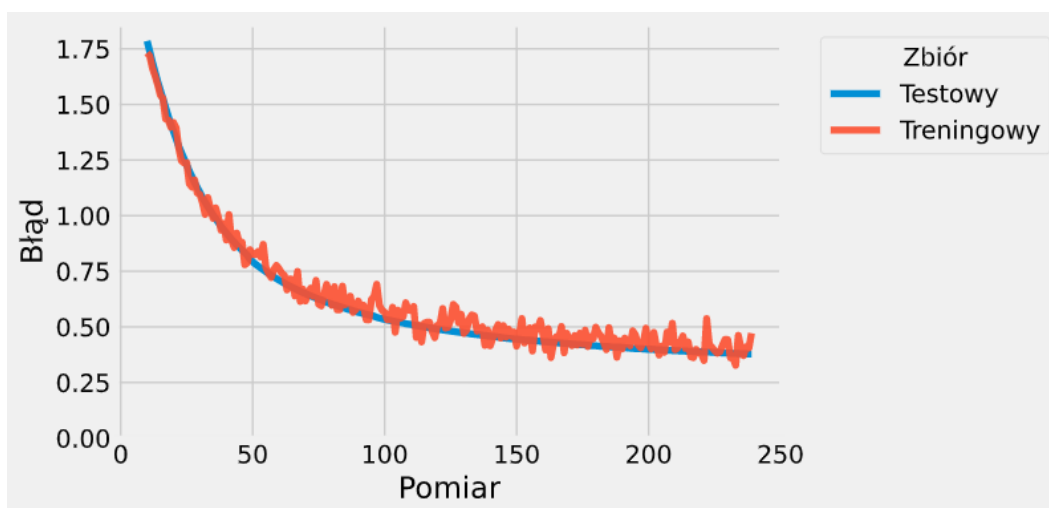
Wykres 11: Dokładność modelu w końcowym etapie uczenia w zależności od użytego optymalizatora dla funkcji Sigmoid



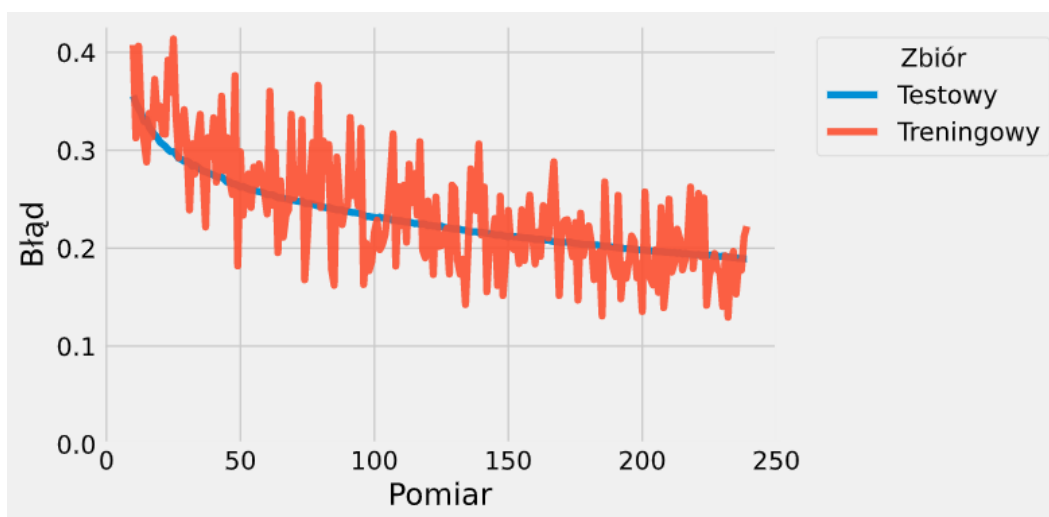
Wykres 12: Zachowanie funkcji błędu dla funkcji ReLU i optymalizatora Momentum



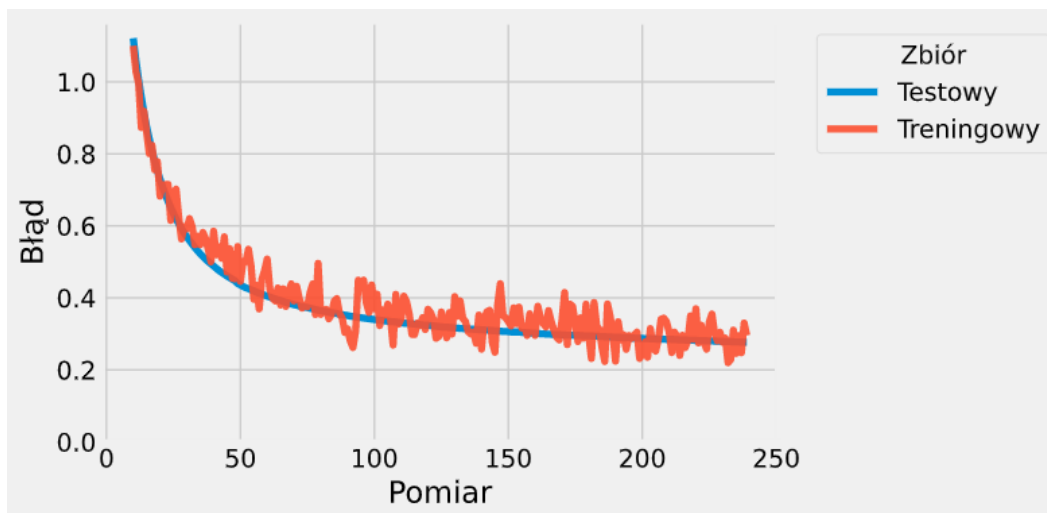
Wykres 13: Zachowanie funkcji błędu dla funkcji ReLU i optymalizatora Nesterov



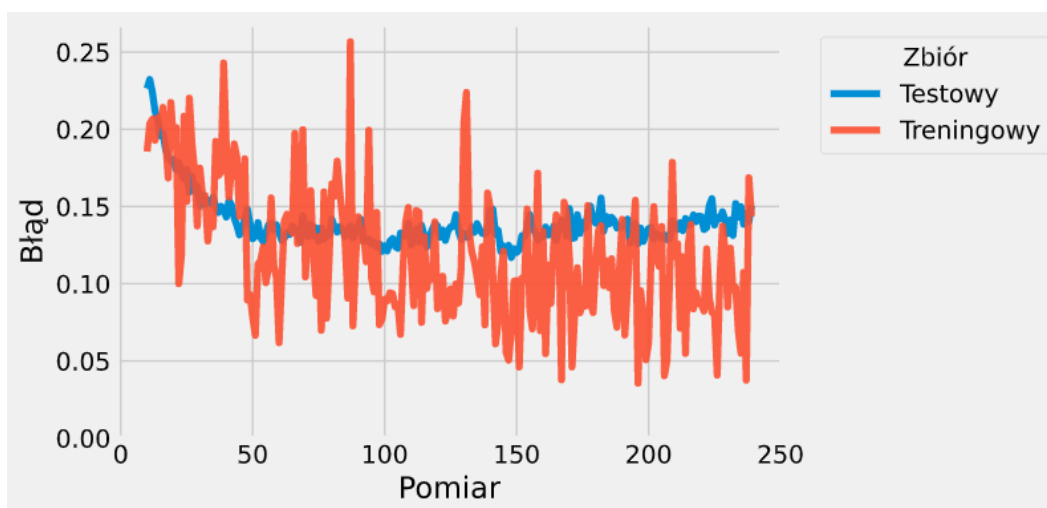
Wykres 14: Zachowanie funkcji błędu dla funkcji ReLU i optymalizatora AdaGrad



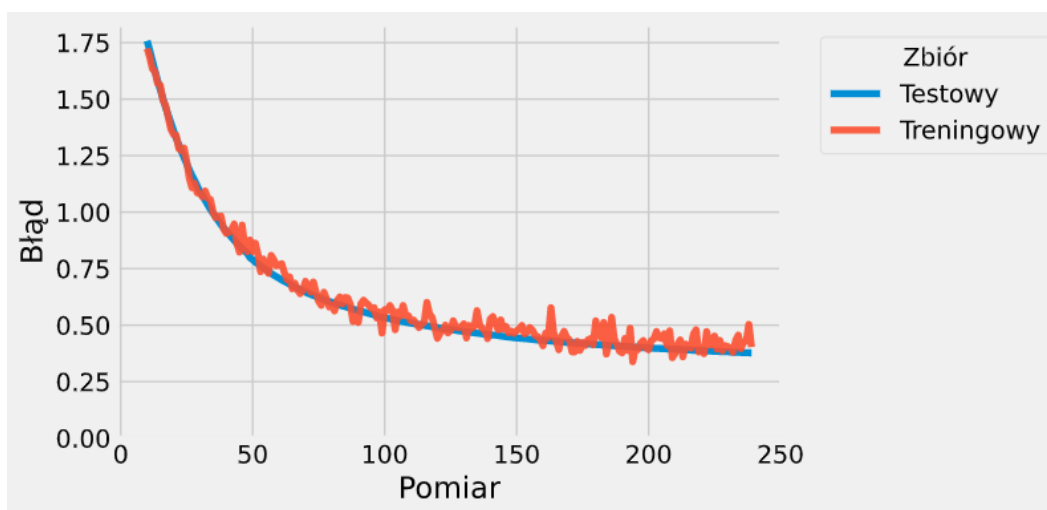
Wykres 15: Zachowanie funkcji błędu dla funkcji ReLU i optymalizatora AdaDelta



Wykres 16: Zachowanie funkcji błędu dla funkcji ReLU i optymalizatora Adam



Wykres 17: Zachowanie funkcji błędu dla funkcji ReLU i optymalizatora SGD



Wykres 18: Zachowanie funkcji błędu dla funkcji Sigmoid z listy 2

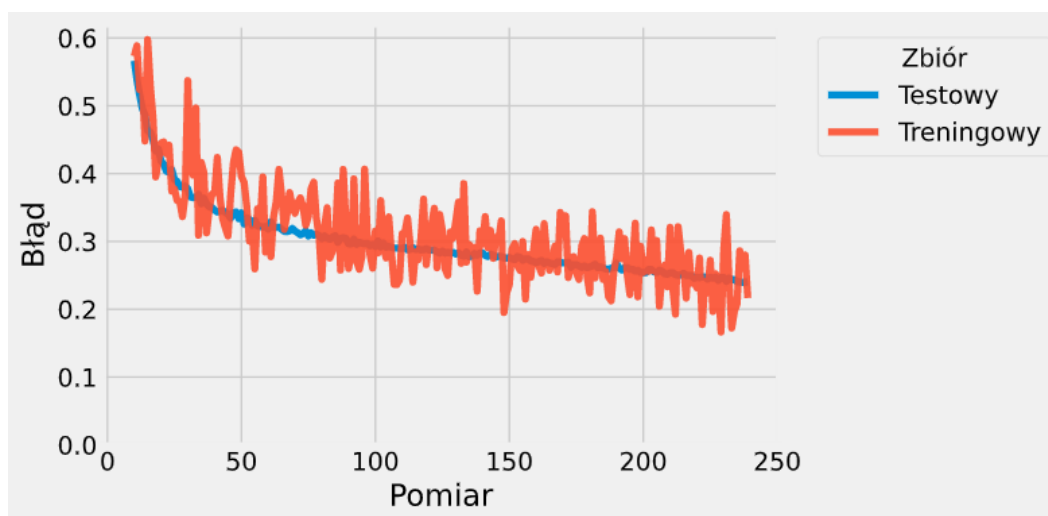


Tabela 2: Średnia maksymalna dokładność w zależności od użytego optymalizatora

Optymalizator	Dokładność [%]	
	ReLU	Sigmoid
Momentum	93.92	89.91
Nesterov	93.91	89.81
AdaGrad	97.72	94.62
AdaDelta	95.28	92.21
Adam	93.86	97.15
SGD	93.86	89.89
Lista 2	97.70	93.29

Wnioski

TODO

2.2 Wpływ inicjalizacji wag na przebieg procesu uczenia

Założenia

Tabela 3: Stałe dla eksperymentu 2

Parametr	Wartość
Optymalizator	SGD

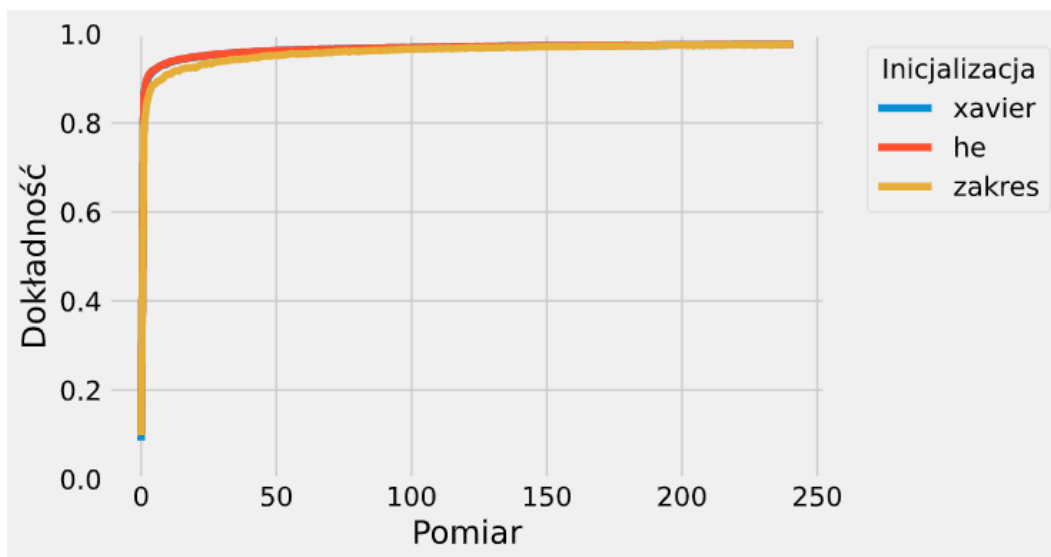
Zmienną w tym eksperymencie był sposób inicjalizacji wag. Użyto metod ze zbioru {Zakres, Xavier, He}

Przebieg

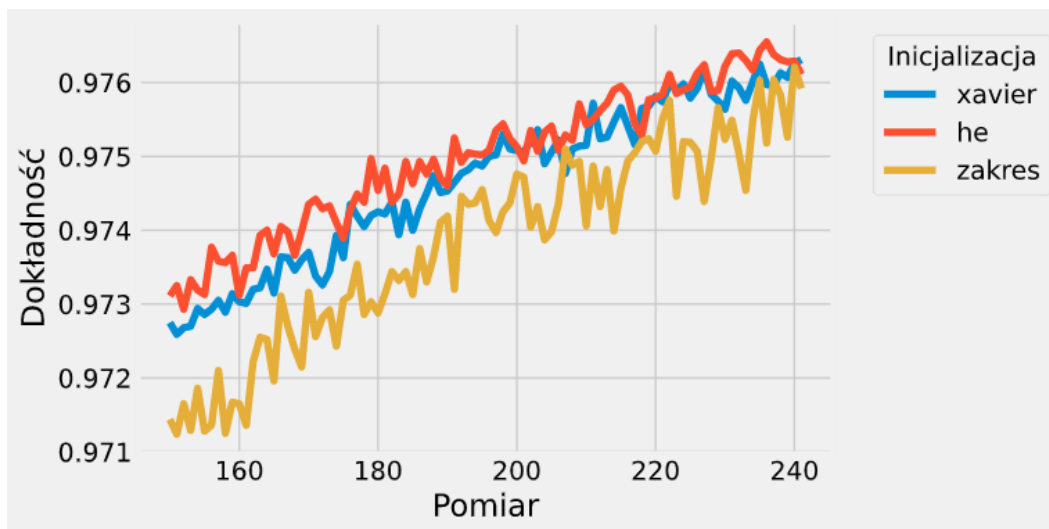
Podczas eksperymentu model został zainicjalizowany 10 razy dla każdej z badanych wartości oraz wyuczony, uzyskane wyniki zostały zapisane w postaci pliku .plk do dalszej analizy. Badania wykonano dla funkcji aktywacji Sigmoid oraz ReLU.

Wyniki

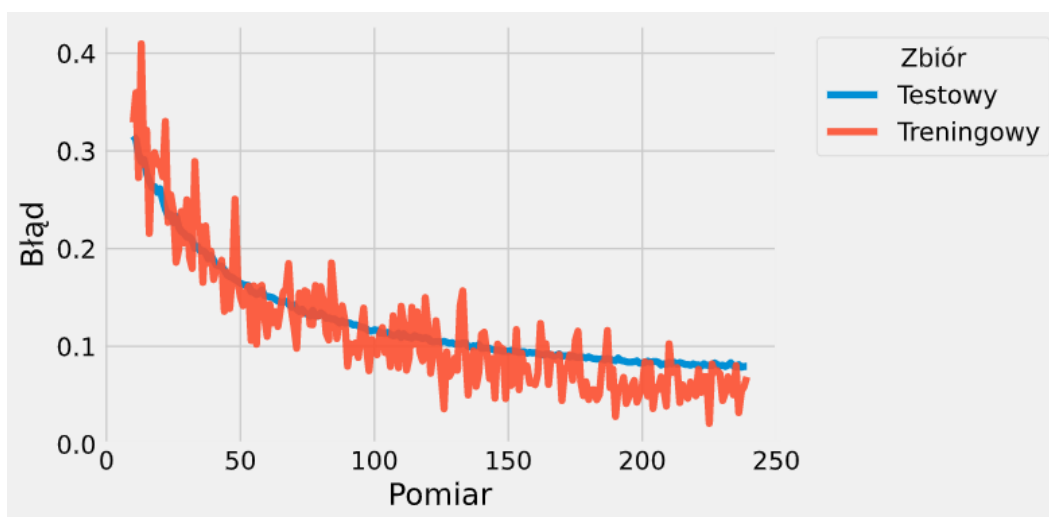
Wykres 19: Dokładność modelu w zależności od sposobu inicjalizacji wag dla funkcji ReLU



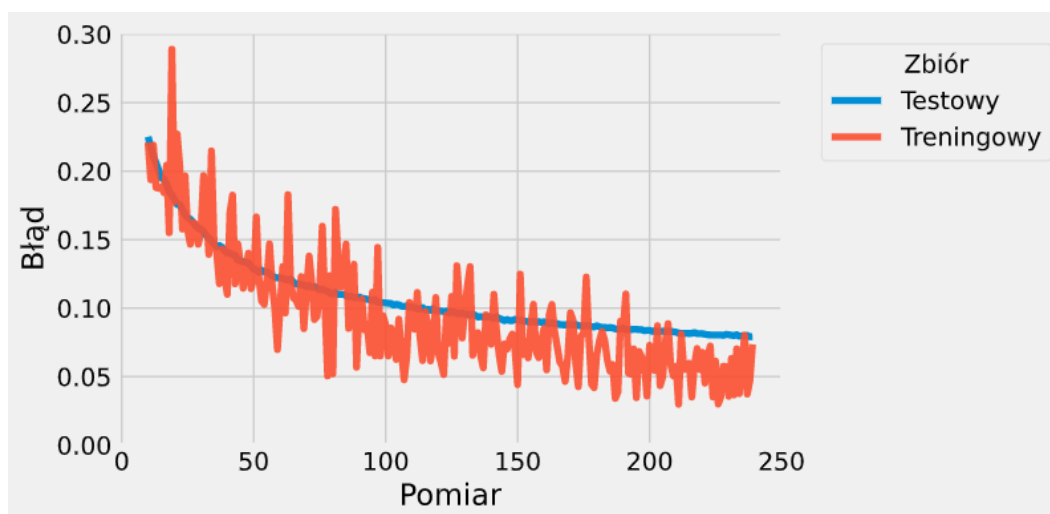
Wykres 20: Dokładność modelu w końcowym etapie uczenia w zależności od sposobu inicjalizacji wag dla funkcji ReLU



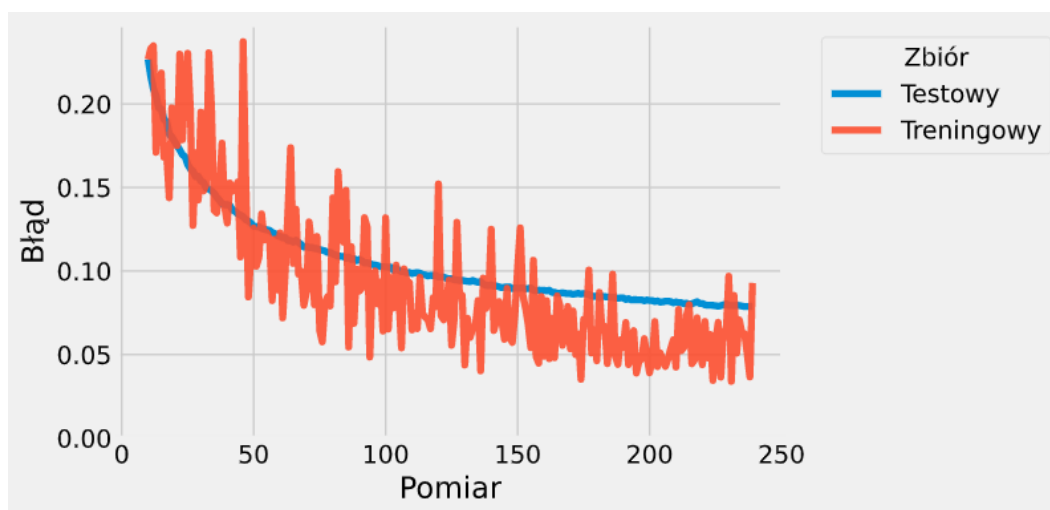
Wykres 21: Zachowanie funkcji błędu dla funkcji ReLU i inicjalizacji z zakresu



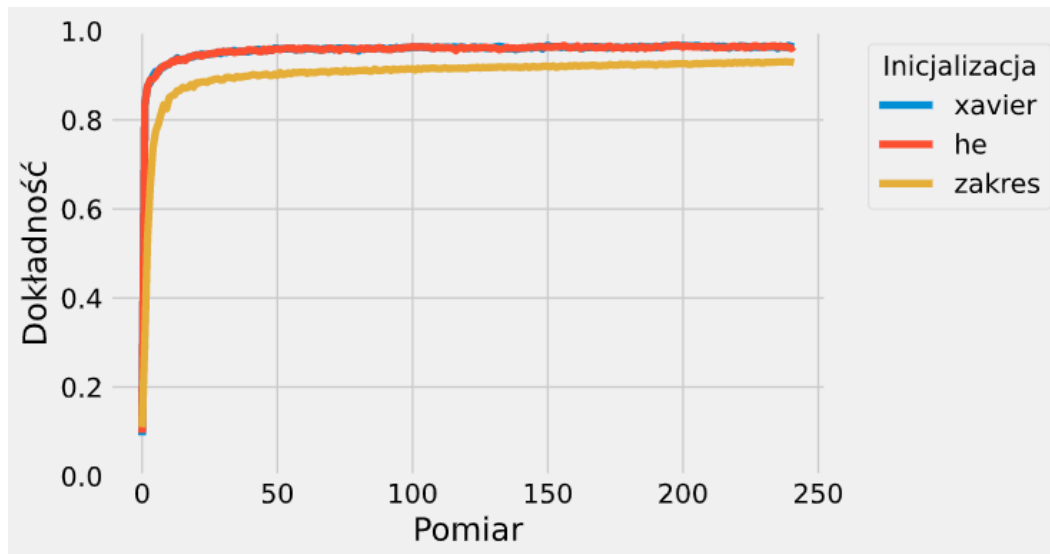
Wykres 22: Zachowanie funkcji błędu dla funkcji ReLU i inicjalizacji Xaviera



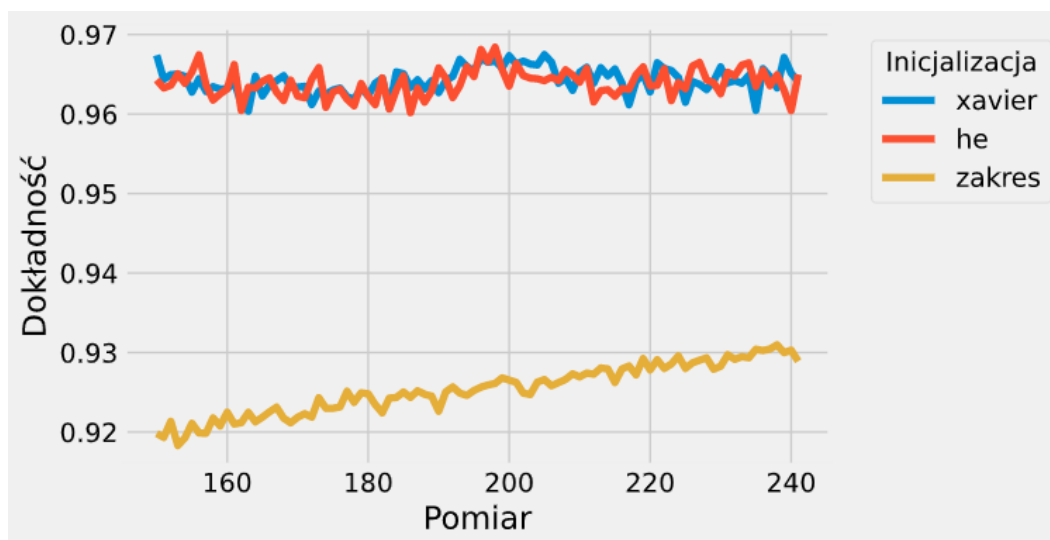
Wykres 23: Zachowanie funkcji błędu dla funkcji ReLU i inicjalizacji He



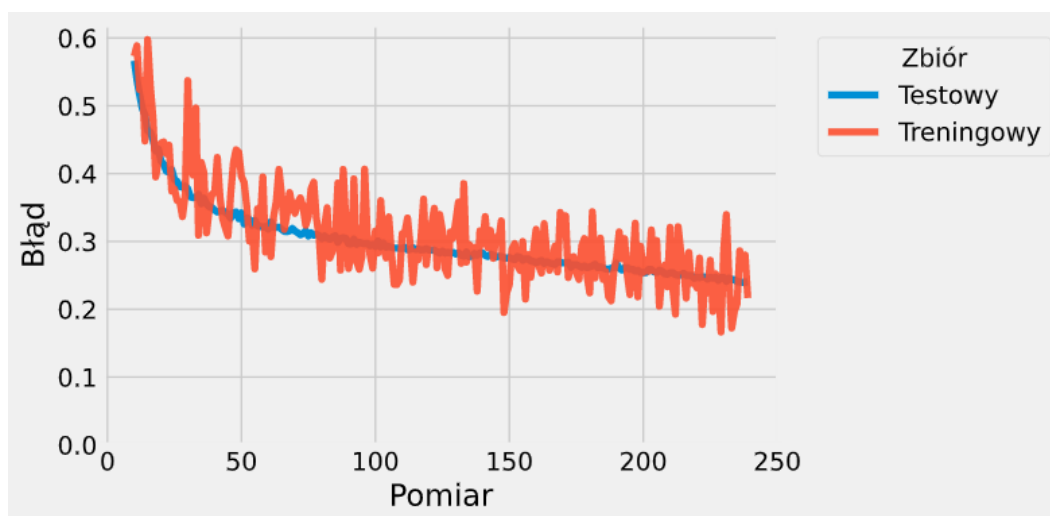
Wykres 24: Dokładność modelu w zależności od sposobu inicjalizacji wag dla funkcji Sigmoid



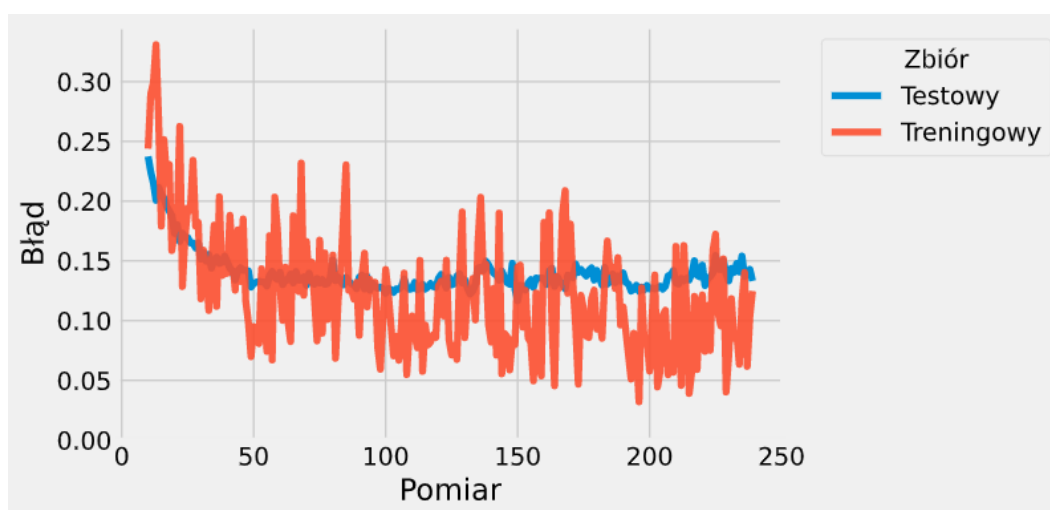
Wykres 25: Dokładność modelu w końcowym etapie uczenia w zależności od sposobu inicjalizacji wag dla funkcji Sigmoid



Wykres 26: Zachowanie funkcji błędu dla funkcji ReLU i inicjalizacji z zakresu



Wykres 27: Zachowanie funkcji błędu dla funkcji ReLU i inicjalizacji Xavier



Wykres 28: Zachowanie funkcji błędu dla funkcji ReLU i inicjalizacji He

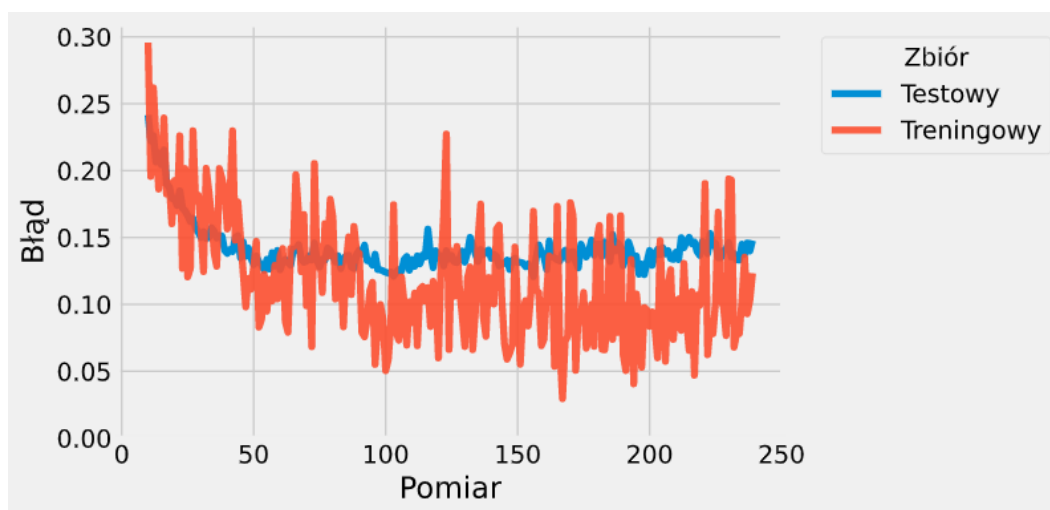


Tabela 4: Średnia maksymalna dokładność w zależności od sposobu inicjalizacji wag

Inicjalizacja	Dokładność [%]	
	ReLU	Sigmoid
Zakres	97.70	93.29
Xavier	97.71	97.18
He	97.73	97.16

Wnioski

TODO

3 Wnioski

-