

Sztuczna inteligencja i systemy ekspertowe

Zadanie: MNIST Dataset Digits - learning

1. Wstęp

Celem zadania było wykorzystanie sieci zaimplementowanej w poprzednim zadaniu do nauki zestawu danych udostępnionego przez MNIST, tj. bazy danych obrazów 28x28 pikseli reprezentujących odręcznie napisane liczby całkowite [0;9].

Model naszej sieci zawsze będzie miał 784 neurony w warstwie wejściowej i 10 neuronów w warstwie wyjściowej. Ilość neuronów jest wynikiem wielkości obrazów 28x28 oraz ilości klas przewidywanych.

2. Badania

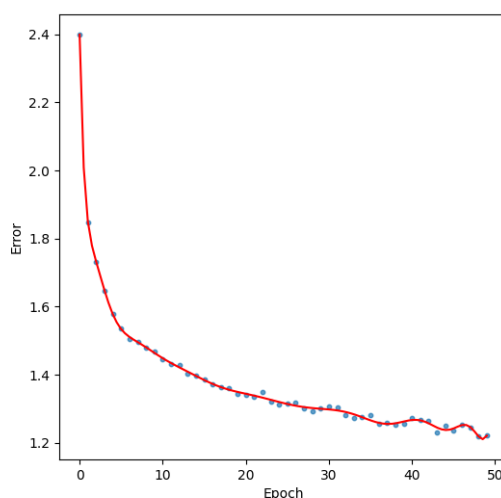
Przetestowaliśmy wydajność nauki na różnych architekturach sieci w celu porównania zachowania zbioru danych biorąc pod uwagę skomplikowanie modelu.

Dane zostały podzielone na treningowe – 60000 przypadków oraz testowe – 10000.

Parametry sieci:

- **learning rate: 0.3,**
- **momentum: 0.8**

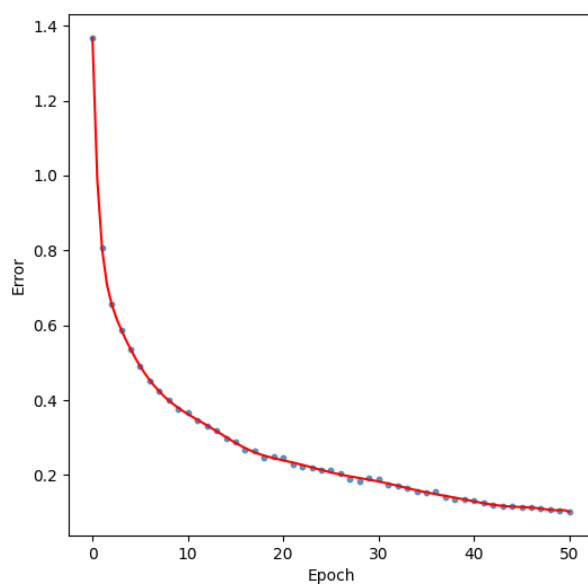
2.1. Model 784 - 8 – 10 porównawczy



Rys. 1. Wykresy błędów podczas nauki sieci (model 784-8-10) względem epoki.

W przypadku powyższego modelu po upływie 50 epok osiągnęliśmy dokładność zgadywania na poziomie 91,61%.

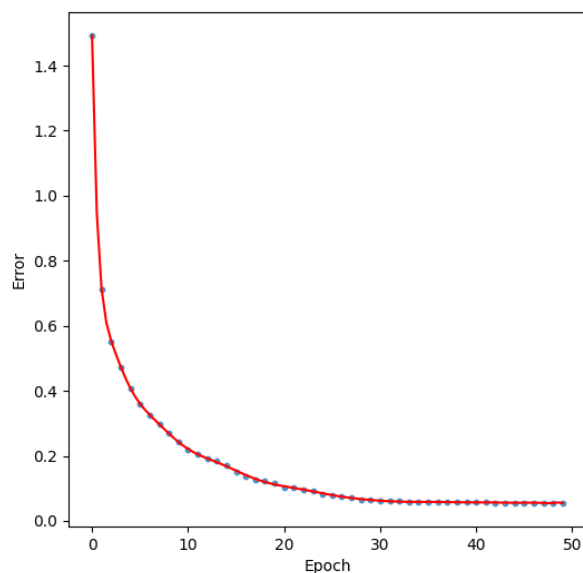
2.2 Model 784 - 60 - 10 porównawczy



Rys. 2. Wykresy błędów podczas nauki sieci (model 784-60-10) względem epoki.

W przypadku powyższego modelu po upływie 50 epok osiągnęliśmy dokładność zgadywania na poziomie 97,25%.

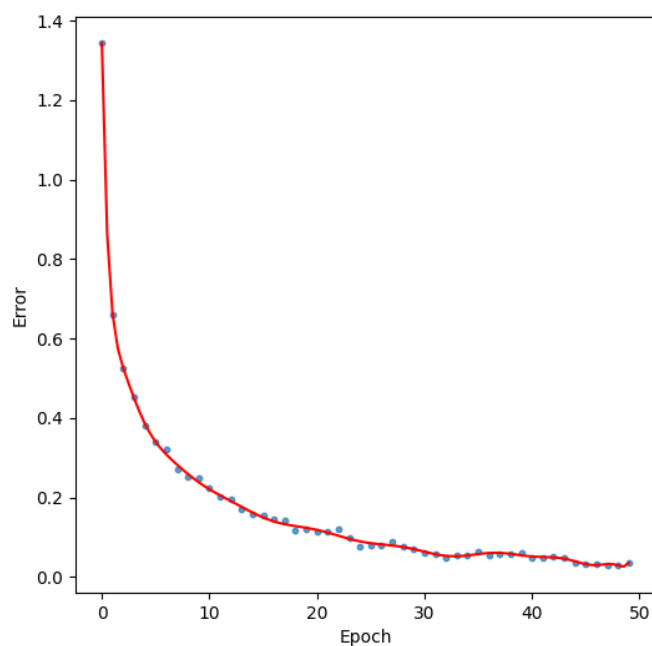
2.3 Model 784 – 120 - 10 porównawczy



Rys. 3. Wykresy błędów podczas nauki sieci (model 784-120-10) względem epoki

W przypadku powyższego modelu po upływie 50 epok osiągnęliśmy dokładność zgadywania na poziomie 97,74%.

2.4 Model 784 - 120 - 80 - 10 porównawczy



Rys. 4. Wykresy błędów podczas nauki sieci (model 784-120-80-10) względem epoki

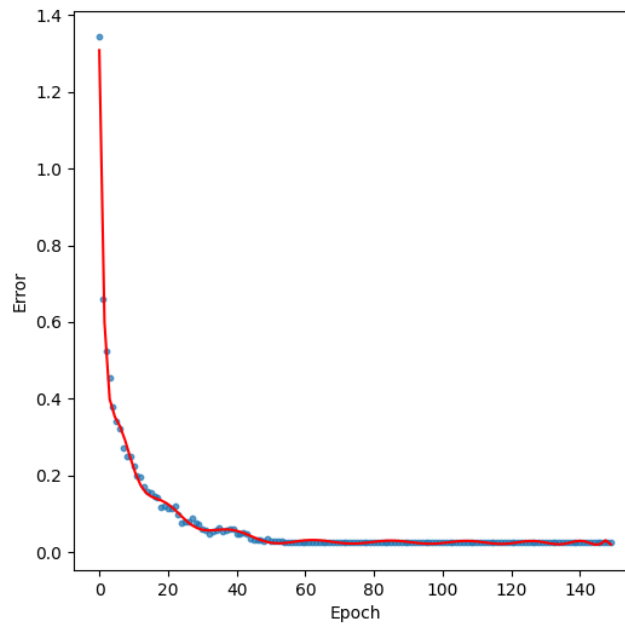
W przypadku powyższego modelu po upływie 50 epok osiągnęliśmy dokładność zgadywania na poziomie 97,63%.

2.5 Porównanie modeli testowych

Tab. 1. Wartości osiągnięte przy nauczaniu poszczególnych modeli sieci

Model	Dokładność zgadywania
784 – 8 – 10	91,61%
784 – 60 – 10	97,25%
784 – 120 – 10	97,74%
784 – 120 – 80 – 10	97,87%

3. Wybrana sieć końcowa (784 – 120 – 80 – 10)



Rys. 5. Wykresy błędów podczas nauki sieci (model 784-120-80-10) względem epoki

W przypadku powyższego modelu po upływie 50 epok osiągnęliśmy dokładność zgadywania na poziomie 97,94%.

4. Wnioski końcowe

- Badania dają nam podstawę do stwierdzenia, że wraz ze wzrostem skomplikowania modelu rośnie wydajność zgadywania na danych testowych.
- Możemy zauważyć, że niewielka ilość epok wystarcza do nauczenia sieci wydajnego zgadywania.
- Po pewnej ilości epok sieć stabilizuje się w zakresie jednego błędu.