

Aleksander Kaźmierczak 251544
Adrian Urbańczyk 252960

Rok akademicki 2024/25
Wtorek 10:30

SZTUCZNA INTELIGENCJA I SYSTEMY EKSPERTOWE – LABORATORIUM

Zadanie 3 – klasyfikacja osób pacjentów zdrowych i prawdopodobnie chorych

Celem projektu było rozwiązanie problemu klasyfikacji osób potencjalnie chorych. Do tego celu użyliśmy sieci MLP z gotowej biblioteki PyTorch. Dane pochodziły z [ogólnodostępnej bazy danych](#). Wstępne przygotowanie danych zostało zrealizowane przy użyciu biblioteki sklearn. Do skalowania użyliśmy StandardScaler.

W bazie danych występowały braki, które zostały uzupełnione przez wstawienie średnich wartości.

Zastosowane ustawienia sieci:

Podział danych treningowych – 90%

Podział danych testowych – 10%

Warstwa wejściowa – 13 wejścia

Pierwsza warstwa ukryta – 30 neuronów

Druga warstwa ukryta – 30 neuronów

Warstwa wyjściowa – 2 wyjścia (one-hot)

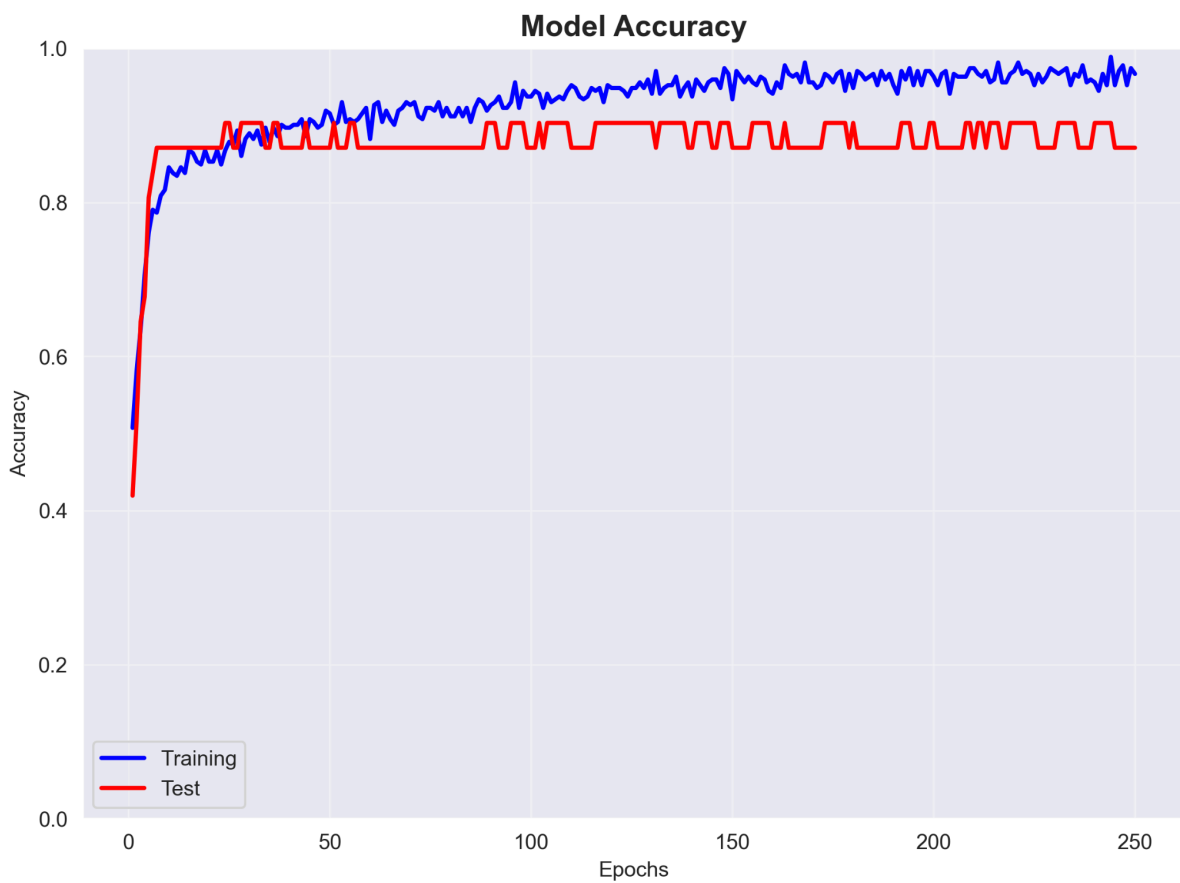
Współczynnik nauki – 0.001

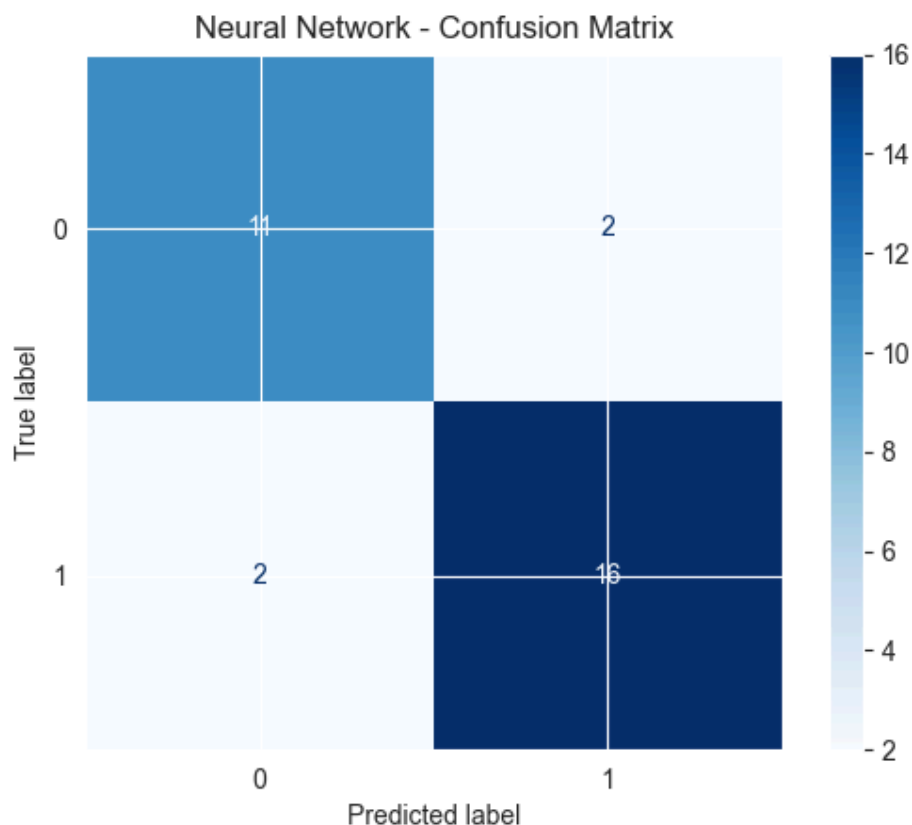
Optymalizator - Adam

Liczba epok – 250

Funkcja aktywacji - ReLU

$$\text{ReLU}(x) = x^+ = \max(0, x) = \frac{x + |x|}{2} = \begin{cases} 0 & \text{dla } x < 0 \\ x & \text{dla } x \geq 0 \end{cases}$$





Accuracy: 0.87
Precision: 0.89
Recall: 0.89
F1 Score: 0.89

Neural Network Final Training Accuracy: 96.69%
Neural Network Final Test Accuracy: 87.10%
Neural Network Final Accuracy: 87.10%

Z powyższych wykresów możemy wywnioskować, że w pewnym momencie (około 20 epoki) nie zwiększa się dokładność.

Powyżej 250 epoki sieć już się przeucza.

Ze statystyk możemy stwierdzić, iż:

- 87% przewidywań było trafnych,
- model przewidział chorobę serca w 89% (precision 89%) - nie myli ze zdrowymi przypadkami,
- model wykrył 89% faktycznych przypadków chorób (recall 89%)
- ogólna jakość wykrywania chorób to 0.89 (średnia harmoniczna precyzji i czułości)