Jarosław Dutka

Inżynieria Obliczeniowa, rok III

WIMIiIP, AGH

**Sprawozdanie nr 3**

„Budowa i działanie sieci wielowarstwowej”

Celem ćwiczenia było poznanie budowy i działania wielowarstwowej sieci neuronowej poprzez uczenie rozpoznawania liter z użyciem algorytmu wstecznej propagacji.

1. **Przebieg ćwiczenia**

W realizacji ćwiczenia wykorzystano program MATLAB wraz z biblioteką Neural Network Toolbox.

W ramach ćwiczeń wygenerowano 10 małych oraz 10 dużych liter alfabetu łacińskiego. Litery wykonano w rozmiarze 4 x 6 pikseli.



Rys. 1.1 Przykładowa litera wraz z numeracją pól.

Do stworzenia wielowarstwowej sieci neuronowej użyto funkcji net = newff(PR, S, {Fa1, Fa2,.. Fan}, FU) , gdzie:

* net - oznacza nazwę tworzonej sieci;
* PR – macierz dwuwymiarowa, w której pierwszym elementem jest minimalna wartość wejścia a drugim maksymalna;
* S – macierz zawierająca ilość neuronów w każdej warstwie;
* Fa – funkcja aktywacji dla każdej warstwy;
* FU – funkcja uczenia sieci, w tym przypadku algorytm wstecznej propagacji (‘traingda’).

Dane uczące zostały zapisane w postaci macierzy tablicy dwuwymiarowej o nazwie WE. Podobnie postąpiono z wynikami, które zapisano w macierzy WY.

Uczenie zostało zrealizowane w funkcją net = train(net, WE, WY).

Do testów stworzono 20 zmiennych tablicowych o wymiarach 4 x 6. Każda z tych zmiennych odpowiada innej literze ze zbioru danych uczących. Stworzono także zmienną wynik, do której zapisywano wynik funkcji sim(net, testX). Funkcja ta oblicza skalę podobieństwa podanej danej do danych wejściowych. Następnie spośród danych zapisanych w „wynik” wyszukano wartość maksymalną i do konsoli wypisano odpowiednią informację.

**Testowanie działania aplikacji**

Każdą próbę przeprowadzono pięciokrotnie i policzono wynik średni.

1. Test I

Ustawiono dwie warstwy sieci, składające się kolejno z 24 i 20 neuronów. W obydwu warstwach funkcją aktywacji była funkcja logarytmiczna.

Dla podanych danych uzyskano średnią ilość epok na poziomie 532,4. Średni czas nauki wynosił 1,2 sekundy.

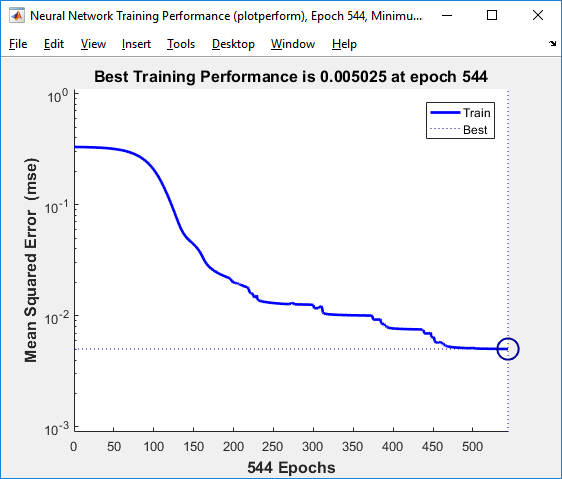
|  |  |  |
| --- | --- | --- |
| **Lp.** | **Ilość epok** | **Czas [s]** |
| 1. | 456 | 2 |
| 2. | 532 | 1 |
| 3. | 522 | 1 |
| 4. | 564 | 1 |
| 5. | 588 | 1 |
| **Średnia** | **532,4** | **1,2** |

Tabela przedstawia

wyniki uzyskane w

poszczególnych

próbach.

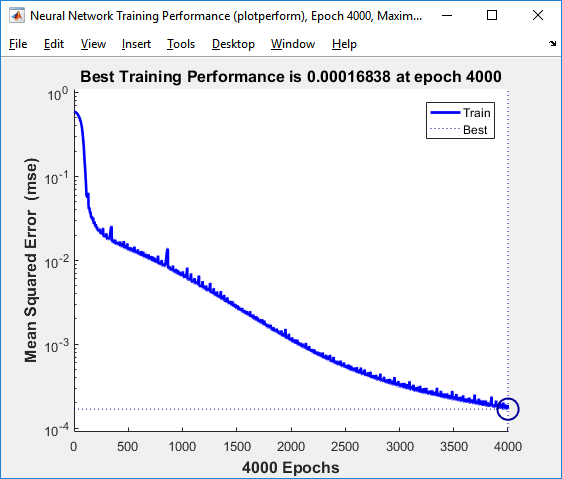


Przykładowy wykres średniego błędu kwadratowego w zależności od ilości epok podczas nauki.

1. Test II

Ustawiono dwie warstwy sieci, składające się kolejno z 24 i 20 neuronów. W obydwu funkcją aktywacji był tangens hiperboliczny.

Przy tym ustawieniu za każdym razem uzyskiwano programowo ustawioną maksymalną ilość epok, równą 4000.

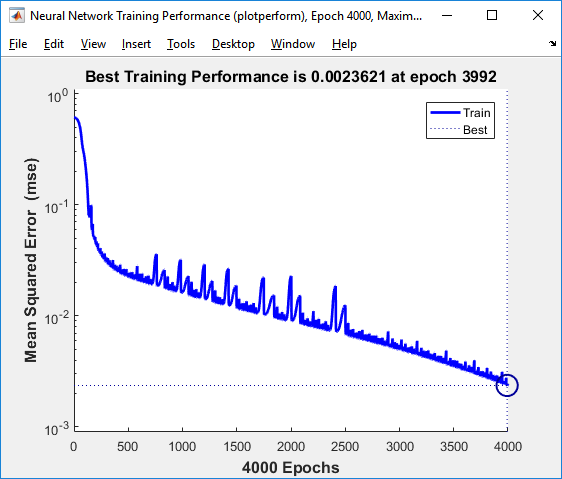


Przykładowy wykres średniego błędu kwadratowego w zależności od ilości epok podczas nauki.

1. Test III

Ustawiono dwie warstwy sieci, składające się kolejno z 24 i 20 neuronów. W pierwszej warstwie funkcją aktywacji była funkcja logarytmiczna a w drugiej tangens hiperboliczny.

W każdej próbie uzyskiwano maksymalną ilość iteracji równą 4000.



Przykładowy wykres średniego błędu kwadratowego w zależności od ilości epok podczas nauki.

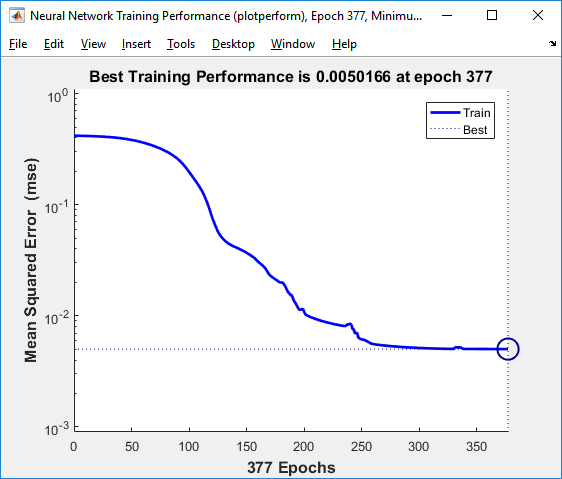
1. Test IV

Ustawiono dwie warstwy sieci, składające się kolejno z 24 i 20 neuronów. W pierwszej warstwie funkcją aktywacji był tangens hiperboliczny a w drugiej funkcja logarytmiczna.

W tej próbie osiągano skończoną ilość epok, przedstawioną w poniższej tabeli.

|  |  |  |
| --- | --- | --- |
| **Lp.** | **Ilość epok** | **Czas [s]** |
| 1. | 455 | 1 |
| 2. | 471 | 1 |
| 3. | 532 | 1 |
| 4. | 620 | 1 |
| 5. | 377 | 1 |
| **Średnia** | **491** | **1** |

Tabela przedstawia średnią ilość epok i czas wykonania programu w teście IV.



Przykładowy wykres średniego błędu kwadratowego w zależności od ilości epok podczas nauki.

**Wnioski**

1. Z testów nr I i IV wynika, że użycie funkcji logarytmicznej jako funkcji aktywacji w pierwszej warstwie daje w efekcie skończoną ilość epok nauki.
2. Najlepszy wynik nauki (Best Training Performance) jaki udało się uzyskać wnosił 0,00017, przy 4000 iteracji. Był on efektem użycia w obydwu warstwach sieci neuronowej tangensa hiperbolicznego jako funkcji aktywacji.
3. Z powyższych testów wynika, że ogromne znaczenie ma nie tylko rodzaj funkcji aktywacji, ale również kolejność ich zastosowania w poszczególnych warstwach.
4. Współczynnik uczenia ma równie istotny wpływa na jakość nauki w sieciach jednowarstwowych i wielowarstwowych.
5. Najmniejsza ilość epok, która wystarczyła do prawidłowego działania programu, którą udało się uzyskać podczas testów wynosiła 377.