Paweł Put, Informatyka Stosowana gr. 3

Sprawozdanie nr 1.

Temat ćwiczenia: Budowa i działanie perceptronu.

Cel ćwiczenia: Celem ćwiczenia jest poznanie budowy i działania perceptronu poprzez implementację oraz uczenie perceptronu realizującego wybraną funkcję logiczną dwóch zmiennych.

Przebieg: Zaimplementowałem perceptron w języku C++ tak, aby uczył się funkcji logicznej OR. Korzystałem przy tym ze stron internetowych:

* <http://edu.pjwstk.edu.pl/wyklady/nai/scb/wyklad2/w2.htm>
* <http://edu.pjwstk.edu.pl/wyklady/nai/scb/wyklad3/w3.htm>

Algorytm uczenia:

1. Inicjujemy wagi losowo
2. Ustawiamy nauczanie perceptronu dla czterech par neuronów: 0, 0; 1, 0; 0, 1 i 1,1 i nadajemy im odpowiednie oczekiwane wartości do wyjścia, czyli 0, 1, 1, 1.
3. Obliczamy wartość wyjścia dla wylosowanych wag i odchylenia standardowego.
4. Jeżeli wartość wyjścia i oczekiwanego wyjścia się zgadza, przechodzimy do następnego przypadku. W przeciwnym razie modyfikujemy wagi połączeń i odchylenie standardowe tak, by działały dla aktualnej pary neuronów i wracamy do pierwszego wypadku.
5. Powtarzamy krok 4 do momentu, aż przejdziemy przez wszystkie pary neuronów nie zmieniając wartości wag i odchylenia standardowego.

Wzór na modyfikowanie wag oraz odchylenia standardowego:

* weight[i] += learnCoefficient \* (expectAnswer - out) \* in[i]
* bias += learnCoefficient \* (expectAnswer - out),

gdzie:

* weight[i] - waga przy i-tym neuronie
* learnCoefficient - współczynnik nauczania podawany przez użytkownika
* expectAnswer - wartość oczekiwana dla danej pary neuronów
* out - wartość wyliczona ze wzoru
* in[i] - wartość i-tego neuronu
* bias - wartość odchylenia

Wartość wyjściową otrzymujemy dodając za każdym razem iloczyn wszystkich wag i neuronów, a następnie dodając wartość odchylenia standardowego. Jeżeli podana wartość jest większa lub równa 0 to na wyjściu otrzymujemy logiczną 1, w przeciwnym wypadku otrzymujemy 0.

Przetestujemy działanie programu dla różnego zestawu danych:

1. Współczynników odchylenia:

* 0,1
* 0,5
* 1

1. Współczynników nauczania:

* 0,01
* 0,1
* 0,5
* 1

Dla współczynnika odchylenia 0,1 i współczynnika nauczania 0,5:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **l.p.** | **waga pierwsza przed nauką** | **waga pierwsza po nauce** | **różnica w pierwszych**  **wagach** | **waga druga przed nauką** | **waga druga po nauce** | **różnica w drugich wagach** | **liczba pętli do nauki** | **wsp.**  **odchylenia**  **na wyjściu** |
| 1 | 0,206 | 0,706 | 0,5 | 0,352 | 0,852 | 0,5 | 5 | -0,4 |
| 2 | 0,065 | 0,565 | 0,5 | 0,135 | 0,635 | 0,5 | 5 | -0,4 |
| 3 | 0,287 | 0,787 | 0,5 | 0,131 | 0,631 | 0,5 | 5 | -0,4 |
| 4 | 0,444 | 0,944 | 0,5 | -0,234 | 0,766 | 1 | 5 | -0,4 |
| 5 | -0,279 | 0,721 | 1 | -0,123 | 0,877 | 1 | 9 | -0,4 |
| **średnia** | **0,145** | **0,745** | **0,6** | **0,052** | **0,752** | **0,7** | **5,8** | **-0,4** |

Dla współczynnika odchylenia 0,5 i współczynnika nauczania 0,5:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **l.p.** | **waga pierwsza przed nauką** | **waga pierwsza po nauce** | **różnica w pierwszych**  **wagach** | **waga druga przed nauką** | **waga druga po nauce** | **różnica w drugich wagach** | **liczba pętli do nauki** | **wsp.**  **odchylenia**  **na wyjściu** |
| 1 | -0,347 | 0,653 | 1 | -0,264 | 0,736 | 1 | 10 | -0,5 |
| 2 | 0,469 | 0,969 | 0,5 | -0,139 | 0,861 | 1 | 8 | -0,5 |
| 3 | -0,253 | 0,747 | 1 | -0,028 | 0,972 | 1 | 10 | -0,5 |
| 4 | -0,074 | 0,926 | 1 | 0,31 | 0,81 | 0,5 | 8 | -0,5 |
| 5 | 0,341 | 0,841 | 0,5 | 0,102 | 0,602 | 0,5 | 6 | -0,5 |
| **średnia** | **0,033** | **0,827** | **0,8** | **-0,004** | **0,796** | **0,8** | **8,4** | **-0,5** |

Dla współczynnika odchylenia 1 i współczynnika nauczania 0,5:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **l.p.** | **waga pierwsza przed nauką** | **waga pierwsza po nauce** | **różnica w pierwszych**  **wagach** | **waga druga przed nauką** | **waga druga po nauce** | **różnica w drugich wagach** | **liczba pętli do nauki** | **wsp.**  **odchylenia**  **na wyjściu** |
| 1 | 0,347 | 0,847 | 0,5 | 0,043 | 0,543 | 0,5 | 7 | -0,5 |
| 2 | -0,405 | 0,595 | 1 | -0,278 | 0,722 | 1 | 11 | -0,5 |
| 3 | -0,137 | 0,863 | 1 | 0,123 | 0,623 | 0,5 | 9 | -0,5 |
| 4 | 0,072 | 0,572 | 0,5 | -0,107 | 0,893 | 1 | 9 | -0,5 |
| 5 | -0,5 | 0,5 | 1 | -0,09 | 0,91 | 1 | 11 | -0,5 |
| **średnia** | **-0,125** | **0,675** | **0,8** | **-0,062** | **0,738** | **0,8** | **9,4** | **-0,5** |

Dla współczynnika odchylenia 0,5 i współczynnika nauczania 0,01:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **l.p.** | **waga pierwsza przed nauką** | **waga pierwsza po nauce** | **różnica w pierwszych**  **wagach** | **waga druga przed nauką** | **waga druga po nauce** | **różnica w drugich wagach** | **liczba pętli do nauki** | **wsp.**  **odchylenia**  **na wyjściu** |
| 1 | 0,489 | 0,489 | 0 | -0,123 | 0,017 | 0,14 | 79 | -0,01 |
| 2 | -0,035 | 0,015 | 0,05 | 0,493 | 0,493 | 0 | 61 | -0,01 |
| 3 | 0,164 | 0,164 | 0 | -0,214 | 0,016 | 0,23 | 97 | -0,01 |
| 4 | -0,388 | 0,012 | 0,4 | -0,242 | 0,018 | 0,26 | 183 | -0,01 |
| 5 | 0,017 | 0,017 | 0 | 0,072 | 0,072 | 0 | 51 | -0,01 |
| **średnia** | **0,049** | **0,683** | **0,09** | **-0,003** | **0,123** | **0,126** | **94,2** | **-0,01** |

Dla współczynnika odchylenia 0,5 i współczynnika nauczania 0,1:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **l.p.** | **waga pierwsza przed nauką** | **waga pierwsza po nauce** | **różnica w pierwszych**  **wagach** | **waga druga przed nauką** | **waga druga po nauce** | **różnica w drugich wagach** | **liczba pętli do nauki** | **wsp.**  **odchylenia**  **na wyjściu** |
| 1 | -0,455 | 0,145 | 0,6 | -0,178 | 0,122 | 0,3 | 24 | -0,1 |
| 2 | 0,283 | 0,283 | 0 | 0,13 | 0,13 | 0 | 6 | -0,1 |
| 3 | 0,161 | 0,161 | 0 | 0,099 | 0,199 | 0,1 | 8 | -0,1 |
| 4 | 0,432 | 0,432 | 0 | 0,481 | 0,481 | 0 | 6 | -0,1 |
| 5 | 0,381 | 0,381 | 0 | 0,176 | 0,176 | 0 | 6 | -0,1 |
| **średnia** | **0,16** | **0,28** | **0,12** | **0,071** | **0,222** | **0,08** | **10** | **-0,1** |

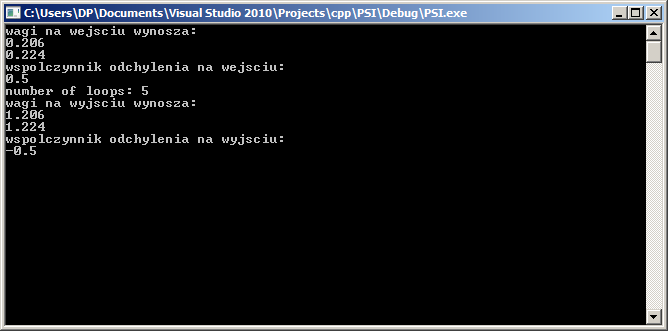
Dla współczynnika odchylenia 0,5 i współczynnika nauczania 0,5:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **l.p.** | **waga pierwsza przed nauką** | **waga pierwsza po nauce** | **różnica w pierwszych**  **wagach** | **waga druga przed nauką** | **waga druga po nauce** | **różnica w drugich wagach** | **liczba pętli do nauki** | **wsp.**  **odchylenia**  **na wyjściu** |
| 1 | -0,347 | 0,653 | 1 | -0,264 | 0,736 | 1 | 10 | -0,5 |
| 2 | 0,469 | 0,969 | 0,5 | -0,139 | 0,861 | 1 | 8 | -0,5 |
| 3 | -0,253 | 0,747 | 1 | -0,028 | 0,972 | 1 | 10 | -0,5 |
| 4 | -0,074 | 0,926 | 1 | 0,31 | 0,81 | 0,5 | 8 | -0,5 |
| 5 | 0,341 | 0,841 | 0,5 | 0,102 | 0,602 | 0,5 | 6 | -0,5 |
| **średnia** | **0,033** | **0,827** | **0,8** | **-0,004** | **0,796** | **0,8** | **8,4** | **-0,5** |

Dla współczynnika odchylenia 0,5 i współczynnika nauczania 1:

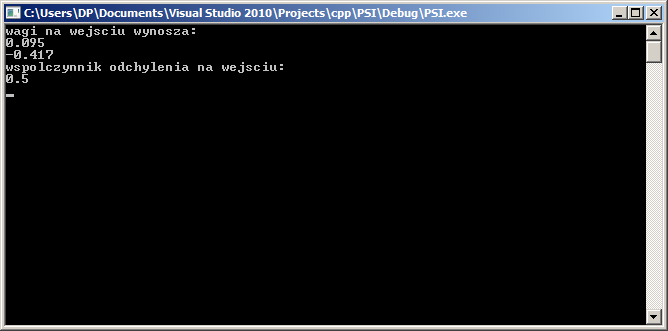
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **l.p.** | **waga pierwsza przed nauką** | **waga pierwsza po nauce** | **różnica w pierwszych**  **wagach** | **waga druga przed nauką** | **waga druga po nauce** | **różnica w drugich wagach** | **liczba pętli do nauki** | **wsp.**  **odchylenia**  **na wyjściu** |
| 1 | 0,194 | 1,194 | 1 | -0,446 | 0,554 | 1 | 5 | -0,5 |
| 2 | -0,287 | 0,713 | 1 | -0,386 | 0,614 | 1 | 5 | -0,5 |
| 3 | -0,147 | 0,853 | 1 | 0,044 | 1,044 | 1 | 5 | -0,5 |
| 4 | -0,032 | 0,968 | 1 | 0,022 | 1,022 | 1 | 5 | -0,5 |
| 5 | 0,092 | 1,092 | 1 | 0,245 | 1,245 | 1 | 5 | -0,5 |
| **średnia** | **-0,036** | **0,964** | **1** | **-0,104** | **0,896** | **1** | **5** | **-0,5** |

Przykładowy screen z wynikami:



Wykresy:

Omówienie wyników:

* Z pierwszego wykresu wynika, że im mniejszy współczynnik odchylenia na wejściu, tym mniej pętli potrzeba wykonać, aby perceptron wykonywał funkcję logiczną OR. Dzieje się tak, gdyż po modyfikacji współczynnika odchylenia wg wzoru: bias += learnCoefficient \* (expectAnswer - out), perceptron nasz potrzebuje, by bias był wartością ujemną. Co za tym idzie, tym szybciej nasz perceptron nauczy się funkcji, im mniejszy będzie bias.
* Z drugiego wykresu wynika, że im większy współczynnik nauczania, tym mniej pętli potrzeba wykonać, aby perceptron wykonywał funkcję logiczną OR. Dzieje się tak, gdyż do modyfikacji współczynnika odchylenia oraz wag neuronów w podanych wyżej wzorach, wszystkie zmienne, oprócz właśnie współczynnika nauczania osiągają 0 lub 1. Dlatego im większy ten współczynnik tym szybciej nasz perceptron będzie dążył do idealnego ustawienia wag i współczynnika odchylenia
* Z tabelek z wynikami wynika, że nasz perceptron do nauki potrzebuje dwóch wag dodatnich oraz współczynnika odchylenia ujemnego. Co więcej zauważamy, że nie wystarczy, by obie wagi były tylko i wyłącznie dodatnie, ale również ich wartości muszą być większe od wartości bezwzględnej współczynnika odchylenia
* Na wielu stronach internetowych pojawia się informacja, że podanej sieci nie można nauczyć funkcji XOR. Postanowiłem to sprawdzić. Oto wyniki pracy: 

Jak widać nasza funkcja nauki wpadła w nieskończoną pętlę. Potwierdza to przypuszczenia, że perceptronu zbudowanego w ten sposób nie da się nauczyć funkcji XOR.

Wnioski:

* Dla jak najlepszych wyników nauczania powinniśmy dać ujemny współczynnik odchylenia, ale blisko zera i dużą wartość współczynnika nauczania
* Z omówienia wyników wnioskuję, że przy ustawieniu błędu odchylenia na ujemnej wartości w okolicy zera i wylosowaniu wag dodatnich większych od wartości bezwzględniej współczynnika nauczania, nasz perceptron powinien być od razu gotowy do pracy. Oto screen potwierdzający moją tezę:

