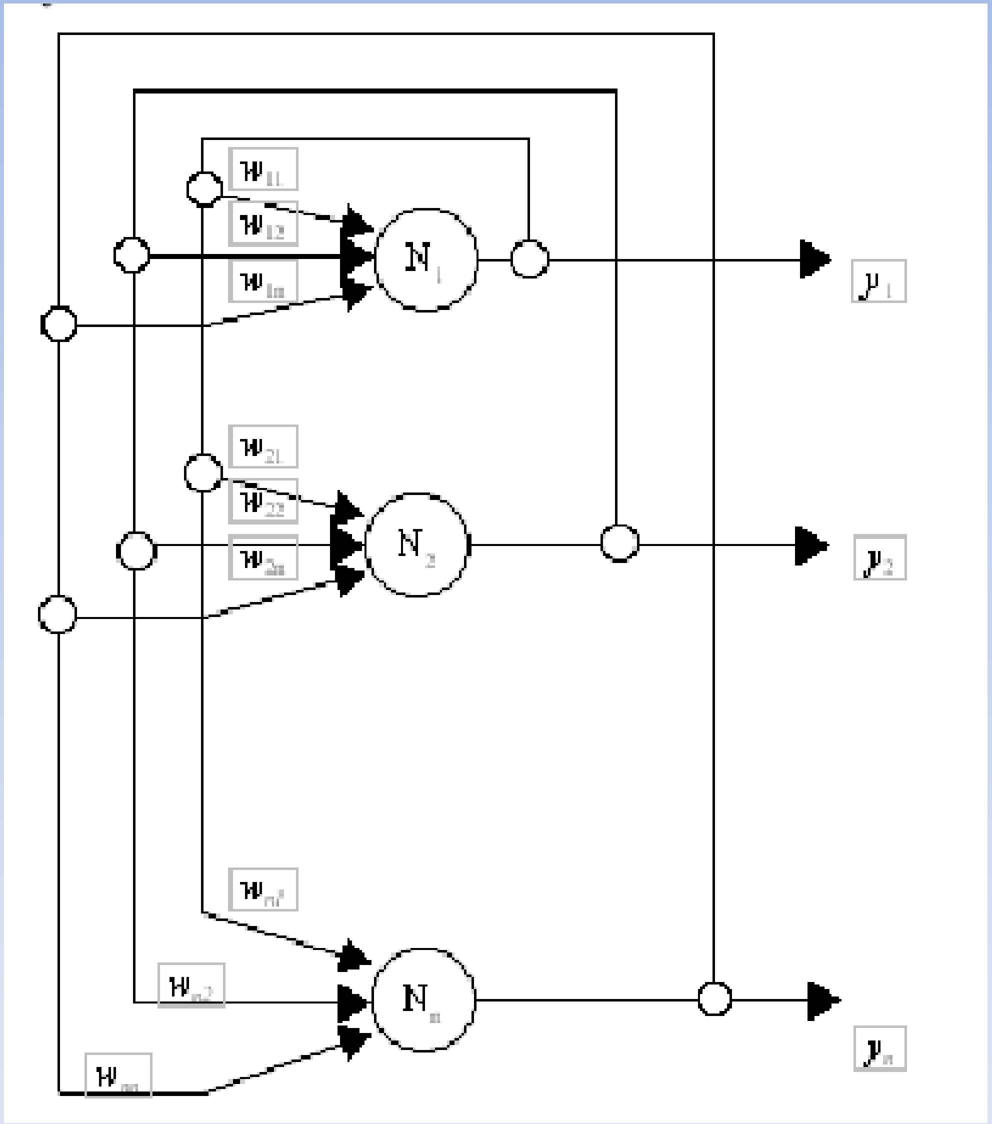
Zastosowanie sieci Hopfielda do rozpoznawania obrazów binarnych

W ramach projektu należy przeprowadzić implementację sieci Hopfielda z uogólnioną regułą uczenia Hebba oraz uwzględnieniem pseudoinwersji macierzy wraz z modułem wirtualizacji uczenia.

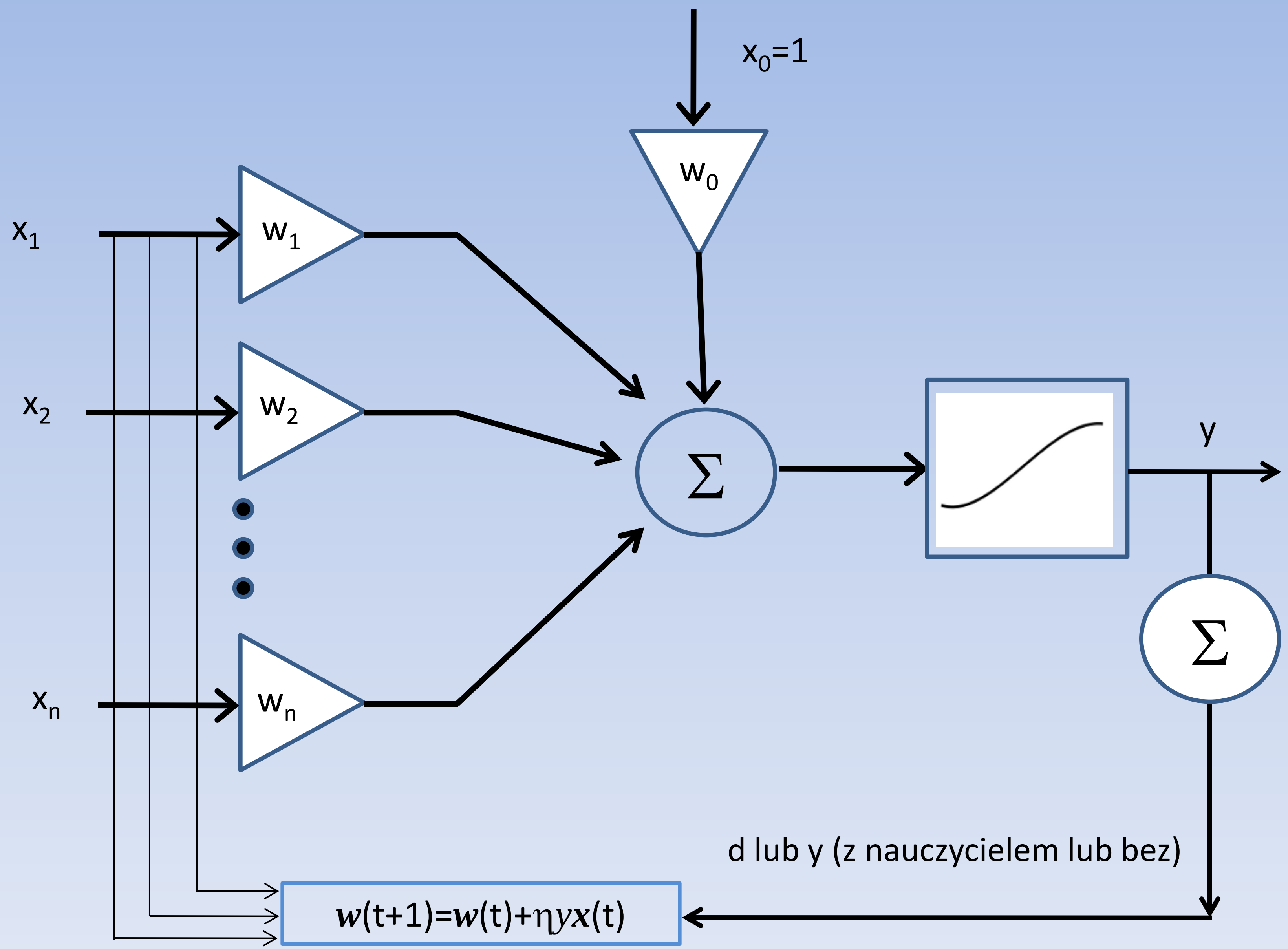
Sieć Hopfielda

Sieć Hopfielda jest przykładem sieci rekurencyjnej. Taka architektura sieci charakteryzuje się sprzężeniem zwrotnym między warstwami sieci. Sieć neuronowa Hopfielda składa się z *n* neuronów, przy czym w danej chwili czasu *t* aktywny jest tylko jeden. Każdy neuron wybierany jest z jednakowym prawdopodobieństwem, a zmiany stanu neuronu następują w dyskretnych chwilach czasu.

Sieć Hopfielda



Model neuronu Hebba



Sieć Hopfielda, zasada działania

Sieci Hopfielda w *trybie uczenia* (czyli doboru wag) jednorazowo przypisuje się odpowiednie wartości (nie stosujemy procedur iteracyjnych jak np. wsteczna propagacja). W *trybie odtworzeniowym* (po nadaniu wag) sygnał wejściowy inicjuje pobudzenie sieci, która dzięki mechanizmowi sprzężenia zwrotnego wielokrotnie przyjmuje na swoje wejście zmodyfikowany sygnał pierwotny, aż do ustabilizowania odpowiedzi. Ta ustabilizowana wartość jest traktowana jako wynik działania sieci.

Ustabilizowana wartość to taka, w której sygnał nie ulega dalszej zmianie.

Dobieranie wag w sieci Hopfielda

Jedną z metod uczenia sieci Hopfielda jest *uogólniona metoda Hebba*. Zgodnie z tą regułą wagi definiowane są następująco

$$w_{ij} = \begin{cases} 0 & \text{dla } i = j, \\ \dfrac{1}{n}\sum_{t=1}^{T} x_i^{(t)} x_j^{(t)} & \text{dla } i \neq j, \end{cases}$$

gdzie

$$x^{(1)}, x^{(2)}, \dots, x^{(T)},$$

jest ciągiem wektorów uczących, a n jest liczbą neuronów w sieci. Taki tryb uczenia sprawia, że wagi przyjmują wartości wynikające z uśrednienia wielu próbek uczących.

Reguła Hebba

W zapisie macierzowym wzory na wagi wg uogólnionej reguły Hebba można zapisać następująco

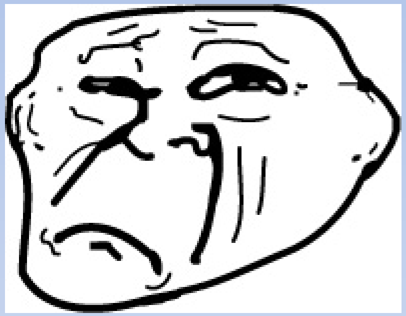
$$W = \frac{1}{n}\left(X \cdot X^{T} - T\, I \right),$$

gdzie I jest macierzą jednostkową , X jest macierzą której kolumnami są wektory ciągu uczącego, X^T oznacza transponowanie macierzy X, stała T jest liczbą wektorów uczących.

Funkcje aktywacji w sieci Hopfielda są dwustanowe. W praktyce używa się funkcji bipolarnej {-1, 1} lub unipolarnej {0, 1}. Podobnie jest z wektorami sygnałów: składają się z współrzędnych bipolarnych lub unipolarnych. Jeżeli stosujemy wektory unipolarne, to uogólniona reguła Hebba przyjmie postać

$$w_{ij} = \begin{cases} 0 & \text{dla } i = j, \\ \frac{1}{n}\sum_{t=1}^{T}(2x_i^{(t)} - 1)(2x_j^{(t)} - 1) & \text{dla } i \neq j, \end{cases}$$

Niestety ta metoda nie jest zbyt skuteczna



Dlatego wykorzystujemy metodę pseudoinwersji

Metoda pseudoinwersji dla sieci Hopfielda

Punktem wyjścia w tej metodzie jest założenie, że przy właściwie dobranych wagach w_{ij} każdy wzorzec $\boldsymbol{x}^{(t)}$ z ciągu uczącego podany na wejście generuje na wyjściu samego siebie, prowadząc zatem do natychmiastowego stanu ustalonego. Mamy więc warunek $W\boldsymbol{x}^{(t)}=\boldsymbol{x}^{(t)}$.

Jeżeli wprowadzimy pomocniczą macierz X, której kolumnami są kolejne wektory uczące

$$X = [x^{(1)}, \ldots, x^{(T)}],$$

to warunki $Wx^{(t)}=x^{(t)}$ można zapisać w postaci następującego równania macierzowego

$$WX = X,$$

gdzie W jest macierzą wag o wymiarach n x n, a X jest macierzą prostokątną o wymiarach n x T.

Metoda pseudoinwersji dla sieci Hopfielda

Rozwiązanie takiego układu (pamiętajmy, że macierz *X jest dana* a szukana jest macierz *W*) jest następującej postaci

$$W = XX^{+},$$

gdzie X^{+} oznacza tzw. *pseudoinwersję* (pseudoodwrotność) macierzy *X*. Jeżeli wektory uczące

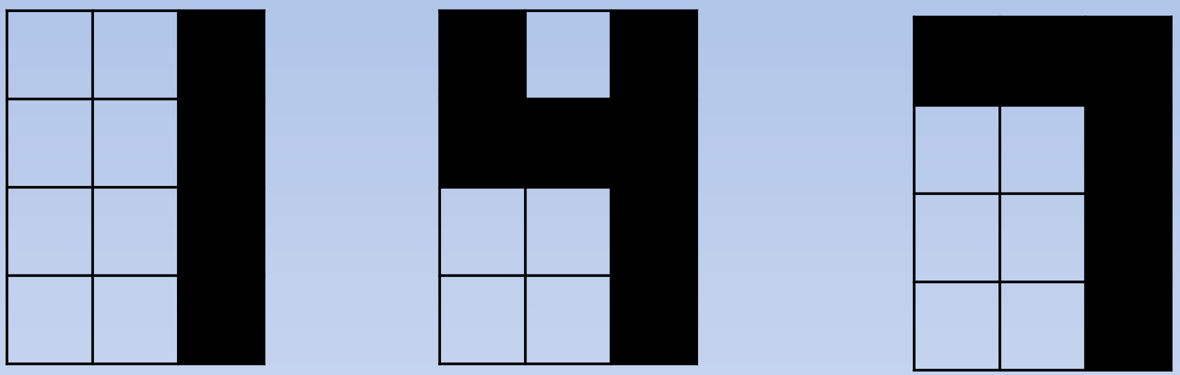
$$x^{(1)}, x^{(2)}, \ldots, x^{(T)}$$

są liniowo niezależne, to macierz *W* może być wyrażona przy pomocy zwykłej odwrotności (nie musimy używać pseudoinwersji):

$$W = X\left(X^{T}X\right)^{-1}X^{T}.$$

Verba docent, exempla trahunt
Przykład

Użyjemy sieci Hopfielda do rozpoznawania trzech cyfr: 1, 4 i 7. Cyfry będą definiowane na matrycy 3x4



Ciąg uczący tworzymy wg kodowania: biały piksel = -1; czarny piksel = +1.

$$x^{(1)} = (-1, -1, +1, -1, -1, +1, -1, -1, +1, -1, -1, +1),$$
$$x^{(2)} = (+1, -1, +1, +1, +1, +1, -1, -1, +1, -1, -1, +1),$$
$$x^{(3)} = (+1, +1, +1, -1, -1, +1, -1, -1, +1, -1, -1, +1).$$

Wektory kodujące cyfry 1, 4, 7

$$x^{(1)} = (-1, -1, +1, -1, -1, +1, -1, -1, +1, -1, -1, +1),$$
$$x^{(2)} = (+1, -1, +1, +1, +1, +1, -1, -1, +1, -1, -1, +1),$$
$$x^{(3)} = (+1, +1, +1, -1, -1, +1, -1, -1, +1, -1, -1, +1).$$

$$x^{(1)} = \begin{bmatrix} -1 \\ -1 \\ +1 \\ -1 \\ -1 \\ +1 \\ -1 \\ -1 \\ +1 \\ -1 \\ -1 \\ +1 \end{bmatrix}, \quad x^{(2)} = \begin{bmatrix} +1 \\ -1 \\ +1 \\ +1 \\ +1 \\ +1 \\ -1 \\ -1 \\ +1 \\ -1 \\ -1 \\ +1 \end{bmatrix}, \quad x^{(3)} = \begin{bmatrix} +1 \\ +1 \\ +1 \\ -1 \\ -1 \\ +1 \\ -1 \\ -1 \\ +1 \\ -1 \\ -1 \\ +1 \end{bmatrix} \quad \rightarrow \quad X = \begin{bmatrix} -1 & 1 & 1 \\ -1 & -1 & 1 \\ 1 & 1 & 1 \\ -1 & 1 & -1 \\ -1 & 1 & -1 \\ 1 & 1 & 1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \\ 1 & 1 & 1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Mnożymy macierz transponowaną X^T przez macierz X i otrzymujemy macierz kwadratową.

$$X^T X = \begin{bmatrix} -1 & -1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & 1 & 1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 1 & 1 \\ -1 & -1 & 1 \\ 1 & 1 & 1 \\ -1 & 1 & -1 \\ -1 & 1 & -1 \\ 1 & 1 & 1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \\ 1 & 1 & 1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 12 & 6 & 8 \\ 6 & 12 & 6 \\ 8 & 6 & 12 \end{bmatrix}.$$

Teraz macierz X^TX musimy odwrócić (jest to możliwe, gdyż wektory $x^{(1)}$, $x^{(2)}$, $x^{(3)}$ są liniowo niezależne).

$$(X^T X)^{-1} = \begin{bmatrix} 12 & 6 & 8 \\ 6 & 12 & 6 \\ 8 & 6 & 12 \end{bmatrix}^{-1} = \begin{bmatrix} 0.1607 & -0.0357 & -0.0893 \\ -0.0357 & 0.1190 & -0.0357 \\ -0.0893 & -0.0357 & 0.1607 \end{bmatrix}$$

$$X(X^TX)^{-1}X^T = \begin{bmatrix} -1 & 1 & 1 \\ -1 & -1 & 1 \\ 1 & 1 & 1 \\ -1 & 1 & -1 \\ -1 & 1 & -1 \\ 1 & 1 & 1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \\ 1 & 1 & 1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0.1607 & -0.0357 & -0.0893 \\ -0.0357 & 0.1190 & -0.0357 \\ -0.0893 & -0.0357 & 0.1607 \end{bmatrix} \begin{bmatrix} -1 & -1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & 1 & 1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 \end{bmatrix}$$

Po wykonaniu mnożeń macierzy uzyskamy ostatecznie macierz wag $W=X(X^TX)^{-1}X^T$ dla przykładowej sieci. Jest to macierz kwadratowa wymiaru 12x12 (liczba wejść do sieci – w przykładzie jest to związane z liczbą pikseli matrycy).

$$W = X(X^T X)^{-1} X^T$$

0,404762	0,190476	0,02381	0,190476	0,404762	0,02381	-0,02381	-0,02381	0,02381	-0,02381	-0,02381	0,02381
0,190476	0,619048	-0,04762	-0,38095	0,190476	-0,04762	0,047619	0,047619	-0,04762	0,047619	0,047619	-0,04762
0,02381	-0,04762	0,119048	-0,04762	0,02381	0,119048	-0,11905	-0,11905	0,119048	-0,11905	-0,11905	0,119048
0,190476	-0,38095	-0,04762	0,619048	0,190476	-0,04762	0,047619	0,047619	-0,04762	0,047619	0,047619	-0,04762
0,404762	0,190476	0,02381	0,190476	0,404762	0,02381	-0,02381	-0,02381	0,02381	-0,02381	-0,02381	0,02381
0,02381	-0,04762	0,119048	-0,04762	0,02381	0,119048	-0,11905	-0,11905	0,119048	-0,11905	-0,11905	0,119048
-0,02381	0,047619	-0,11905	0,047619	-0,02381	-0,11905	0,119048	0,119048	-0,11905	0,119048	0,119048	-0,11905
-0,02381	0,047619	-0,11905	0,047619	-0,02381	-0,11905	0,119048	0,119048	-0,11905	0,119048	0,119048	-0,11905
0,02381	-0,04762	0,119048	-0,04762	0,02381	0,119048	-0,11905	-0,11905	0,119048	-0,11905	-0,11905	0,119048
-0,02381	0,047619	-0,11905	0,047619	-0,02381	-0,11905	0,119048	0,119048	-0,11905	0,119048	0,119048	-0,11905
-0,02381	0,047619	-0,11905	0,047619	-0,02381	-0,11905	0,119048	0,119048	-0,11905	0,119048	0,119048	-0,11905
0,02381	-0,04762	0,119048	-0,04762	0,02381	0,119048	-0,11905	-0,11905	0,119048	-0,11905	-0,11905	0,119048

Dziękuję za uwagę

Bibliografia:

1. Klemens GOLAN, Systemy wspomagania decyzji
2. Radosław MATUSIK, Metody uczenia sieci neuronowej Hopfielda