
**Funkcja XOR;
Dwuwarstwowa sieć jednokierunkowa.**

**Cwiczenia zestaw 2
Termin oddania: 29.11 2012**

➤ **Zaimplementuj dwu wejściową funkcję logiczną XOR wg. przykładu:**

S. Osowski, „Sieci neuronowe do przetwarzania informacji”,
Officyna Wydawnicza Politechniki Warszawskiej 2000, strona 40-42

➤ **Zaimplementuj dwuwarstwową sieć jednokierunkową wg. przykładu 2.1.**

J. Zurada, M. Barski, W. Jędruch, „Sztuczne sieci neuronowe”,
PWN 1996, strona 44-47

➤ **Użyj zaimplementowanej sieci do separacji sygnału i tła z przykładu (1) zestaw 1.**

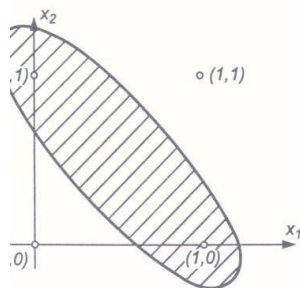
ronu, jeśli należy do klasy liniowo separowalnych (np. problem logiczny $\neg D$, OR).

Ograniczone możliwości sieci jednowarstwowych [101] zostaną zademonstrowane na przykładzie realizacji dwuwęściowej funkcji logicznej XOR. Dla roszczenia przyjęta będzie funkcja aktywacji w postaci skoku jednostkowego. Zbiór danych uczących dla problemu logicznego XOR jest przedstawiony w tablicy 3.1

Tablica 3.1
Zbiór danych uczących dla
problemu XOR

x_1	0	0	1	1
x_2	0	1	0	1
d	0	1	1	0

two pokazać, że w tym przypadku nie można przeprowadzić jednej liniowo separującej przestrzeni danych na dwie podprzestrzenie, z których

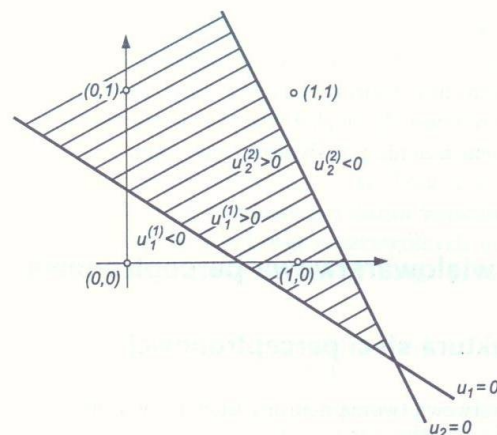


rys. 3.2. Ilustracja nieseparowalności liniowej danych uczących, odwołujących się do problemu logicznego XOR

Jednowarstwowy perceptron nie jest więc w stanie zrealizować nawet tak nieskomplikowanej funkcji, jaką jest XOR. Problem można łatwo rozwiązać, rozszerzając sieć neuronową. W tym celu zwiększamy liczbę neuronów w warstwie o jeden, dobierając wagi obu neuronów w ten sposób, aby realizowały podział obszaru na części:

$$u_1 = w_{11}x_1 + w_{12}x_2 + w_{10} > 0 \text{ i } u_1 = w_{11}x_1 + w_{12}x_2 + w_{10} < 0$$

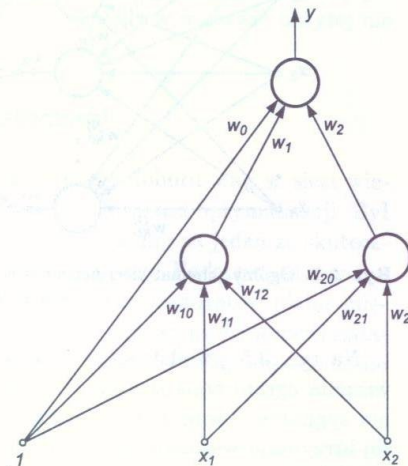
w zależności od wektora wejściowego $\mathbf{x} = [x_1, x_2]^T$ (pierwszy neuron) oraz $u_2 = w_{21}x_1 + w_{22}x_2 + w_{20} > 0$ i $u_2 = w_{21}x_1 + w_{22}x_2 + w_{20} < 0$ (drugi neuron). Dobór wag powinien zapewnić podzielenie obszaru takie, jak to pokazano na rys. 3.3.



Rys. 3.3. Rozwiązanie problemu nieliniowej separowalności przez złożenie dwóch separowalności liniowych

Dokładając na wyjściu dodatkową warstwę zawierającą jeden neuron, można zrealizować sumę logiczną odpowiadającą części wspólnej

zbiorów $u_1 > 0, u_2 > 0$. Ostateczną postać sieci neuronowej realizującej funkcję XOR przedstawiono na rys. 3.4. Warto zauważyć, że wprowadzenie dodatkowej warstwy ukrytej sieci pozwoliło rozwiązać problem należący do klasy zadań nieseparowalnych liniowo. Każdy neuron w warstwie ukrytej wprowadza dodatkowy podział liniowy obszaru, przy czym granica podziału na część $u_i > 0$ i $u_i < 0$ jest uzależniona od wartości wag neuronu. Warstwa wyjściowa dokonuje wyboru odpowiedniej kombinacji liniowej (np. sumy logicznej) podobszarów, na które został podzielony obszar danych wejściowych przez neurony warstwy ukrytej.



Rys. 3.4. Struktura sieci neuronowej rozwiązującej problem XOR

do rozwiązania określonego problemu. Wybór architektury takiej sieci jest prosty. Liczbę neuronów wejściowych określa wymiar wektora wejściowego x , a liczba neuronów wyjściowych jest zadawana przez wymiar wektora d . Uczenie sieci odbywa się zwykle w trybie z nauczycielem i jest dokładnym odpowiednikiem uczenia pojedynczego neuronu.

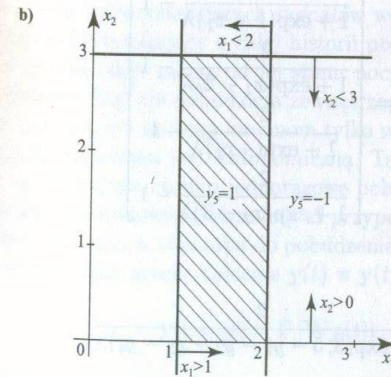
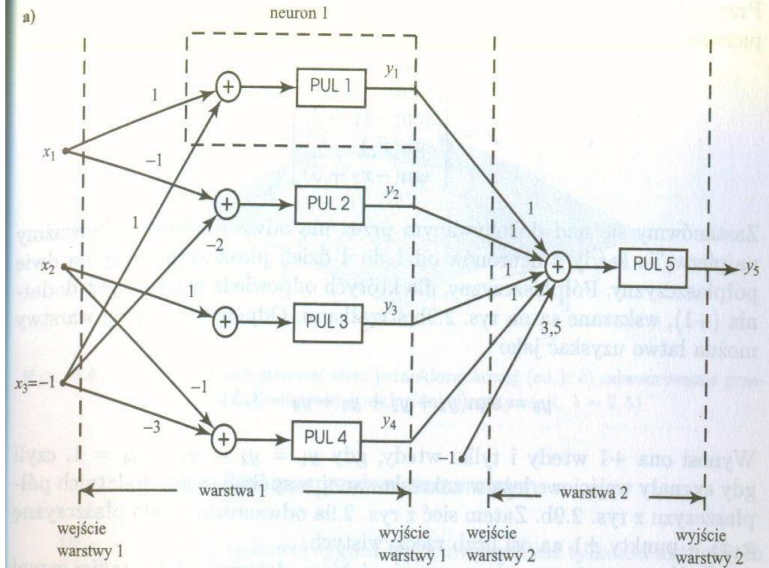
Przykład 2.1

Przeanalizujemy tu dwuwarstwową sieć jednokierunkową, złożoną z neuronów o binarnej funkcji aktywacji, jak w (2.3b), i pokazaną na rys. 2.9. Znajdziemy sygnał wyjściowy y_5 dla zadanej sieci i danego obrazu wejściowego. Każda z warstw sieci opisana jest równaniem (2.12a), przy czym dla pierwszej warstwy:

$$y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ -1 \end{bmatrix}, \quad W_1 = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & -1 & -3 \end{bmatrix},$$

dla drugiej zaś:

$$y = [y_5], \quad x = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}, \quad W_2 = [1 \ 1 \ 1 \ 1 \ 3,5].$$



Rys. 2.9. Przykład dwuwarstwowej sieci jednokierunkowej: a) schemat, b) odwzorowanie przestrzeni dwuwymiarowej (dyskretna funkcja aktywacji)

Przy założeniu bipolarnych binarnych funkcji aktywacji wektor wyjściowy pierwszej warstwy obliczamy jako

$$\mathbf{y} = \begin{bmatrix} \operatorname{sgn}(x_1 - 1) \\ \operatorname{sgn}(-x_1 + 2) \\ \operatorname{sgn}(x_2) \\ \operatorname{sgn}(-x_2 + 3) \end{bmatrix}$$

Zastanówmy się nad dokonywanym przez nią odwzorowaniem. Zauważmy najpierw, że każdy z neuronów od 1 do 4 dzieli płaszczyznę x_1x_2 na dwie półpłaszczyzny. Półpłaszczyzny, dla których odpowiedź neuronu jest dodatnia (+1), wskazane są na rys. 2.9b strzałkami. Odpowiedź drugiej warstwy można łatwo uzyskać jako

$$y_5 = \operatorname{sgn}(y_1 + y_2 + y_3 + y_4 - 3, 5).$$

Wynosi ona +1 wtedy i tylko wtedy, gdy $y_1 = y_2 = y_3 = y_4 = 1$, czyli gdy sygnały wejściowe leżą w zakreślanej, wspólnej części dodatnich półpłaszczyzn z rys. 2.9b. Zatem sieć z rys. 2.9a odwzorowuje całą płaszczyznę x_1x_2 w punkty ± 1 na osi liczb rzeczywistych.

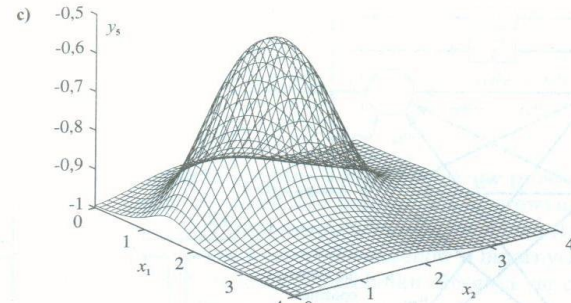
Spójrzmy teraz na odwzorowanie, jakiego dokonuje sieć o takiej samej architekturze, lecz z neuronami o charakterystykach sigmoidalnych. Biorąc ciągłą, bipolarną funkcję aktywacji typu (2.3a), uzyskujemy dla pierwszej warstwy

$$\mathbf{y} = \begin{bmatrix} \frac{2}{1 + \exp(1 - x_1)\lambda} - 1 \\ \frac{2}{1 + \exp(x_1 - 2)\lambda} - 1 \\ \frac{2}{1 + \exp(-x_2)\lambda} - 1 \\ \frac{2}{1 + \exp(x_2 - 3)\lambda} - 1 \end{bmatrix}$$

i dla drugiej warstwy

$$y_5 = \frac{2}{1 + \exp(3, 5 - y_1 - y_2 - y_3 - y_4)\lambda} - 1.$$

Sieć z neuronami ciągłymi realizuje odwzorowanie pokazane na rys. 2.9c. Jak widać, choć istnieje pewne podobieństwo do działania neuronu dyskretnego, to tym razem odwzorowanie jest znacznie bardziej złożone. Przypomina ono całą płaszczyznę x_1x_2 przedział $(-1, 1)$. Z podobnym odwzorowaniem mieliśmy już do czynienia wcześniej, omawiając zagadnienie aproksymacji funkcji (rys. 1.3).



Rys. 2.9. Przykład dwuwarstwowej sieci jednokierunkowej (cd.): c) odwzorowanie przestrzeni dwuwymiarowej (ciągła funkcja aktywacji, $\lambda = 2, 5$)