

Sztuczne Sieci Neuronowe

Wykład 8

Sieci rezonansowe

wykład przygotowany na podstawie.

R. Tadeusiewicz, "Sieci Neuronowe", Rozdz. 6. Akademicka Oficyna Wydawnicza RM,
Warszawa 1993.

Wprowadzenie

Sieci wielowarstwowe sigmoidalne tworzą rodzinę sieci neuronowych jednokierunkowych, w których przepływ sygnałów odbywa się tylko w jedną stronę, od wejścia do wyjścia.

Oddzielną grupę stanowią sieci, w których istnieje sprzężenie zwrotne między wyjściem i wejściem

Są to tzw. **sieci rekurencyjne (rezonansowe)**.

Wspólną cechą tych sieci jest często istnienie symetrycznych powiązań synaptycznych między neuronami ($W_{ij} = W_{ji}$) oraz duża ich liczba w stosunku do liczby neuronów.

Wprowadzenie

Podstawową cechą wyróżniającą je z rodziny sieci neuronowych są **zależności dynamiczne** na każdym etapie działania.

Zmiana stanu jednego neuronu przenosi się przez masowe sprzężenie zwrotne na całą sieć, wywołując stan przejściowy, kończący się określonym stanem ustalonym, na ogół innym niż stan poprzedni.

Jeżeli funkcję aktywacji neuronu oznaczy się tak jak poprzednio poprzez $\phi(\mathbf{e})$, przy czym \mathbf{e} jest sumą wagową pobudzeń, zwaną również sygnałem aktywacji, to sygnał wyjściowy neuronu

$$y_i = \phi(e_i) = \phi \left(\sum W_{ij} x_j \right)$$

oznacza jego stan.

Wprowadzenie

Biorąc pod uwagę, że przy masowym sprzężeniu zwrotnym pobudzeniami dla neuronu są sygnały wejściowe innych neuronów, zmiana stanu neuronów jest opisana układem równań różniczkowych nieliniowych

$$\tau_i \frac{de_i}{dt} = \sum W_{ij} \phi(e_j) - e_i - b_i$$

Dla $i=1,2,\dots,N$, przy czym b_i jest wartością progową, wynikającą z zewnętrznego źródła. Współczynnik τ_i jest pewną stałą liczbową, a jego interpretacja jest analogiczna do stałej czasowej w równaniach opisujących stan dynamiczny w obwodach RC. Stan neuronu uzyskuje się z rozwiązania równania różniczkowego jako $y=\phi(e)$.

Podstawowy schemat działania sieci

Jedną z bardziej znanych zasad budowy rezonansowych sieci neuronowych jest

Adaptive Resonance Theory

Teoria sieci ART oparta jest na pracach Grossberga, a konkretnie metody wywodzące się z tej teorii są chronione patentem (Univ. Boston). W literaturze wyróżnia się sieć ART1 (do analizy obrazów binarnych) i ART2 (do analizy obrazów analogowych).

Każda z tych sieci składa się **z dwóch jednakowo liczących warstw** neuronów – pierwszej (dolnej) rejestrującej sygnały wejściowe (cechy rozpoznawania obrazów) i drugiej analizującej te cechy na zasadzie budowy ważonej sumy wejść i wypracowywującej decyzje na zasadzie rywalizacji pomiędzy neuronami.

Nie buduje się sieci ART o większej liczbie warstw.

Podstawowy schemat działania sieci

Liczba neuronów pierwszej (dolnej) warstwy uzależniona jest od liczby cech wyróżnionych w rozpoznawanych obiektach. Np. informujące o jasności obrazu w określonym punkcie. Sygnał wejściowy może być więc interpretowany jako swojego rodzaju siatkówka, binarny raster o wymiarach $n_1 \times n_2$. Przy takim założeniu liczba neuronów $n = n_1 n_2$.

Liczba neuronów drugiej warstwy jest w zasadzie dowolna, ustala się ją biorąc pod uwagę zadanie realizowane przez sieć.

Sieć ART jest siecią klasyfikującą bodźce lub rozpoznającą obrazy. Jej uczenie zachodzi bez nauczyciela, więc ilość wyjść musi być większa od liczby przewidywanych klas. Przykładowo, do rozpoznawania dużych liter alfabetu (26 klas) należy zbudować sieć z około 40 neuronów w warstwie wyjściowej.

Podstawowy schemat działania sieci

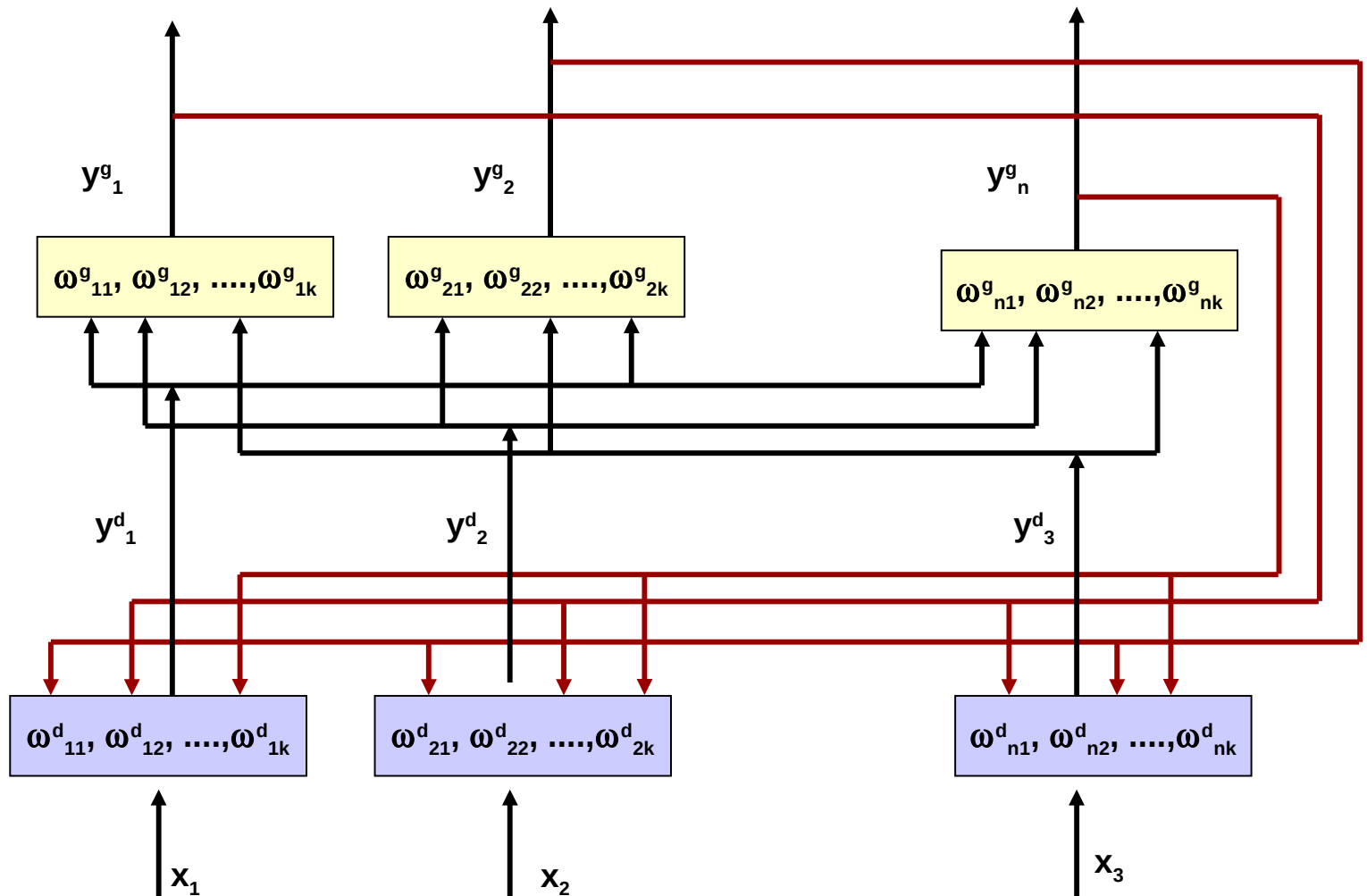
W górnej warstwie tylko jeden neuron ma sygnał wyjściowy, różny od zera. Jest to ten którego suma ważonych sygnałów wejściowych osiągnęła największą wartość.

Taki schemat działania odpowiada omówionym już koncepcjom sieci MADALINE czy PERCEPTON. Element rywalizacji wystąpił w koncepcjach sieci Kohonena.

Na czy polega nowość?...

Wadą sieci MADALINE czy PERCEPTON było **niestabilne zachowanie**. W trakcie procesu uczenia neuron już “nauczony” na jeden sygnał może zostać “przestawiony” i zacząć być czuły na inny. Aby temu zapobiec w sieci ART wprowadzono **sprzężenie zwrotne**. Zwycięski neuron z warstwy wyjściowej może dzięki nim wzmacniać sygnały w tych neuronach warstwy wejściowej od których otrzymał pobudzenie.

Schemat sieci ART



Uczenie sieci ART

Proces uczenia sieci ART definiowany jest prawem zwanym **prawem Webera** i zaprezentujemy go w formie iteracyjnej.

Wprowadźmy najpierw oznaczenia:

⇒ Zakładamy że wszystkie sygnały są binarne

tzn. $x_i \in \{0,1\}$, $y_i^d \in \{0,1\}$ i $y_i^g \in \{0,1\}$

⇒ Definiujemy współczynniki wagowe przy odpowiednich połączeniach

ω_{ij}^ξ , jako współczynnik połączenia między neuronami i, j ,

i = index neuronu **w warstwie pierwszej**,

j = index neuronu **w warstwie drugiej**

w sieci ART neurony są numerowane oddzielnie w każdej warstwie.

Jeżeli $\xi = g$ to odpowiednia waga znajduje się przy połączeniu wiodącym do góry (czyli na drodze między warstwą pierwszą a drugą), natomiast $\xi = d$ oznacza że współczynnik znajduje się przy połączeniu w dół, czyli połączeniu sprzężenia zwrotnego (między warstwą drugą a pierwszą) w szczególności ω_{ii}^ξ oznacza połączenie neuronów o tych samych numerach a nie sprzężenie zwrotne neuronu z samym sobą.

Uczenie sieci ART

Reguła dla wag warstwy dolnej:

$$\omega_{ij}^{d(k+1)} = \begin{cases} 1 & \text{gdy } y_i^{d(k)} = 1 \text{ i } y_j^{g(k)} = 1 \\ \omega_{ij}^{d(k)} & \text{gdy } y_j^{g(k)} = 0 \\ 0 & \text{gdy } y_i^{d(k)} = 0 \text{ i } y_j^{g(k)} = 1 \end{cases}$$

gdzie index k jest numerem kroku, y_j^g oznacza sygnał wyjściowy warstwy górnej, y_j^d oznacza sygnał wyjściowy warstwy dolnej.

Warto zauważyć, że podana reguła uczenia nie jest regułą typu Hebba ze względu na **czynnik zapominania** (ostatni wiersz we wzorze).

Z powodu rywalizacji w warstwie wyjściowej jedynie nieliczne współczynniki wagowe podlegają uczeniu.

Wartości początkowe nadane współczynnikom wagowym **nie są losowe, są jednakowe** i wynoszą $\omega_{ij}^{d(1)} = 1$.

Uczenie sieci ART

Reguła dla wag warstwy górnej:

$$\omega_{ij}^{g(k+1)} = \begin{cases} \omega_{ij}^{g(k)} + \delta_1 & \text{gdy } y_i^{d(k)} = 1 \text{ i } y_j^{g(k)} = 1 \\ \omega_{ij}^{g(k)} & \text{gdy } y_j^{g(k)} = 0 \\ \omega_{ij}^{g(k)} - \delta_2 & \text{gdy } y_i^{d(k)} = 0 \text{ i } y_j^{g(k)} = 1 \end{cases}$$

Wprowadza ona mniej radykalne zmiany w wartościach wag. Przyrosty δ_1 , δ_2 nadawane współczynników wagowym mogą być stałe, ale lepsze wyniki uczenia uzyskuje się przyjmując te wartości jako zależne od liczby pobudzonych w pierwszej warstwie neuronów. Liczbę tę można wyrazić jako $\mathbf{p}^{(k)} = \sum_{i=1}^n y_i^{d(k)}$ ponieważ $y_i^{d(k)}$ są binarne.

Mając wyznaczoną liczbę $\mathbf{p}^{(k)}$ można ustalać δ_1 , δ_2 ze wzorów

$$\delta_1 = \eta (1 - \omega_{ij}^{g(k)}) - \omega_{ij}^{g(k)} (\mathbf{p}^{(k)} - 1), \quad \delta_2 = \omega_{ij}^{g(k)} \mathbf{p}^{(k)}$$

parametr η jest dobierany empirycznie.

Uczenie sieci ART

Dobór parametru η :

⇒ wartości $\eta > 1$ i większe sprzyjają tworzeniu przez sieć nowych wzorców kategorii

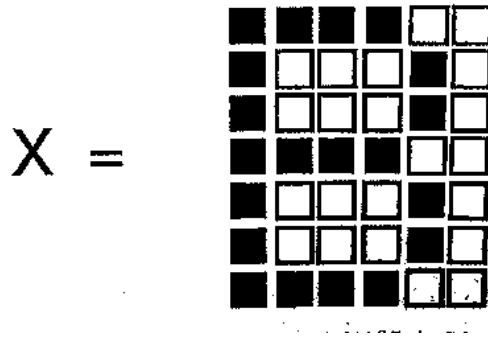
⇒ wartości $\eta < 1$ skłaniają sieć do uogólniania kategorii już wcześniej wypracowanych

Wartości początkowe przyjmowane dla wag $\omega_{ij}^{g(1)}$ przyjmowane są losowo z przedziału $[0, 1/n]$.

W wyniku tak radykalnej procedury adaptacji wag, możliwe jest dopasowywanie sieci do wielu wzorców różniących się od siebie, a ponadto sieć wykazuje pewną (ograniczoną) zdolność do uśredniania sygnałów wejściowych i tworzenia uogólnionych wzorców rozpoznawanych obiektów. Ta własność sieci **ART** jest podobna do analogicznych cech wykazywanych przez sieć **Madaline**, **Perceptron** i **Backpropagation**.

Zasada działania sieci ART

Rozważmy przykład sieci rozpoznającej litery. Niech sygnał wejściowy podany do neuronu ma postać

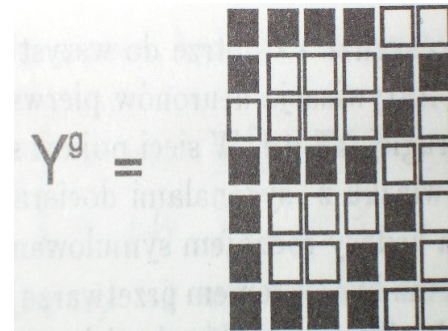
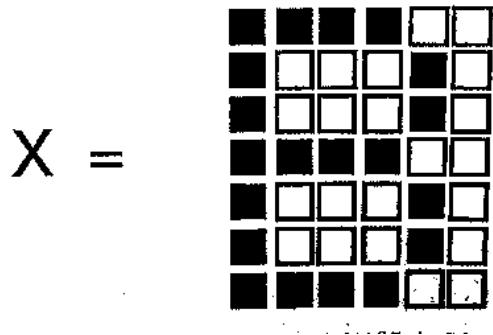


Zaciemnione elementy rastra wskazują elementy przesyłające sygnał wejściowy $x_i = 1$, puste elementy odpowiadają wartościom $x_i = 0$.

W wyniku pobudzenia pewien neuron zostaje pobudzony co można uznać za sukces (poprawne rozwiązanie). Sieć funkcjonuje bez nauczyciela nie jest więc możliwe wskazanie “a priori” tego neuronu. Po pierwszym skojarzeniu pewnego neuronu z pewnym sygnałem wejściowym sieć utrzymuje to skojarzenie w sposób stabilny i zdeterminowany.

Początkowo odwzorowanie może nie być całkowicie poprawne.

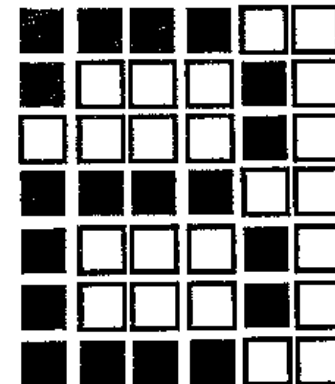
Zasada działania sieci ART



Wówczas obszar, w którym spotykają się sygnały wyjściowe Y^g z wejściowymi X ulega wzmocnieniu na zasadzie teorii rezonansu.

Obszar ten ma w omawianym przykładzie postać

$$Y^g \cap X =$$



I jest lepiej dopasowany do realizacji postawionego przed siecią zadania klasyfikacji od pierwszego obrazu pobudzeń “zwycięskiego” neuronu.

Zasada działania sieci ART

W pierwszym przybliżeniu można przyjąć że obszar $Y^g \cap X$ odpowiada obszarowi, w którym sygnały pierwszej warstwy są niezerowe, a więc rozkład pobudzeń pierwszej warstwy można zapisać

$$Y^d = Y^g \cap X$$

Działanie mechanizmu uczenia może w pewnym przybliżeniu być więc opisane następująco:

- ⇒ modyfikacji podlegają wagi połączeń ω_{iz}^g dochodzących do zwycięskiego elementu warstwy wejściowej
- ⇒ modyfikacji podlegają wagi połączeń ω_{iz}^d biorących początek w zwycięskim neuronie.
- ⇒ wszystkie pozostałe połączenia pozostają niezmiennione.

Modyfikacja wag polega (z grubsza) na tym, że wagi ω_{iz}^d odpowiadające połączeniom z pobudzonymi elementami pierwszej warstwy ($Y^d = Y^g \cap X$) przyjmują wartości 1, a wszystkie pozostałe 0. Jest to dość “drastyczna” modyfikacja. Podobnie modyfikowane są wagi ω_{iz}^g . Ich wartości są zerowane dla wszystkich niepobudzonych elementów warstwy wejściowej, względnie są ustawiane na pewną wartość zależną od liczby pobudzonych elementów w przypadku połączeń biorących początek w tych pobudzonych neuronach.

Rola i struktura układu kontrolnego

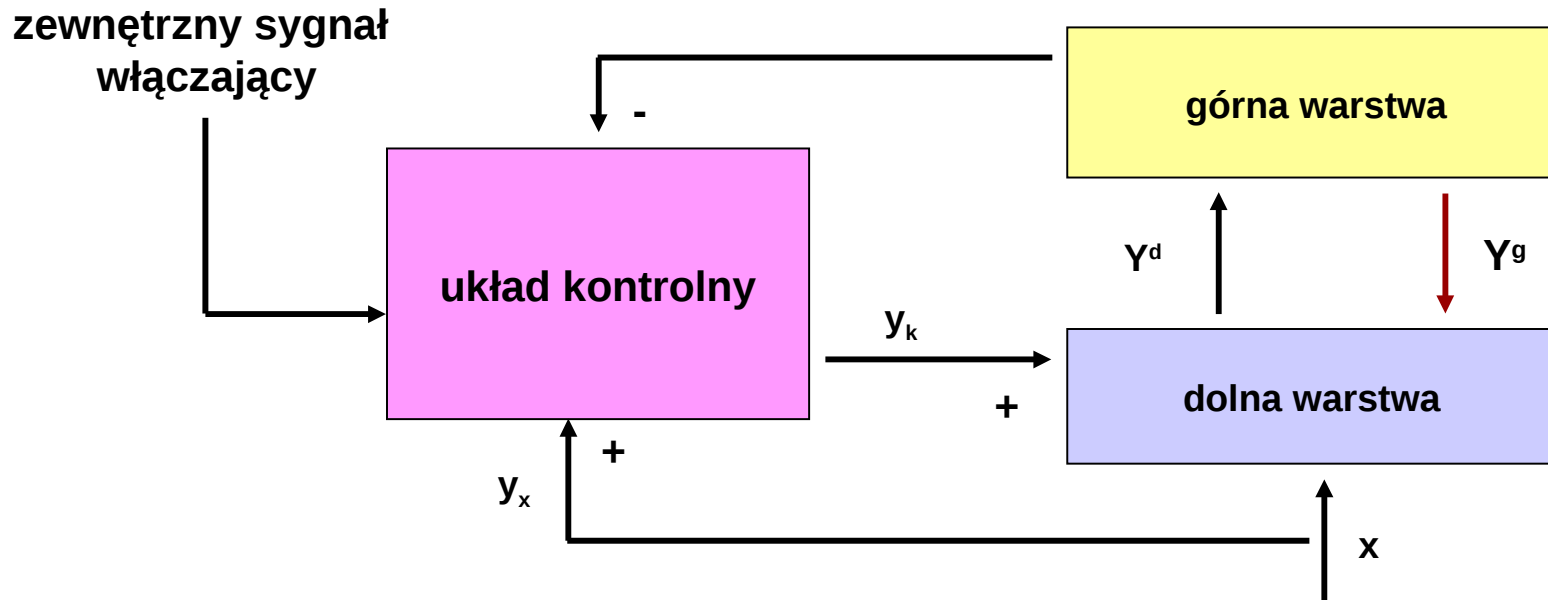
Może zaistnieć taka sytuacja, kiedy na wejściu sieci nie ma żadnego sygnału ($X=0$), ale na którymś z wyjść pojawi się na chwile sygnał $y_j^g \neq 0$. Na skutek sprzężenia zwrotnego sygnał Y^g dotrze do wszystkich neuronów pierwszej warstwy, wywołując ich aktywizację. Aktywizacja neuronów pierwszej warstwy powoduje pobudzenie neuronów warstwy drugiej Y^g .

W sieci pojawi się *złożony proces dynamiczny*, nie mający literalnie żadnego związku z sygnałami docierającymi do sieci z zewnątrz i z jej podstawowymi zadaniami, a będący rodzajem symulowanej “halucynacji”.

W sieci, która ma być systemem przetwarzającym informacje w sposób powtarzalny i ukierunkowany na określony cel, takie zjawiska są niedopuszczalne.

Dlatego elementem sieci ART jest zawsze *układ kontrolny* (attentional gain control unit), który uniemożliwia pierwszej warstwie sieci reagowanie na sygnały sprzężenia zwrotnego w przypadku braku sygnałów wejściowych.

Układ kontrolny sieci ART



Działanie układu kontrolnego polega na tym, że jego sygnał dodatkowo pobudza (albo zwiększa czułość) neuronów dolnej warstwy. Bez tego dodatkowego sygnału neurony te nie są w stanie reagować na sygnały górnej warstwy Y^g i pętla “halucynacji” ulega przerwaniu.

Układ kontrolny sieci ART

Realizacja układu kontrolnego zazwyczaj oznacza dobudowanie do sieci dodatkowego układu złożonego z trzech neuronów.

Pierwszy z neuronów nazywa się detektorem wejścia i bada stan wszystkich wejść X . Jego działanie można opisać jako *alternatywę* sygnałów wejściowych:

$$y_x = \begin{cases} 1 & \text{gdy } \sum_{i=1}^n x_i > 0 \\ 0 & \text{gdy } \sum_{i=1}^n x_i \leq 0 \end{cases}$$

W podobny sposób działa drugi z neuronów, badający aktywność Y^g drugiej warstwy neuronów.

$$y_g = \begin{cases} 1 & \text{gdy } \sum_{i=1}^n y_i^g > 0 \\ 0 & \text{gdy } \sum_{i=1}^n y_i^g \leq 0 \end{cases}$$

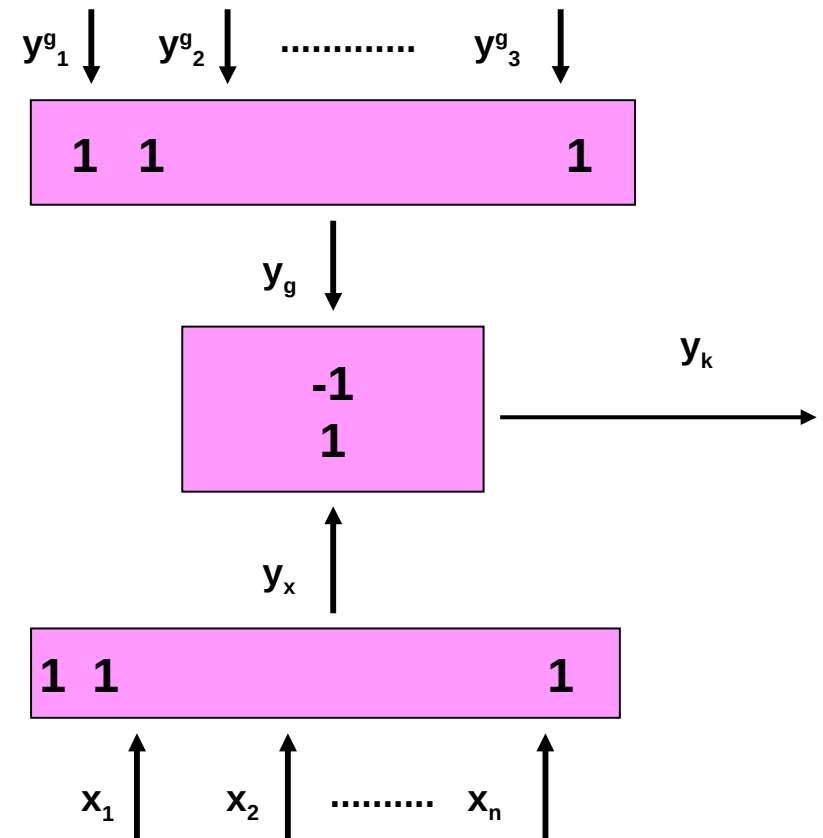
Progową, nieliniową charakterystykę ma też trzeci neuron tworzący sygnał wejściowy y_k układu kontrolnego, z tym jednak, że ma on zaledwie dwa wyjścia.

$$y_k = \begin{cases} 1 & \text{gdy } y_x - y_g > 0 \\ 0 & \text{gdy } y_x - y_g \leq 0 \end{cases}$$

Układ kontrolny sieci ART

Strukturę układu kontrolnego można przedstawić następująco:

Sygnal y_k doprowadzony jest do wszystkich elementów dolnej (wejściowej) warstwy sieci i współuczestniczy w kształtowaniu ich sygnałów.



Układ kontrolny sieci ART

Sygnal y_k doprowadzony jest do wszystkich elementów dolnej (wejściowej) warstwy sieci i współuczestniczy w kształtowaniu ich sygnałów wyjściowych zgodnie z równaniem

$$y_i^d = \begin{cases} 1 & \text{gdy } x_i + \sum_{j=1}^k \omega_{ij}^d y_j^g + y_k > \Theta \\ 0 & \text{gdy } x_i + \sum_{j=1}^k \omega_{ij}^d y_j^g + y_k \leq \Theta \end{cases}$$

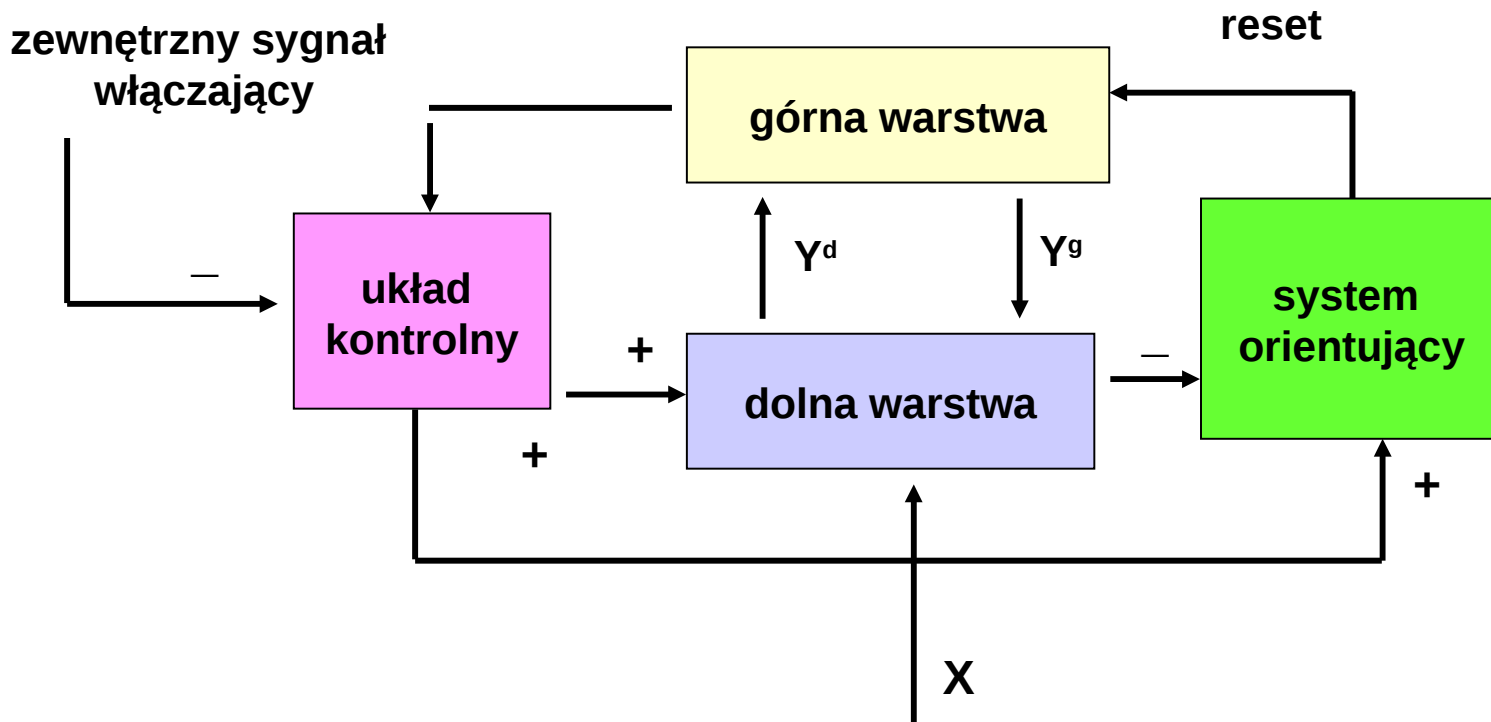
Równanie to odwzorowuje strukturę połączeń elementów sieci, jednak z punktu widzenia aktywnych sygnałów występujących w niej podane wyżej równanie jest całkowicie równoważne prostszej zależności

$$y_i^d = \begin{cases} 1 & \text{gdy } x_i + \omega_{iz}^d y_z^g + y_k > \Theta \\ 0 & \text{gdy } x_i + \omega_{iz}^d y_z^g + y_k \leq \Theta \end{cases}$$

biorąc pod uwagę że $y_i^g = 0$ dla wszystkich $j \neq z$. Prog Θ przyjmuje się tak ($\Theta = 1.5$), aby zadziałała reguła “dwa z trzech”, tzn. do zadziałania neuronu potrzebna jest obecność dwóch spośród trzech sygnałów wejściowych.

Układ orientujący

Innym oryginalnym elementem sieci ART jest system *orientujący* (orienting system), którego celem jest *sterowanie precyzją odwzorowania* poszczególnych kategorii w sieci ART.



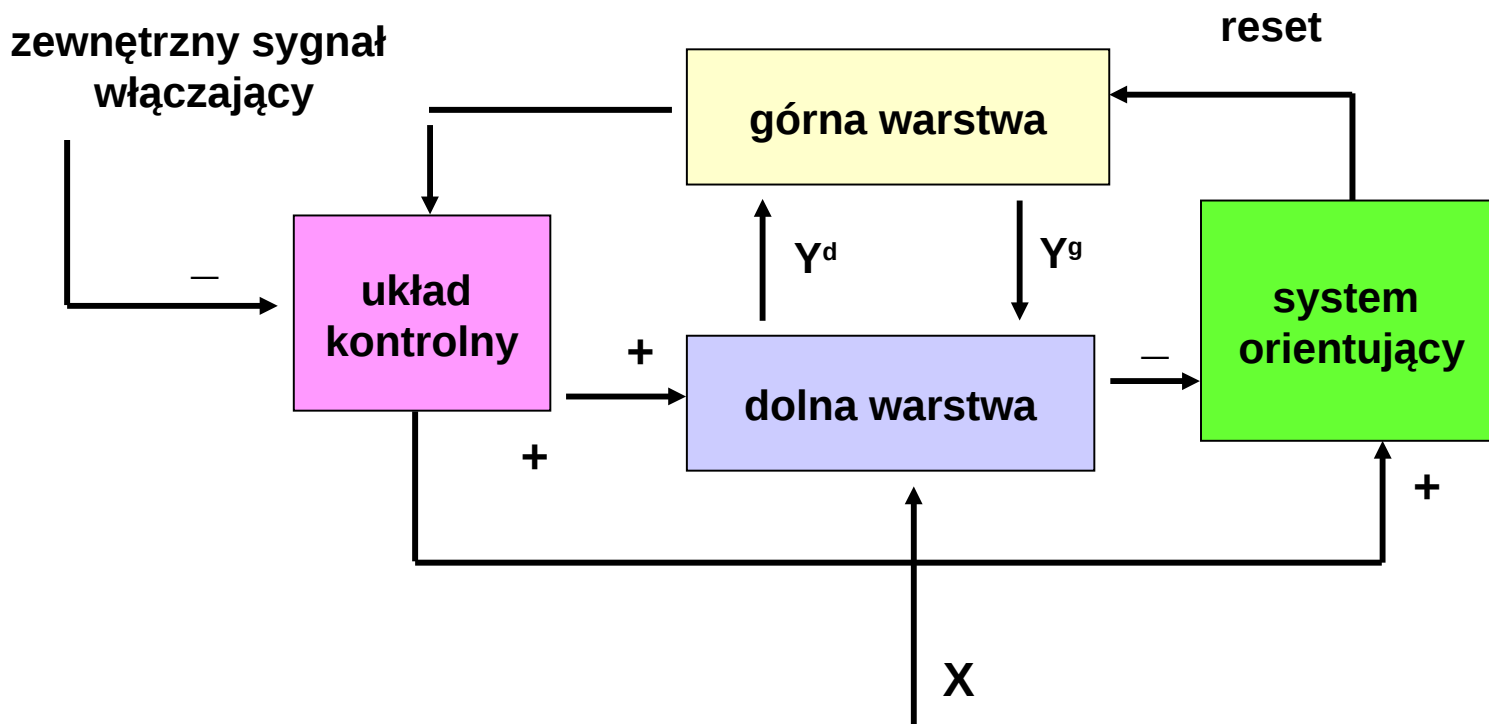
Układ orientujący

System orientujący jest pobudzany przez sygnał wejściowy i hamowany przez sygnał globalnej aktywności dolnej warstwy sieci. System ten pozostaje nieaktywny w przypadku gdy sygnałowi wejściowemu X towarzyszy reakcja w postaci pobudzenia dolnej warstwy sieci Y^d , gdyż wtedy sieć po prostu rozpoznaje jeden w wcześniej zapamiętanych wzorców wejściowego sygnału i żadne dodatkowe działania nie są potrzebne.

Założmy, że na wejściu pojawił się nowy sygnał X. Sieć nie zna go, zatem nie dochodzi do *rezonansu* i pobudzenie dolnej warstwy jest niewielkie. Jest to przedstawiciel *nowej kategorii* sygnałów które także powinny być w sieci zapamiętane, zatem trzeba spowodować jego rejestrację w postaci odpowiedniej reprezentacji w górnej warstwie. przy czym powinna to być reprezentacja *inna* niż w przypadku wszystkich wcześniej zapamiętanych bodźców.

Układ orientujący

Połączenia warstw górnej i dolnej spowodują, że nowy bodziec wejściowy pobudzi pewne neurony w górnej warstwie, a te przez sprzężenia zwrotne będą oddziaływały na neurony warstwy dolnej.



Układ orientujący

Brak korelacji pomiędzy uprzednio zapamiętanymi wzorcami sygnałów a nowym sygnałem jest powodem, że wynikowe pobudzenie warstwy dolnej zmaleje – nie pojawi się efekt rezonansu. Ta sytuacja aktywizuje **system orientujący**. Wysyła on do neuronów warstwy wyjściowej (górnej) krótkotrwały sygnał blokujący (reset wave), który powoduje, że wszystkie aktywne neurony warstwy górnej zostają zablokowane.

Wyłącza to w praktyce oddziaływanie sygnału Y^g na dolną warstwę i powoduje, że rozkład pobudzeń na dolnej warstwie może się kształtować w sposób swobodny, bez wprowadzających deformacji oddziaływań zwrotnych.

Prowadzi to do wzrostu poziomu sygnału Y^d i silnego pobudzenia tych elementów warstwy górnej które nie podlegały blokadzie; jeden z tych sygnałów staje się “zwycięzcą” i neuron odpowiadający temu sygnałowi staje się detektorem nowej klasy. Następuje korekta wag połączeń synaptycznych i sieć nabiera zdolności rozpoznawania nowej klasy.

Układ orientujący

System orientujący pełni rolę *zastępczego nauczyciela*. Uczenie sieci ma charakter uczenia “bez nauczyciela”, w ciągu uczącym brak jawnie zadawanych wzorców sygnałów wyjściowych z sieci. Zachowanie się nie jest przypadkowe gdyż system orientujący klasyfikuje pojawiające się wzorce wejściowych sygnałów na “znane” i “nowe”, dzięki czemu poszczególne neurony warstwy wyjściowej nabywają zdolności do rozpoznawania i sygnalizowania różnych wzorców. Działanie systemu orientującego zależne jest od *parametru gotowości v* (vigilance) określającego, jak duża musi być niezgodność między sygnałem wejściowym X , a sygnałem zwrotnym Y^g aby można mówić o braku rezonansu czyli aby dochodziło do aktywacji sygnału blokowania (reset).

⇒ Jeżeli v jest duży, małe niezgodności są tolerowane i łatwo dochodzi do wykrycia nowej kategorii.

⇒ Małe wartości v czynią sieć bardzo skłoną do uśredniania wejściowych sygnałów i generacji bardziej “zagregowanych” reprezentacji.

Wybór konkretnej wartości v jest kwestią swobodnej decyzji użytkownika.

Formalny opis systemu orientującego

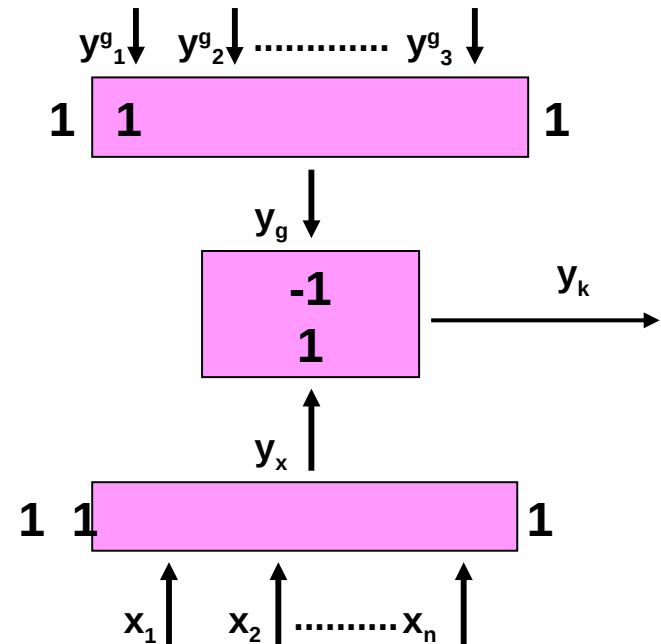
Warunek na generowanie sygnału blokującego można zapisać jako :

$$v \sum_{i=1}^n x_i > \sum_{i=1}^n y_i^d$$

i opisać działanie **systemu orientującego** wzorem.

$$r = \begin{cases} 1 & \text{gdy } v \sum x_i > \sum y_i^d \\ 0 & \text{gdy } v \sum x_i \leq \sum y_i^d \end{cases}$$

System ten daje się zrealizować przy pomocy systemu trzech neuronów o budowie zbliżonej do budowy układu kontrolnego.

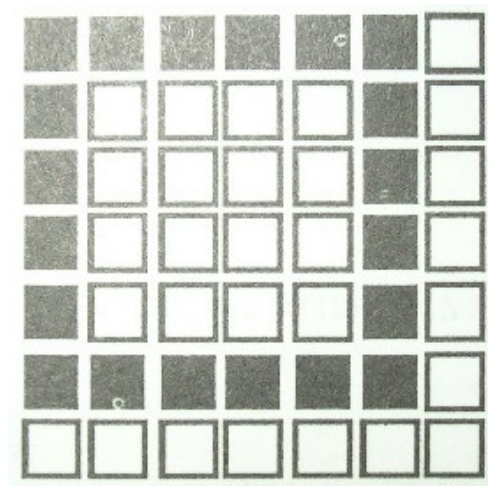
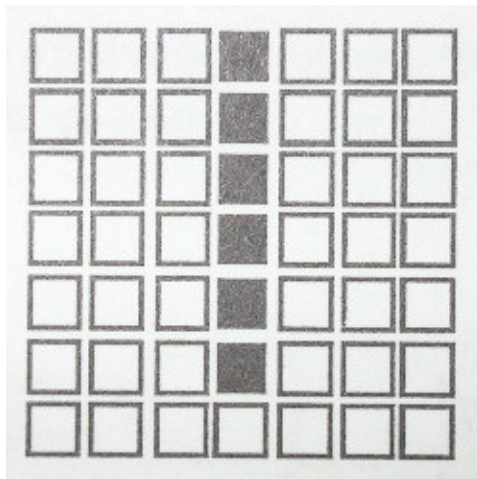


Mankamenty sieci ART

W istocie system orientujący posługuje się swoistym modelem zadania rozwiązywanego przez sieć. Model ten zakłada, że sygnały należące do tej samej klasy mają większość cech wejściowych (składowych wektora X) identycznych, natomiast duża liczba różniących się cech wejściowych jest charakterystycznym wyróżnikiem sytuacji, kiedy pojawia się obiekt z nowej, nieznannej klasy. To założenie może być prawdziwe ale może nie być.

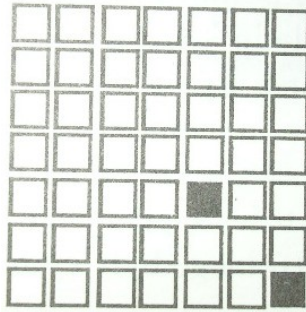
Przykład:

Rozważmy sieć, która już nauczyła się rozpoznawać wzorce liter I oraz O. Sygnały wejściowe X dla tych rozpoznawalnych klas mają postać.

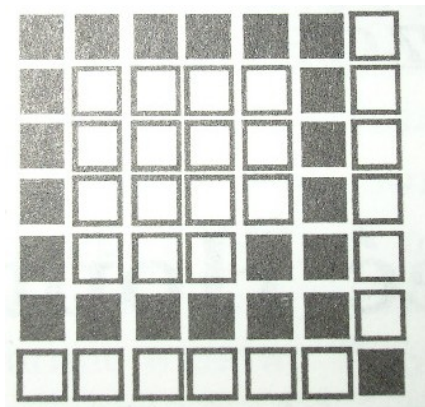
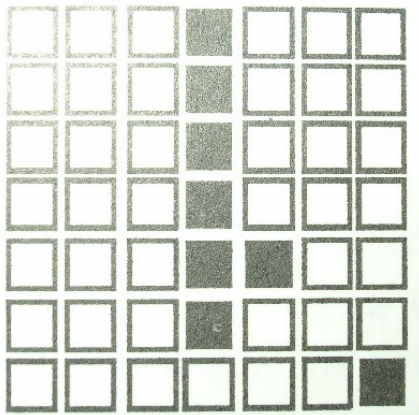


Mankamenty sieci ART

Przyjmijmy teraz, że rozpatrujemy te same obrazy, ale zakłócone dodaniem dwóch pikseli według następującego wzoru.



Zakłócenie ma charakter drobny i nie powinno wpływać na kategoryzacje obrazów.



Mankamenty sieci ART

Sieć w takiej sytuacji jest kompletnie bezradna. Manewrowanie współczynnikiem v nie zapewni sensownego działania sieci, gdyż udział **“szumowych”** pixeli w obrazie litery I jest znacznie większy, niż ich udział w obrazie litery O, a tymczasem właśnie w tym drugim przypadku powinno się zbudować nową klasę, a w pierwszym przypadku należy uznać, że mamy do czynienia ze znanym uprzednio bodźcem - tylko obecnie prezentowanym w formie zakłóconej.

Niezdolność sieci ART do rozróżniania drobnych, ale istotnych zmian rozpoznawanego obrazu od ewidentnych zakłóceń jest jej oczywistą wadą. Trzeba jednak stwierdzić, że nie jest to wada wyłącznie tej właśnie sieci; w istocie takie problemy w mniej lub bardziej istotny sposób pojawiają się przy użyciu dowolnej sieci wykorzystującej **uczenie bez nauczyciela**. Z tego względu, w zastosowaniach mających rozwiązywać zagadnienia praktyczne preferuje się sieci realizujące zasadę **“supervised learning”**, na przykład ze wsteczną propagacją błędów.

Zestaw pytań do testu

1. Jak jest podstawowa zasada sieci rezonansowych?
2. Z ilu warstw składa się sieć ART?
3. Jakie jest zadanie układu kontrolnego w sieci ART?
4. Jakie jest zadanie układu orientującego w sieci ART.?
5. Jaka jest podstawowa wada sieci ART.?