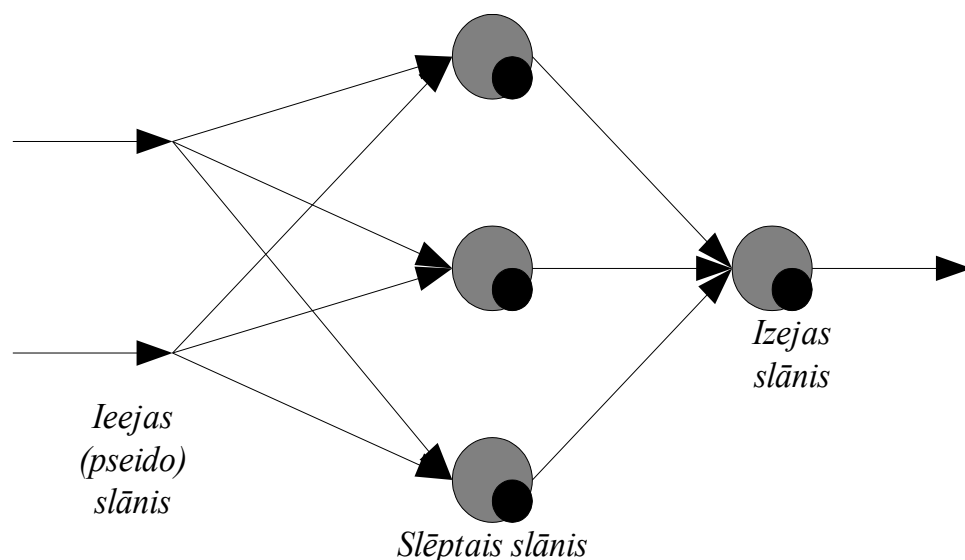


M. Daudzslāņu perceptrons ar kļūdu atgriezeniskās izplatīšanās apmācības metodi (Multilayer Perceptron; Backpropagation)

Daudzslāņu perceptrons vienslāņu perceptrona paplašinājums, kurā bez izejas slāņa ir arī vēl vismaz viens “īstais” slānis. Daudzslāņu perceptrona domājamais pielietojums, neatšķiras no vienslāņu perceptrona domājamā, atšķirība ir tikai spējā izpildīt uzstādīto uzdevumu – daudzslāņu perceptrons ir no skaitļošanas viedokļa daudz spēcīgāks, jo spēj risināt nelineāras klasifikācijas problēmas. Daudzslāņu perceptrons bija zināms jau 20. gs. 60.-jos, gados, tikai tajā laikā un vēl ilgi pēc tam nebija atklāts tā apmācības algoritms. Par “Backpropagation” algoritma atklājējiem uzskata Rummelhartu, Hintonu un Viljamsu (Rummelhart, Hinton, Williams), kas par to paziņoja 1986. gadā.

M1. DAUDZSLĀŅU PERCEPTRONA UZBŪVE UN DARBĪBAS PRINCIPI.

1. Vismaz trīs slāņi – ieejas slānis (*input layer*), vismaz viens slēptais slānis (*hidden layer*) un izejas slānis (*output layer*). Ieejas slānis, tāpat kā vienslāņa perceptronam, uzskatāms par neīstu jeb pseido-slāni.
2. Slāņi izvietoti rindā (lineāri) viens aiz otra. Katram slānim, izņemot ieejas slāni, ir iepriekšējais slānis, un katram slānim, izņemot izejas slāni, ir nākošais slānis. Divu blakus stāvošo slāņu neironi ir savienoti katrs ar katru. Tādējādi katrā neironā ir tik daudz svaru, cik iepriekšējā slānī neironu (neskaitot papildus svaru).
3. Neirona uzbūve kopumā atbilst vienslāņa perceptrona neironam. Vienīgā atšķirība ir tā, ka tiek lietota nelineāra (sigmoidāla) aktivizācijas funkcija. Izņēmums varētu būt vienīgi izejas slānis. Lietojot lineāru aktivizācijas funkciju, daudzslāņu perceptrons neiegūtu tās priekšrocības, kas tam ir, salīdzinot ar vienslāņa perceptronu.
4. Daudzslāņu perceptrona darbības principi ir līdzīgi kā vienslāņa perceptronam. Neironu tīkla darbināšana notiek pa slāņiem pēc kārtas – sākot ar vērtību uzstādīšanu ieejas slānī un beidzot ar izejas slāņa darbināšanu. Katra slāņa darbināšana atbilst vienslāņa perceptrona darbināšanai.



Attēls M1. Daudzslāņu perceptrona grafs ar 2 ieejām, 3 neironiem slēptajā slānī un vienu izeju. Katrā slēptā un izejas slāņa neironā ir papildus svars – nobīde.

M2. DAUDZSLĀŅU PERCEPTRONA APMĀCĪBA AR KĻŪDU ATGRIEZENISKĀS IZPLATĪŠANĀS (ERROR BACK-PROPAGATION) METODI.

M2.1. APMĀCĪBAS PROCESS.

Dots neironu tīkls $\{O_1[n_1], W_2[n_2][n_1+1], O_2[n_2], \dots, W_m[n_m][n_{m-1}+1], O_m[n_m]\}$, kurā ir n_1 ieejas un n_m izejas un $m>2$. (Piemērs, kad $m=3$, $n_1=2$, $n_2=3$ un $n_3=1$, parādīts attēlā M1.) Tiek izmantota sigmoidālā funkcija.

Tāpat kā vienslāņa perceptronam, dots testa piemēru kopums P , kurā ir piemēri skaitā r . Katrs testa piemērs tiek identificēts kā $\{P_s, D_s\}$, kur $s=1..r$. Katra ieejas parauga P_s garums ir n_1 , bet izejas parauga (vēlamās atbildes) D_s garums ir n_m .

Apmācības process ir tāds pats kā vienslāņa perceptronam un notiek šādi:

1. Visu svaru inicializācija ar gadījuma vērtībām intervālā $[-x, +x]$, kur $0<x<1$. Parasti x vērtība ir tuvāka nullei nekā 1, piemēram, 0.3.
2. Apmācības procesa veikšana pēc kārtas uz visiem apmācības piemēriem, atkārtojot to tik ilgi, kamēr tiek sasniegts konverģences nosacījums – vai nu tīkla kopējā kļūda kļūst pietiekoši maza vai arī iterāciju skaits kļūst pārāk liels (formulas P3, P6, P7).

Katrā apmācības procesa solī tiek veiktas šādas darbības:

1. Tīkla inicializācija ar parauga P_s vērtībām.
2. Tīkla darbināšana ar doto paraugu pēc kārtas visiem slāņiem, sākot ar 2. slāni un beidzot ar izejas slāni (formulas P1, P2 (P2a)).
3. Svaru korekcija pēc kārtas visiem slāņiem pretējā secībā – sākot ar izejas slāni (formulas P3, P4, P5) un tad pārējiem slāņiem (formulas M1, P4, P5) – beidzot ar 2. slāni. Svarīgi, ka izejas slānim apmācība notiek savādāk nekā pārējiem.

Svara korekcijas aprēķināšana neironam (formulas P4 vispārinājums):

$$\Delta w_{ji}(t+1) = \eta h o_i \delta_j + \alpha \Delta w_{ji}(t) \quad (M1)$$

kur $\Delta w_{ji}(t+1)$ – j -tā neirona i -tā svara izmaiņa,

$\Delta w_{ji}(t)$ – j -tā neirona i -tā svara izmaiņa iepriekšējā solī,

η (eta) – apmācības koeficients $[0..1]$,

$h = \frac{1}{g}$ – līknes slīpuma koeficients,

o_i – i -tais ienākošais signāls (ieejas slāņa i -tā neirona izeja),

δ_j (delta j) – t.s. lokālais gradients – apmācības formulas komponente, kuru atšķirīgi rēķina izejas slānim un pārējiem slāņiem (M2a, M2b),

α (alfa) – inerces faktors (momentum factor).

Lokālā gradienta komponentes aprēķins izejošajā slānī.

$$\delta_j = e_j o_j (1 - o_j) \quad (M2a)$$

kur e_j – j -tā neirona izrēķinātā kļūda,

o_j – j -tā neirona izrēķinātā izejas vērtība.

Lokālā gradienta komponentes aprēķins pārējos (slēptajos) slāņos.

$$\delta_j = o_j (1 - o_j) \sum_{k=1}^n \delta_k w_{kj} \quad (M2b)$$

kur o_j – j -tā neirona izrēķinātā izejas vērtība,

δ_k (delta k) – nākošā slāņa k -tā neirona lokālais gradients,

w_{kj} – nākošā slāņa k -tā neirona j -tais svars,

n – neironu skaits nākošajā slānī.

M2.2. KĻŪDU ATGRIEZENISKĀS IZPLATĪŠANĀS METODES PAĀTRINĀŠANA.

Šeit ir uzskaitītas dažas heuristikas, kas var tikt izmantotas apmācības algoritma konverģences paātrināšanā (pēc [Haykin, 1999]).

1. Katram maināmam neironu tīkla parametram (svaram) būtu jābūt savam individuālam apmācības koeficientam.
2. Katram apmācības koeficientam būtu jānodrošina iespēja tikt izmainītam apmācības procesa laikā.
3. Ja potenciālā svara izmaiņa kārtējā solī ir ar to pašu (algebrisko) zīmi kā iepriekšējā solī vai vairākos iepriekšējos soļos, tad apmācības koeficients dotajam svaram jāpalielina.
4. Ja potenciālā svara izmaiņa kārtējā solī ir ar pretēju (algebrisko) zīmi nekā iepriekšējā solī vai vairākos iepriekšējos soļos, tad apmācības koeficients dotajam svaram jāsamazina.
5. Neironu tīkla slāņos, kas ir vairāk uz sākuma galu, apmācības koeficientam jābūt lielākam.

M3. DAUDZSLĀŅU PERCEPTRONA NOVĒRTĒJUMS UN ĪPAŠĪBAS.

Apmācība ar kļūdu atgriezeniskās izplatīšanās metodi (Error backpropagation) uzskatāma par standarta algoritmu daudzslāņu perceptrona apmācīšanā, kas tiek uzskatīts par etalonu, pret kuru salīdzina citus algoritmus. Algoritms savu nosaukumu ieguvis no tā, ka kļūdas funkcijas parciālie atvasinājumi tiek noteikti, veicot izejas neironu kļūdas signāla izplatīšanu apgrieztā virzienā no slāņa uz slāni (*back-propagating the error signals*).