

R. Radiālo bāzes funkciju (RBF) tīkls

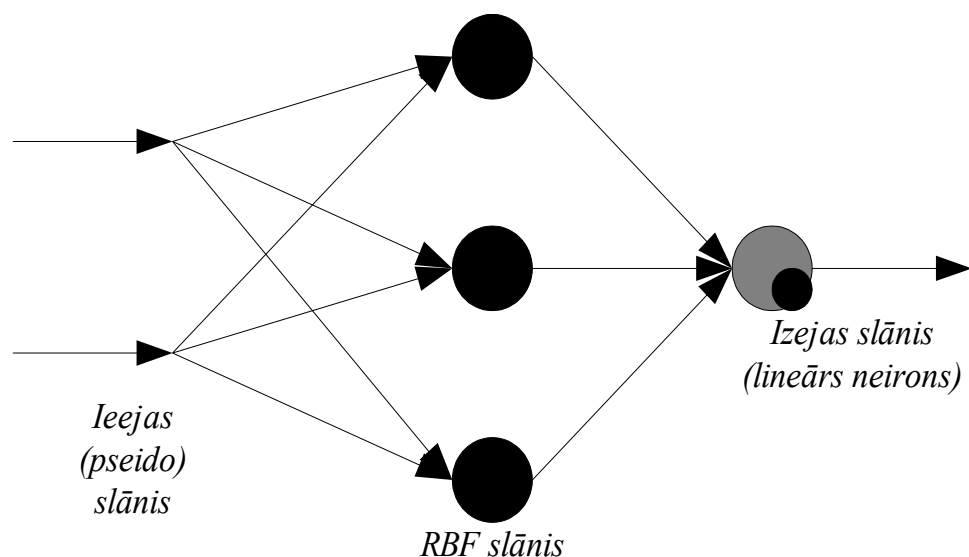
Radiālo bāzes funkciju (*Radial-Basis Function, RBF*) tīkls pieder tīkliem ar apmācību ar skolotāju un pēc pielietojuma un funkcionalitātes ir līdzīgs daudzslāņu perceptronam.

Radiālās bāzes funkcijas un RBF tīklus aprakstījuši dažādi autori: *Powell* (1985), *Light* (1992), *Poggio* (1989), *Girosi* (1989), *Light* (1992).

R1. RBF TĪKLA UZBŪVE UN DARBĪBAS PRINCIPI.

R1.1. RBF TĪKLA GALVENĀS KONSTRUKTĪVĀS ĪPAŠĪBAS.

1. Līdzīgi kā daudzslāņu perceptronam ar vienu slēpto slāni – ir viens ieejas slānis (pseido slānis) un divi īstie slāņi, kas izvietoti lineāri viens aiz otra, un blakus esošo slāņu neironi ir savienoti katrs ar katru.
2. Atšķirībā no daudzslāņu perceptrona īstie slāņi ir dažādi pēc uzbūves un darbības principa. 1. īstais (slēptais) slānis ir t.s. radiālo bāzes funkciju slānis (RBF slānis), bet otrs (izejas) slānis ir perceptrona tipa slānis, kurā parasti izmanto lineāru aktivizācijas funkciju.
3. RBF slāņa neironu specifiskās aktivizācijas funkcijas (Gausa funkcija) dēļ tīkls ieguvīs šādu nosaukumu.
4. RBF slāņa ieejas un izejas vērtības, kā arī svāri ir robežās $[0; +1]$.
5. Runājot par RBF tīkliem, bieži tiek izmantota nevis neironu tīklu terminoloģija, bet gan matemātiskā terminoloģija, jo arī termins “radiālās bāzes funkcijas” nāk no skaitļu analīzes. Pašus RBF slāņa neironus bieži sauc par radiālajām bāzes funkcijām, bet to svarus – par šo funkciju centriem.



Attēls R1. RBF tīkla grafs ar 2 ieejām, 3 neironiem RBF slānī un vienu izeju.

R1.2. RBF TĪKLA DARBĪBAS PRINCIPI.

1. RBF slānis apmācības dēļ būtu uzskatāms nevis par konkrētu modeli, bet gan par šablonu, kurā var ievietot dažādus algoritmus.
2. Apmācība notiek divos etapos. Vispirms tiek apmācīts RBF slānis, pēc tam izejas slānis.
3. Apmācības paraugu kopums tāpat kā perceptronam sastāv no pāriem – piemērs un vēlamā vērtība.
4. RBF tīkla apmācība notiek bez skolotāja (tātad, parasti neizmantojot vēlamās vērtības) un ir iespējams izmantot dažādus (vienkāršākus vai sarežģītākus) algoritmus. Vispārīgi runājot, RBF tīkla apmācība ir ievadparaugu klāsterizācija.
5. Izejas slānis ir perceptrons un tā tas arī tiek apmācīts. Vispirms katrs ievadparaugs tiek darbināts RBF slānī, pēc tam ar RBF slāņa izejām un vēlamajām vērtībām tiek darbināts perceptrona algoritms izejas slānī.
6. Paraugu klasificēšana notiek vienā solī un salīdzinoši ātri – tāpat kā perceptronam.

R2. APMĀCĪTA RBF TĪKLA DARBINĀŠANA.

Tīkla inicializācija ar ievadvērtībām.

$$O_1 \leftarrow X$$

kur $X[m]$ – paraugs,

$O_i[m]$ – ieejas slāņa neironu izejas.

Parauga atpazīšana notiek vienā solī – vispirms tiek darbināts RBF slānis, tad izejas slānis.

Katram izejas slāņa neironam notiek aprēķini – formulas K1, K2 un P1, P2 (nodaļa P).

Summēšanas funkcija RBF slānī.

Summēšanas funkcija nosaka attālumu starp ievadvērtību virkni un neirona svaru vērtību virkni. Formulā R1 tiek izmantots t.s. Eiklīda attālums, un summēšanas funkcija ir ļoti līdzīga Kohonena tīkla summēšanas funkcijai.

$$NET_j = \|W_j - O\| = \sqrt{\sum_{i=1}^m (w_{ji} - o_i)^2} \quad (R1)$$

kur NET_j – j -tā neirona summēšanas vērtība,
 w_{ji} – j -tā neirona i -tais svars,
 o_i – ieejas slāņa i -tā neirona izejas vērtība.

Aktivizācijas funkcija RBF slānī.

Par aktivizācijas funkciju tiek izmantota Gausa funkcija.

$$o_j = e^{-\frac{NET_j^2}{2\sigma^2}} \quad (R2)$$

kur NET_j – j -tā neirona summēšanas funkcija,
 σ (*sigma*) – līknes slīpuma (liekuma) faktors.

Izejas (perceptrona) slāņa darbināšana.

Izejas slānis ir perceptrona slānis, tādēļ to darbina, izmantojot perceptrona summēšanas un aktivizācijas funkcijas (sk. formulas P1 un P2, P2b, P2c nodaļā). Izejas slānī parasti tiek izmantota lineāra aktivizācijas funkcija.

Rezultāts.

Pēc tīkla darbināšanas (uz paraugu X) tīkla dotā atbilde atrodas uz izejas slāņa neironu izejām O_2 .

R3. RBF TĪKLA APMĀCĪBA.

R3.1. APMĀCĪBAS PROCESA VISPĀRĪGS APRAKSTS.

Dots neironu tīkls $\{O_1[n_1], W_2[n_2][n_1], O_2[n_2], W_3[n_3][n_2+1], O_3[n_3]\}$, 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 R1.)

Tāpat kā 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 notiek 2 etapos:

1. RBF slāņa apmācība (sk. sadaļu R3.2).
2. Izejas slāņa apmācība (sk. sadaļu R3.3).

R3.2. RBF SLĀŅA APMĀCĪBA.

RBF slāņa apmācībā nelieto konkrētu algoritmu. RBF slānis ir šablons vai ietvars, kurā var ievietot dažādus algoritmus. RBF slāņa apmācības mērķis ir nevis gūt spēju atpazīt vai klasificēt, bet savā ziņā tieši otrādi – “sajaukt” ieejas datus tā, lai pēc sajaukšanas tie būtu atdalāmi, izmantojot lineāras metodes (piemēram, perceptrona slāni). Šai sajaukšanai dažos gadījumos der arī klāsterizācija, kas atbilst Kohonena tīklā realizētajai.

Pielietotie algoritmi ir diezgan dažādi – no ļoti vienkāršiem līdz sarežģītiem.

Triviālā metode.

Katrs ieejas paraugs tiek noteikts par svaru vērtību vektoru, tādējādi RBF slānī vajadzīgi tik neironi, cik ir apmācības piemēru. Jau šeit var redzēt, ka, lai ar šo metodi apmācītu tīklu atpazīt 2-argumenta Būla funkciju, RBF slānī nepieciešami 4 neironi, kas jau tik vienkāršā gadījumā ir, varētu teikt, daudz vairāk, nekā nepieciešams daudzslāņu perceptronam.

Klāsterizācijas metode.

Ja ir daudz apmācības paraugu, tad triviālā metode nav pielietojama tīri tehnisku ierobežojumu dēļ. Tādēļ var lietot kādu no klāsterizācijas metodēm, lai uzstādītu svarus RBF slānī. Viens no vienkāršākajiem paņēmieniem ir izmantot Kohonena tīkla algoritmu RBF slāņa apmācībai.

R3.3. IZEJAS SLĀŅA APMĀCĪBA.

Izejas slāņa apmācības process ir identisks vienslāņa perceptrona apmācībai un notiek šādi:

1. Visu izejas slāņa 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 izejas slāņa apmācības procesa solī tiek veiktas šādas darbības:

1. Tīkla inicializācija ar parauga P_s vērtībām.
2. RBF slāņa darbināšana.
3. Izejas slāņa darbināšana.
4. Izejas slāņa apmācība identiski vienslāņa perceptronam (formulas P3, P4 vai P4b, P5).

R4. RBF TĪKLA NOVĒRTĒJUMS UN ĪPAŠĪBAS.

Ar RBF tīklu palīdzību var risināt tās pašas problēmas, ko ar daudzslāņu perceptroniem ar kļūdu atgriezeniskās izplatīšanās metodi. Lūk, RBF tīkla salīdzinājums ar daudzslāņu perceptronu (pēc [Haykin, 1999]).

1. RBF tīkls pamatvariantā ir tīkls ar vienu slēpto slāni (kopā 3 slāņi), turpretī daudzslāņu perceptronam tie var būt vairāki.
2. Daudzslāņu perceptrona neironi visos slāņos parasti ietilpst kopējā neironu tīkla modelī, bet RBF tīklā RBF slāņa neironi ir pilnīgi atšķirīgi no izejas slāņa neironiem un tiem ir pilnīgi atšķirīgs darbības uzdevums.
3. RBF tīkls RBF slānis ir nelineārs, bet izejas slānis – lineārs, turpretī daudzslāņu perceptronam parasti visi slāņi ir nelineāri.
4. RBF slāņa neironiem ir atšķirīga ne tikai aktivizācijas funkcija (Gausa funkcija), bet arī summēšanas funkcija (Eiklīda attālums starp ieejas vektoru un svaru vektoru).
5. Daudzslāņu perceptroni konstruē nelineāru attēlojumu globālas aproksimācijas, turpretī RBF tīkli – lokālas. Tas nozīmē to, ka, lai aproksimētu nelineāru attēlojumu, daudzslāņu

perceptronā nepieciešams mazāk parametru (svaru) nekā RBF tīklā tās pašas precizitātes panākšanai.

R5. TEORĒTISKAIS PAMATOJUMS.

RBF tīkla darbība pamatojas uz Kavera teorēmas (*Cover's theorem on the separability of patterns*; Cover, 1965).

A complex pattern-classification problem cast in a high-dimensional space nonlinearly is more likely to be linearly separable than in a low-dimensional space.

/Ja sarežģītu paraugu klasifikācijas problēmu ievieto daudzdimensiju telpā, tad ir vairāk iespējams, ka tā ir lineāri atdalāma nekā, to ievietojot telpā ar mazāku dimensiju skaitu./