

• Clase de funcții învățabile de RPF

Proprietate 3 Clasa funcțiilor booleene de dimensiunea d (adică $B = \{ f : \{0,1\}^d \rightarrow \{0,1\} \}$) poate fi învățată de o RPFM cu două straturi functionale și cu funcții de transfer ~~ideale~~ ^{ideale} pe hebbende în pragul de activare zero.

Denumire Să numărăm că card(B) = 2^{2^d} și că o funcție booleană este complet specificată dacă pentru fiecare ~~într-un~~ $x \in \{0,1\}^d$ este specificat (dintă 2^d astfel de forme - patterns-) este dat $f(x) \in \{0,1\}$ corespunzător.

Mai reamintim și rezultatul lui Lewis și Oates (1967) (vezi pag 19 cursul 1) care afirmează că o rețea cu un singur strat funcțional (doar cu stratul de ieșire) nu poate învăța decât cel mult $\left[2^{2^d} / d! \right]$ funcții.

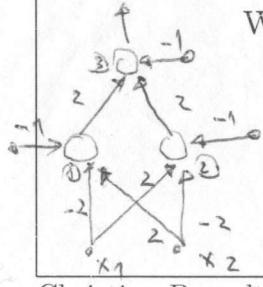
Representing Arbitrary Boolean Functions

Let $y = f(x_1, \dots, x_n)$ be a Boolean function of n variables.

- (i) Represent $f(x_1, \dots, x_n)$ in disjunctive normal form. That is, determine $D_f = K_1 \vee \dots \vee K_m$, where all K_j are conjunctions of n literals, i.e., $K_j = l_{j1} \wedge \dots \wedge l_{jn}$ with $l_{ji} = x_i$ (positive literal) or $l_{ji} = \neg x_i$ (negative literal).
- (ii) Create a neuron for each conjunction K_j of the disjunctive normal form (having n inputs — one input for each variable), where

$$w_{ji} = \begin{cases} 2, & \text{if } l_{ji} = x_i, \\ -2, & \text{if } l_{ji} = \neg x_i, \end{cases} \quad \text{and} \quad -\theta_j = n - 1 + \frac{1}{2} \sum_{i=1}^n w_{ji}.$$

- (iii) Create an output neuron (having m inputs — one input for each neuron that was created in step (ii)), where



$$w_{(n+1)k} = 2, \quad k = 1, \dots, m,$$

and

$$-\theta_{n+1} = 1.$$

Ex: $x_1 \text{ XOR } x_2 \equiv (x_1 \wedge x_2) \vee (\neg x_1 \wedge \neg x_2)$

x_1	x_2	K_1	K_2	$x_1 \text{ XOR } x_2$
0	0	0	0	0
0	1	1	0	1
1	0	0	1	1
1	1	0	0	0

19

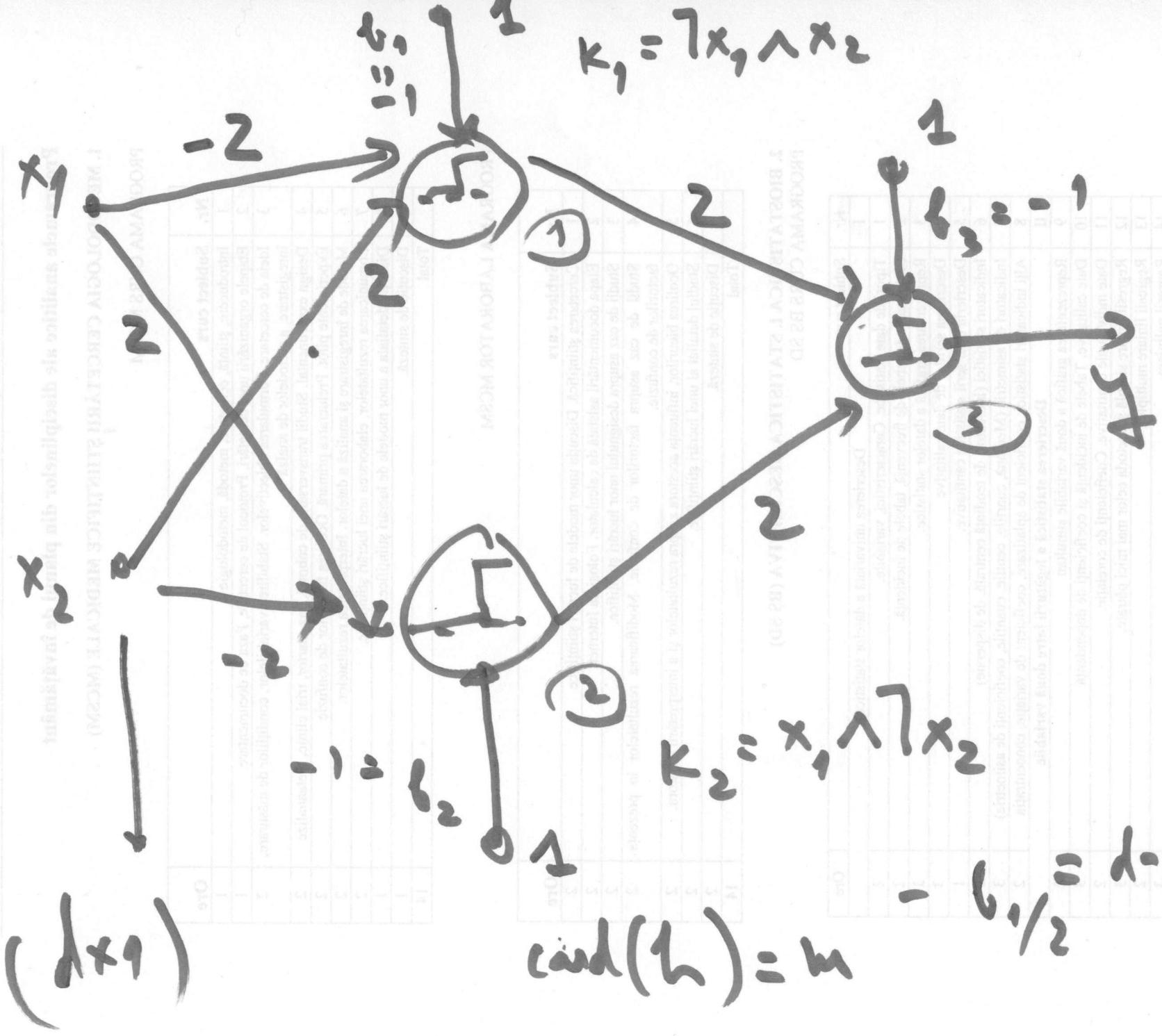
$$\sum q_i$$

$$d - \frac{b_1}{2} = d - 1 + \frac{1}{2}$$

$$\text{card}(L) = m$$

$$K_2 = x_1 \wedge x_2$$

$$K_1 = x_1 \wedge x_2$$



$$\mathcal{B} = \left\{ f: \{0,1\}^3 \rightarrow \{0,1\} \right\}$$

$f \in \boxed{\dots}$

handlung (x)

$$= \begin{cases} 1 \times 30 \\ 0 \times <0 \end{cases}$$

x_1	$x_0 R$	x_2	G	f	G	f	G	f	$f \equiv u$
0	-4	0	-1	0	-4	0	-4	0	0
1	1	1	-3	0	1	1	1	1	1
1	-3	0	1	1	1	1	1	1	1
0	-3	0	-3	0	-4	0	0	0	0

$K_1 = 7x_0 \wedge x_2$ $K_2 = x_1 \wedge 7x_2$ $K_1 \vee K_2 = D_f$

struktur ausweg (h) struktur der logik
(out put)

Observații (1) Percepționii de pe traiul ascuns din construcție demonstrativă de mai sus ilustrează conceptul de "sablon"-templat - ; este arătată că un sablon pt. vectorii de intrare "nerecunoscând", adică neștiind ~~de~~ un exemplu - un "1" - decât datele vectorul de intrare - input pattern - coincide cu vectorul sablon-template pattern -

(2) Construcția de mai sus este de ~~o interes~~^{o valoare} practică redusă cau "memorializă" doar o funcție booleană neavând posibilitatea să generalizeze ~~adica~~^{adica} să recunoască noi vectori în afara schimbului de învățare (deoarece topologia retelei depinde de fără selecție de învățare este exhaustivă)

(3) Desvoltând desunție de la pag. 19-20 curs 1 privind posibilități de dicotomizare a unui multibil de puncte din \mathbb{R}^d de către o RPM (în cursul 1 se arată că acesta este ~~posibil~~^{imposibil} face fără eroare ~~fără~~^{de către} RPM cu un unușor nivel funcțional doar dacă multibilul este liniar separabil), lucru întotdeauna posibil dacă $n \leq d+1$, Nilsson în 1965 (citat de Bishop 1995) a arătat că acesta este întotdeauna posibil fără eroare pt. o RPM cu domeniu ~~statornic~~ nile funcționale, arănd pe nivelul ascuns $n-1$ percepționii. Baebu în 1988 (citat de Bishop 1995) a demonstrat rezultatul în sensul că a redus numărul de percepționii astfel încât în kernel să aibă numărul de percepționii de pe nivelul ascuns de $[n/d]$ (unde $[\cdot]$ este funcție partea întreagă). dacă cele n puncte din selecția de învățare sunt în \mathbb{R}^d , dacă cele n puncte din selecția de învățare sunt în același plan (probabilitatea ca $d+1$ puncte să fie în același plan este nulă).

Propoziția 4 (proprietatea de universalitate a RPM) \forall RPM cu trei nivele funcționale poate approxima cu o eroare dată oricât de mică orice funcție continuă scalară sau multivariată (adică $\text{Dom}(f) \subseteq \mathbb{R}^k$ sau $\text{Dom}(f) \subseteq \mathbb{R}^d$) scalară sau vectorială (adică $f(x) \in \mathbb{R}$ sau $f(x) \in \mathbb{R}^k$) cu domeniul de definiție în același număr apărări printr-dimensionale.

Sketch de demonstrație (după Bishop 1995 care-i citează pe Jones 1990; Blum și Li 1991).

Se consideră că $f: A \subseteq \mathbb{R}^2 \rightarrow B \subseteq \mathbb{R}$ fără a pierde prea multă generalitate.

Se presupune cunoscută următoarele rezultate:

- pt. un $\epsilon > 0$ dat, dacă altfel oricare

$$f(x_1, x_2) \approx \sum_{\ell} \sum_{s} A_{s\ell} \cos(\ell x_1) \cos(s x_2)$$

Potă fi ~~definită~~^{aproximată de} o serie Fourier cu o eroare ϵ .

Sau scriind produsul de cosinusuri ca sumă de cosinuzuri rezultă:

$$f(x_1, x_2) \approx \sum_{\ell} \sum_{s} \frac{A_{s\ell}}{2} [\cos(\ell x_1 + s x_2) + \cos(\ell x_1 - s x_2)]$$

- pt. un $\delta > 0$ dat, dacă altfel oricare

$$\cos(z) \approx g_0 + \sum_{i=1}^N (g_i - g_0) H(z - z_i)$$

Potă fi aproximată de o sumă de funcții scăză cu o eroare δ (M. lui Weierstrass) unde $g_i = \cos(z_i)$, $i = 0, N$ iar H este funcția Heaviside în prag de activare zero.

Propoziția este demonstrată dacă se construiește o RPTM cu 3 nivele funcționale ce implementează seria:

$$(4.1) \quad \sum_{s=1}^S \sum_{\ell=1}^L \frac{A_{s\ell}}{2} [2g_0 + \sum_{n=1}^N H(\ell x_1 + s x_2 - z_n) + H(\ell x_1 - s x_2 - z_n)]$$

Care din considerențele amintite mai sus, putem să numărăm câteva baterii de termeni, aproape atât de liniile funcției continue f .

Trebuie reținută următoarea topologie:

- 2 intrări - nivelul de intrare;
- $2N$ baterii de căte ~~$S+L$~~ perceptri pe primul nivel funcțional (primul nivel ascuns);
- 2 baterii de căte $S+L$ perceptri pe al doilea nivel funcțional (al doilea nivel ascuns);
- un perceptor pe nivelul de ieșire (al treilea nivel funcțional).

Inteligibilitate artificială

(14)

Curs 2

În construcție de mai sus: - N - numărul de puncte pentru descrierea corectă a funcției f . Acest număr este ~~detinut~~ impreună cu punctele de sprijin $\{z_m\}_{m=1}^{2^N}$; $\{y_m\}_{m=1}^N$;

- S și L - limitele maximă de variație a vîrșilor și a liniei de evoluție în serie Fourier a lui f . Aceste numere sunt date, astfel încât eroarea de aproximare a lui f să fie $\leq \varepsilon$ fixat, impreună cu coeficientul A_{SL} .

Graful RPM este următorul:
~~Functia de baterie de percepție utilizat precizare fizică~~

- cele două intrări sunt conectate la toți perceptorii de pe primul nivel funcțional.
- perceptorii de pe primul nivel funcțional sunt conectați la toți perceptorii de pe al doilea strat funcțional. Analog perceptorii de pe al doilea strat funcțional sunt conectați la toți perceptorii de pe al treilea strat funcțional.
- perceptorii de pe al doilea strat funcțional sunt conectați la perceptorii de pe al treilea strat funcțional.

Perceptorii de pe același strat sunt conectați între ei și fluxul grafului este de la nivelul de intrare la cel de ieșire.

Funcțiile de integrare ale tuturor receptorilor sunt identice și produsul scalar din spațiu corespunzător (R^2 pt. primul strat funcțional, R^N pt. al doilea strat asimilat și $R^{S \times L}$ pt. al treilea strat asimilat).

Funcțiile de transfer sunt:

- funcție δ_L cu pragul de activare $\{z_m\}_{m=1}^N$ pentru cele ~~de~~ baterii de receptorii de pe primul strat asimilat;
- funcție liniară pentru ceilalți receptorii.

Matricele coeficienților de pondere sunt: $(2(S \times L))$

$$W(1) = \left(w_k^{(1)} \right)_{k=1, n_1}^{i=1, 2}; \quad W(2) = \left(w_k^{(2)} \right)_{k=1, n_2}^{i=1, n_1+1}; \quad W(3) = \left(w_k^{(3)} \right)_{k=1}^{i=1, n_2}$$

Inteligenta artificiala

'curs 2'

(15)

unde

$$w_k^i(1) = \begin{cases} l \text{ dacă } i=1 \\ s \text{ dacă } i=2 \text{ și } k = m + (l-1)N + (s-1)N \cdot L \\ \rightarrow \text{dacă } i=2 \text{ și } k = S \cdot L \cdot N + m + (l-1)N + (s-1)N \cdot L \end{cases}$$

$n=1, N$
 $s=1, S$
 $l=1, L$
 $k=1, 2(S \cdot L \cdot N)$

Se observă că percepționii de pe primul strat au sensuri deplasare

$$w_k^i(2) = \begin{cases} g_n - g_{n-1} \text{ dacă } k = m + (l-1)N + (s-1)N \cdot L \\ \text{ sau } k = S \cdot L \cdot N + m + (l-1)N + (s-1)N \cdot L \\ \text{dacă } i = S \cdot L + 1 \\ \text{și } (k) \in \overline{1, n_2} \end{cases}$$

$i = (s-1)L + l \quad n=1, N$
 $\text{ sau } \quad s=1, S$
 $k = S \cdot L + (s-1)L + l \quad l=1, L$

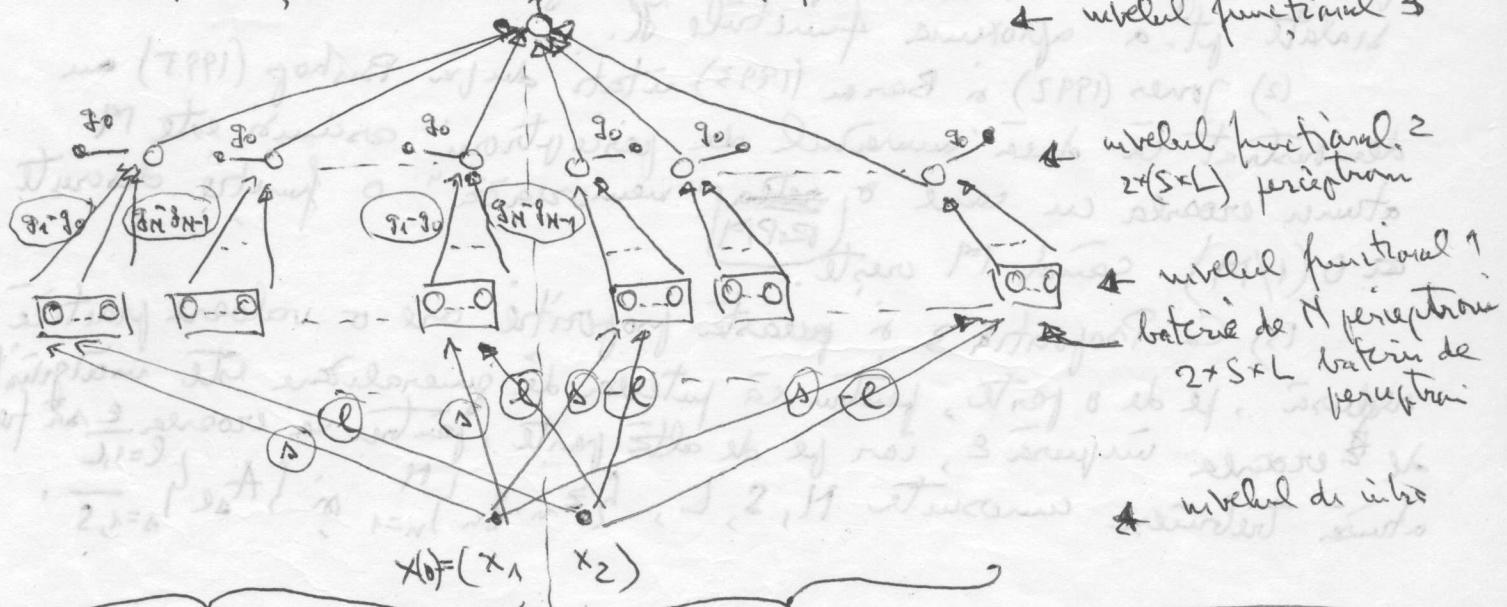
Se observă că percepționii de pe al doilea strat au sensuri deplasare identice, egale cu g_0 .

$$w_1^i(3) = \frac{Ase}{2} \quad \text{cu } i = \overline{1, n_2} \text{ de forma } i = (s-1)L + l$$

$\text{ sau } \quad i = S \cdot L + (s-1)L + l$

~~Topologia rețelei este ilustrată în figura 10.~~

→ nivelul funcțional 3



5 (Fig 10) Topologia rețelei se implementează o funcție continuă

Din construcție rețea este simetrică. Se studiază funcția implementată de rețea care notăm cu S :

$$X(1) = \left(\underbrace{H_1(x_1 + x_2), \dots, H_1(x_1 + Lx_2), \dots, H_1(Sx_1 + Lx_2), \dots, H_1(Sx_1 + (S-1)Lx_2)}_{\text{prima batere de } M \text{ percepționi}}, \dots, \underbrace{H_2(x_1 + (S-1)Lx_2), \dots, H_2(Sx_1 + (S-1)Lx_2)}_{\text{a } (S-1)L+1 \text{ baterii de } M \text{ percepționi}} \right)$$

a $S \cdot L$ baterie
de M percepționi

prima batere

de M percepționi

-1.

M percepționi

→ (x1, x2)

$$x(2) = \left(g_0 + \sum_{n=1}^N (g_n - g_{n-1}) H_2 \frac{(x_1 + l x_2)}{2} \right), \dots, g_0 + \sum_{n=1}^N (g_n - g_{n-1}) H_2 \frac{(s x_1 + l x_2)}{2}, \dots, g_0 + \sum_{n=1}^N (g_n - g_{n-1}) H_2 \frac{(2 x_1 + l x_2)}{2}$$

primul nivel perceptiv
de pe al 2-lea nivel fizișor

al (s+1) L + l perceptiv
de pe al 2-lea nivel fizișor

al SL perceptiv
de pe al 2-lea nivel fizișor

$$x(3) = \sum_{S=1}^S \sum_{L=1}^L \frac{A_{SL}}{2} \left[g_0 + \sum_{n=1}^N (g_n - g_{n-1}) H_2 \frac{(s x_1 + l x_2)}{2} \right]$$

la subrețea
stângă a rețelei

Analog

$$x(3) = \sum_{S=1}^S \sum_{L=1}^L \frac{A_{SL}}{2} \left[g_0 + \sum_{n=1}^N (g_n - g_{n-1}) H_2 \frac{(s x_1 - l x_2)}{2} \right]$$

la subrețea
dreaptă a rețelei

În perceptivoul de pe nivelul de ieșire se obține înălță $x(3)|_S$,
 $x(3)|_D$ deci rețeaua implementează seria (4.1)

q.e.d.

Obs. (1) Proprietatea de aproximare universală a RPM-urilor rămâne valabilă dacă funcțiile de transfer H sunt înlocuite cu funcții sigmoidice (de tip "logistic sau tangenta hiperbolică") convenabile și ale cărui funcții de transfer sunt H .

(2) Jones (1992) și Barron (1993) citati după Bishop (1995) au demonstrat că dacă numărul de perceptri orizontali este M atunci eroarea cu care θ este "memorată" o funcție crescentă și $O(1/M)$ când M crește.

(3) Ca Propoziția 3 și această proprietate are o valoare practică redusă, pe de o parte, pentru că puterea de generalizare este marginală și crește împreună cu θ , iar pe de altă parte, pentru că eroarea este de $O(1/M)$ atunci trebuie cunoști $N, S, L, \{g_n\}_{n=1}^N$ și $\{A_{SL}\}_{S=1, L=1}^{S, L}$.

Inteligenta artificială

- în secolul XIX și în secolul XX începutul secolului XXII sunt o
o tehnologie care este în curs de dezvoltare și care
continuă să aducă progrese în ceea ce privește
rețelelor neuronale vor suflare din pătruncile de
bază (fonte ~~proveniente~~ de abordare):
- (a) soluția lui Kolmogorov la unul din problemele lui
Hilbert privind reprezentarea funcțiilor continue multivariante;
 - (b) Teorema Stone-Weierstrass că aproape orice funcție
continuă poate fi exprimată ca o sumă de funcții continue;
 - (c) aproximări trigonometrice, polinomiale, spline, Fourier
a unei funcții de mai multe variabile.

(*) continuă cu "Teorema Kolmogorov" (pag 16)

Teoremă (Kolmogorov/Slepchenko). Fie $D \subset \mathbb{R}^d$ un集 de numere
așa că $\sum_i \lambda_i x_i = 0$ și $x_i \neq 0$. Fie $D = [0, 1]^d$,
atunci (\exists) o funcție continuă monotonă cresătoare $f: D \rightarrow D$
având următoarele proprietăți: (i) $f: [0, 1]^d \rightarrow \mathbb{R}$ continuă
și $d \geq 2$ (ii) f funcție continuă și constantele $\{a_j\}$,
 b_j și a_i .

$$f(x_1, \dots, x_d) = \sum_{j=1}^{2d+1} g(b_j + \sum_{i=1}^d \lambda_i h(x_i + j a_i))$$

$$\text{cu } h_j(x) = h(x + j a_d)$$

În fapt, dacă se supuneam pe $x \in \mathbb{R}^d$ în $b_j + \sum_{i=1}^d \lambda_i h(x_i + j a_d) \in \mathbb{R}^{2d+1}$
(aceasta se numește Whitney) atunci o funcție
continuă de d variabile reale poate fi scrisă ca

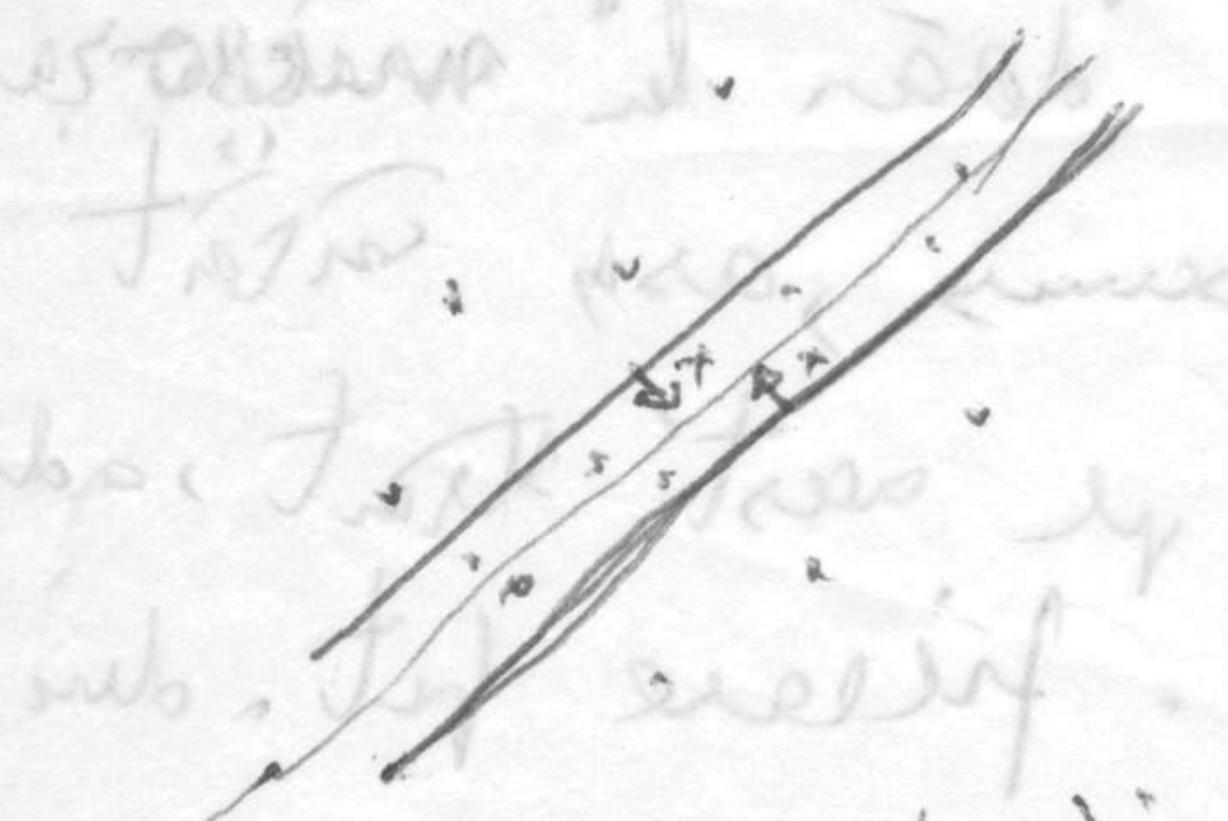
O lună de $2d+1$ puncte care să
adă și picăne din cele $2d+1$ argumente este o
variabilă reală arbitrală (adică
se supune să fie orice număr
real).

Teorema 3.1 (a construcție sandwichă, menită să demonstreze că dimensiunea lui \mathbb{B} este n)
 Matheron (1990) The Mathematical Foundations of Geostatistics
 Macchiesi, Morgan (Caufmann Pub., San Mateo, CA). Orice multime E de n vectori din \mathbb{R}^d și multimea generală pot fi diotomizate în $E^+ = \{x \in E \mid t=1\}$ și $E^- = \{x \in E \mid t=-1\}$ de către un plan în \mathbb{R}^{d+1} .

$E^- = \{x \in E \subset \mathbb{R}^n | t = -\lambda\}$ și $\overline{x} \in$ stratul de
 $\pi(x)[\frac{n}{d}]$ perceptroni \Leftrightarrow (faza de transfer) pe stratul de ieșire.

as among the
Toti, periepithemis reticulata and the priaful zero

Din Făcă o permută din generalitatea presupunerii
 $\text{card}(\mathbb{E}^+) \leq \text{card}(\mathbb{E})$. + notăti generale cu \underline{x} , în
 Se impărtă vectorii din \mathbb{E}^+ , notăti generale cu \underline{x}^+ , în
 grupuri de căte d vectori. ~~Se construiește hiperplanul~~
 unui planar de d vectori). ~~construiește hiperplane~~
~~care conține aceste căte un suport~~ Reunim părțile grupurilor
 se construiesc hiperplanele care conțin (posibil din ipoteza
 de poziție generală); pe acesta notăm cu $(\underline{w}, \underline{z})$.
 Se impărtă $(\underline{w}, \underline{z})$ în două hiperplane paralele, de orientare
 pe spațiu $(\underline{w}, \underline{z} + \epsilon)$ și $(\underline{w}, -\underline{z} + \epsilon)$ care formează ~~sandwich~~
 și $\underline{w} + \epsilon$ considerat (vezi figura). În
 se continuă ^{doar} grupul de d vectori considerat
 care să nu se va alge $\epsilon > 0$



(w,z) Fig. Gastrocnie sandwich

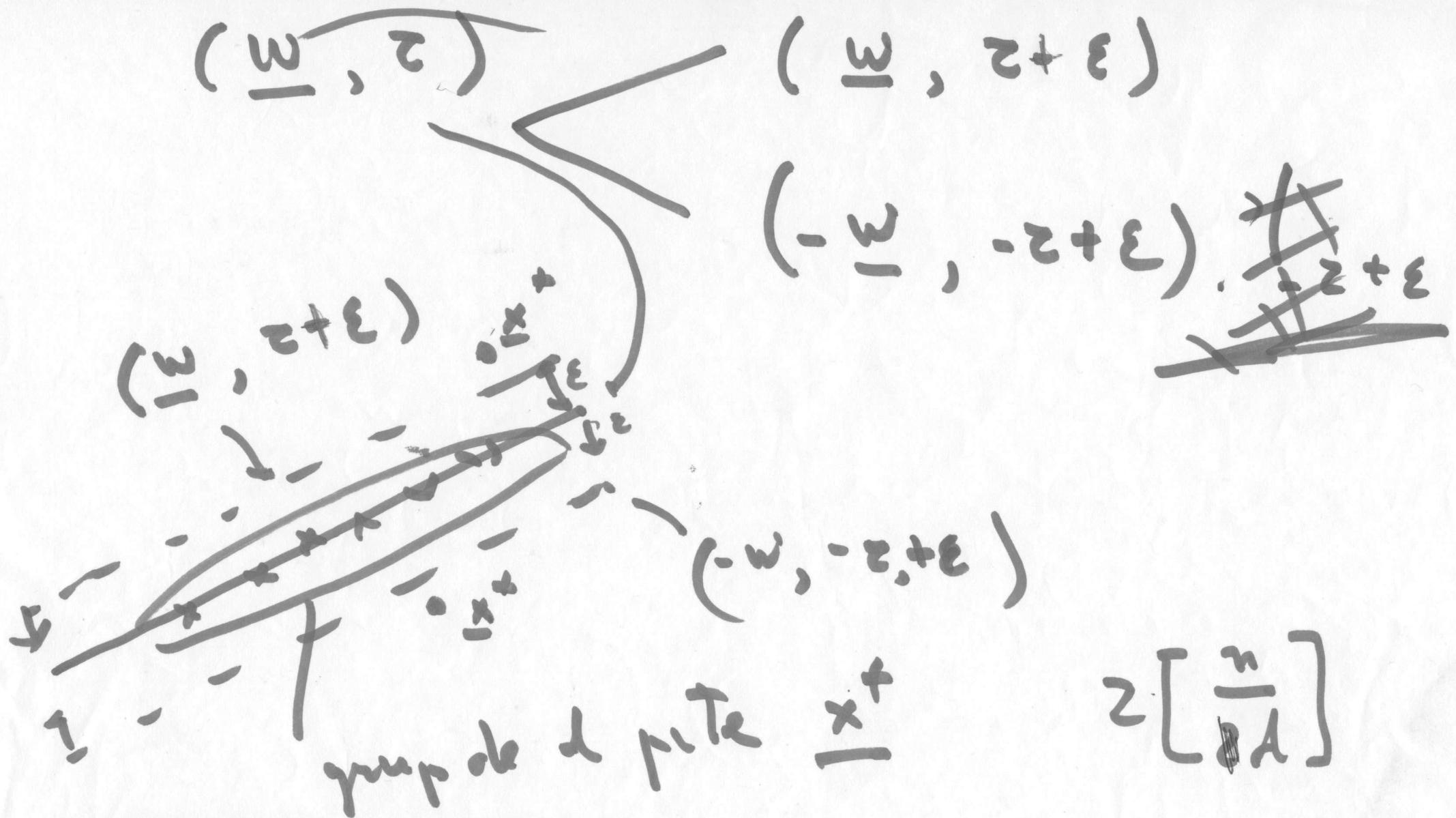
Dacă hiperplanele (\underline{w}, c_1) și $(-\underline{w}, -c_1)$ sunt astfel orientate încât față lor pozitivă să fie orientarea către grupul de vectori din E^+ exprimă între ele atunci ~~aceste~~ rezultă perceptronul ~~corespondă~~ acestor hiperplane vor fi egale în 1 (în ~~atunci~~ din construcție, rezulătorul acestor perceptriuni $\neq 0$ unde al vector din E^+ nu apartine grupului exprimat între hiperplane vor fi zero).

În concluzie, nu există deosebită diferență între teoria clasică și teoria apăsării logică OR sau (\exists) care se bazează pe o reprezentare a realității în termeni de vectori și în termeni de clasificare. Această diferență se manifestă în modul în care se exprimă teoremele de clasificare. Teorema lui Fisher este de exemplu:

$$\text{L}_2 \text{ (f. elegante și algebrică)}, \text{ în care realizabilitatea} \\ \text{unei sujekte percepționă obiecte stratificate în zone} \\ (\text{practic acesta aplică funcția logitică OR circular} \\ \text{percepționă obiecte stratificate în clase}) \text{ și Toti vectorii} \\ \underline{x} \in E^+ \text{ vor fi clasificati } \underline{\text{împreună}} \text{ în clasa } C_1 \text{ și} \\ t=1 \text{ și Toti vectorii } \underline{x} \in E^- \text{ vor fi clasificati } \underline{\text{împreună}} \text{ în clasa } C_2 \text{ în } t=0.$$

Observation

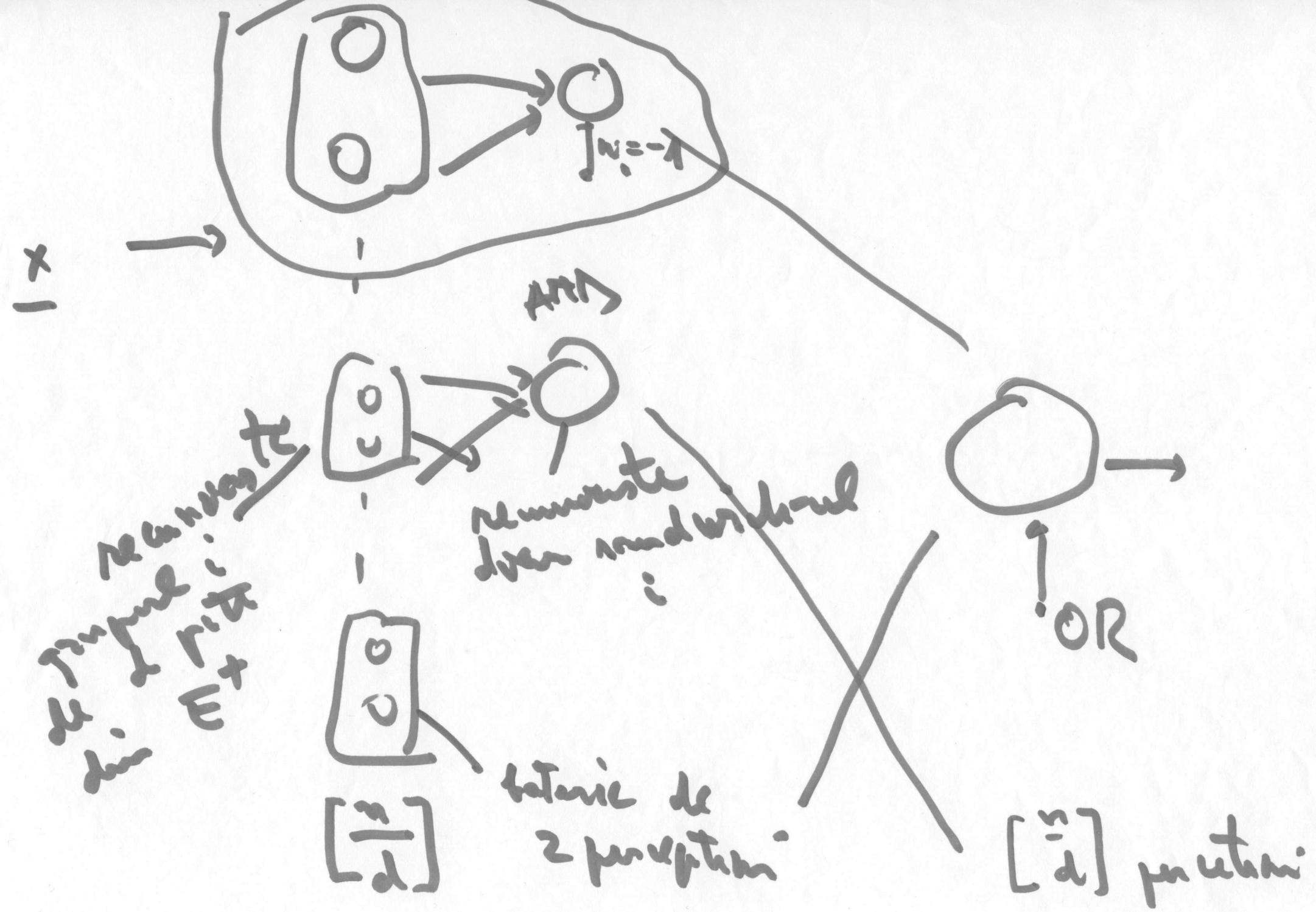
- (1) Teorema 3.1 susține că într-un spațiu cu măsură Riemanniană Ω există o mulțime denumerită de puncte x_1, x_2, \dots, x_n care sunt în poziție generală. Într-o altă formulare, teorema spune că există un număr finit de puncte x_1, x_2, \dots, x_n care sunt în poziție generală.



Hillsm 1990

1/2

- 2 -



Complexitatea implementarii Th. Sanduich

nr. de parantezi
(parantezi si deplasare)

$$P \leq \underbrace{(d+1)2 \cdot \frac{n}{d}}_{\text{deplasare}} + \underbrace{3 \frac{n}{d}}_{\substack{\text{"} \\ (2^{t_1}) \\ \text{deplasare}}} + \underbrace{\frac{n}{d} + 1}_{\text{strat III}}$$

$$\frac{n}{d} (2d+2+4) + 1 = \frac{2^n}{d} (d+3) + 1$$

$$\left(1 + \frac{n}{d}\right) + \left(\frac{2}{d} + \frac{4}{d}\right) + \frac{2^n}{d} (d+3)$$

$$nS \approx n\left(\frac{e}{d} + 1\right) \approx S_n \cdot \left(1 + \frac{2}{d} + \frac{4}{d}\right)$$

$$= 2n \left(1 + \frac{3}{d}\right) + 1$$

$$\approx 2n$$

$$\approx \frac{d}{3} \cdot \frac{2^n}{d}$$

$$\begin{aligned} nS &\approx n\left(\frac{e}{d} + 1\right) \approx S_n \cdot \left(1 + \frac{2}{d} + \frac{4}{d}\right) \\ bS &\approx n\left(b + \frac{4}{d}\right) \approx S_n \cdot \left(1 + \frac{2}{d}b + \frac{4}{d}\right) \\ eS &\approx 1 + \frac{2}{d}bS + \frac{4}{d} \end{aligned}$$

Inteligenta artificială

Curs 2

- Rețele de percepție multistat pentru regimuri de decizie corecte (vezi carte pg 79-82)

Următoarele întrebări sunt justificate:

- dându-se o RPM cu o arhitectură fixă cât de complexă topologic pot fi multivile pe care aceasta le poate separa?;
- dându-se două multivile disjuncte, care este arhitectura optimă a unei RPM care le săpare cu o eroare de măclarare minimă (eventual nul?).

Răspunsul la aceste întrebări este dat de următoarea:

Propoziție 5 (formulată ca o conjectură de Hopfmann în 1987) \rightarrow

(i) Fie o RPM cu următoarele arhitecturi:

- trei straturi funcționale ($\gamma = 3$);

- n_1 - numărul de hiperplane ce delimită multivile C_1 și

- n_2 - numărul de multivile disjuncte ce alcătuiesc C_1

$$\text{adică } C_1 = \bigcup_{i=1}^{n_2} R_{2i}.$$

$$- n_3 = 1.$$

- fiecare perceptor de pe stratul 2 (ce corespunde către mulți multivile R_{2i}) este conectat doar cu același perceptor de pe stratul 1 (ce corespunde hiperplanelor ce delimită multivile R_{2i} (adică cele două straturi ascunse nu sunt total interconectate));

- punctile de transfer ale perceptorilor sunt puncte Heaviside cu profil de activare nul, ~~exceptie perceptorul de la mijloc care are profil de activare -0.5~~.

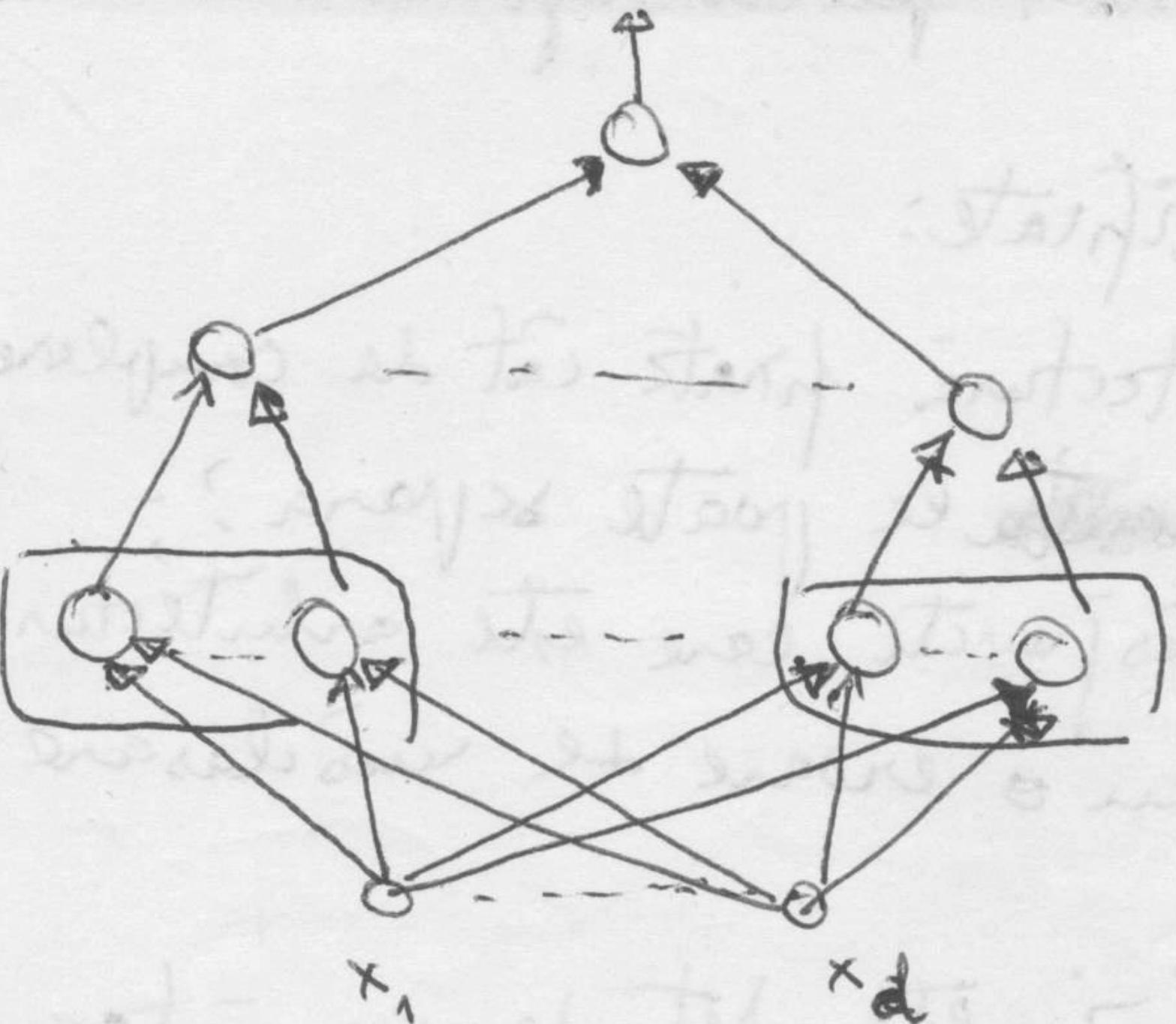
(ii) O RPM cu trei nivele funcționale poate recunoaște și distanțele la un R^d cu eroare de clasare nulă dacă $C_2 = \bigcup_{i=1}^{n_2} R_{2i}$ unde n_2 este fapt încă R_{2i} sunt ~~sub~~ multivile disjuncte, inclusiv sau deschise, delimitate de hiperplane (nu neapărat și convexe).

(iii) Date în rețea de la punctul (ii) punctile de transfer sunt următoarele puncte și anumite parametrii $\beta = \{1, 0\}$ atin multivile ce alcătuiesc regimul ~~Heaviside~~ C_2 pot fi delimitate de suprafețe de puncte continue.

Schemă Schematic, arhitectura mulii astfel de rețele

este date în fig. 12

VERSO



$\leftarrow m_3 = 1$ stratul functional de ieșire

$\leftarrow m_2$ - perceptron ce recunoaște o sau o subperă multime din periferie lui C²

$\leftarrow m_2$ - baterie de percepatori
care baterie continuând astăzi percepțiile hiperplane de liniuță care

\leftarrow stratul de intrare

Fig 12. Arhitectura unui rețelui folosită pt. a demonstra că ~~RPM~~ RPM în trei nivele funcționale și fiabil de transfer și pot recunoaște diuștonii ale șipășilor în topologie arbitrară de complexe. Pentru claritatea imaginii deplasarea perceptorilor nu este indicată.

Inteligenta artificiala

18

Curs 2

Dacă $n_2 = 1$, adică G_2 este formată dintr-o singură mulțime, atunci există cel puțin de al doilea strat ascuns este superfluo.

Ponderile de pe stratul 1 și stratul 2 și de pe stratul 2 și stratul 3 sunt atât egale cu zero. Ponderile de tulcare de pe stratul de intrare pe stratul 1 sunt atât egale cu coeficienții hiperplanelor de delimitare reprezentate $\{R_{2i}\}$ încă deplasările perceptronului de pe stratul 1 în termenii liberi ai acelorași hiperplane. Deplasările perceptronului de pe stratul 2 și 3 sunt egale cu zero.

(Se observă, însă, că ieșirea perceptronului de pe stratul 1

Pentru fiecare regim R_{2i} ecuațiile $\{H_{2i}^k(x) = 0\}_{k=1}^{k_i}$ ale hiperplanelor și o delimitare vor fi astfel sănătoasă încât $(x) \in R_{2i} \Rightarrow H_{2i}^k(x) > 0 \quad (\forall) \quad k=1, k_i \quad (\sum_{i=1}^{k_i} k_i = m_1)$. Deplasările perceptronului de pe stratul 2 sunt atât egale cu $1 - k_i$, $i=1, n_2$ iar deplasările perceptronului de pe stratul de ieșire este egală cu zero.

• Să verificăm cum funcționalitate RPM:

$$\text{există un perceptor } k(x_1, y_1) \text{ de pe stratul 1 și } \\ \text{există } x \in C, \text{ pentru că } (H_{2i}^k(x) \leq 0) \Rightarrow (H_{2i}^k(x) \leq 0) \quad (\forall) \quad k=1, k_i$$

(deci funcția de integrare a tuturor perceptronilor este produsul scalar în \mathbb{R}^d pt. stratul 1, \mathbb{R}^{k_i} pt. stratul 2 și \mathbb{R}^{n_2} pt. stratul 3) iar funcția de transfer este funcție liniarizabilă.

Perceptorul responsabil

- dacă $x \in C$, atunci în fiecare baterie n_2 baterii de căte k_i perceptorii săi (de pe stratul i cu $i=1, n_2$) va exista cel puțin un perceptor $k (= 1 - k_i)$ așa că $H_{2i}^k(x) \leq 0$, $(\forall) i=1, n_2$ și deci $f_k(H_{2i}^k(x)) = 0$, adică ieșirea să fie zero.

În această situație funcția de integrare a tuturor perceptorilor de pe stratul 2 va fi negativă [adică $G_2(x) = n_2(x) + 1 - k_i$]

-/-

Veti vede

-

significativă. Să considerăm o combinație liniară de două astfel de suprafeți în care o doară și jumătate are aceeași orientare la prima, dar nu altă W_0 ; să obținem figura 3.2 (b) din carte). Dacă orientările punctelor sunt ortogonale, primă combinație liniară ale perceptronilor corespunzătoare suprafețelor unei regiuni R_{2i} se obține și cu ajutorul de tipul celui din fig. 3.2 (c) din carte. Crestile de la marginile unor poli vor fi "lăsată" de perceptronul corespondator multivinții R_{2i} din al doilea strat ascuns (vezi fig. 3.2 (d) din carte). Intrările perceptronului de ieșire vor fi, în cazul în care $x \in R_{2i}$ sub forma unei "picături" ca în fig. 3.2 (d).

q.e.d.

Observații: a) Perceptronul de pe stratul 2 simulează funcție de ieșire SI logic (transmitem un impuls dacă datele toate intrările sunt active) în timp ce perceptronul de pe stratul 3 simulează funcție de ieșire SAV logic (transmitem un impuls dacă cel puțin o intrare este activă).

b) Cu toate că demonstrația este construcțivă, valoarea practică și rezultatelor este scăzută, cînd nevoie să se cunoască a priori și frontiere regiunii R_2 .

c) Din demonstrație se poate că o rețea cu date străbunți fizionomice poate separa regiuni de decizie arbitrar de complexe, exemplul din ff. 3.3. din carte [arată că acest lucru nu este totuști adevărat (demonstrația date de Gibson & Cowan 1990; Blum & Li 1991 citată de Bishop 1995)].

d) O sinteză grafică a rezultatelor propoz. 5 este prezentată în fig. 13.

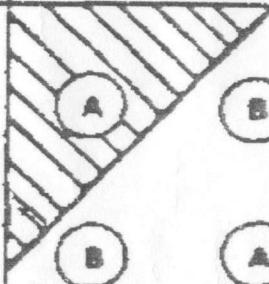
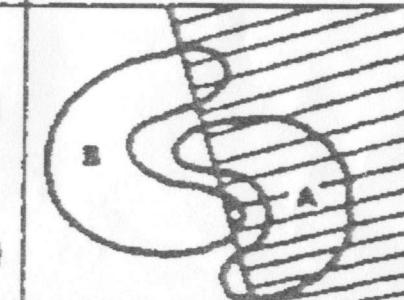
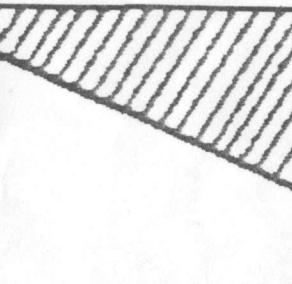
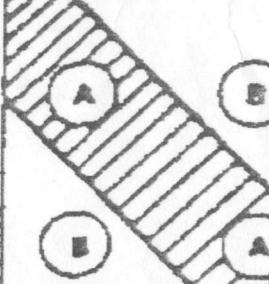
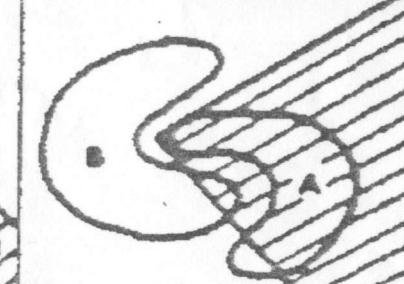
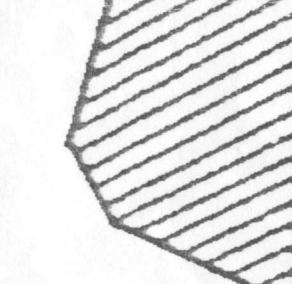
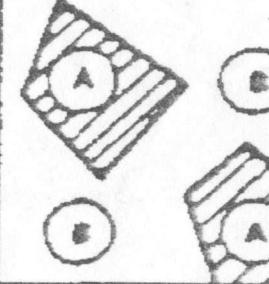
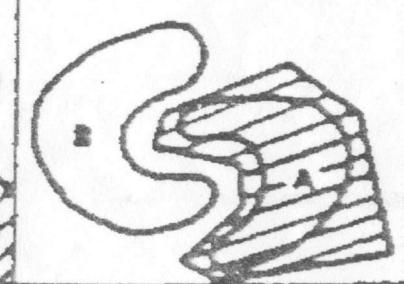
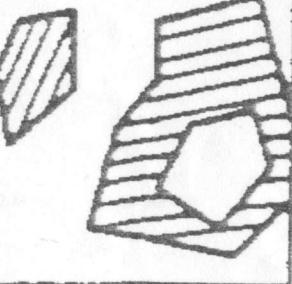
Structura	Tipuri de regiuni de decizie	Problema lui SAU exclusiv	Regiuni rinichiforme	Cea mai generala forma a regiunilor
Un singur strat	Semiplan delimitat de hiperplan			
Doua straturi	Regiuni convexe inchise / deschise			
Trei straturi	Arbitrare (complexitatea depinde de numarul de neuroni)			

Fig. 13. Tipuri de dicotonetări a lui \mathbb{R}^2 și pot fi reamenajate ^{exact} de o RPM ~~cu numeroase dendrese~~