Subiecte date la examenul de I.A, mai 2008, gr 241

(1p of)

1 I. (2p) a) Specificați două proceduri de estimare a lui $R_{real}(h)$. Ce proprietăți are estimatorul furnizat de procedură?

b) Specificați două reguli de învățare robuste (pt. perceptron). Justificați

2 II. (3p) Construiți o RPM standard care să implementeze forma normal disjunctivă a fcției $f: \{0,1\}^2 \to \{0,1\}$ cu $f(x,y) = [(2x+y) \bmod 2]$.

ocazie

III (4p) a) Scrieți ecuațiile de actualizare a ponderilor din alg. backprop pt. o rețea ff cu fcția de transfer tanh, fcția de eroare criteriul corelației folosind ~~alg.~~ metoda ~~of~~ gradientului descendent

Nu pctul b) b) Scrieți în MATLAB subrutina

function [w, b] = algRos (w, b, S)

ce implementează alg Rosenblatt pt. un perceptron cu funcția de transfer signum, antrenat 'off-line', pe o mulțime de antrenare $S=[p;t]$ cu $p \in \mathcal{M}_{d \times n}$ și $t \in \mathcal{M}_{1 \times n}$ cu rata de învățare 0.5. Intrările corespund lui $w^{(t)}$, $b^{(t)}$ iar ieșirile lui $w^{(t+1)}$, $b^{(t+1)}$.

10 pcte

Timp de lucru 1:30 h

– Soluții 1 –

I. (2p)

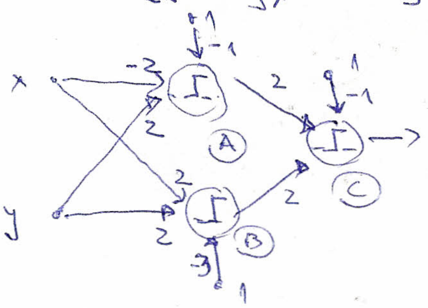
1p a) Estimația lui $R_{real}(h)$ este notată $\hat{R}_{real}(h)$ și calculată prin variante ale metodei validării încrucișate (cross-validation).
Variante ale acestei metode sunt: • n-fold cross-validation
• leave one out method

1p b). Regula lui May care nu permite „apropierea de zero"
• Regula lui Butz care evită „oscilațiile" în planul lui $\underline{w}$
• sau regulile AHK I – III.

II (3p)

1p $f(x,y) = [(2x+y) \bmod 2] = K_1 \vee K_2 = (\neg x \wedge y) \vee (x \wedge y)$ forma normal disjunctivă

1p desenarea rețelei



1p demonstrarea corectitudinii

x	y	f(x,y)	K_1	K_2	$K_1 \vee K_2$	A: G	A: hard	B: G	B: hard	C: G	C: hard
0	0	0	0	0	0	−1	0	−3	0	−1	0
0	1	1	1	0	1	1	1	−1	0	1	1
1	0	0	0	0	0	−3	0	−1	0	−1	0
1	1	1	0	1	1	−1	0	1	1	1	1

III (4p)

a) 2p

$$J(\underline{w}^{(1)}, \underline{w}^{(2)}) = -\sum_{\ell=1}^{L} d_\ell f_0(net_\ell)$$ vezi Aceea din fig 3.1.

$net_\ell = \sum_{j=0}^{J} w_{\ell j}^{(1)} z_j$, $z_j = f_h(net_j)$, $net_j = \sum_{i=0}^{n} w_{ji}^{(2)} x_i$, $f_0(x) = f_h(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$

1p $\underline{W}^{(1)} = \begin{pmatrix} w_{10}^{(1)}, \dots, w_{1J}^{(1)} \\ \vdots \\ w_{L0}^{(1)}, \dots, w_{LJ}^{(1)} \end{pmatrix}$, $\underline{W}^{(2)} = \begin{pmatrix} w_{10}^{(2)}, \dots, w_{1n}^{(2)} \\ \vdots \\ w_{J0}^{(2)}, \dots, w_{Jn}^{(2)} \end{pmatrix}$

mulțimea de antrenare $S = \{(\underline{x}, \underline{d}) \mid \underline{x} = (x_1, \dots, x_n)', \underline{d} = (d_1, \dots, d_L)\}$

Actualizarea ponderilor pe stratul de ieșire

$$\Delta w_{\ell j}^{(1)} = -\rho_0 \frac{\partial J(\underline{W}^{(1)}, \underline{W}^{(2)})}{\partial w_{\ell j}^{(1)}} = \rho_0 d_\ell f_0'(net_\ell) z_j \qquad \ell = \overline{1,L},\ j = \overline{0,J}$$

cu $\rho_0 > 0$ rată de învățare și $f_0'(x) = 1 - f^2(x)$

→ vezi pag 2

→ (Continuare pag 1)

III

a) Actualizarea ponderilor pe stratul ascuns

$$\Delta w_{ji}^{(2)} = -\rho_h \frac{\partial J}{\partial w_{ji}^{(2)}} \qquad j=\overline{1,J} \quad i=\overline{0,n}$$

$$\frac{\partial J}{\partial w_{ji}^{(2)}} = \frac{\partial J}{\partial z_j} \cdot \frac{\partial z_j}{\partial net_j} \cdot \frac{\partial net_j}{\partial w_{ji}^{(2)}}$$

1p

cu $\dfrac{\partial net_j}{\partial w_{ji}^{(2)}} = x_i$; $\dfrac{\partial z_j}{\partial net_j} = f_h'(net_j)$

$$\frac{\partial J}{\partial z_j} = \frac{\partial}{\partial z_j}\left[-\sum_{l=1}^{L} d_l f_o(net_l)\right] = -\sum_{l=1}^{L} d_l \frac{\partial f_o}{\partial z_j} = \sum_{l=1}^{L} d_l f_o'(net_l) w_{lj}^{(1)}$$

$$\Rightarrow \Delta w_{ji}^{(2)} = \rho_h \left[\sum_{l=1}^{L} d_l f_o'(net_l) w_{lj}^{(1)}\right] f_h'(net_j) x_i$$

cu $\rho_h > 0$ rata de învățare și $f_h'(x) = 1 - f^2(x)$

b)

```
function [w,b]= algRos (w, b, S)
[d1,n]= size (S);
d1 = d-1;
p = S[1:d,:];
t = S[d1,:];
a = hardlims (w*p, b)
e = (t-a)/2;
dw = e*p';
db = e* ones(n,1);
w = w + dw;
b = b + db;
```