

10p

2p

I. Definiți, în cazul învățării supervizate, mulțimea de antrenare și mulțimea de validare. Desenați și explicați în acest caz graficul erorilor de antrenare și de validare. Precizați momentul "early stopping".

acord

3p

II. Arătați că pentru o RPN cu funcție de transfer liniară și funcție de eroare suma pătratelor erorilor, antrenată backprop, originea spatelui de prudență este un pit. Staționar și funcție de eroare. (se folosesc notatunile din fig 3.1 curs)

III. Fie funcția booleană XOR

a) Scrieți în cod Matlab matricea p a vectorilor din date de def. și vectorul t a val. funcției XOR.

b) Reprezentați graficul RPN ce implementează funcția XOR folosind topologia "sandwich". Justificați răspunsul.

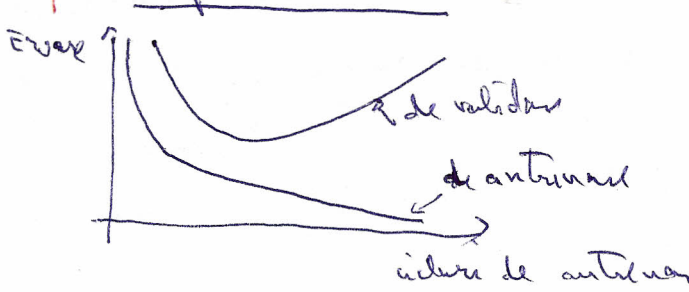
c) Scrieți instrucțiunile Matlab ce ~~realizează~~ implementează rețeaua de la punctul b)

10p

Țiimp de lucru 1230 h

10

Soluții subiectele NR 1

- 2p I. 0.5p** S.n. multime de antrenare partea din multimea de învățare care este utilizată pt. calculul ponderilor.
- 0.5p** S.n. multime de validare partea din multimea de învățare, complementară în multimea de antrenare utilizată pt. evaluarea erorii de generalizare.
- 0.5p** Grăful erori
- 
- Prin antrenare repetată, eroarea pe multimea de antrenare scade (fenomen 'overfitting') în timp ce eroarea de generalizare, estimată pe multimea de validare, crește după care începe să scadă din nou.
- 0.5p** Momentul 'early stopping' este cel în care este atins minimul erorii de validare.

- 3p II. 1p** Pt. staționare $(w_0^{(1)}, w_0^{(2)}) \equiv$ pt. pt. care $\frac{\partial E}{\partial w_{ij}^{(1)}} \bigg|_{(w_0^{(1)}, w_0^{(2)})} = 0 = \frac{\partial E}{\partial w_{ji}^{(2)}}$ cu notată pt. rețeaua din fig 3.1.
- 1p** Pt. straturilor de ieșire

$$\frac{\partial E}{\partial w_{ij}^{(1)}} = - (d_e - y_e) \underbrace{f'_0(\text{net}_e)}_{=1} z_j \quad \text{pt } e=1, \overline{L} \quad \text{cu } z_j = f_h(\text{net}_j) \quad \text{și } \text{net}_j = \sum_{i=0}^n w_{ji}^{(2)} x_i$$

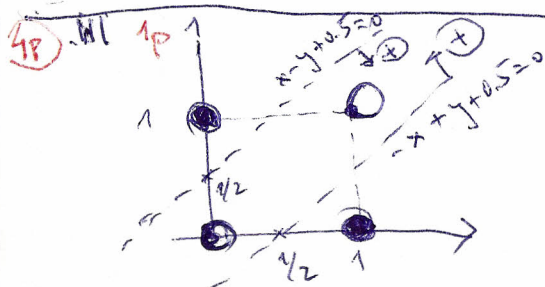
Cum $w_{ji}^{(2)} = 0, j=1, \overline{J}$ și $i=0, \overline{n}$ din ip.

$$\Rightarrow \text{net}_j = 0 \Rightarrow z_j = 0 \text{ (cui } f_h(x) = x \Rightarrow \frac{\partial E}{\partial w_{ij}^{(1)}} \bigg|_0 = 0$$

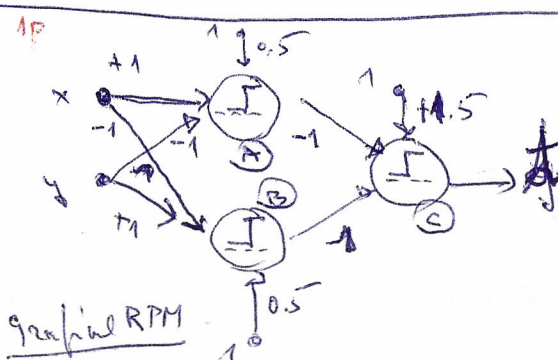
1p Pt. straturilor ascunde

$$\frac{\partial E}{\partial w_{ji}^{(2)}} = \left[\sum_{e=1}^L (d_e - y_e) \underbrace{f'_0(\text{net}_e)}_{=1} w_{ij}^{(1)} \right] \underbrace{f'_h(\text{net}_j)}_{=1} x_i \quad \text{cu } i=0, \overline{n}$$

$$\text{Cum } w_{ij}^{(1)} = 0, e=1, \overline{L}, j=1, \overline{J} \text{ din ip} \Rightarrow \frac{\partial E}{\partial w_{ji}^{(2)}} \bigg|_0 = 0$$



Th. 'sandwich'



Grăful RPM

2p justificare

	x	y	XOR	A	B	C
0	0	0	0	1/2	1/2	1
0	1	1	0	1/2	1/2	1
1	0	1	1	1/2	1/2	0
1	1	0	1	1/2	1/2	0