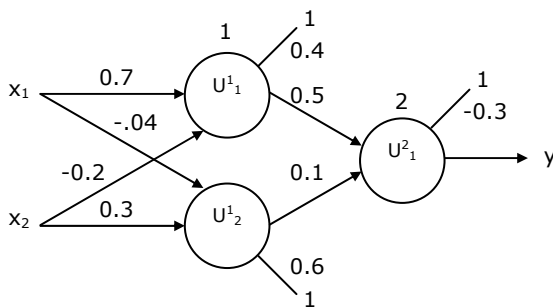


XARXES MULTI-CAPA

PROBLEMA 1

Estem entrenant la xarxa BPN de la figura amb un coeficient d'aprenentatge $\eta = 0.25$.

- Troba els pesos quan es presenta a la xarxa el patró d'entrenament $\{(0, 1), 1\}$, suposant una funció d'activació sigmoïdal binària.
- Troba els pesos quan es presenta a la xarxa el patró d'entrenament $\{(-1, 1), 1\}$, suposant una funció d'activació sigmoïdal bipolar.



PROBLEMA 2

Volem entrenar una xarxa MLP per classificar quatre classes diferents, definides pels centroides $\{(0, 1, 1), (1, 1, 0), (0, 0, 1) \text{ i } (1, 0, 1)\}$ en l'espai tridimensional de solucions. Proposa una arquitectura de xarxa per resoldre el problema.

- Defineix el tipus (bipolar, binària...) i número d'entrades i sortides de la xarxa.
- Utilitzaries els patrons d'entrenament donats tal i com estan? Perquè?

PROBLEMA 3

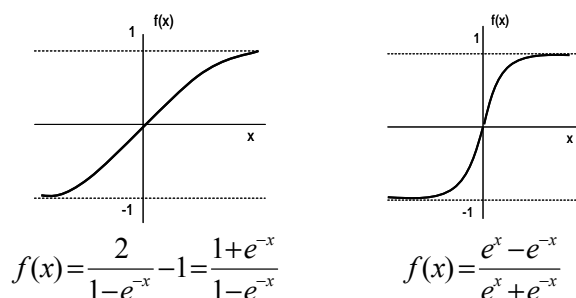
Volem entrenar una xarxa BPN per classificar quatre classes diferents, definides pels vectors $\{(0, 1, 100), (1, 1, 0), (0, 1, 0) \text{ i } (1, 0, 100)\}$ en l'espai tridimensional de solucions. Creus que la xarxa aprendrà bé? Per què? Proposa una solució senzilla per solucionar aquest problema.

PROBLEMA 4

La estandardització i la normalització relativa ens traslladen el conjunt d'entrades a l'interval $[0, 1]$, però l'estandardització ens proporciona a més una distribució de dades amb mitjana $m=0$ i variància $\sigma=1$ que la normalització relativa no ens garanteix. Té algun avantatge d'aquesta distribució?

PROBLEMA 5

La sigmoïdal bipolar i la tangent hiperbòlica són algunes de les funcions que s'utilitzen habitualment en xarxes neurals. Ambdues són contínues, no lineals, derivables i amb derivada relativament senzilla. Se t'ocorre alguna altra propietat que compleixin i que les faci indicades per xarxes neurals?



PROBLEMA 6

Volem implementar l'algorisme BackPropagation en una arquitectura multi-processorador. És possible fer-ho? Quins problemes planteja?

PROBLEMA 7

Volem accelerar la convergència de l'algorisme BackPropagation amb un coeficient progressiu d'aprenentatge $\eta_i = \eta_{i-1} \cdot \lambda$, on $\lambda = 0.95$, $\eta_0 = 1$ i i correspon al número de èpoques. Quins problemes presenta aquesta modificació? Com es pot millorar?

PROBLEMA 8

Més enllà de malgastar memòria i temps de procés, quin sentit podria tenir utilitzar coeficients d'aprenentatge individuals per cada pes de la xarxa?

PROBLEMA 9

Suposem una xarxa de tres capes ben inicialitzada (e.g. Nguyen-Widrow), amb entrades i sortides estandarditzades i entrenada per BackPropagation amb coeficient adaptatiu maximal. Aprendran tots els pesos per igual? Per què? Com es pot arreglar?

PROBLEMA 10

Per què no es pot utilitzar l'error absolut de la xarxa com un criteri de terminació eficaç? I l'error relatiu?

PROBLEMA 11

Per què és convenient un 3^{er} conjunt de validació (apart d'un conjunt d'entrenament i un de test) de cara a millorar la generalització quan apliquem terminació per aturada anticipada?

PROBLEMA 12

Estem entrenant una xarxa i es produeix memorització. Provem l'aturada anticipada, però no es soluciona problema. Provem de subdimensionar la xarxa, però llavors no apren bé. Se'ns ocorre introduir soroll a la xarxa per millorar l'aprenentatge. Millorarà?

PROBLEMA 13

Volem modificar l'algorisme BackPropagation perquè re-dimensioni automàticament la xarxa i eviti la memorització. Proposa una solució senzilla i computacionalment poc costosa per aconseguir-ho. Quan és aconsellable aplicar una tècnica d'aquesta mena?