Examen de los temas 5 a 7 de Aprendizaje Automàtico

ETSINF, Universitat Politècnica de València, 16 de enero de 2014

Apellidos: **Nombre:**

Cuestiones (2 puntos; tiempo estimado: 30 minutos)

Marca cada recuadro con una única opción de entre las dadas.

1 $\boxed{\text{A}}$ Si la función de acción no lineal en la capa de salida de un perceptrón de dos capas fuese lineal, las fórmulas que permiten modificar los pesos de dicha capa de salida en el algoritmo BackProp verifican que (solo una respuesta es correcta):

A) $\Delta\theta^2_{ij} = \rho\ (t_i - s^2_i)\ s^1_j$
B) $\Delta\theta^2_{ij} = \rho\ (t_i - s^2_i)\ g(\phi^2_i)\ (1 - g(\phi^2_i))\ s^1_j$
C) $\Delta\theta^2_{ij} = \rho\ (t_i - s^2_i)\ g(\phi^2_i)\ s^1_j$
D) $\Delta\theta^2_{ij} = \rho\ (t_i - s^2_i)\ (1 - g(\phi^2_i))\ s^1_j$

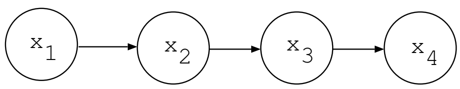
2 $\boxed{\text{B}}$ En una presentación de las N muestras de aprendizaje mediante el algoritmo de retropropagación del error, indicar qué afirmación es la correcta:

A) Los pesos se modifican una sola vez tanto en la versión "online" o incremental como en la versión batch.
B) Los pesos se modifican N veces en la versión "online" o incremental y una en la versión batch.
C) Los pesos se modifican N veces en la versión "online" y también en la versión batch.
D) Los pesos no se modifican en la versión "online" o incremental y una en la versión batch.

3 $\boxed{\text{B}}$ En el perceptrón multicapa, la parálisis de la red se produce cuando (solo una respuesta es correcta)

A) En test, los pesos son muy grandes
B) En entrenamiento, los valores de las combinaciones lineales de los nodos son muy grandes
C) En test, los valores de las combinaciones lineales de los nodos son muy grandes
D) En entrenamiento, los pesos son nulos

4 $\boxed{\text{A}}$ En la red bayesiana lineal

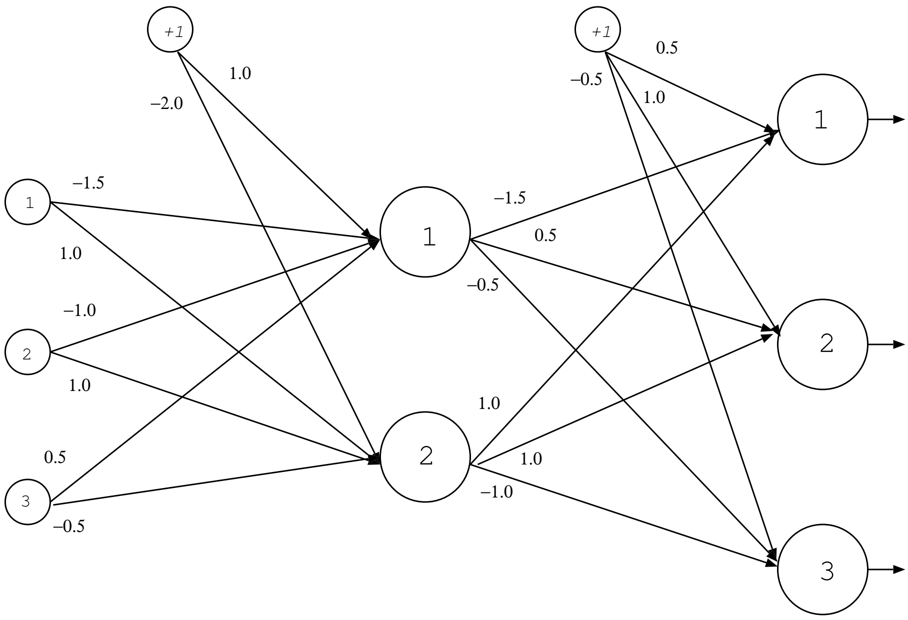


¿cuál de las relaciones siguientes es correcta?

A) $P(x_2 \mid x_1, x_3, x_4) \ = \ P(x_2 \mid x_1, x_3)$
B) $P(x_2 \mid x_1, x_3, x_4) \ = \ P(x_2 \mid x_1)$
C) $P(x_2 \mid x_1, x_3, x_4) \ = \ P(x_2 \mid x_3, x_4)$
D) $P(x_2 \mid x_1, x_3, x_4) \ = \ P(x_2 \mid x_2)$

Problema 1 (4 puntos; tiempo estimado: 30 minutos)

En el perceptrón multicapa de la figura con funciones sigmoid en todos los nodos



se aplica el algoritmo de retropropagación del error para la muestra de entrenamiento $(\boldsymbol{x}, \boldsymbol{t})$ con $\boldsymbol{x} = (1, 0, -2)$ y $\boldsymbol{t} = (0, 1, 0)$ con un factor de aprendizaje $\rho = 1.0$ y en la fase ascendente se obtienen las siguientes salidas (las datos numéricos están redondeados a la primera cifra decimal para facilitar el cálculo)

- Capa oculta: $\begin{array}{ll} \phi_1^1 = -1.5 & s_1^1 = 0.2 \\ \phi_2^1 = 0.0 & s_2^1 = 0.5 \end{array}$
- Capa de salida: $\begin{array}{ll} \phi_1^2 = 0.2 & s_1^2 = 0.5 \\ \phi_2^2 = 0.6 & s_2^2 = 0.7 \\ \phi_3^2 = -0.6 & s_3^2 = 0.3 \end{array}$

a) Calcular el error que se observan en el nodo 2 de la capa oculta

b) Calcular los pesos finales del nodo 1 al nodo 2 de la capa de salida y del nodo 1 al nodo 2 de la capa de oculta después de una iteración

c) Si las funciones de activación de la capa de salida fuesen lineales y las salidas de la capa de salida para $\boldsymbol{x}$ fueran $s_1^2 = 0.6$ $s_2^2 = 0.6$ $s_3^2 = 0.4$ y las salidas de la capa de oculta fueran $s_1^1 = 0.2$ $s_2^1 = 0.5$, calcular los pesos finales del nodo 1 al nodo 2 de la capa de salida después de una iteración

a) Errores en la capa salida:

$\delta_1^2 = (t_1 - s_1^2)\ g'(\phi_1^2) = (t_1 - s_1^2)\ s_1^2\ (1 - s_1^2) = -0.125$

$\delta_2^2 = (t_2 - s_2^2)\ g'(\phi_2^2) = (t_2 - s_2^2)\ s_2^2\ (1 - s_2^2) = 0.063$

$\delta_3^2 = (t_3 - s_3^2)\ g'(\phi_3^2) = (t_3 - s_3^2)\ s_3^2\ (1 - s_3^2) = -0.063$

Errores en el nodo 2 de la capa oculta:

$\delta_2^1 = \left(\delta_1^2\ \theta_{12}^2 + \delta_2^2\ \theta_{22}^2 + \delta_3^2\ \theta_{32}^2\right)\ g'(\phi_2^1) = \left(\delta_1^2\ \theta_{12}^2 + \delta_2^2\ \theta_{22}^2 + \delta_3^2\ \theta_{32}^2\right)\ s_2^1\ (1 - s_2^1) = 0.00025$

b) Modificaciones de los pesos:

Peso del nodo 1 de la capa oculta al nodo 2 de la capa de salida:

$\Delta\theta_{21}^2 = \rho\ \delta_2^2\ s_1^1 = 0.0126 \Rightarrow \theta_{21}^2 = 0.5 + 0.0126 = 0.5126$

Peso del nodo 1 de la capa de entrada al nodo 2 de la capa oculta:

$\Delta\theta_{21}^1 = \rho\ \delta_2^1\ x_1 = 0.00025 \Rightarrow \theta_{21}^1 = 1.0 + 0.00025 = 1.00025$

c) Caso de funciones de activación lineal en la capa de salida:

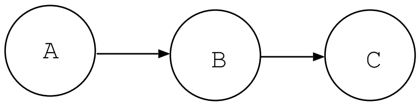
$\delta_2^2 = (t_2 - s_2^2) = 1 - 0.6 = 0.4$

Peso del nodo 1 de la capa oculta al nodo 2 de la capa de salida:

$\Delta\theta_{21}^2 = \rho\ \delta_2^2\ s_1^1 = 0.08 \Rightarrow \theta_{21}^2 = 0.5 + 0.08 = 0.58$

Problema 2 (4 puntos; tiempo estimado: 30 minutos)

En la red bayesiana de la figura, las variables A, B y C toman valores en el conjunto $\{0, 1\}$ y las distribuciones de probabilidad asociadas son:



$$\begin{aligned} P(A = 1) &= 0.3 & & \\ P(B = 1 \mid A = 1) &= 0.2 & P(B = 1 \mid A = 0) &= 0.9 \\ P(C = 1 \mid B = 1) &= 0.1 & P(C = 1 \mid B = 0) &= 0.4 \end{aligned}$$

a) Calcular $P(B = 0 \mid A = 1, C = 0)$

b) Calcular $P(A = 1 \mid B = 1, C = 1)$

$$P(A, B, C) = P(A)\ P(B \mid A)\ P(C \mid B)$$

a) Cálculo de $P(B = 0 \mid A = 1, C = 0)$:

$$\begin{aligned} P(B = 0 \mid A = 1, C = 0) &= \frac{P(A = 1, B = 0, C = 0)}{P(A = 1, C = 0)} \\ &= \frac{P(A = 1)P(B = 0 \mid A = 1)P(C = 0 \mid B = 0)}{P(A = 1) \sum_{b \in 0,1} P(B = b \mid A = 1)P(C = 0 \mid B = b)} \\ &= \frac{0.3 \cdot 0.8 \cdot 0.6}{0.3(0.8 \cdot 0.6\ +\ 0.2 \cdot 0.9)} = \frac{0.144}{0.198} = 0.727 \end{aligned}$$

b) Cálculo de $P(A = 1 \mid B = 1, C = 1)$. Como A y C son independientes cuando se conoce B:

$$\begin{aligned} P(A = 1 \mid B = 1, C = 1) &= P(A = 1 \mid B = 1) \\ &= \frac{P(A = 1, B = 1)}{P(B = 1)} \\ &= \frac{\sum_{c \in \{0,1\}} P(A = 1, B = 1, C = c)}{\sum_{a,c \in \{0,1\}} P(A = a, B = 1, C = c)} \\ &= \frac{P(A = 1)\ P(B = 1 \mid A = 1)\ \sum_{c \in \{0,1\}}\ P(C = c \mid B = 1)}{\sum_{a \in \{0,1\}} P(A = a)\ P(B = 1 \mid A = a)\ \sum_{c \in \{0,1\}} P(C = c \mid B = 1))} \\ &= \frac{P(A = 1)\ P(B = 1 \mid A = 1)}{\sum_{a \in \{0,1\}} P(A = a)\ P(B = 1 \mid A = a)} \\ &= \frac{0.3\ 0.2}{0.3\ 0.2 + 0.7\ 0.9} = \frac{0.06}{0.69} = 0.086957 \end{aligned}$$