

OPENCOURSEWARE
REDES DE NEURONAS ARTIFICIALES
Inés M. Galván – José M. Valls



Preguntas y ejercicios para Evaluación: Tema 2

1. Responda brevemente a las siguientes preguntas:
 - a) ¿Qué representan los pesos y el umbral del ADALINE?
 - b) ¿Qué representan los pesos y el umbral del Perceptron Simple?

Respuesta

- a) Los coeficientes de la recta que aproxima a los patrones dados
 - b) Los coeficientes del hiperplano que separa los patrones en dos clases
2. La ley de aprendizaje del Perceptron Simple cuando la salida deseada ($d(x)$) y la salida obtenida ($s(x)$) no coinciden, viene dada por $w_i(t+1) = w_i(t) + \alpha \cdot d(x) \cdot x_i$, donde x_i es la coordenada i del patrón de entrada x , $d(x)$ es la salida deseada para el patrón x y α es la razón de aprendizaje. Supóngase que se decide utilizar como ley de aprendizaje para el Perceptron Simple la siguiente expresión:

$$w_i(t+1) = w_i(t) + \alpha \cdot (d(x) - s(x)) \cdot x_i,$$

donde $s(x)$ es la salida del Perceptron simple para el patrón de entrada x . ¿Se trata de la misma ley de aprendizaje anterior? Razone su respuesta

Respuesta

Ambas ecuaciones de aprendizaje modifican los pesos siguiendo la misma ley. Si el patrón está bien clasificado, con ambas leyes no se modifican los pesos (para la segunda $(d(x) - s(x))$ vale 0). Si el patrón no está bien clasificado, con ambas leyes los pesos sufren una modificación (positiva o negativa) proporcional a la entrada; la diferencia está en que para la primera la modificación viene dada $\pm \alpha \cdot x_i$ y para la segunda $\pm 2 \alpha \cdot x_i$. Por tanto, son equivalentes

3. Describa brevemente el proceso de aprendizaje del ADALINE.

Respuesta:

Paso 1: Inicializar los pesos y umbral de forma aleatoria

Paso 2: Presentar un patrón de entrada

Paso 3: Calcular la salida, compararla con la deseada y obtener la diferencia: $(d^p - y^p)$

Paso 4: Para todos los pesos y para el umbral, calcular:

$$\Delta_p w_j = \gamma (d^p - y^p) x_j \quad \Delta_p \theta = \gamma (d^p - y^p)$$

Paso 5: Modificar los pesos y el umbral del siguiente modo:

$$w_j(t+1) = w_j(t) + \Delta_p w_j$$

$$\theta(t+1) = \theta(t) + \Delta_p \theta$$

Paso 6: Repetir los pasos 2, 3, 4 y 5 para todos los patrones de entrenamiento (1 ciclo)

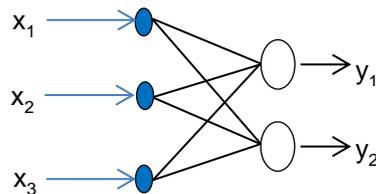
Paso 7: Repetir los pasos 2, 3, 4, 5 y 6 tantos ciclos hasta cumplir el criterio de parada

4. Dado un problema de clasificación no lineal, en ocasiones es posible encontrar una solución al problema combinando varios Perceptrones Simples, como es el caso de la función XOR. ¿Por qué se dice entonces que el Perceptron Simple es apropiado únicamente para problemas lineales?

Respuesta

Porque la ley de aprendizaje del PS no puede aplicarse a la combinación de varios perceptrones, por lo que los pesos habría que ajustarlos manualmente, no disponiendo de un procedimiento automático.

5. Dada la siguiente arquitectura de red con función de activación $f(x)=x$, escriba la ecuación de la salida de la red y la ley de aprendizaje para actualizar los pesos y los umbrales de la red. ¿Qué tipo de problemas podrían ser abordados por esta arquitectura de red?



Respuesta

Salidas:

$$y_1 = x_1 \cdot w_{11} + x_2 \cdot w_{21} + x_3 \cdot w_{31} + \theta_1$$

$$y_2 = x_1 \cdot w_{12} + x_2 \cdot w_{22} + x_3 \cdot w_{32} + \theta_2$$

Pesos. Nombrando w_{ij} el peso de la entrada i a la salida j ($i=1,2,3; j=1,2$), la ley viene dada por:

$$w_{ij}(t+1) = w_{ij}(t) + \alpha(y_j - d_j) \cdot x_i$$

$$\theta_j(t+1) = \theta_j(t) + \alpha(y_j - d_j)$$

Problemas de regresión lineal con dos variables de salida. También problemas de clasificación lineal, interpretando las salidas como clases. Es decir, en un problema por ejemplo de 2 clases, las salidas deseadas serían (0,1) y (1,0). Entonces, dado un patrón de entrada, se calcula el índice para el cual se obtiene el máximo de (y_1, y_2) . Ese índice será la clase

6. Describa las diferencias y similitudes del PS y del ADALINE

Respuesta

Similitudes: Misma arquitectura a diferencia de la función de activación de la salida. N entradas y una salida

Diferencias: La ley de aprendizaje, en el ADALINE los pesos se corrigen en base al error evaluado en la salida de la red. En el PS, los pesos se corrigen si el patrón está mal clasificado, moviendo el hiperplano para clasificar correctamente el patrón.

El ADALINE sirve para resolver problemas de regresión lineal, mientras que el perceptron para problemas de clasificación lineal.

7. Dado los problemas de clasificación mostrados en la figura 1 y los problemas de regresión mostrados en la figura 2, dibuje posibles soluciones que podrían proporcionar el PS para los problemas de clasificación y el Adaline para los problemas de regresión. ¿Qué influye en la obtención de una determinada solución y no otra?

Figura 1. Problemas de Clasificación

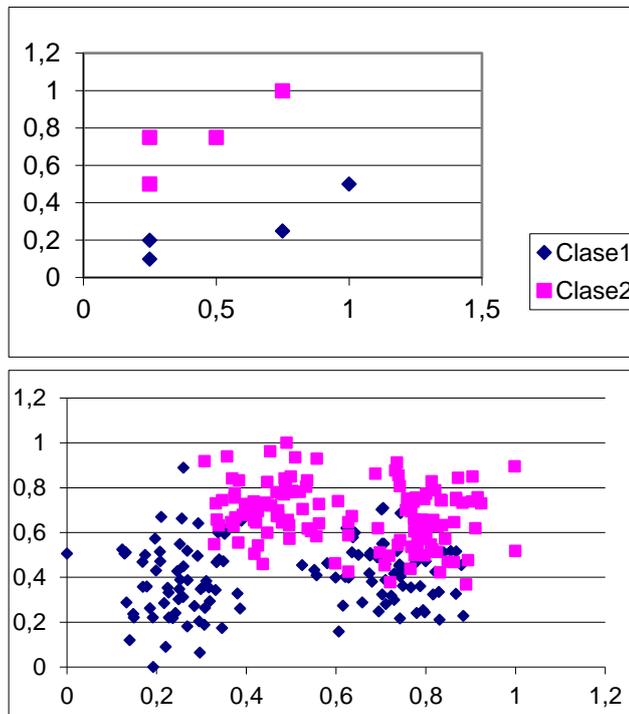
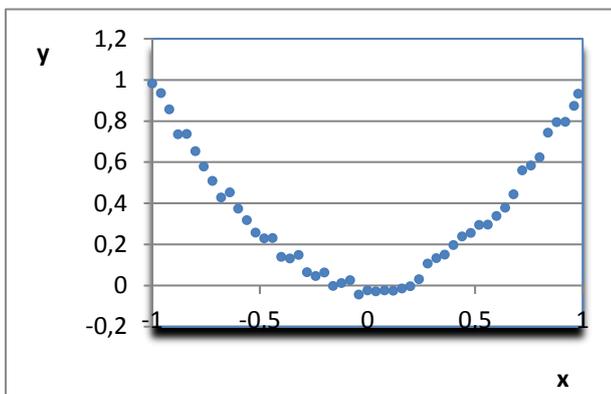
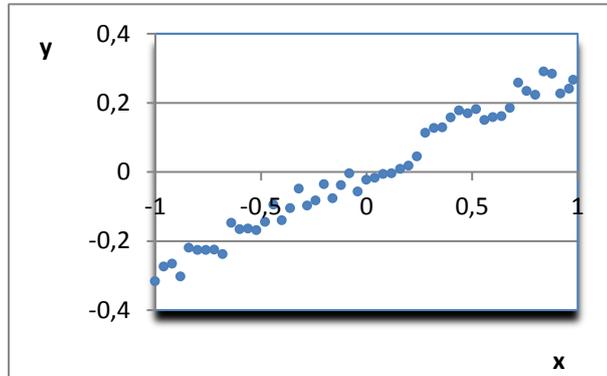
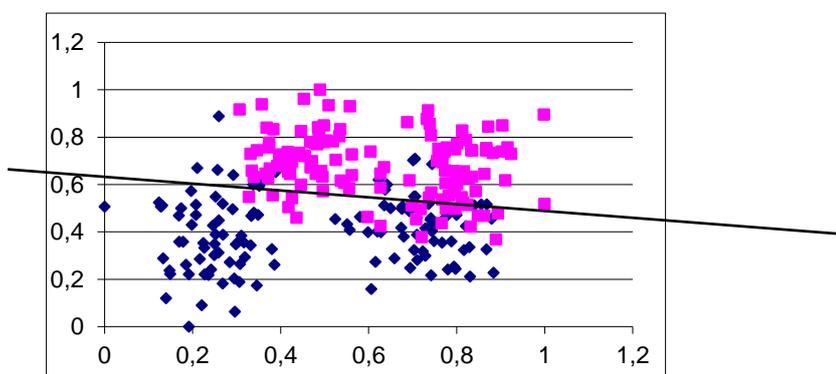
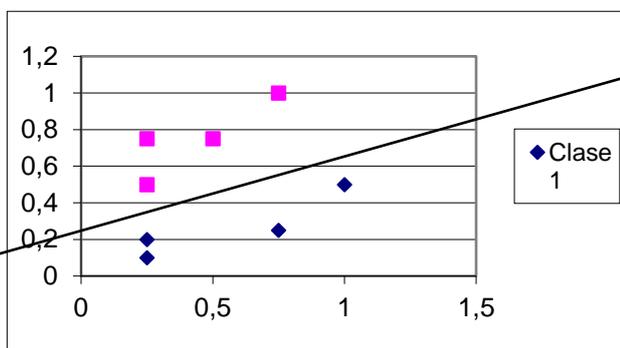


Figura 2. Problemas de Regresión

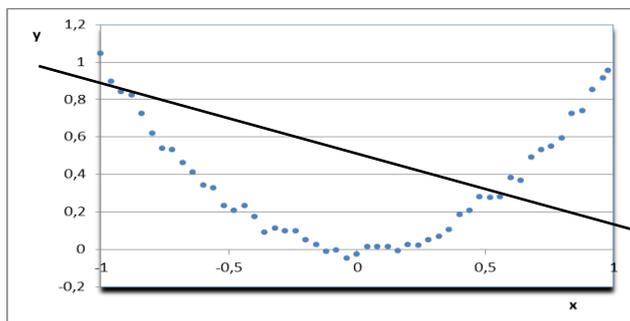
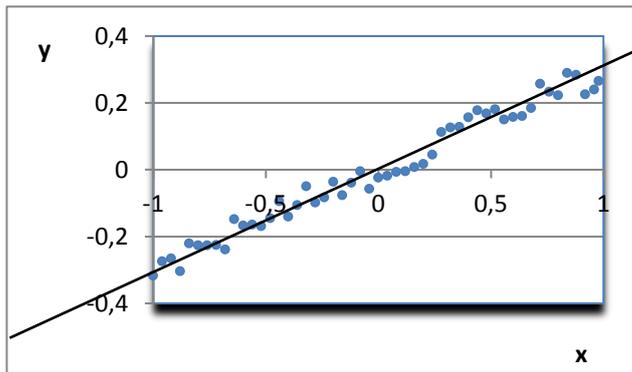


Respuesta

Problemas de Clasificación



Problemas de Regresión



Los pesos iniciales y el orden de presentación de los patrones

8. Dado los siguientes patrones de entrenamiento y test, construya un PS para separarlos en dos clases, partiendo de los siguiente valores iniciales:

$W1=0; W2=1; u=-0.4.$

¿El PS construido clasifica correctamente los patrones de test?

PATRONES DE ENTRENAMIENTO			PATRONES DE TEST		
0,25	0,1	$d(x)=-1$	0,5	1	$d(x)=1$
0,25	0,2	$d(x)=-1$	0,5	0,6	$d(x)=1$
1	0,5	$d(x)=-1$			
0,75	0,25	$d(x)=-1$			
0,25	0,75	$d(x)=1$			
0,25	0,5	$d(x)=1$			
0,5	0,75	$d(x)=1$			
0,75	1	$d(x)=1$			

Respuesta:

Se utiliza la siguiente ley de aprendizaje: $w_j(t+1) = w_j(t) + d(x) \cdot x_j$

Patrón de entrada (0.25, 0.1), la salida del PS es $y=f(-0.3)=-1$. Patrón bien clasificado

Patrón de entrada (0.25, 0.2), la salida del PS es $y=f(-0.2)=-1$. Patrón bien clasificado

Patrón de entrada (1, 0.5), la salida del PS es $y=f(0.1)=1$. Patrón mal clasificado. Se modifican los pesos

$$W1=0+(-1)*1=-1 \quad W2=1+(-1)*0.5=0.5 \quad u=-0.4+(-1)=-1.4.$$

Con estos nuevos pesos la salida para (1, 0.5) es $y=f(-2.5)=-1$. Patrón bien clasificado

Patrón de entrada (0.75, 0.25), la salida del PS con los nuevos pesos es $y=f(-2.025)=-1$.

Patrón bien clasificado

Patrón de entrada (0.25, 0.75), la salida del PS con los nuevos pesos es $y=f(-1.275)=-1$.

Patrón mal clasificado

Se modifican los pesos

$$W1=-1+1*0.25=-0.75 \quad W2=0.5+1*0.75=1.25 \quad u=-1.4+1=0.4.$$

Con estos nuevos pesos la salida para (0.25, 0.75), es $y=f(0.35)=1$. Patrón bien clasificado

Patrón de entrada (0.25, 0.5), la salida del PS con los nuevos pesos es $y=f(0.0375)=1$.

Patrón bien clasificado

Patrón de entrada (0.5, 0.75), la salida del PS con los nuevos pesos es $y=f(0.1625)=1$.

Patrón bien clasificado

Patrón de entrada (0.75, 0.1), la salida del PS con los nuevos pesos es $y=f(0.2875)=1$.

Patrón bien clasificado

Es necesario hacer una nueva pasada por todos los patrones, para ver si patrones que antes estaban bien clasificados, lo siguen estando.

Con los últimos pesos obtenidos, el PS acierta para todos los patrones. Por tanto la ecuación del hiperplano (PS) que separa los patrones de entrenamiento en dos clases es:

$$-0.75*x1+1.25*x2+0.4=0$$

Respecto a los patrones de test, se puede decir que:

Patrón (0.5, 1), la salida del PS es $y=f(0.47)=1$. Patrón bien clasificado

Patrón (0.5, 0.6), la salida del PS es $y=f(-0.025)=-1$. Patrón mal clasificado