Redes Neuronales

Preceptrón y Adeline

carlos.andres.delgado@correounivalle.edu.co

Carlos Andrés Delgado S.

Facultad de Ingeniería. Universidad del Valle

Septiembre de 2017

Contenido

Contenido

Preceptrón

Definición

- Fue introducido por Rossenblat a finales de los año 50
- Se inspira en los procesos de aprendizaje de los animales (ejemplo la visión), en los cuales la información va atravesando diferentes capas de neuronas
- Es un modelo unidireccional, compuesto por dos capas de neuronas, una de entrada y otra de salida
- La operación de este tipo puede darse con n neuronas de entrada y m de salida

Universidad del Valle

Preceptrón

Definición

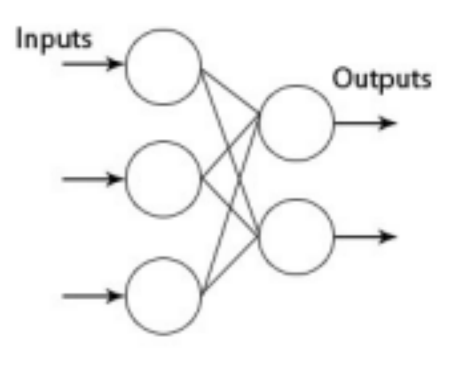

Figura: Modelo de Preceptrón, tomado de
`http://neuroph.sourceforge.net/tutorials/Perceptron.html`

Preceptrón

Definición

- Las neuronas de entrada no realizan ningún computo
- Se consideran señales discretas 0 o 1
- La operación para n neuronas de entrada y m de salida puede considerarse así:

$$y_i = H(\sum_{j=1}^{n} w_{ij}x_j - \Theta_i), \forall i, 1 \leq i \leq m$$

Donde H(x) es la función escalón.

Universidad del Valle

Preceptrón

Definición

- El preceptrón permite clasificar dos conjuntos linealmente separables en un plano o hiperplano
- La respuesta de la neurona es 1 si pertenece a la clase o 0 si no pertenece

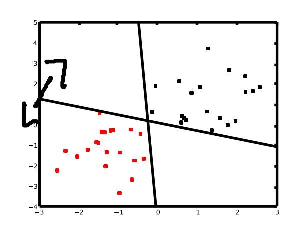

Figura: Conjunto linealmente separable, tomado de `https://en.wikipedia.org/`

Preceptrón

Ejemplo

- Sea una neurona tipo perceptron con entrada x_1 y x_2
- Entonces la operación se define como:

$$y = H(w_1 x_1 + w_2 x_2 - \Theta)$$

Preceptrón

Ejemplo

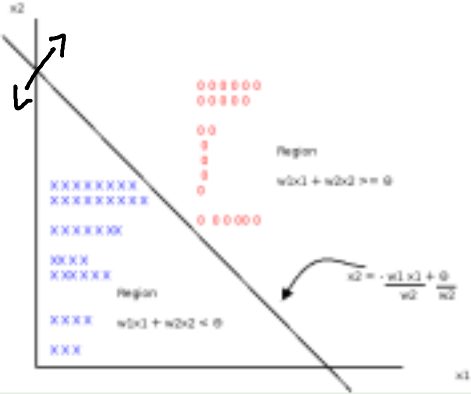

Figura: Regiones de decisión del plano, tomado de [Brio and Molina, 2005]

Preceptrón

Definición

Como se puede ver se divide el plano en dos regiones. Como se puede ver se requiere que el problema a solucionar tenga **solución lineal**

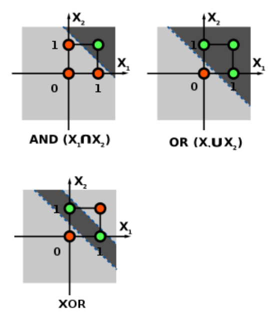

Figura: Casos de compuertas `https://en.wikipedia.org/`

Preceptrón

Algoritmo de aprendizaje

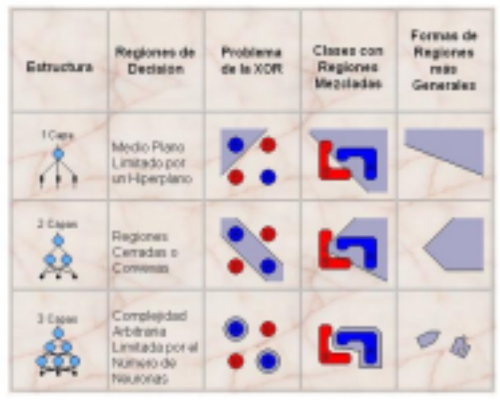

Figura: Regiones de decisión perceptron [Lippmann, 1988]

Preceptrón

Algoritmo de aprendizaje

Vamos a trabajar el perceptrón de una capa

- Se basa en la corrección de errores
- Vamos a introducir una tasa de aprendizaje ϵ: Indica el ritmo de aprendizaje
- Dados unos patrones x^u, salidas obtenidas y^u y salidas deseadas t^u
- Los pesos iniciales son aleatorios entre -1 y 1. Se utiliza la función de activación signo y entradas $\{-1, 1\}$.
- Se examina cada patrón y aplicamos la relación de cambio:

$$\Delta w_{ij}^u(t) = \epsilon.(t_i^u - y_i^u)x_j^u \left\{ \begin{array}{l} 2\epsilon \\ -2\epsilon \end{array} \right.$$

A esto se le conoce como **regla del perceptrón**

Preceptrón

Algoritmo de aprendizaje

La función signo (sgn) se define así:

$$sgn(n) = \begin{cases} 1 & si \quad n > 0 \\ 0 & si \quad n = 0 \\ -1 & si \quad n < 0 \end{cases}$$

Preceptrón

Algoritmo de aprendizaje

Para comprender el preceptrón se mostrará en una forma gráfica

$$y_i^u(t) = sgn(\sum_{j=1}^{n} w_{ij} x_j^u - \Theta_i) = sgn(||w_i||.||x^u||cos(\phi))$$

Preceptrón

Algoritmo de aprendizaje

$$\Delta w = \epsilon\,\underbrace{(t-y)}\,x$$

1   -1

-1   1

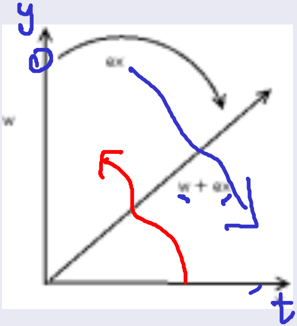

Figura: Aplicación regla perceptron [Brio and Molina, 2005]

Preceptrón

Algoritmo de aprendizaje

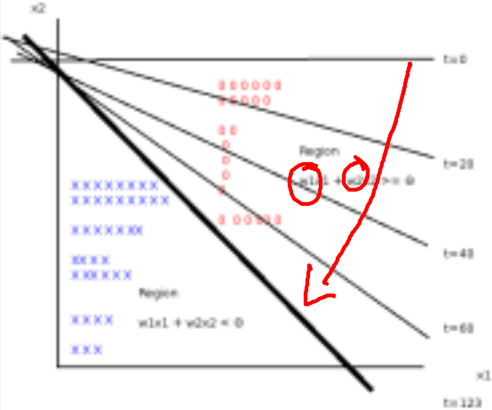

Figura: Aplicación iterativa de las reglas de decision [Brio and Molina, 2005]

Preceptrón

Algoritmo de aprendizaje

La idea es etiquetas con -1 y 1 dos regiones en el espacio

1. Inicializar los pesos aleatoriamente entre [-1 y 1]
2. Para el estado t. Calcular:

$$y^u(k) = sgn(\sum_{j=1}^{n}(w_j(t)x_j)$$

3. Corregir pasos sinápticos (Si $t_j^u \neq y_j^u$)

$$w_j(t+1) = w_j(t) + \epsilon[t_j^u - y_j^u]x^u, j = 1, 2, .., n$$

4. Para si no se han modificado los pesos en los últimos p patrones o se ha llegado a un número de iteraciones especificado.

Preceptrón

Algoritmo de aprendizaje

Miremos la compuerta AND

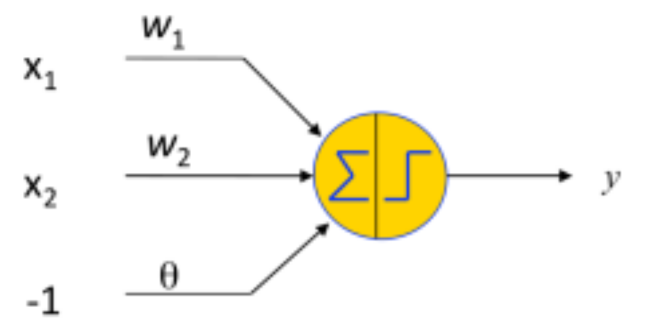

Figura: Perceptrón compuesta AND

Preceptrón

Algoritmo de aprendizaje

Miremos la compuerta AND

Entrada	**Salida** $t^u(k)$
(-1,-1)	-1
(-1,1)	-1
(1,-1)	-1
(1,1)	1

Cuadro: Función AND con lógica función signo

Universidad del Valle

Preceptrón

Algoritmo de aprendizaje

1. Inicialización de pesos. Elegimos $\epsilon = 0{,}5$

   -1

   $$w_1 = 0{,}4, w_2 = -0{,}2, \Theta = 0{,}6$$

2. Con t $= 1$, patrón (-1,-1), $y^u = sgn(-0{,}8) = -1$. Esta bien ya que esperamos -1. $p = 1$
3. Para t $= 1$, patrón (-1, 1), $y^u = sgn(-1{,}2) = -1$. Esta bien. $p = 2$

Preceptrón

Algoritmo de aprendizaje

4. Para t = 1, patrón (1,-1), $y^u = sgn(0) = 0$ Esta mal, ya que esperamos -1. Actualizamos pesos

$$w_1(2) = w_1(1) + \epsilon[t^u(1) - y^u(1)]x_1 = 0,4 + 0,5[-1-0](1)$$
$$w_1(2) = -0,1$$
$$w_2(2) = -0,2 + 0,5[-1-0](-1) = 0,3$$
$$\Theta(2) = 0,6 + 0,5[-1-0](-1) = 1,1$$

5. Ya que actualizamos, ahora $t = 2$, y revisamos.

Preceptrón

Algoritmo de aprendizaje

$$w_1 = -0{,}1, w_2 = 0{,}3, \Theta = 1{,}1$$

6. Para t $= 2$; patrón (-1,-1), $y^u = sgn(-1{,}1) = -1$. Correcto
   $p = 1$
7. Para t $= 2$; patrón (-1,1), $y^u = sgn(-0{,}7) = -1$. Correcto
   $p = 2$
8. Para t $= 2$; patrón (1,-1), $y^u = sgn(-1{,}5) = -1$. Correcto
   $p = 3$

Universidad del Valle

Preceptrón

Algoritmo de aprendizaje

$$w_1 = -0{,}1, w_2 = 0{,}3, \Theta = 1{,}1$$

9 Para t = 2; patrón (1,1), $y^u = sgn(-0{,}9) = -1$. Incorrecto

$$w_1(3) = -0{,}1 + 0{,}5[1 - (-1)](1) = 0{,}9$$
$$w_2(3) = 0{,}3 + 0{,}5[1 - (-1)](1) = 0{,}7$$
$$\Theta(3) = 1{,}1 + 0{,}5[1 - (-1)](-1) = 0{,}1$$

10 Ya que actualizamos, ahora $t = 3$, y revisamos.

Preceptrón

Algoritmo de aprendizaje

$$w_1 = 0{,}9, w_2 = 0{,}7, \Theta = 0{,}1$$

11. Para t = 3; patrón (-1,-1), $y^u = sgn(-1{,}7) = -1$. Correcto $p = 1$
12. Para t = 3; patrón (-1,1), $y^u = sgn(-0{,}3) = -1$. Correcto $p = 2$
13. Para t = 3; patrón (1,-1), $y^u = sgn(0{,}1) = 1$. Incorrecto

Universidad del Valle

Preceptrón

Algoritmo de aprendizaje

$$w_1 = 0{,}9, w_2 = 0{,}7, \Theta = 0{,}1$$

14 Para t = 3; patrón (1,-1), $y^u = sgn(0{,}1) = 1$. Incorrecto

$$\begin{aligned} w_1(4) &= 0{,}9 + 0{,}5[-1 - 1](1) = -0{,}1 \\ w_2(4) &= 0{,}7 + 0{,}5[-1 - 1](-1) = 1{,}7 \\ \Theta(4) &= 0{,}1 + 0{,}5[-1 - 1](-1) = 1{,}1 \end{aligned}$$

15 Ya que actualizamos, ahora $t = 4$, y revisamos.

Universidad del Valle

Preceptrón

Algoritmo de aprendizaje

$$w_1 = -0{,}1, w_2 = 1{,}7, \Theta = 1{,}1$$

16 Para t = 4; patrón (-1,-1), $y^u = sgn(-2{,}7) = -1$. Correcto $p = 1$

17 Para t = 4; patrón (-1,1), $y^u = sgn(0{,}7) = -1$. Incorrecto

$$\begin{aligned} w_1(5) &= -0{,}1 + 0{,}5[-1 - 1](-1) = 0{,}9 \\ w_2(5) &= 1{,}7 + 0{,}5[-1 - 1](1) = 0{,}7 \\ \Theta(5) &= 1{,}1 + 0{,}5[-1 - 1](-1) = 2{,}1 \end{aligned}$$

Ahora con $t = 5$

Preceptrón

Algoritmo de aprendizaje

$$w_1 = 0{,}9, w_2 = 0{,}7, \Theta = 2{,}1$$

18 Para t = 5; patrón (1,1), $y^u = sgn(-0{,}5) = -1$. Incorrecto, diferencia salidas 2.

$$w_1(6) = 0{,}9 + 0{,}5[2](1) = 1{,}9$$
$$w_2(6) = 0{,}7 + 0{,}5[2](1) = 1{,}7$$
$$\Theta(6) = 2{,}1 + 0{,}5[2](-1) = 1{,}1$$

Ahora con $t = 6$

Universidad del Valle

Preceptrón

Algoritmo de aprendizaje

$$w_1 = 1{,}9, w_2 = 1{,}7, \Theta = 1{,}1$$

19. Para t = 6; patrón (-1,-1), $y^u = sgn(-4{,}7) = -1$. Correcto $p = 1$
20. Para t = 6; patrón (-1,1), $y^u = sgn(-1{,}3) = -1$. Correcto $p = 2$
21. Para t = 6; patrón (1,-1), $y^u = sgn(-0{,}9) = 1$. Correcto $p = 3$
22. Para t = 6; patrón (1,1), $y^u = sgn(2{,}5) = 1$. Correcto $p = 4$

Termina.

Preceptrón

Ejercicio

1. Aplica el algoritmo para la compuerta AND, con $w_1 = 0{,}7, w_2 = -1{,}8, \Theta = 1{,}4, \epsilon = 0{,}5$
2. Aplica el algoritmo para la compuerta OR, con $w_1 = 0{,}7, w_2 = -1{,}8, \Theta = 1{,}4, \epsilon = 0{,}5$

Contenido

Universidad del Valle

Adeline

Definición

- Fue introducido por Widrow [Widrow and Hoff, 1988], [Widrow and Winter, 1988] entre 1959 y 1988.
- Es de respuesta lineal a diferencia del perceptron
- Puede trabajar con entradas continuas
- Se incorpora un elemento adicional llamado **bias** u **umbral** Θ. La cual se suma a la entrada (usualmente es -1)
- Utiliza mínimos cuadrados para el cálculo del error.
- Se tiene una tasa de aprendizaje ϵ

Universidad del Valle

Adeline

Regla LMS

- Si las entradas tiene vectores de entrada ortogonales, se podrá llegar a asociaciones perfectas
- La salida de la neurona es $y = \sum_{j=1}^{n} x_j w_j + \Theta$. Ya que la función de activación es lineal
- El cambio se basa en el cálculo del LMS para los patrones de entrada. En este caso el error cuadrático $Err = \frac{1}{2} \sum_{j=0}^{n} (t^u(j) - y^u(j))^2$
- Lo que se busca es modificar los valores de forma iterativa mediante la regla del descenso del gradiente:

$$\Delta_p w_j = -\epsilon \frac{\partial Err^P}{\partial w_j}$$

Adeline

Algoritmo de aprendizaje

1. Inicialice los pesos aleatorios
2. Para cada patrón, actualice los pesos a razón de:

$$\Delta w_j = \alpha(t^u - y^u).x_j$$

3. Puede detenerse cuando todos los patrones cumplen la salida deseada o bien se han cumplido cierto número de iteraciones.

Ejercicio

En el código solucionar el problema del codificador binario-decimal.

x_1	x_2	y
0	1	1
1	0	2
1	1	3

Cuadro: Codificador binario a decimal

Nota: No se puede trabajar bajo este enfoque la entrada (0,0) ya que sin importar el peso la salida siempre será Θ.

Referencias I

Brio, B. and Molina, A. (2005).
Redes neuronales y sistemas difusos.
Textos universitarios. Alfaomega.
Pages 41–63.

Lippmann, R. P. (1988).
An introduction to computing with neural nets.
SIGARCH Comput. Archit. News, 16(1):7–25.

Widrow, B. and Hoff, M. E. (1988).
Neurocomputing: Foundations of research.
chapter Adaptive Switching Circuits, pages 123–134. MIT Press, Cambridge, MA, USA.

Universidad del Valle

Referencias II

Widrow, B. and Winter, R. (1988).
Neural nets for adaptive filtering and adaptive pattern recognition.
Computer, 21(3):25–39.

¿Preguntas?

Próximo tema:
Preceptrón multicapa