INTELIGENCIA ARTIFICIAL Grupo 2, semestre 2020-2
Ejercicio sobre Perceptrón Unicapa

I. Introducción

Este documento explica en detalle la solución de un ejercicio sobre la aplicación de la técnica inteligente denominada Redes Neuronales para resolver un problema de ingeniería. De esta manera se pretende complementar lo presentado sobre el mismo tema, tanto en el aspecto teórico como en el práctico, en la décima sesión del presente.

II. Problema

Se desea implementar el operador booleano NAND mediante la red neuronal denominada Perceptrón Unicapa.

Responda:

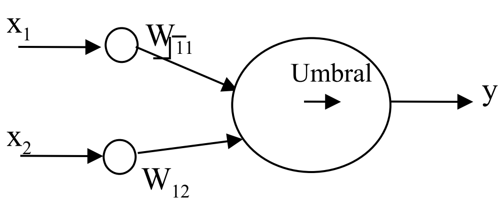
a) Representar el problema mediante la red indicada. Use un esquema mostrando las capas de la red, las neuronas en cada una de ellas y sus funciones de propagación, de activación y de salida. Mostrar además los pesos de la red.

b) Se sabe que los valores de los pesos de la red al final de una determinada época del entrenamiento son: $w_{11} = -0.3$ $w_{12} = -0.2$ Umbral = -0.7
Se pregunta si con esos valores ya está resuelto el problema o es necesario continuar con el proceso de aprendizaje. Si la respuesta fuese continuar con el aprendizaje, efectuar una época más. Para tal efecto, considere una tasa de aprendizaje $\eta = 0.1$.
<u>Nota</u>: recordar que se denomina época a la presentación de todos los ejemplos usados para el entrenamiento

c) Las clases presentes en el problema, ¿son linealmente separables? Si la respuesta fuese positiva, mostrar en un gráfico x1-x2 los valores de entrada de la función lógica y la recta que separa las clases con su ecuación correspondiente.

Solución del problema:

a) Representación del problema mediante la red Perceptrón Unicapa



Dado que en el problema se tiene dos variables de entrada y una de salida, se usarán dos neuronas de entrada y una de salida. Los pesos están indicados según la convención mostrada en clase w_{ji}

Regla de propagación

$$u_j = w_{ji}\, x_i$$

donde:
u_j : entrada neta o impulso total de la neurona j
x_i : i-ésima entrada para la neurona j

Regla de Activación

$f(u_j) = 1$, si $u_j \geq \text{Umbral}_j$
$f(u_j) = 0$, si $u_j < \text{Umbral}_j$

donde:
$f(u_j)$: función de activación de la neurona j
Umbral_j : valor Umbral de la neurona j

Regla de Salida

$y_j = f(u_j)$

Luego del artificio se tiene:

Regla de propagación

$$u'_j = w_{0j}\, x_0 + \sum w_{ji}\, x_i$$

donde:
u'_j : entrada neta modificada para incluir al Umbral como un peso más de la red
w_{j0} : valor Umbral de la neurona j (antiguo Umbral_j)
x_0 : valor constante e igual a -1

Regla de Activación

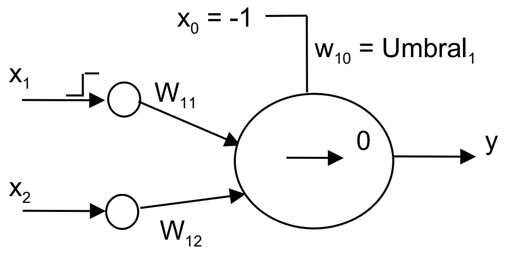
$f(u'_j) = 1$, si $u'_j \geq 0$
$f(u'_j) = 0$, si $u'_j < 0$

donde:
$f(u'_j)$: función de activación de la neurona j que después del artificio usa como argumento el valor u'_j (antiguo u_j)

Regla de Salida

$y_j = f(u'_j)$

Y la topología de la red, después del artificio, queda así:



b)
El problema es un problema de clasificación en la que hay dos clases presentes: la clase de los "1" y la clase de los "0'. La red Perceptrón Unicapa se aplica para este tipo de problemas y si las clases presentes son linealmente separables, el teorema de convergencia asegura que la red obtendrá una solución al problema en un número finito de pasos para cualquier elección inicial de pesos (ver punto c) para un mayor análisis).

Sabemos por el enunciado del problema que

$w_{11} = -0.3$ $w_{12} = -0.2$ Umbral $= W_{10} = -0.7$

Sabemos que para que la red aprenda es necesario que ella entrene. Y para el entrenamiento se requieren de ejemplos de comportamiento, los cuales los obtendremos a partir de la tabla de verdad del operador booleana que se quiere implementar con la red neuronal.

En seguida se presenta la tabla de verdad del AND, luego la tabla de verdad del AND negado y finalmente la codificación de éste a valores binarios de modo que se pueda trabajar con la red neuronal.

Tabla de verdad del AND

s_1	s_2	t
F	F	F
F	V	F
V	F	F
V	V	V

Tabla de verdad del NAND ($\neg$AND)

s_1	s_2	t
F	F	V
F	V	V
V	F	V
V	V	F

Tabla de verdad del NAND codificada

s_1	s_2	t
0	0	1
0	1	1
1	0	1
1	1	0

Adicionalmente sabemos que $x_0 = -1$ y que la tasa de aprendizaje es $\eta = 0.1$

Analizaremos la Época 1 para ver si con esos valores se ha resuelto el problema :

EPOCA 1

Aplicaremos el algoritmo de entrenamiento del Perceptrón Unicapa sobre el conjunto de entrenamiento a usar.

Sabemos que: $u'_1 = w_{10}\, x_0 + w_{11}\, x_1 + w_{12}\, x_2$ y que $x_i = s_i$ luego $x_1 = s_1$, $x_2 = s_2$

EJM. 1

s_1	s_2	t
0	0	1

$x_1 = 0$, $x_2 = 0$

$u = (-0.7)(-1) +(-0.3)(0) + (-0.2)(0) = 0.7$

$f(u) = f(0.7) =1$ luego $y = 1$;

como $t =1$ $\therefore$NO SE MODIFICAN LOS PESOS

EJM. 2

s_1	s_2	t
0	1	1

$x_1 = 0$, $x_2 = 1$

$u = (-0.7)(-1) +(-0.3)(0) + (-0.2)(1) = 0.5$

$f(u) = f(0.5) =1$ luego $y = 1$;

como $t =1$ $\therefore$NO SE MODIFICAN LOS PESOS

EJM. 3

s_1	s_2	t
1	0	1

$x_1 = 1$, $x_2 = 0$

$u = (-0.7)(-1) +(-0.3)(1) + (-0.2)(0) = 0.4$

$f(u) = f(0.4) =1$ luego $y = 1$;

como t =1 ∴NO SE MODIFICAN LOS PESOS

EJM. 4

s_1	s_2	t
1	1	0

$x_1 = 1$, $x_2 = 1$

$u = (-0.7)(-1) +(-0.3)(1) + (-0.2)(1) = 0.2$

$f(u) = f(0.2) =1$ luego $y = 1$;

como t =0 ∴SE MODIFICAN LOS PESOS

Para la actualización de los pesos se usará la regla de aprendizaje:

$$w_{ji} \text{ (nuevo)} = w_{ji} \text{ (anterior)} + \eta\, (t_j - y_j)\, x_i$$
$$w_{j0} \text{ (nuevo)} = w_{j0} \text{ (anterior)} + \eta\, (t_j - y_j)\, x_0$$

Como tenemos una sola neurona en la salida j =1

Actualizando:

Para i=1

$$w_{11} \text{ (nuevo)} = w_{11} \text{ (anterior)} + \eta\, (t_1 - y_1)\, x_1$$
$$w_{12} \text{ (nuevo)} = w_{12} \text{ (anterior)} + \eta\, (t_1 - y_1)\, x_2$$

Reemplazando valores:

$$w_{11} \text{ (nuevo)} = -0.3 + 0.1(0 - 1)1 = -0.4$$
$$w_{12} \text{ (nuevo)} = -0.2 + 0.1(0 - 1)1 = -0.3$$
$$w_{10} \text{ (nuevo)} = -0.7 + 0.1(0 - 1)(-1) = -0.6$$

Al final de la presentación del ejemplo 4 los nuevos valores de los pesos son:

$w_{11} = -0.4$, $w_{12} = -0.3$, $w_{10} = -0.6$

La condición de parada del algoritmo es que la red acierte en todos los ejemplos de entrenamiento. Como la red falló en el ejemplo 4, será necesario entrenar, según lo solicitado en el enunciado, una época más. En general se debe continuar entrenando hasta que consiga el mapeo entradas-salida requerido.

IMPORTANTE

Luego de actualizar los pesos, en los subsiguientes ejemplos se deberán utilizar los valores actualizados. La actualización se da cuando la red falla en su estimativa del valor target, lo cual puede suceder en cualquiera de los ejemplos que se presentan; en este problema se hizo en el ejemplo 4 pero pudo haber fallado la red en cualquiera de los ejemplos.

EPOCA 2

EJM. 1

s_1	s_2	t
0	0	1

$x_1 = 0$, $x_2 = 0$

$u = (-0.6)(-1) +(-0.4)(0) + (-0.3)(0) = 0.6$

$f(u) = f(0.6) =1$ luego $y = 1$;

como t =1 ∴NO SE MODIFICAN LOS PESOS

EJM. 2

s_1	s_2	t
0	1	1

$x_1 = 0$, $x_2 = 1$

$u = (-0.6)(-1) +(-0.4)(0) + (-0.3)(1) = 0.3$

$f(u) = f(0.3) =1$ luego $y = 1$;

como t = 1 ∴NO SE MODIFICAN LOS PESOS

EJM. 3

s_1	s_2	t
1	0	1

$x_1 = 1$, $x_2 = 0$

$u = (-0.6)(-1) +(-0.4)(1) + (-0.2)(0) = 0.2$

$f(u) = f(0.2) =1$ luego $y = 1$;

como t =1 ∴NO SE MODIFICAN LOS PESOS

EJM. 4

s_1	s_2	t
1	1	0

$x_1 = 1$, $x_2 = 1$

$u = (-0.6)(-1) +(-0.4)(1) + (-0.3)(1) = -0.1$

$f(u) = f(-0.1) = 0$ luego $y = 0$;

como $t = 0$ $\therefore$ NO SE MODIFICAN LOS PESOS

En toda la época la red no falló en ejemplo alguno (acertó en todos los ejemplos de entrenamiento), lo cual significa que se ha cumplido la condición de parada del algoritmo de entrenamiento. Los pesos con que la red implementa el operador NAND son:

$$w_{11} = -0.4$$
$$w_{12} = -0.3$$
$$w_{10} = -0.6$$

c)

A partir de la función de activación:

$$u'_1 \geq 0$$

se puede determinar que para el valor de $u'_1 = 0$, se tiene el punto para el cual se separan las clases. Por encima de la recta definida para dicho los puntos pertenecerán a la clase de los "0" y por debajo, a la clase de los "1".

$$u'_1 \geq 0$$

Luego podremos determinar la ecuación de la recta así:

$$u'_1 = w_0 x_0 + w_{11} x_1 + w_{12} x_2 = 0$$
$$- w_0 + w_{11} x_1 + w_{12} x_2 = 0$$

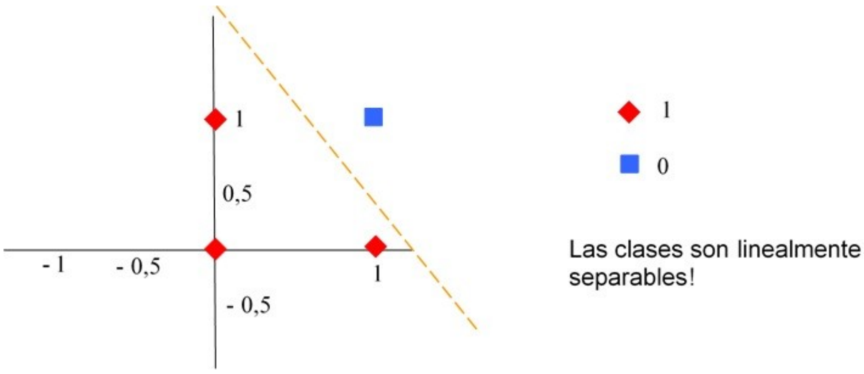
Manipulando algebraicamente la ecuación se tiene:

$$w_{12} x_2 = w_0 - w_{11} x_1$$

Despejando x_2 , se obtiene la ecuación de la recta que separa las clases presentes.

$$x_2 = \frac{w_0}{w_{12}} - \frac{w_{11}}{w_{12}} x_1 \quad , \quad \text{entonces} \quad x_2 = \frac{-0.6}{-0.3} - \frac{-0.4}{-0.3} x_1 = 2 - \frac{4}{3} x_1$$

Notar que la pendiente es negativa, lo cual es coherente con la recta que separa las clases.



RAMP