Actividades resueltas

Redes Neuronales: Perceptron

Descripción de la actividad

Se tiene una base de datos que contiene tres registros. Cada uno de las instancias o ejemplos tiene dos atributos de entrada binarios (x_1, x_2) y un atributo de salida o clase también binaria (y) que toman los valores +1 y -1.

	x_1	x_2	y
E1	-1	-1	-1
E2	-1	+1	+1
E3	+1	-1	+1

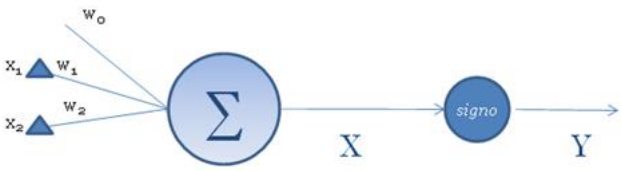
A partir de esta base de datos se quiere construir un modelo predictor utilizando un **perceptrón**, la red neuronal más simple. Este perceptrón utiliza como función de activación la función signo y tiene una tasa de aprendizaje de 0.75. Para iniciar el algoritmo de aprendizaje del perceptrón se utilizan unos pesos escogidos aleatoriamente con valores $w_0 = 0.5$, $w_1 = -0.5$ y $w_2 = 0.5$.

Aplica el algoritmo de aprendizaje del perceptrón e indica claramente los pesos obtenidos al finalizar la fase de aprendizaje del perceptrón.

Una vez entrenado el modelo, imagina que se presenta a la entrada del perceptrón una nueva instancia de clase desconocida y con valores de los atributos de entrada $x_1 = +1$ y $x_2 = +1$. Calcula la clase (y) en la que se clasificaría esta nueva instancia. Describe claramente los pasos que has seguido para realizar la clasificación

Resolución de la actividad

El perceptrón se representa gráficamente de la siguiente forma:



La entrada ponderada (X) se puede calcular con la siguiente fórmula:

$$X = w_0 + \sum_{i=1}^{2} x_i \, w_i \; = w_0 + x_1 \cdot w_1 + x_2 \cdot w_2$$

Como función de activación se utiliza la función signo, por lo tanto, la salida del perceptrón (Y) se describe con la siguiente fórmula:

$$Y = \begin{cases} +1 & si\ X \geq 0 \\ -1 & si\ X < 0 \end{cases}$$

La regla de aprendizaje del perceptrón es la siguiente, considerando que la tasa de aprendizaje es α = 0.75:

$$w_i(t+1) = \; w_i(t) + \; \alpha \cdot \; x_i(t) \cdot \; e(t) = \; w_i(t) + \; \alpha \cdot \; x_i(t) \cdot \; \big(y_d(t) - y(t)\big)$$

$$= \; w_i(t) + \; 0.75 \cdot \; x_i(t) \cdot \; (y_d(t) - y(t))$$

Entonces los pesos se calcularán con las siguientes fórmulas:

$$w_1(t+1) = \; w_1(t) + \; 0.75 \cdot \; x_1(t) \cdot \; (y_d(t) - y(t))$$

$$w_2(t+1) = \; w_2(t) + \; 0.75 \cdot \; x_2(t) \cdot \; (y_d(t) - y(t))$$

Se debe tener en cuenta que para el cálculo del umbral w_0 la fórmula es la siguiente:

$$w_0(t+1) = \; w_0(t) + \; 0.75 \cdot \; (y_d(t) - y(t))$$

Una vez se ha definido la arquitectura de la red neuronal, la función de activación y la regla de aprendizaje se puede aplicar el algoritmo de aprendizaje del perceptrón.

En el primer paso se asignan valores aleatorios en el intervalo [-0.5, 0.5] al umbral w_0 y a los pesos w_1 y w_2. Según el enunciado $w_0 = 0.5$, $w_1 = -0.5$ y $w_2 = 0.5$.

En el segundo paso se activa el perceptrón utilizando la primera instancia E1, es decir excitando la red con $x_1 = -1$ y $x_2 = -1$. Se espera que la salida deseada sea $y_d = -1$, el valor de la clase en los datos de entrenamiento.

Aplicando la fórmula de la entrada ponderada:

$$X = w_0 + x_1 \cdot w_1 + x_2 \cdot w_2 = 0.5 + (-1) \cdot (-0.5) + (-1) \cdot 0.5 = 0.5$$

Como la entrada ponderada es mayor o igual a cero ($X \geq 0$), al aplicar la función signo se obtiene la salida real del perceptrón que será:

$$y = +1$$

Observamos que el perceptrón comete un error porque la salida real no es igual a la salida deseada $e(t) = y_d(t) - y(t) = (-1) - (+1) = -2$ y a partir de este error vamos a actualizar los pesos.

En el tercer paso se actualizan los pesos:

$$w_0(t+1) = w_0(t) + 0.75 \cdot \left(y_d(t) - y(t)\right) = 0.5 + 0.75 \cdot (-2) = -1$$

$$w_1(t+1) = w_1(t) + 0.75 \cdot x_1(t) \cdot \left(y_d(t) - y(t)\right) = (-0.5) + 0.75 \cdot (-1) \cdot (-2) = 1$$

$$w_2(t+1) = w_2(t) + 0.75 \cdot x_2(t) \cdot \left(y_d(t) - y(t)\right) = 0.5 + 0.75 \cdot (-1) \cdot (-2) = 2$$

Quedan disponibles datos de entrenamiento, por lo que se retorna al segundo paso donde se activa de nuevo el perceptrón utilizando la segunda instancia E2 ($x_1 = -1$ y $x_2 = +1$) y se espera que la salida deseada sea $y_d = +1$.

Aplicando la fórmula de la entrada ponderada:

$$X = w_0 + x_1 \cdot w_1 + x_2 \cdot w_2 = (-1) + (-1) \cdot 1 + (+1) \cdot 2 = 0$$

Como la entrada ponderada es mayor o igual a cero ($X \geq 0$), al aplicar la función signo se obtiene la salida real del perceptrón que será:

$$y = +1$$

En este caso el perceptrón no comete un error porque la salida real es igual a la salida deseada $e(t) = y_d(t) - y(t) = (+1) - (+1) = 0$ y por lo tanto vamos a observar que al actualizar los pesos estos no van a variar.

En el tercer paso se actualizan los pesos:

$$w_0(t+1) = w_0(t) + 0.75 \cdot \left(y_d(t) - y(t)\right) = (-1) + 0.75 \cdot 0 = -1$$
$$w_1(t+1) = w_1(t) + 0.75 \cdot x_1(t) \cdot \left(y_d(t) - y(t)\right) = 1 + 0.75 \cdot (-1) \cdot 0 = 1$$
$$w_2(t+1) = w_2(t) + 0.75 \cdot x_2(t) \cdot \left(y_d(t) - y(t)\right) = 2 + 0.75 \cdot (+1) \cdot 0 = 2$$

Quedan disponibles datos de entrenamiento, por lo que se retorna al segundo paso donde se activa de nuevo el perceptrón utilizando la tercera instancia E3 (x_1 = +1 y x_2 = -1) y se espera que la salida deseada sea y_d = +1.

Aplicando la fórmula de la entrada ponderada:

$$X = w_0 + x_1 \cdot w_1 + x_2 \cdot w_2 = (-1) + (+1) \cdot 1 + (-1) \cdot 2 = -2$$

Como la entrada ponderada es menor que cero ($X < 0$), al aplicar la función signo se obtiene la salida real del perceptrón que será:

$$y = -1$$

El perceptrón comete un error porque la salida real no es igual a la salida deseada $e(t) = y_d(t) - y(t) = (+1) - (-1) = 2$ y a partir de este error vamos a actualizar los pesos.

En el tercer paso se actualizan los pesos:

$$w_0(t+1) = w_0(t) + 0.75 \cdot \left(y_d(t) - y(t)\right) = (-1) + 0.75 \cdot 2 = 0.5$$
$$w_1(t+1) = w_1(t) + 0.75 \cdot x_1(t) \cdot \left(y_d(t) - y(t)\right) = 1 + 0.75 \cdot (+1) \cdot 2 = 2.5$$
$$w_2(t+1) = w_2(t) + 0.75 \cdot x_2(t) \cdot \left(y_d(t) - y(t)\right) = 2 + 0.75 \cdot (-1) \cdot 2 = 0.5$$

En este punto hemos utilizado ya todos los datos de entrenamiento para activar el perceptrón, pero no hemos alcanzado la convergencia esperada. Es decir que los pesos no son estables, han ido cambiando a lo largo de las iteraciones cuando se ha ido activando el perceptrón con los diferentes datos de entrenamiento. Para asegurar la convergencia debemos comprobar que los pesos no cambian (o lo que es lo mismo el error es cero) para una ronda completa en la que se active el perceptrón con todos los datos de entrenamiento. Por lo tanto, vamos a volver a utilizar los datos de entrenamiento para que el perceptrón aprenda hasta que en una ronda el error sea cero para todos los datos de entrenamiento, es decir los pesos no cambien y se alcance la convergencia.

Se retorna al segundo paso donde se activa de nuevo el perceptrón utilizando la primera instancia E1 (x_1 = -1 y x_2 = -1) y se espera que la salida deseada sea y_d = -1.

Aplicando la fórmula de la entrada ponderada:

$$X = w_0 + x_1 \cdot w_1 + x_2 \cdot w_2 = 0.5 + (-1) \cdot 2.5 + (-1) \cdot 0.5 = -2.5$$

Como la entrada ponderada es menor que cero ($X < 0$), al aplicar la función signo se obtiene la salida real del perceptrón que será:

$$y = -1$$

El perceptrón no comete un error porque la salida real es igual a la salida deseada $e(t) = y_d(t) - y(t) = (-1) - (-1) = 0$ y por lo tanto los pesos no varían. Entonces los pesos en el tercer paso de esta iteración son:

$$w_0(t+1) = w_0(t) = 0.5$$
$$w_1(t+1) = w_1(t) = 2.5$$
$$w_2(t+1) = w_2(t) = 0.5$$

Se retorna al segundo paso donde se activa de nuevo el perceptrón utilizando la segunda instancia E2 (x_1 = -1 y x_2 = +1) y se espera que la salida deseada sea y_d = +1.

Aplicando la fórmula de la entrada ponderada:

$$X = w_0 + x_1 \cdot w_1 + x_2 \cdot w_2 = 0.5 + (-1) \cdot 2.5 + (+1) \cdot 0.5 = -1.5$$

Como la entrada ponderada es menor que cero ($X < 0$), al aplicar la función signo se obtiene la salida real del perceptrón que será:

$$y = -1$$

El perceptrón comete un error porque la salida real no es igual a la salida deseada $e(t) = y_d(t) - y(t) = (+1) - (-1) = 2$ y a partir de este error vamos a actualizar los pesos.

En el tercer paso se actualizan los pesos:

$$w_0(t+1) = w_0(t) + 0.75 \cdot \left(y_d(t) - y(t)\right) = 0.5 + 0.75 \cdot 2 = 2$$

$$w_1(t+1) = w_1(t) + 0.75 \cdot x_1(t) \cdot \left(y_d(t) - y(t)\right) = 2.5 + 0.75 \cdot (-1) \cdot 2 = 1$$

$$w_2(t+1) = w_2(t) + 0.75 \cdot x_2(t) \cdot \left(y_d(t) - y(t)\right) = 0.5 + 0.75 \cdot (+1) \cdot 2 = 2$$

Se retorna al segundo paso donde se activa de nuevo el perceptrón utilizando tercera instancia E3 (x_1 = +1 y x_2 = -1) y se espera que la salida deseada sea y_d = +1.

Aplicando la fórmula de la entrada ponderada:

$$X = w_0 + x_1 \cdot w_1 + x_2 \cdot w_2 = 2 + (+1) \cdot 1 + (-1) \cdot 2 = 1$$

Como la entrada ponderada es mayor o igual a cero ($X \geq 0$), al aplicar la función signo se obtiene la salida real del perceptrón que será:

$$y = +1$$

El perceptrón no comete un error porque la salida real es igual a la salida deseada $e(t) = y_d(t) - y(t) = (+1) - (+1) = 0$ y por lo tanto los pesos no varían. Entonces los pesos en el tercer paso de esta iteración son:

$$w_0(t+1) = w_0(t) = 2$$

$$w_1(t+1) = w_1(t) = 1$$

$$w_2(t+1) = w_2(t) = 2$$

En este punto hemos utilizado ya dos veces todos los datos de entrenamiento para activar el perceptrón, pero no hemos alcanzado la convergencia esperada. En esta segunda ronda los pesos han ido cambiando a lo largo de las iteraciones. Por lo tanto, vamos a volver a utilizar por tercera vez los datos de entrenamiento para que el perceptrón aprenda y ver si se alcanza la convergencia. El criterio de parada sería que en una ronda el error sea cero para la activación de la red neuronal con todos los datos de entrenamiento o lo que es lo mismo que los pesos no cambien.

Se retorna al segundo paso donde se activa de nuevo el perceptrón utilizando la primera instancia E1 (x_1 = -1 y x_2 = -1) y se espera que la salida deseada sea y_d = -1.

Aplicando la fórmula de la entrada ponderada:

$$X = w_0 + x_1 \cdot w_1 + x_2 \cdot w_2 = 2 + (-1) \cdot 1 + (-1) \cdot 2 = -1$$

Como la entrada ponderada es menor que cero ($X < 0$), al aplicar la función signo se obtiene la salida real del perceptrón que será:

$$y = -1$$

El perceptrón no comete un error porque la salida real es igual a la salida deseada $e(t) = y_d(t) - y(t) = (-1) - (-1) = 0$ y por lo tanto los pesos no van a variar. Entonces los pesos en el tercer paso de esta iteración son:

$$w_0(t+1) = w_0(t) = 2$$
$$w_1(t+1) = w_1(t) = 1$$
$$w_2(t+1) = w_2(t) = 2$$

Se retorna al segundo paso donde se activa de nuevo el perceptrón utilizando la segunda instancia E2 (x_1 = -1 y x_2 = +1) y se espera que la salida deseada sea y_d = +1.

Aplicando la fórmula de la entrada ponderada:

$$X = w_0 + x_1 \cdot w_1 + x_2 \cdot w_2 = 2 + (-1) \cdot 1 + (+1) \cdot 2 = 3$$

Como la entrada ponderada es mayor o igual a cero ($X \geq 0$), al aplicar la función signo se obtiene la salida real del perceptrón que será:

$$y = +1$$

El perceptrón no comete un error porque la salida real es igual a la salida deseada $e(t) = y_d(t) - y(t) = (+1) - (+1) = 0$ y por lo tanto los pesos no van a variar. Entonces los pesos en el tercer paso de esta iteración son:

$$w_0(t+1) = w_0(t) = 2$$
$$w_1(t+1) = w_1(t) = 1$$
$$w_2(t+1) = w_2(t) = 2$$

Se retorna al segundo paso donde se activa de nuevo el perceptrón utilizando la tercera instancia E3 ($x_1 = +1$ y $x_2 = -1$) y se espera que la salida deseada sea $y_d = +1$.

Aplicando la fórmula de la entrada ponderada:

$$X = w_0 + x_1 \cdot w_1 + x_2 \cdot w_2 = 2 + (+1) \cdot 1 + (-1) \cdot 2 = 1$$

Como la entrada ponderada es mayor o igual a cero ($X \geq 0$), al aplicar la función signo se obtiene la salida real del perceptrón que será:

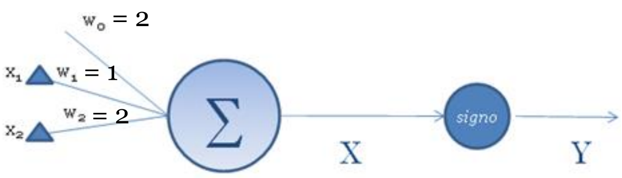
$$y = +1$$

El perceptrón no comete un error porque la salida real es igual a la salida deseada $e(t) = y_d(t) - y(t) = (+1) - (+1) = 0$ y por lo tanto los pesos no van a variar. Entonces los pesos en el tercer paso de esta iteración son:

$$w_0(t+1) = w_0(t) = 2$$
$$w_1(t+1) = w_1(t) = 1$$
$$w_2(t+1) = w_2(t) = 2$$

En este punto hemos utilizado por tercera todos los datos de entrenamiento para activar el perceptrón y observamos que hemos alcanzado la convergencia. Esto se debe a que en esta tercera ronda los pesos no han ido cambiando a lo largo de las iteraciones o lo que es lo mismo el error ha sido cero para la activación de la red neuronal con todos los datos de entrenamiento. Entonces se cumple el criterio de parada y finalizamos el algoritmo de aprendizaje del perceptrón.

El perceptrón resultante de la fase de entrenamiento tiene un valor del umbral $w_0 = 2$ y un valor de los pesos $w_1 = 1$ y $w_2 = 2$.



Por último, en el enunciado se nos indica que una vez entrenado el perceptrón lo utilicemos para clasificar una instancia de clase desconocida y con valores de los atributos de entrada $x_1 = +1$ y $x_2 = +1$.

Vamos a calcular de la entrada ponderada al excitar el perceptrón con $x_1 = +1$ y $x_2 = +1$:

$$X = w_0 + x_1 \cdot w_1 + x_2 \cdot w_2 = 2 + (+1) \cdot 1 + (+1) \cdot 2 = 5$$

Como la entrada ponderada es mayor o igual a cero ($X \geq 0$), al aplicar la función signo se obtiene que la salida del perceptrón será:

$$y = +1$$

Por lo tanto, la nueva instancia de clase desconocida de clasificaría en la clase +1.