Redes Neuronales

Redes Neuronales competitivas I: Aprendizaje no supervisado

carlos.andres.delgado@correounivalle.edu.co

Carlos Andrés Delgado S.

Facultad de Ingeniería. Universidad del Valle

Octubre de 2017

Contenido

Universidad del Valle

Contenido

Introducción

Definiciones

- Existe una capa llamada capa competitiva
- la capa competitiva se conecta totalmente con los nodos de entrada
- Se utilizan conexiones inhibitorias, las cuales tienen dos estados: encendido y apagado.
- El potencial sináptico de la neurona i para una entrada $[x_1, x_2, ...x_n]$ está dado por la expresión:

$$h_i = \underbrace{w_{i1}x_1 + w_{i2}x_2 + ... + w_{iN}x_n}_{\text{entrada neta}} - \Theta_i$$

Donde $\Theta_i = \frac{1}{2}(w_{i1}^2 + w_{i2}^2 + ... + w_{iN}^2)$

Introducción

Definiciones

- Se utilizan en problemas de clasificación con una clase (Es o no es de una clase)
- Tenemos aprendizaje supervisado y no supervisado

Contenido

Universidad del Valle

Contenido

Aprendizaje no supervisado

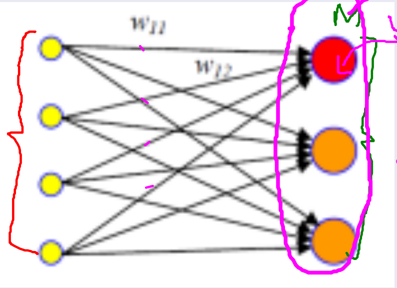
Definiciones

- Una red competitiva es constituida por N señales de entrada y M unidades de proceso (neuronas artificiales)
- Para cada patrón de entrada sólo se activa una neurona (Aquella que es ganadora)

#clases

3 { 0, 1, 2

M=3



Aprendizaje no supervisado

Definiciones

- El estado de la unidad de proceso i es una variable booleana y_i
- La variable y_i viene dada por la expresión:

$$y_i = \begin{cases} 1 & \text{si} \quad h_i = \max_k\{h_1, h, 2, ..., h_M\} \\ 0 & \text{si} \quad \text{otro caso} \end{cases}$$

- Cada entrada a la red es un vector $[x_1, x_2, ...x_n]$ y para el cual sólo se activa una neurona, permaneciendo las restantes desactivadas

Aprendizaje no supervisado

¿Como se determinan los pesos?

- Se utiliza un proceso de aprendizaje no supervisado
- Se busca que se active sólo la neurona cuya vector de pesos sinápticos sea el más similar a la entrada
- Por esto, los pesos sinápticos son la mejor representación del conjunto de patrones
- Se busca demostrar que la unidad ganadora es aquella cuyo valor de pesos sinápticos es que más se parece al vector de entrada

Aprendizaje no supervisado

¿Como se determinan los pesos?

- Se utiliza la distancia euclidiana entre los pesos y la entrada:

$$d(x, w_i) = \sqrt{(x_1 - w_{i1})^2 + ... + (x_N - w_{iM})^2}$$

Aprendizaje no supervisado

Teorema

- Si r es la neurona ganadora entonces cumple:

$$\underline{d(x, w_r)} \leq d(x, w_k), 1 \leq k \leq M$$

- Ahora vamos a introducir el concepto de error cuadratico:

$$\sum_{i=1}^{M} \sum_{j=1}^{P} a_{ij} (d(x(j), w_i))^2$$

Donde: $a_{ij} = \begin{cases} 1 & \text{si} \quad x(j) \text{Es de clase C} \\ 0 & \text{si} \quad \text{otro caso} \end{cases}$

Aprendizaje no supervisado

Teorema

- En este punto se introduce la regla de aprendizaje
- Para cada iteración k se busca determianr los nuevos valores en los vectores de aprendizaje por regla del gradiente descendiente:

$$\nabla w_i(k) = -\epsilon \frac{\partial E}{\partial w_i(k)}$$

- Realizando el álgebra del caso se llega a que la regla de aprendizaje es: $\nabla w_i(k) = \begin{cases} \epsilon(k)(x(k) - w_i(k)) & \text{si} \quad i = r \\ 0 & \text{si} \quad i \neq r \end{cases}$

Universidad del Valle

Aprendizaje no supervisado

Teorema

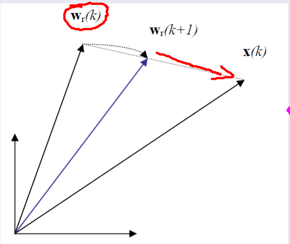
- Para realizar el proceso de aprendizaje se van a determinar los vectores de pesos sinápticos
- Se realiza un proceso iterativo, el cual busca minimizar el error cuadrático, mediante el método de gradiente descendiente
- La regla de aprendizaje puede ser de dos formas:
  - Actualizar los pesos cada vez que se introduce un patrón de entrada, el cual es aprendizaje individualizado
  - Actualizar los pesos después de introducir todos los patrones, que es el conocido como aprendizaje por lotes

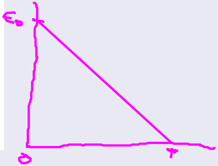
Universidad del Valle

Aprendizaje no supervisado

Teorema

- Los pesos sinápticos solo se actualizan en las neurona ganadora





- Se introduce un factor de aprendizaje ϵ_0 el cual se modifica en cada iteración k, así $\epsilon(k) = \epsilon_0(1 - \frac{k}{T})$ Donde T es el número total de iteraciones hasta concluir el aprendizaje.

Contenido

Aprendizaje individualizado

Procedimiento

1. Elegir como vectores de pesos sinapticos iniciales M patrones de entrenamiento aleatorios y poner $k = 1$. Se tienen M neuronas.

   $M \times N$

2. Elegir un patrón de entrenamiento
3. Calcular los potenciales iniciales $h_1(k), h_2(k), ..., h_n(k)$
4. Escoger el mayor potencial sináptico $h_r(k) = (max)(h_1(k), h_2(k), ..., h_n(k))$
5. Actualizar w_r (r es la neurona ganadora) así:

$$w_r(k+1) = w_r(k) + \epsilon(k)(x(k) - w(k))$$

Aprendizaje individualizado

Procedimiento

6. Se actualiza la taza de aprendizaje $\epsilon(k) = \epsilon_0(1 - \frac{k}{T})$
7. Si $k = T$ parar, en otro caso $k = k + 1$ e ir a paso 1. El valor T puede ser especificado o usando un valor de aprendizaje pequeño.

Contenido

Procedimiento

1. Elegir como vectores de pesos sinápticos iniciales M patrones de entrenamiento aleatorios y poner $k = 1$. Se tienen M neuronas.
2. Calcular los potenciales sinápticos para cada patrón de entrenamiento $h_1(k), h_2(k), ..., h_n(k)$
3. Determinar la neurona ganadora r $h_r(k) = (max)(h_1(k), h_2(k), ..., h_n(k))$ para cada patrón de entrada. Poner
$$a_y = \begin{cases} 1 & \text{si} \quad i \text{es la unidad ganadora para x(j)} \\ 0 & \text{si} \quad \text{otro caso} \end{cases}$$

Universidad del Valle

Aprendizaje por lotes

Procedimiento

4. Actualizar cada w_i así:

$$w_i(k+1) = w_i(k) + \epsilon(k)\overbrace{\sum_{j=1}^{p} a_y (x(k) - w(k))}^{\text{patrones}}$$

5. Se actualiza la taza de aprendizaje $\epsilon(k) = \epsilon_0(1 - \frac{k}{T})$
6. Si $k = T$ parar, en otro caso $k = k + 1$ e ir a paso 1. El valor T puede ser especificado o usando un valor de aprendizaje pequeño.

Referencias I

Curso de modelos computacionales.
http://www.lcc.uma.es/~munozp/.
Accessed: Ocubre-2017.

Du, K.-L. and Swamy, M. N. S. (2010).
Neural Networks in a Softcomputing Framework.
Springer Publishing Company, Incorporated, 1st edition.

Heaton, J. (2008).
Introduction to Neural Networks with Java.
Heaton Research.

¿Preguntas?

Próximo tema:
Redes competitivas II: Mapas autoorganizados y aprendizaje supervisado

Universidad del Valle