

Sistemas de inteligencia artificial

TP3: Perceptron simple y multicapa

Docentes: Rodrigo Ramele, Juliana Gambini,
Juan Santos, Paula Oseroff,
Eugenia Piñeiro, Santiago Reyes

Primer cuatrimestre de 2022

1. Implemente el algoritmo de perceptron simple con función de activación escalón y utilíselo para aprender los siguientes problemas:

- Función lógica 'Y' con entradas
 $x = \{\{-1, 1\}, \{1, -1\}, \{-1, -1\}, \{1, 1\}\},$
y salida esperada
 $y = \{-1, -1, -1, 1\}.$
- Función lógica 'O exclusivo' con entradas
 $x = \{\{-1, 1\}, \{1, -1\}, \{-1, -1\}, \{1, 1\}\},$
y salida esperada
 $y = \{1, 1, -1, -1\}.$

¿Qué puede decir acerca de los problemas que puede resolver el perceptron simple escalón en relación a la resolución de los problemas que se le pidió que haga que el perceptron aprenda?

2. Implemente el algoritmo de perceptron simple lineal y perceptron simple no lineal y utilíselos para aprender el problema especificado en los archivos TP3-ej2-Conjunto-entrenamiento.txt y TP3-ej2-Salida-deseada.txt.
 - Evalúe la capacidad del perceptron simple lineal y perceptron simple no lineal para aprender la función cuyas muestras están presentes en los archivos indicados.
 - Evalúe la capacidad de generalización del perceptron simple no lineal utilizando, de los datos provistos, un subconjunto de ellos para entrenar y otro subconjunto para testear.
 - ¿Cómo podría escoger el mejor conjunto de entrenamiento?

- ¿Cómo podría evaluar la máxima capacidad de generalización del perceptron para este conjunto de datos?

3. Implemente un perceptron multicapa y utilíselo para aprender los siguientes problemas:

- 1) Función lógica 'O exclusivo' con entradas
 $x = \{\{-1, 1\}, \{1, -1\}, \{-1, -1\}, \{1, 1\}\},$
 y salida esperada
 $y = \{1, 1, -1, -1\}.$
- 2) Discriminar si un número es par, con entradas dadas por el conjunto de números decimales del 0 al 9 (usar archivo TP3-ej3-mapa-de-pixeles-digito-decimales.txt) representados por imágenes de 5 x 7 pixeles.
 Entrene con un subconjunto de los dígitos y utilice el resto para testear a la red.
 ¿Qué podría decir acerca de la capacidad para generalizar de la red?
- 3) Construya un perceptron muticapa con la misma entrada usada en 2) pero con 10 unidades de salidas de modo que cada salida represente a un dígito (por ejemplo, si se le presenta a la red la imagen del 7, el estado de la unidad de salida número dígito 7 esté en 1 y las restantes en 0).
 Una vez que la red haya aprendido, use patrones de entrada correspondientes a los dígitos de entrenamiento pero con sus píxeles afectados por ruido (por ejemplo, con probabilidad 0,02, intercambie el valor de los bits en la imagen del dígito) y evalúe los resultados.
- (Opcional) Contemple que la función de activación de las unidades de salida sea la función softmax() e interprete dicha activación como una probabilidad.