

Ejercicios para el examen

CC50Q - Teoría de la Información y Redes Neuronales

Pedro Ortega <peortega@dcc.uchile.cl>
Francisco Claude <fclaude@dcc.uchile.cl>

9 de julio de 2006

Redes neuronales

Ejercicio 1: Considere una neurona de dos entradas con $w_0 = 1$, $w_1 = 0,5$, $w_2 = -0,5$ y una función de activación escalón f :

$$f(x) = \begin{cases} 1 & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$

a) Calcule las salidas $y(\vec{x}^i, \vec{w})$ de la neurona ante las entradas $\{\vec{x}^i\} = \{(0; 1), (0,5; 2), (3; 1)\}$. b) Dibuje la función representada por la neurona. c) Ahora, considere que las salidas deseadas son $\{d^i\} = \{1, 1, 0\}$. Calcule el error cuadrático que comete la neurona, dado por

$$E = \sum_{i=1}^N (d^i - y(\vec{x}^i, \vec{w}))^2$$

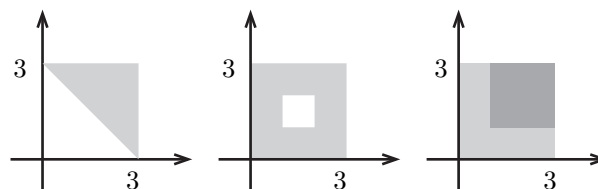
d) Proponga un nuevo set de parámetros \vec{w} para la neurona, de manera que la clasificación se efectúe correctamente.

Ejercicio 2: Considere una neurona de dos entradas con $w_0 = 0$ *fijo* (constante), y w_1, w_2 libres para ser optimizados. La neurona tiene una función de activación $f(x) = x$. Sea $\{\vec{x}^i\} = \{(-1; 1), (1; -1)\}$ el conjunto de datos y $\{d^i\} = \{1, 0\}$ sus respectivas salidas deseadas. Considere el error absoluto,

$$E = \sum_{i=1}^N |d^i - y(\vec{x}^i, \vec{w})|$$

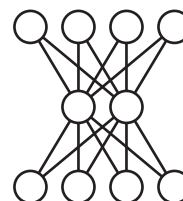
Dibuje la superficie de error en el espacio de los parámetros w_1, w_2 .

Ejercicio 3: Utilizando sólo neuronas con activación escalón (ver ej. 11) o activación lineal ($f(x) = x$), construya redes neuronales multicapa que implementen las funciones ilustradas a continuación:



donde las áreas valen: blanco=0, claro=1, oscuro=-1.

Ejercicio 4: Considere la red neuronal de dos capas con función de activación escalón (en ambas capas) ilustrada a continuación.



Ud. no sabe cómo se escogieron los pesos de la red, pero si la alimenta con entradas

$(0, 1, 1, 1), (1, 1, 0, 0), (-1, 1, -1, 0), (-1, -1, -1, -1)$

responde con salidas

$(0, 1, 0, 0), (1, 1, 0, 0), (1, 1, 1, 1), (0, 0, 1, 1)$

respectivamente. ¿Cuánto valen las salidas de la capa oculta?

Ejercicio 5: Considere una neurona de 1 entrada y $w_0 = 1, w_1 = 1$ y función de activación logística. Ud. tiene el conjunto de datos $\{x^i\} = \{1, 2\}$ y salidas deseadas $\{d^i\} = \{0, 1\}$. Ejecute dos iteraciones del algoritmo de aprendizaje *descenso por el gradiente* con un factor de aprendizaje $\mu = 0,5$ utilizando el criterio de error cuadrático.

Ejercicio 6: Considere una neurona con función de activación *sigmoide logística* y criterio de error *entropía cruzada*, dado por

$$E = - \sum_{i=1}^N (d^i \ln y^i + (1 - d^i) \ln(1 - y^i))$$

donde $y^i \equiv y(\vec{x}, \vec{w})$. Derive una regla de aprendizaje basada en el *descenso por el gradiente*.

Ejercicio 7: Ud. trabaja en una fábrica de plásticos. El proceso de producción, bastante difícil, contempla un total de 10 parámetros: la máquina que realiza la mezcla posee 5 parámetros (temperatura, velocidad de mezcla, etc.) y los restantes 5 parámetros fijan los porcentajes de cada químico presente en la solución. Según las características del plástico deseado (un total de 5 valores reales), un operador experimentado fija los parámetros del proceso “al ojo”.

Diga qué haría Ud. para automatizar el proceso de producción, de manera que, dadas las características del plástico deseado, se puedan calcular automáticamente los parámetros del proceso.

Ejercicio 8: Considere un sistema de cancelación de ruido de dos micrófonos y un módulo de cancelación, como aquel indicado en la figura. El primer micrófono registra la señal de interés $s(t)$ más un ruido ambiental $n(t)$ ponderado (el factor de ponderación a es desconocido). El segundo micrófono, ubicado en otro lugar, registra sólo al ruido ambiental $n(t)$. El cancelador de ruido F entrega dos resultados: una salida filtrada (idealmente $\hat{s}(t) \approx s(t)$), y una señal de error $e(t)$.

La función implementada por el cancelador de ruido es

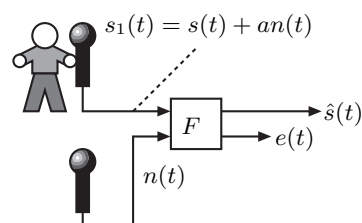
$$\hat{s}(t) = s_1(t) - wn(t)$$

$$e(t) = \hat{s}^2(t)$$

donde w es el único parámetro del módulo F .

Muestre que minimizando la esperanza $E[e(t)]$ se tiende a purificar la señal $s_1(t)$, i.e. cancelar su ruido ambiental. Proponga una regla de aprendizaje *on-line*, de manera que el parámetro w inicializado aleatoriamente converja a uno que produzca una señal purificada $\hat{s}(t)$.

Hint: $E[s(t)n(t)] = 0$ cuando $s(t)$ y $n(t)$ no están correlacionados. Recuerde que un algoritmo *on-line* optimiza una esperanza.



Pregunta 9: Se desea interpolar una función mediante un polinomio. Se tienen N datos de ejemplo $\{x^{(1)}, x^{(2)}, \dots, x^{(N)}\}$ y $m+1$ parámetros ajustables $(w_i, i \in \{0, 1, \dots, m\})$. El modelo planteado es el siguiente:

$$y(\mathbf{w}, x) = \sum_{i=0}^m w_i x^i$$

$$E(\mathbf{w}) = \sum_{n=1}^N (y(\mathbf{w}, x^{(n)}) - t^{(n)})^2$$

donde $t^{(n)}$ es la salida esperada para el ejemplo número n .

1. Derive una regla de aprendizaje mediante descenso por el gradiente.
2. Plantee un sistema de ecuaciones que permita obtener \mathbf{w} de la siguiente forma:

$$\mathbf{A}\mathbf{w} = \mathbf{b}$$

donde \mathbf{A} es una matriz y \mathbf{b} un vector.

3. Determine cuantos ejemplos son necesarios para que el entrenamiento se adapte de forma óptima.

Mapas Autoorganizativos

Ver ejemplos en
<http://sund.de/netze/applets/som/som1/> y en
<http://www.cis.hut.fi/research/som-research/worldmap.html>.

Redes de Hopfield

Ver ejemplo en <http://suhep.phy.syr.edu/courses/modules/MM/sim/hopfield.html>.

Ejercicio 10: Para la red de Hopfield ilustrada a continuación, dibuje el diagrama de transiciones posibles y luego determine los puntos fijos de la red.

