Ejercicios para el Examen
CC50Q - Teoría de la Información y Redes Neuronales

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

19 de noviembre de 2005

Los Ejercicios de esta guía no aparecen en el libro "Information Theory, Inference and Learning Algorithms", a menos que diga explícitamente lo contrario.

Redes Neuronales Multicapa

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

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

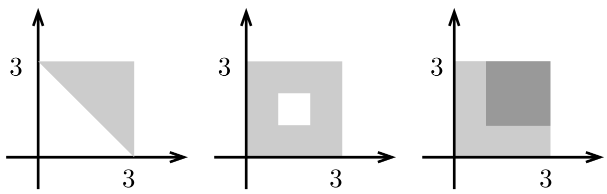
d) Proponga un nuevo set de parámetros $\vec{w}$ para la neurona, de manera que la clasificación se efectúe correctamente.

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,

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

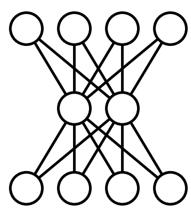
Dibuje la superficie de error en el espacio de los parámetros w_1, w_2.

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 (con alta probabilidad) 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

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

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 una red neuronal multicapa implementada por módulos funcionales. Uno de éstos módulos $y_i = f(x_i)$ consiste en la función tangente hiperbólica, dada por

$$y_i = \tanh(x_i) = \frac{e^{2x_i}-1}{e^{2x_i}+1}, \qquad i = 1,\ldots,N$$

Demuestre que su (única) función de retropropagación del error es

$$\dot{x}_i = (1-y_i^2)\dot{y}_i, \qquad i = 1,\ldots,N.$$

Ejercicio 9: Análogamente al ejercicio anterior, demuestre que las funciones de retropropagación del error de las funciones

1. *Activación Escalamiento (Lineal):*
$$y_i = w_i x_i$$
2. *Activación Gaussiana:*
$$y_i = \exp\left(-\frac{x_i^2}{w_i}\right)$$
3. *Activación Softmax:*
$$y_i = \frac{\exp x_i}{\sum_{j=1}^{N} \exp x_j}$$

son

1. *Activación Escalamiento (Lineal):*
$$\dot{x}_i = w_i \dot{y}_i$$
$$\dot{w}_i = x_i \dot{y}_i$$
2. *Activación Gaussiana:*
$$\dot{x}_i = -\frac{2x_i y_i}{w_i}\dot{y}_i$$
$$\dot{w}_i = \frac{x_i^2}{w_i^2} y_i \dot{y}_i$$
3. *Activación Softmax:*
$$\dot{x}_i = \sum_{j=1}^{N} y_j(\delta_j^i - y_i)\dot{y}_j$$

Mapas Autoorganizativos

Ver ejemplos en:

- `http://websom.hut.fi/websom/`,
- `http://davis.wpi.edu/~matt/courses/soms/applet.html`,
- `http://sund.de/netze/applets/som/som1/`
- `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.