Cuestiones RNA

Teoría de Autómatas y Lenguajes Formales II. Curso 2007-08.

3º curso de Ingeniería Técnica en Informática de Sistemas

Escuela Técnica Superior de Ingeniería Informática

Universidad de Valladolid

Enero 2000

Se desea probar un perpectrón multicapa (MLP) con una sola capa oculta como estimador de funciones. En este caso particular, la función en cuestión es una escalar de varias variables. Matemáticamente, el funcionamiento deseado de la red se puede expresar como $f : \Re^n \rightarrow \Re$, donde no existe restricción alguna respecto a los valores alcanzables, tanto para las variables de entrada, como para la salida. Así pues, para el entrenamiento del MLP se dispone de una tabla con valores concretos $\left(x_1^v, x_2^v \ldots x_n^v\right) \rightarrow y^v$.

A partir de estos datos se pide:

a) Determinar el tamaño de la capa de entrada y de la de salida
b) Modificaciones en el funcionamiento del MLP habitual para adaptarlo a este problema.
c) Sin llegar a desarrollar en sus expresiones últimas, exponer cómo sería el aprendizaje.
d) Indicar un método válido para medir la eficiencia de esta red en la tarea encomendada.

Solución:

a) N0=n y N[H]=1
b) Para la capa de salida: F(x)=x
c) $\Delta_p w_{ij}^1 = -\gamma \frac{\partial E^p}{\partial y_{1i}^p} \frac{\partial y_{1i}^p}{\partial w_{ij}^1} = -\gamma \frac{\partial E^p}{\partial y_{1i}^p} y_{0j}^p; \quad \Delta_p w_{ij}^0 = -\gamma \frac{\partial E^p}{\partial y_{0i}^p} \frac{\partial y_{0i}^p}{\partial w_{ij}^0} = -\gamma \frac{\partial E^p}{\partial y_{0i}^p} F'(u_{0i}^p) I_j^p$
d) Media extendida a los errores relativos de cada muestra

Junio 2000

Responder razonadamente a las siguientes cuestiones:

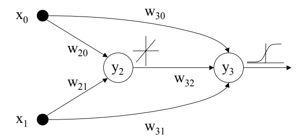
a) Diferencias entre reconocimiento de patrones estáticos y dinámicos.
b) Por las características del MLP y de su algoritmo de aprendizaje, como cabría calificarlo:
   (1) Recurrente o no recurrente
   (2) Supervisado o no supervisado
   (3) Heteroasociativo o autoasociativo.
   Obsérvese que estas características no son mutuamente excluyentes.
c) Definir el término momento en la modificación iterativa de pesos y dar la justificación fundamental para su incorporación, en la práctica, como parte del algoritmo de aprendizaje del MLP.
d) En general, a falta de más datos, conociendo únicamente el valor de una entrada particular y la salida de la neurona correspondiente, cómo se modificaría el peso asociado a dicha conexión entre la entrada y la salida conocida.

Solución:

a) Aparición de la variable tiempo explícitamente
b) (No recurrente) Supervisado (Heteroasociativo).
c) $\Delta_{p-1} w_{ij}^h$.Evitar caer en mínimos locales y el comportamiento oscilatorio.
d) Regla de Hebb: $\Delta_p w_{ij} \ \alpha \ y_i^p x_j^p$

Junio 2001 (I)

Dada la siguiente red neuronal, dedúzcase las expresiones matemáticas para la actualización de sus pesos conforme al algoritmo de retropropagación del error. Tal y como se indica en la figura, la salida de las neuronas tienen funciones de activación distintas. En la primera se trata de la función F(x) = x; en la otra, se utiliza la sigmoide.



Si se quisiera utilizar este sistema como un clasificador de dos categorías solamente, ¿cómo deberían ser las salidas deseadas y qué criterio se aplicaría a la salida de la red para identificarla con una clase concreta?

Por el contrario, si esta red se pretendiera utilizarla como aproximador de funciones universal (y = f(x)), ¿qué modificación necesaria habría que realizar en su funcionamiento para conseguirlo?

Junio 2000 (II)

$$y_2^p = F(u_2^p); \quad \text{donde} \quad u_2^p = w_{20} x_0^p + w_{21} x_1^p + w_{22} = y_2^p$$

$$y_3^p = F(u_3^p); \quad \text{donde} \quad u_3^p = w_{30} x_0^p + w_{31} x_1^p + w_{32} y_2^p + w_{33}$$

$$\Delta_p w_{ij} = -\gamma \frac{\partial E^p}{\partial w_{ij}}; \quad \text{donde} \quad E^p = \frac{1}{2}\left(d^p - y_3^p\right)^2$$

$$\Delta_p w_{3j} = -\gamma \frac{\partial E^p}{\partial y_3^p} \frac{\partial y_3^p}{\partial u_3^p} \frac{\partial u_3^p}{\partial w_{3j}} = \gamma (d^p - y_3^p) y_3^p (1 - y_3^p) X_j^p;$$

$$\text{donde} \quad X_{0|1}^p = x_{0|1}^p; \quad X_2^p = y_2^p; \quad X_3^p = 1$$

$$\Delta_p w_{3j} = \gamma \delta_3^p X_j^p; \quad \text{donde} \quad \delta_3^p = (d^p - y_3^p) y_3^p (1 - y_3^p)$$

Junio 2000 (III)

$$\Delta_p w_{2j} = -\gamma \frac{\partial E^p}{\partial y_3^p} \frac{\partial y_3^p}{\partial y_2^p} \frac{\partial y_2^p}{\partial w_{2j}} = \gamma (d^p - y_3^p) y_3^p (1 - y_3^p) w_{32} I_j^p$$

$$= \gamma \delta_3^p w_{32} I_j^p = \gamma \delta_2^p I_j^p; \quad \text{donde} \quad \delta_2^p = \delta_3^p w_{32}; \; I_{0|1}^p = x_{0|1}^p; \; I_2^p = 1$$

- Clasificador de dos categorías:
  Fijar un valor de decisión el intervalo de salida, por ejemplo: 0.5.
  Salidas deseadas próximas a 0 y a 1 respectivamente

  $$\text{Si} \begin{cases} y_3^p > 0.5 \rightarrow Clase\#1 \\ y_3^p \leq 0.5 \rightarrow Clase\#0 \end{cases}$$

  $$\text{Se podría fijar una zona de indeterminación (GAP)} \begin{cases} y_3^p > 0.6 \rightarrow Clase\#1 \\ y_3^p < 0.4 \rightarrow Clase\#0 \\ 0.4 \leq y_3^p \leq 0.6 \rightarrow \text{Indeterminación} \end{cases}$$

- Aproximador universal de funciones $y = f(x)$
  $y_3^p = F(u_3^p) = u_3^p$ La respuesta no estaría acotada

Septiembre 2002

Respóndase razonadamente a las siguientes cuestiones:

a) En la Regla de Hebb, ¿qué signo debería tener la constante de proporcionalidad?

b) En un perceptrón multicapa ¿se podría sustituir la función de activación sigmoide por la arcotangente (atan(x))?. En caso afirmativo, realícese las transformaciones matemáticas para que, desde el punto de vista externo, la respuesta de la red siguiera siendo la misma.

c) De acuerdo con el teorema de convergencia del perceptrón simple, ¿se puede asegurar que siempre hay una solución? En cualquier caso, de existir, ¿ésta sería única?

d) Supóngase un problema de clasificación. Para su resolución se dispone de un perceptrón multicapa sin ninguna capa oculta y con una función activación lineal (F(x)=x) en las neuronas de la capa de salida. Diséñese un algoritmo de aprendizaje basado en el método de mínimos cuadrados (regresión lineal). Nota: no se pide obtener las fórmulas detalladas, sino la descripción de los pasos y resultados obtenidos.

Solución:

a) Positiva
b) Desplazar (pi/2) y reescalar dividiendo por pi.
c) No se puede asegurar la existencia y de existir, en general, no es única.
d) Aplicar el cálculo directamente de la regresión lineal

Septiembre 2003

En la siguiente tabla aparecen cuatro muestras (A, B, C, D), cada una de las cuales viene dada por un vector de dos componentes (X_1, X_2) y su salida deseada (d(X)). A la derecha aparecen los pesos iniciales de un perceptrón simple compuesto por una sola neurona. Calcúlese la evolución de los pesos según la regla de aprendizaje característica de esta red.

	X_1	X_2	d(X)
A	2	3	1
B	-3	1	-1
C	3	1	1
D	1	-1	-1

$W_1=8$
$W_2=-1$
$\Theta=-3$