uc3m | Universidad Carlos III de Madrid

OPENCOURSEWARE
REDES DE NEURONAS ARTIFICIALES
Inés M. Galván – José M. Valls

Preguntas y Ejercicios para Evaluación: Tema 6

1. ¿Qué particularidad poseen las redes recurrentes frente a las redes no recurrentes?

   Respuesta

   Que poseen conexiones entre las neuronas en cualquier dirección y no solo de la capa de entrada a la salida, como en el caso de las redes feedforward. Esto permite que la red pueda almacenar en su estructura estados anteriores, haciendo que la salida de la red no solo depende de la entrada, sino de estados o entradas anteriores, pudiendo procesar así información temporal

2. Señale las afirmaciones correctas (puede haber más de una):
   a) El algoritmo de retropropagación puede utilizarse para entrenar una red totalmente recurrente
   b) Si una red recurrente se desarrolla en el tiempo, entonces el algoritmo de retropropagación puede utilizarse para entrenar la red recurrente
   c) Las redes parcialmente recurrentes pueden entrenarse con el algoritmo de retropropagación
   d) La red Hopfield puede entrenarse con el algoritmo de retropropagación

   Respuesta

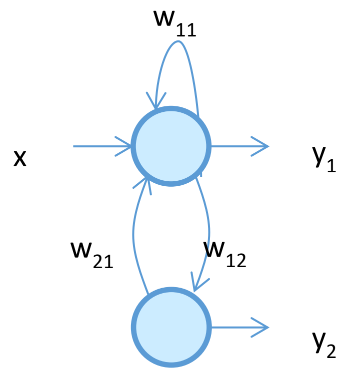
   b) y c)

3. Señale las afirmaciones correctas (puede haber más de una):
   a) Las neuronas de contexto en las redes de Jordan y Elman memorizan los valores de los patrones de entrada en el instante anterior
   b) La diferencia entre la red de Jordan y Elman está en los valores que memorizan las neuronas de contexto
   c) El parámetro μ en la red de Jordan activa o desactiva la activación de las neuroans de salida de la red
   d) En la red de Joradn hay tantas neuronas de contexto como salidas y en la red de Elman tantas como neuronas ocultas

   Respuesta

   b) y d)

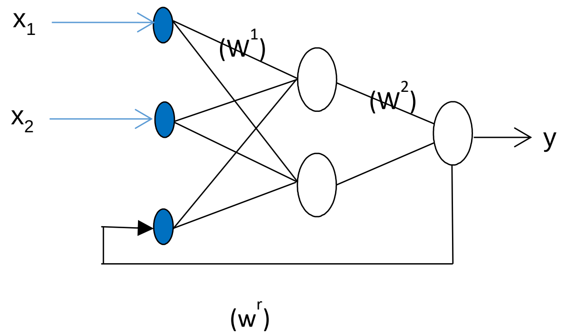
4. Dada la siguiente arquitectura de red, escriba la ecuación para calcular las activaciones de las neuronas de la red.



Respuesta

$$y_1(t) = f(x + w_{11}y_1(t-1) + w_{21}y_2(t-1))$$
$$y_2(t) = f(w_{12}y_1(t-1))$$

5. Dada la arquitectura de red que se muestra en la figura:



a) Indique las expresiones para calcular las activaciones de la red, considerando que las neuronas en la capa de entrada tienen función de activación lineal (f(x)=x) y el resto de las neuronas de la red, función sigmoidal.
b) ¿Es posible utilizar el algoritmo de retropropagación para adaptar los pesos de la red?
c) Si la conexión recurrente de la red no llevara ningún peso asociado, ¿se trataría de una red recurrente?

Respuesta

a) Las dos primeras neuronas de entrada propagan x1 y x2 respectivamente. La activación de la tercera neurona de entrada vienen dada por:

$$c(t) = w^r y(t-1)$$

La activación de las neuronas ocultas viene dada por:

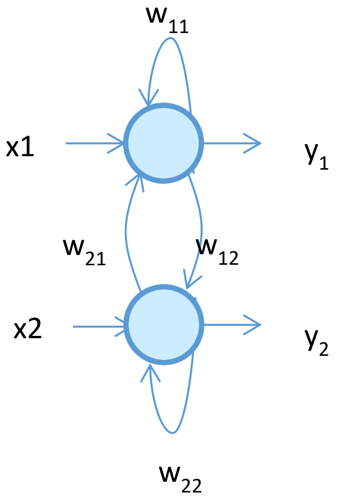
$$a_1(t) = f(w_{11}^1 x_1 + w_{21}^1 x_2 + w_{31}^1 c(t) + u_1)$$
$$a_2(t) = f(w_{12}^1 x_1 + w_{22}^1 x_2 + w_{32}^1 c(t) + u_2)$$

La activación de la neurona de salida es: $y(t) = f(w_1^2 a_1(t) + w_2^2 a_2(t) + v)$, siendo ($u_j$) los umbrales de las neuronas ocultas y v el umbral de la neurona de salida.

b) Para los pesos W^1 y W^1 sí se podría utilizar el algoritmo de retropropagación, pero no para el peso w^r, ya que al derivar la salida con respecto a ese peso, aplicando la regla de la cadena, tendríamos que calcular $\frac{\partial c(t)}{\partial w^r}$, la cual y teniendo en cuenta la expresión de c(t) sería $y(t-1)\frac{\partial y(t-1)}{\partial w^r}$, resultando por tanto un sistema dinámico.

c) Se trataría de una red parcialmente recurrente, concretamente la red de Jordan con parámetro µ igual a 1. Red que procesa información temporal, pero su aprendizaje puede realizarse con el algoritmo de retropropagación ya que las conexiones recurrentes no llevan un peso asociado.

6. Extienda en el intervalo [1,3] la siguiente red recurrente:



Respuesta

$$y_1(t) = f(x_1(t-1) + w_{11} y_1(t-1) + w_{21} y_2(t-1))$$

$$y_2(t) = f(x_2(t-1) + w_{22} y_2(t-1) + w_{12} y_1(t-1))$$

Entonces

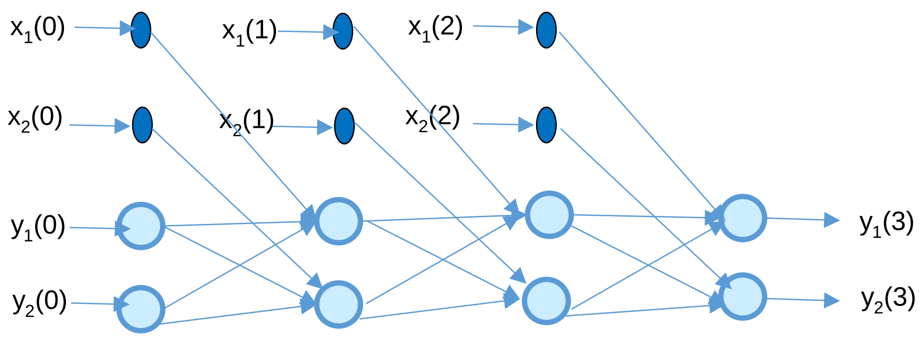
$$y_1(1) = f(x_1(0) + w_{11} y_1(0) + w_{21} y_2(0)),\ y_2(1) = f(x_2(0) + w_{22} y_2(0) + w_{12} y_1(0))$$

$$y_1(2) = f(x_1(1) + w_{11} y_1(1) + w_{21} y_2(1)),\ y_2(2) = f(x_2(1) + w_{22} y_2(1) + w_{12} y_1(1))$$

$$y_1(3) = f(x_1(2) + w_{11} y_1(2) + w_{21} y_2(2)),$$

$$y_2(3) = f(x_2(2) + w_{22} y_2(2) + w_{12} y_1(2))$$

Por tanto, la arquitectura desarrollada en el tiempo sería:



7. Considérense las siguientes secuencias de datos, donde X(t) representa el nivel de la marea en Gijón medida cada 12 horas e Y(t) la presión atmosférica medida cada 6 horas.
X(t): -14.00 -18.00 -11.00 -26.00 5.00 -8.00 43.00 32.00 ………
Y(t): 1022.20 1016.40 1012.90 1013.30 1016.90 1018.90 1018.20 1017.50 1012.60 1010.30 1006.80 1004.00 1001.80 996.50 998.70 … … …

Se sabe que el nivel de la marea depende de la evolución anterior y de la presión atmosférica, pero se desconoce el histórico necesario para encontrar esta dependencia. Se pide:

a) ¿Con que modelo(s) de redes abordaría el problema?
b) Genere los ficheros de datos para el entrenamiento de las redes.
c) Una vez entrenada la red, indique como utilizar la red para predecir la marea con 12 horas de antelación y con 24 horas de antelación.

Respuesta

a) Con redes recurrentes. El modelo a aproximar se puede plantear como X(t+1)=REDRECURRENTE(X(t),Y(t))
Se observa que solo se utiliza un valor anterior, pero las recurrencias de las red permitirán encontrar relaciones con valores anteriores de X(t) y de Y(t).
El paso de tiempo es 12 horas, ya que es el muestreo que disponemos para la serie X(t)
b) Disponemos de los siguientes datos:
X(0h): -14.00
X(12h): -18.00
X(24h): -11.00
X(36h): -26.00
…..

Y(0h): 1022.20
Y(6h): 1016.40
Y(12h): 1012.90
Y(18h): 1013.30
Y(24h): 1016.90
Y(30h): 1018.90
Y(36h): 1018.20
…..

Entonces los patrones son:

Entrada1: X(t)	Entrada3: Y(t)	Salida Deseada: X(t+1)
X(0h)	Y(0h)	X(12h)
X(12h)	Y(12h)	X(24h)
X(24h)	Y(24h)	X(36h)
…	…	…
…	…	…

Es decir:

Entrada1: X(t)	Entrada3: Y(t)	Salida Deseada: X(t+1)
-14	1016.9	-18
-18	1012.90	-11
-11	1016.90	-26
…	…	…
…	…	…

c) Con 12 horas de antelación, bastaría presentarle a la red el valor de la marea y de la presión atmosférica en ese instante, es decir:
X(t+12)=REDRECURRENTE(X(t),Y(t))

Con 24 horas de antelación, con el modelo propuesto habría que calcular la predicción en t+12 y posteriormente en t+24, es decir:
X(t+12)=REDRECURRENTE(X(t),Y(t))
X(t+24)=REDRECURRENTE(X(t+12),Y(t+12))
El valor de Y(t+12) no se conoce, por lo que habría que aproximarlo. Como se trata de presione atmosféricas, cuyos valores no cambian de manera radical, se puede asumir que Y(t+12)=Y(t), valor que sí se conoce en el instante t.

8. Disponemos de los valores diarios de cotización del euro frente al dólar desde el 1 de enero de 2002 hasta el 31 de Diciembre del 2002. Se quiere construir un modelo utilizando redes de neuronas para poder predecir la cotización futura del euro con un horizonte de predicción 1, es decir, al día siguiente. Se toma como hipótesis que el valor de la serie en un determinado día dependerá de los valores que ha tomado en los 7 días anteriores. Plantear el problema con redes de neuronas.

Respuesta:
Como se parte de la hipótesis que el valor de la serie en un determinado día dependerá de los valores que ha tomado en los 7 días anteriores, el problema se puede abordar con un PM o una RBR, utilizando como entradas el valor de cotización del euro en los 7 días anteriores, es decir:
Cotización (D) =RED (Cotización (D-1), Cotización (D-2), …, Cotización (D-7))

El problema puede también plantearse con una red recurrente, sin utilizar la información pasada como entrada, es decir:
Cotización (D) =REDRECURRENTE (Cotización (D-1))

O utilizando la información pasada, pero permitiendo que la red recurrente puede encontrar otras relaciones temporales no presentes en los 7 días anteriores