

Predicción de índices de futuros financieros mediante redes neuronales

JOAN BOSCH¹, LLUÍS GARRIDO¹, SERGIO GÓMEZ²

1) Dept. d'Estructura i Constituents de la Matèria,
Facultat de Física, Universitat de Barcelona,
Diagonal 647, E-08028 Barcelona.
Tel: (93) 402 11 91 Fax: (93) 402 11 98
e-mail: garrido@ecm.ub.es

2) Dept. d'Enginyeria Informàtica,
E.T.S.E., Universitat Rovira i Virgili,
Crt. de Salou s/n, E-43006 Tarragona.
Tel: (977) 55 96 81 Fax: (977) 55 97 10
e-mail: sgomez@etse.urv.es

Resumen

En este artículo estudiamos la aplicación de Redes Neuronales Artificiales a la predicción, basada únicamente en datos históricos, del futuro financiero español Bono 10, teniendo en cuenta las comisiones y la dispersión de los precios.

Published in *Swaps & productos derivados* **27** (1997) 19–21.

1 Introducción

La aplicación de redes neuronales a la predicción de series temporales ha atraído la atención de mucha gente relacionada con los mercados financieros de todo el mundo. Sin embargo, el hecho de que la evolución de los mercados dependa de multitud de variables, muchas de las cuales son difíciles de cuantificar, hace que sea complicado encontrar situaciones en las que únicamente un análisis numérico de los datos históricos permita realizar buenas predicciones. En este artículo mostramos nuestros resultados de aplicar redes neuronales a la predicción dentro de un mismo día del *Bono 10*, que es el futuro más contratado en *MEFF Renta Fija*, el mercado de futuros financieros español.

Una red neuronal artificial (ver por ejemplo [1, 2] para una introducción) es un modelo de computación inspirado en nuestros conocimientos sobre neurociencia, es decir, el estudio de las neuronas de nuestro sistema nervioso, aunque sin tratar de ser biológicamente realistas en detalle. En los últimos años estos modelos han experimentado un gran desarrollo gracias al descubrimiento de su excelente comportamiento en problemas de reconocimiento de patrones, predicción y clasificación, entre otros. Las redes neuronales artificiales son mecanismos matemáticos que aprenden a reconocer o clasificar patrones y, tal como lo hace nuestro propio cerebro, dicho aprendizaje no descansa sobre un modelo preconcebido sino que busca las correlaciones existentes entre las variables del problema que se está estudiando.

Para el presente problema hemos escogido una red neuronal multicapa entrenada mediante el algoritmo de la retropropagación [3, 4]. Una arquitectura típica de dichas redes es la representada en la figura 1. Las neuronas (o unidades) están agrupadas en capas. Existe una primera capa por la que entra la información, una o varias capas ocultas que la procesan, y una capa de salida que proporciona los resultados de la red. Durante la fase de entrenamiento se le presenta a la red un conjunto de posibles valores de entrada juntamente con sus resultados de salida deseados, y el algoritmo de retropropagación busca automáticamente las correlaciones entre dichas entradas y sus salidas respectivas. Cuando el aprendizaje ha finalizado, la red puede ser aplicada sobre datos que no ha visto previamente, realizando sus propias predicciones.

En las redes neuronales multicapa se asocia un peso $W_{ij}^{(\ell)}$ a la conexión (sinapsis) existente entre la unidad j de la capa $\ell - 1$ y la unidad i de la capa ℓ . La salida de cada unidad $I_i^{(\ell)}$ se evalúa utilizando la expresión

$$I_i^{(\ell)} = F \left(\sum_j W_{ij}^{(\ell)} I_j^{(\ell-1)} + \theta_i^{(\ell)} \right),$$

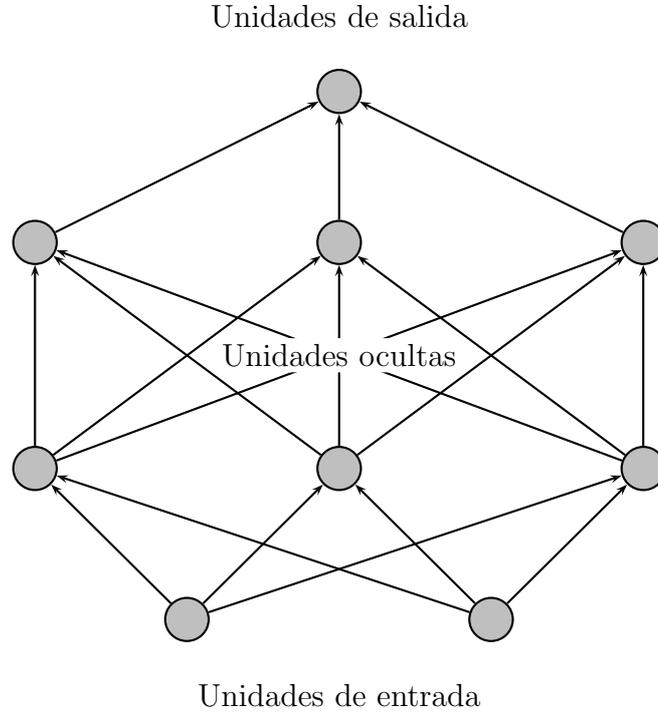


Figura 1: Esquema de una red neuronal multicapa.

donde $\theta_i^{(\ell)}$ es el llamado umbral de la unidad i de la capa ℓ . La función F se conoce como función de activación, y habitualmente se escoge como la función identidad ($F(x) = x$) o como una función sigmoideal ($F(x) = 1/(1 + e^{-x})$). En este último caso es conveniente escalar las entradas y las salidas deseadas entre los valores 0 y 1.

Para entrenar una red neuronal con L capas se necesita un conjunto de datos de entrenamiento $\{(\mathbf{I}^{(0)\mu}, \mathbf{O}^{(L)\mu}), \mu = 1, \dots, N\}$, donde $\mathbf{O}^{(L)\mu}$ representa la salida deseada correspondiente a la entrada $\mathbf{I}^{(0)\mu}$. El error de la red para estos datos es

$$E = \frac{1}{2} \sum_{\mu=1}^N \sum_i (I_i^{(L)\mu} - O_i^{(L)\mu})^2,$$

donde $\mathbf{I}^{(L)\mu}$ representa la salida de la red después de presentarle una entrada $\mathbf{I}^{(0)\mu}$. La retropropagación intenta minimizar este error modificando la intensidad de los pesos y de los umbrales de forma iterativa, en el proceso conocido como de entrenamiento de la red neuronal. Existen otros métodos de entrena-

miento, pero la retropropagación es, con diferencia, el más utilizado. De hecho, la retropropagación es un caso particular del bien conocido algoritmo de descenso por el gradiente, que está basado en una modificación de los pesos según la regla

$$\Delta W_{ij}^{(\ell)} = -\alpha \frac{\partial E}{\partial W_{ij}^{(\ell)}}.$$

Es importante tener presente que la aplicación de las redes neuronales no se reduce únicamente a su entrenamiento con los datos disponibles: uno de los principales problemas es escoger previamente las variables más relevantes, ya que un exceso de variables puede introducir ruido que oculte las más significativas, mientras que su defecto puede provocar una falta de información.

2 El problema de predicción

Cada día MEFF Renta Fija proporciona a los inversores, y hace disponible a través de Internet (<http://www.meff.es>), datos sobre la evolución del mercado. Como en este estudio nos hemos interesado por las predicciones dentro de un mismo día, hemos estudiado los ficheros de información tic a tic, que incluyen la fecha, la hora, el tipo de contrato, el precio y el volumen de cada operación realizada. Como hemos mencionado anteriormente, nos hemos concentrado en el futuro Bono 10.

Teniendo en cuenta que cada día se realizan varios miles de operaciones, hemos decidido condensar la información más relevante de cada periodo de media hora en un reducido número de parámetros. Estos son:

- la fecha y la hora de inicio de dicho periodo de media hora;
- el precio medio de las operaciones dentro de dicha media hora (pr);
- el número de operaciones realizadas;
- el volumen negociado (vo);
- la volatilidad del precio (vt).

Como ya hemos dicho, hemos perseguido una predicción dentro de un mismo día, ya que los sucesos durante las noches pueden alterar significativamente el comportamiento del mercado. Por lo tanto, hemos escogido alimentar nuestras

redes neuronales con los datos disponibles entre la apertura del mercado y las 11:00h, y realizar una predicción del precio del Bono 10 a las 16:00h; el objetivo es saber si el precio a esta última hora es superior o inferior al que tenía a las 11:00h.

3 El modelo neuronal

Para validar la bondad de nuestros modelos, hemos escogido el año 1995 como periodo de test. Algunas de las incógnitas eran la frecuencia con la que haría falta re-entrenar la red, y hasta cuanto tiempo atrás habría que retroceder para obtener datos útiles de entrenamiento. Si el mercado cambia su comportamiento frecuentemente, la utilización de datos demasiado viejos sesgaría las predicciones, mientras que si los datos utilizados son únicamente los más recientes, se perdería información. Hemos encontrado que, en nuestro caso, un histórico de seis meses es el que proporciona mejores resultados. Además, un re-entrenamiento de la red neuronal cada tres meses es suficiente (re-entrenamientos diarios no mejoran la predicción, y tres meses es exactamente el tiempo que pasa entre dos fechas de vencimiento consecutivas del Bono 10). Otro problema era decidir entre predecir la diferencia de precios entre las 16:00h y las 11:00h, o predecir únicamente su signo. En nuestras simulaciones obtuvimos predicciones más estables en el segundo caso.

Sea pr_d^h el precio medio entre las h horas en punto y media hora después de las h horas, d días antes de la fecha a predecir, y sea pr_{d1-d2}^h la media de pr_d^h para d entre $d1$ y $d2$. El mismo criterio es aplicable a vt_d^h (volatilidades) y a vo_d^h (volúmenes).

Para tener una referencia con la que comparar los resultados de la red neuronal, hemos utilizado un modelo lineal sencillo que consiste en lo siguiente: si $pr_0^{10:30} > pr_0^{9-11}$ (el precio en las dos primeras horas está subiendo) entonces compramos un futuro a las 11:00h y lo vendemos a las 16:00h, y en caso contrario ($pr_0^{10:30} < pr_0^{9-11}$) vendemos un futuro a las 11:00h y compramos otro a las 16:00h. El resultado de aplicar esta técnica sobre los datos del año 1995 es un ganancia de

4,50 tics por día y por contrato,

donde 1 tic equivale a 1000 pesetas. Este valor ha sido evaluado tomando como referencia los precios medios $pr_0^{15:30}$ y $pr_0^{10:30}$.

Nuestro primer objetivo ha sido intentar reproducir el resultado anterior utilizando redes neuronales, repitiendo las simulaciones diversas veces para calcular la oscilación de los beneficios esperados. Entrenando la red con una salida esperada igual a $\text{sign}(\text{pr}_0^{15:30} - \text{pr}_0^{10:30})$ a partir de la entrada $\text{pr}_0^{10:30} - \text{pr}_0^{9-11}$ se obtiene un beneficio a lo largo de 1995 de

$$4,55 \pm 0,95 \text{ tics por día y por contrato,}$$

Esto significa que la red ha sido capaz de aprender el “truco” del modelo lineal. No obstante, hay que destacar la gran dispersión de las ganancias esperadas.

Para mejorar estos resultados, introducimos ahora en nuestra red las siguientes variables de entrada:

1. datos del día a predecir:

- (a) $\text{pr}_0^{10:30} - \text{pr}_0^{9-11}$
- (b) $\text{pr}_0^9 - \text{pr}_0^{9-11}$, $\text{pr}_0^{9:30} - \text{pr}_0^{9-11}$, $\text{pr}_0^{10} - \text{pr}_0^{9-11}$
- (c) $\text{vt}_0^{9:00}$, $\text{vt}_0^{9:30}$, vt_0^{10} , $\text{vt}_0^{10:30}$
- (d) $\text{vo}_0^{9:00}$, $\text{vo}_0^{9:30}$, vo_0^{10} , $\text{vo}_0^{10:30}$

2. datos históricos a corto plazo (tres días):

- (a) $\text{pr}_1^9 - \text{pr}_{1-3}^9$, $\text{pr}_1^{10:30} - \text{pr}_{1-3}^{10:30}$, $\text{pr}_1^{15:30} - \text{pr}_{1-3}^{15:30}$

3. datos históricos a medio plazo (10 días):

- (a) $\text{pr}_1^9 - \text{pr}_{1-10}^9$, $\text{pr}_1^{10:30} - \text{pr}_{1-10}^{10:30}$, $\text{pr}_1^{15:30} - \text{pr}_{1-10}^{15:30}$
- (b) $\text{pr}_1^9 - \text{pr}_{1-30}^9$, $\text{pr}_1^{10:30} - \text{pr}_{1-30}^{10:30}$, $\text{pr}_1^{15:30} - \text{pr}_{1-30}^{15:30}$

La salida deseada sigue siendo $\text{sign}(\text{pr}_0^{15:30} - \text{pr}_0^{10:30})$.

La fase de aprendizaje de esta red neuronal de 21 unidades de entrada y una única unidad de salida ha sido realizada en dos pasos: primero se reproduce la red equivalente al modelo lineal fijando los pesos de las entradas adicionales a cero, y en segundo lugar se congelan los pesos obtenidos con el aprendizaje anterior y se aprenden los pesos que anteriormente estaban fijados a cero. El resultado obtenido es de

$$4,88 \pm 0,21 \text{ tics por día y por contrato.}$$

Los beneficios son, en promedio, ligeramente mejores a los anteriores, pero lo que sí ha experimentado una mejora substancial es su dispersión: la red neuronal no ha mejorado demasiado las ganancias pero sí ha reducido dramáticamente su incertidumbre. La figura 2 muestra los beneficios acumulados (en tics) a lo largo del año 1995.

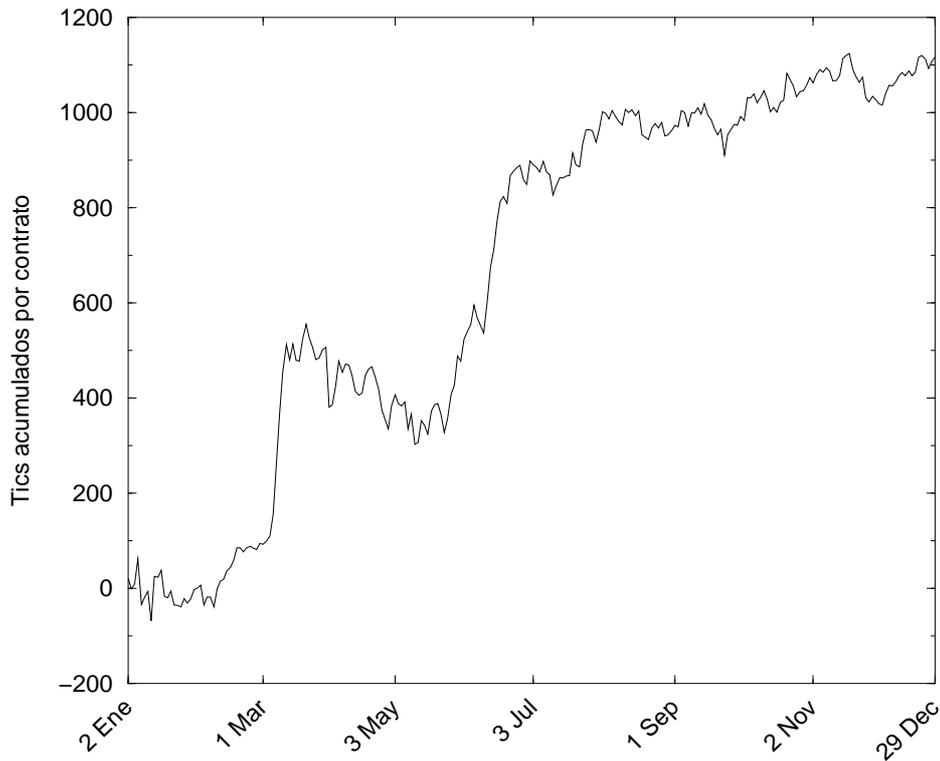


Figura 2: Beneficio acumulado durante 1995.

4 Evaluación del beneficio real

Cuando se dice que en un día particular se han ganado 4,88 tics por contrato, se está refiriendo al valor de la diferencia $pr_0^{15:30} - pr_0^{10:30}$. Sin embargo, $pr_0^{15:30}$ y $pr_0^{10:30}$ son precios promedio sobre intervalos de media hora. Es necesaria una evaluación más realista que tenga en cuenta que $pr_0^{10:30}$ podría ser bastante diferente del precio real con el que podemos comprar o vender a las 11:00h. Considerando que los contratos sean comprados o vendidos aleatoriamente en un intervalo de diez minutos alrededor de la hora deseada (es decir, 10:55–11:05 y 14:55–16:05 res-

pectivamente), y utilizando directamente los datos tic a tic, se puede comprobar que el beneficio se reduce en

$-3,0 \pm 0,3$ tics por día y por contrato.

Finalmente, la otra fuente de disminución del beneficio es la comisión que hay que pagar al operador cada vez que se realiza una operación. Esta tarifa se puede estimar en 0,075 tics por contrato. Como cada día se han de realizar dos operaciones por contrato, la reducción debida a este efecto es de

$-0,15$ tics por día y por contrato.

Combinando todos estos efectos el beneficio final que se obtiene es de

$1,7 \pm 0,4$ tics por día y por contrato.

Habría que mencionar que no se han considerado mecanismos de limitación de las pérdidas, ni eliminación de días difíciles de clasificar, ni ningún otro tipo de truco que pudiera mejorar los resultados. Únicamente nos hemos basado en los resultados proporcionados directamente por la red neuronal aplicada sobre todos los datos, sin excepciones.

5 Conclusiones

Hemos aplicado redes neuronales a la predicción de la evolución dentro de un mismo día del futuro financiero Bono 10, obteniendo un beneficio teórico de

$4,88 \pm 0,21$ tics por día y por contrato,

que se reduce a

$1,7 \pm 0,4$ tics por día y por contrato

cuando se tienen en cuenta las comisiones y la dispersión de los precios.

6 Agradecimientos

Este trabajo de investigación ha sido financiado por Tecno-Glia Ambroj S.L. bajo contrato 2447 con la Fundació Bosch i Gimpera de la Universitat de Barcelona.

Referencias

- [1] J.A. HERTZ, A. KROGH and R.G. PALMER, *Introduction to the theory of neural computation*, Addison-Wesley, Redwood City, California (1991).
- [2] B. MÜLLER and J. REINHARDT, *Neural networks: an introduction*, Springer-Verlag, Berlin (1991).
- [3] D.E. RUMELHART, G.E. HINTON and R.J. WILLIAMS, Learning representations by back-propagating errors, *Nature* **323** (1986) 533.
- [4] P. WERBOS, *Beyond regression: new tools for prediction and analysis in the behavioral sciences*, Ph.D. thesis, Harvard University (1974).