

Estrategias técnicas en el IGM basadas en redes neuronales.

Christian González Martel (Universidad de Las Palmas de Gran Canaria)

Fernando Fernández Rodríguez (Universidad de Las Palmas de Gran Canaria)

En este trabajo analizamos las predicciones sobre los rendimientos en el Índice General de la Bolsa de Madrid., obtenidas mediante una red neuronal feedforward.

Dicho análisis se realiza desde el doble punto de vista estadístico y económico.

Para contrastar, desde el punto de vista estadístico si los aciertos en la predicción del signo de los rendimientos son significativas, empleamos el contraste de Pesaran y Timmermann.

Desde el punto de vista económico, las predicciones se evalúan por medio de su capacidad de generar beneficios al ser empleadas en una estrategia técnica comparando los rendimientos obtenidos con la simple estrategia "buy and hold". Finalmente se procede a calcular la tasa de beneficio neto ideal y el ratio de Sharpe.

1 Introducción

La posibilidad de predecir los cambios de los precios de los bursátiles, ha constituido un tema recurrente en la literatura financiera. La literatura académica ha tomado partido durante los últimos treinta años por la Hipótesis de los Mercados Eficientes (HME). Según esta hipótesis ninguna técnica de predicción puede ser empleada, con éxito, para obtener "beneficios fáciles" en el mercado. En su forma débil, la HME establece que los precios de los activos reflejan, en cada instante, toda la información contenida en el pasado de las series y el logaritmo de los precios se comporta como si fuese un camino aleatorio, es decir

$$\log P_t = \log P_{t-1} + \mu + e_t$$

donde μ es la constante del modelo y e_t la perturbación independiente e idénticamente distribuida (iid).

Definiendo el rendimiento como $r_t = \log\left(\frac{P_t}{P_{t-1}}\right)$, resultará que tales rendimientos son (iid) (ver Fama, 1970).

Los contrastes empíricos sobre la HME han sido numerosos en la literatura. Como trabajo pionero podría citarse a Alexander (1961) que aplicó una *regla de filtro* en la que se compra un activo cuando su precio se incrementa en un x% y se vende cuando su precio cae un x%. Alexander concluye que en los índices Standard y Poor y Dow-Jones existen tendencias predecibles. Fama (1965) y Fama y Blume (1966) presentaron un análisis empírico más detallado de dichas reglas, corregidas por dividendos y costes de transacción, llegando a la conclusión de que tales reglas no mejoran la simple estrategia de "comprar y mantener".

Desde finales de los ochenta se ha prestado especial atención en contrastar la posibilidad de predecir los rendimientos bursátiles. Tal es el caso de diversos trabajos como Fama y French (1988), Lo y Mackinlay (1988) y Fama (1991).

Recientes estudios sobre los rendimientos de los activos han puesto de manifiesto la posibilidad de predecir hasta cierto grado en muchos mercados. Pesaran y Timmermann (1994) presentan evidencias sobre cierta posibilidad predictiva del exceso de rendimientos en carteras de los índices Standard y Poor 500 y el Dow Jones. Brock et al (1992), aunque no investigan la capacidad de generar beneficios de las reglas del Análisis Técnico en el índice Dow Jones, estudian las fuentes de la predecibilidad encontrada. Brock et al (1992) encuentran claros síntomas de estructura no lineal que se manifiesta en la naturaleza asimétrica de los rendimientos y en que los rendimientos obtenidos de señales de compra son mayores y menos volátiles que los obtenidos tras una señal de venta.

Gençay (1998) mide la capacidad de generar beneficios de simples reglas técnicas basadas en modelos de redes neuronales. Encuentran que tales reglas proporcionan beneficios significativos frente a la estrategia de comprar y mantener. Además concluyen que las predicciones sobre los signos de los rendimientos en tales modelos de redes son estadísticamente significativos.

Más recientemente, Fernández-Rodríguez, Sosvilla-Rivero y García-Artiles (1999) aplican técnicas locales no paramétricas de predicción, conocidas como "nearest-neighbours" en el índice Nikkei 255. Encuentran, usando técnicas de filtro, que durante los episodios no alcistas del mercado, la capacidad de generar beneficios de las predicciones no paramétricas es superior a la estrategia de "comprar y mantener".

En el mercado español se han producido, en este sentido, algunos trabajos sobre el Índice General de la Bolsa de Madrid (IGBM). Olmeda y Pérez (1995) estudian la posibilidad de mejorar la capacidad predictiva, de un modelo de camino aleatorio por medio del uso de redes neuronales.

Por otra parte, Fernández-Rodríguez, Sosvilla-Rivero y García-Artiles (1997) analizan la capacidad de generar beneficios de las predicciones "nearest-neighbours".

En este trabajo se pretende analizar las predicciones sobre los rendimientos en el Índice de la Bolsa General de Madrid realizadas con un modelo no paramétrico basado en una red neuronal. Emplearemos varios puntos de vista en el análisis. Por un lado, analizaremos la bondad predictiva del modelo; prestaremos atención tanto en la significatividad estadística de los aciertos en la predicción de los signos, como de la capacidad de generar beneficios de nuestras predicciones. Por otro, estudiaremos separadamente la capacidad predictiva del modelo en situaciones alcistas y bajistas que se presentan en el IGBM.

La estructura del trabajo es la siguiente. En la sección 2 se describirá el modelo empleado para realizar las predicciones. En la sección 3 se contrastará la bondad predictiva y se comentarán los resultados obtenidos. Se finalizará comentando las conclusiones de este trabajo en la sección 4.

2 Modelo

Las redes neuronales simulan una estructura computacional en paralelo, donde ciertas unidades individuales, llamadas neuronas, están interconectadas. Estas neuronas están organizadas en capas: la capa de entrada, la capa oculta o intermedia y la capa de salida.

En este trabajo se empleará una red feedforward basada en la regla de aprendizaje de retropropagación. Para un estudio de las redes feedforward y recurrentes ver Kuan and White (1994).

Las redes feedforward son redes que transforman las entradas o inputs en salidas o outputs con señales que viajan en un solo sentido. Cada unidad de la capa oculta y de la capa de salida contiene una función de transferencia o también llamada función de activación, que transforma la señal recibida. La capa de entrada no tiene función de transferencia pero se utiliza para distribuir las señales de entrada a la red.

Cada conexión tiene un peso o ponderación, que modifica las señales que pasan a través de dicha conexión.

Para estudiar el comportamiento de una red feedforward vamos a considerar una red con tres capas: la capa de entrada, la capa oculta que contiene k unidades y la de salida con una neurona.

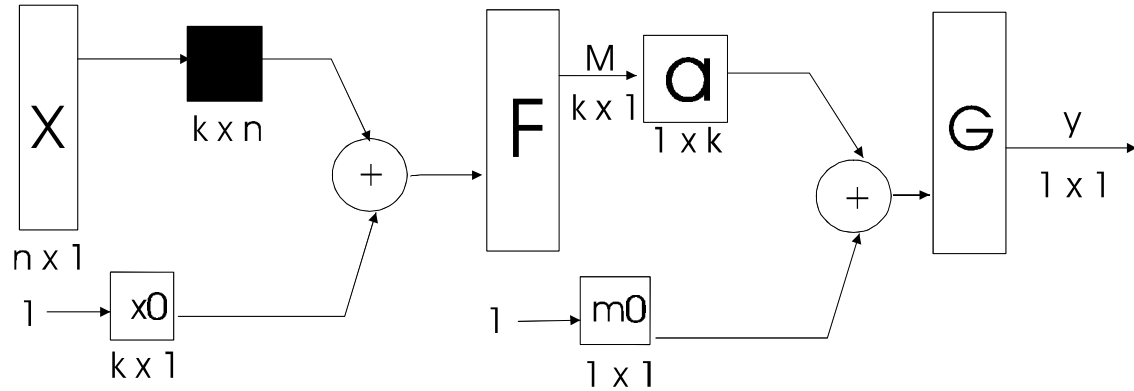


Fig. 1: Red neuronal de tres capas

En la figura se representa, esquemáticamente, una red neuronal de tres capas. La capa de entrada se pueden representar por el vector $X = (x_0, \dots, x_n)'$, donde x_0 es una entrada que siempre va a ser igual a uno y representa la desviación para la primera capa.

La capa oculta se representa por el vector $M = (m_1, \dots, m_k)'$. La salida de la red está representada por y . Cualquier unidad de la capa oculta recibe la suma ponderada de todas las entradas más la desviación y produce la salida,

$$m_j = F\left(\sum b_{ji}x_i\right), \quad j = 1, 2, \dots, k \quad i = 0, 1, \dots, n$$

donde F es la función de transferencia de la capa oculta, x_i es la entrada i -ésima y b_{ji} es el peso de la conexión que va de la i -ésima unidad de entrada a la j -ésima unidad de la capa oculta.

De la misma forma, la unidad de salida recibe como señal la suma ponderada de las señales de salida de la capa intermedia y produce el resultado:

$$y = G\left(\sum a_j m_j\right), \quad j = 0, 1, 2, \dots, k$$

donde G es la función de transferencia de la capa de salida, a_j es el peso de la conexión que va desde la unidad j -ésima de la capa intermedia y donde la unidad m_0 representa una desviación que siempre da una señal igual a uno, por lo que obtenemos

$$y = G\left(a_0 + \sum a_j F\left(\sum b_{ji} x_i\right)\right)$$

La red feedforward que emplearemos en este trabajo utiliza la regla de aprendizaje "backpropagation" o "retropropagación". Esta regla mide los errores que existen entre las salidas de la red y los valores reales, y a través de una técnica de vector gradiente descentente intenta ajustar los parámetros con el fin de encontrar el mínimo error cuadrático medio.

En este trabajo utilizaremos una red de tres capas con 9 entradas, que son los nueve retardos anteriores, $(r_{t-1}, r_{t-2}, \dots, r_{t-9})$, del rendimiento de las acciones que componen el IGM. Emplearemos una capa oculta con cuatro unidades en donde se trabajará con la función logarítmica y una capa de salida con una única neurona con la función tangente hiperbólica. La salida será la posición recomendada estimada y está comprendida entre -1 y 1 . Este dato es el que se comparará con la posición recomendada real y en el proceso de aprendizaje se minimizará el error cuadrático medio entre la estimación proporcionada por la red y el valor real. El modelo que utilizamos en este trabajo es:

$$y_t = G\left(a_0 + \sum_{j=1}^4 a_j F\left(b_{0j} + \sum_{i=1}^9 b_{ji} r_{t-i}\right)\right)$$

donde F sería la función logarítmica y G la función tangente hiperbólica.

La serie corresponde al Índice General de Madrid y contiene 6931 observaciones diarias, de precios al cierre, para el periodo comprendido entre el 4 de enero de 1966 y el 15 de octubre de 1997, lo que constituye una muestra representativa del volumen de transacciones realizadas en este mercado de valores.

Con el objetivo de comparar los resultados, en mercados alcistas y bajistas hemos realizado las predicciones de tres submuestras de 250 observaciones cada una, que corresponden a un tramo alcista, otro bajista y uno estable de la serie. En cada caso, en el entrenamiento de la red, se utilizará el conjunto de observaciones anteriores a cada uno de dichos. Tras realizar el proceso de aprendizaje, empleando sólo la información pasada, se realizarán las predicciones en cada submuestra. En cada caso se tiene un periodo de predicción de 250 observaciones diarias (aproximadamente un año). Lo que constituye un periodo suficientemente largo para evitar el "fisgoneo".

En las figuras 2,3 y 4 se presentan los diferentes tramos de la serie, en cuyas 250 últimas observaciones hemos realizado la predicción.

Existen diversos criterios para valorar la bondad de las predicciones realizadas por medio de una red neuronal. Algunos de ellos son estadísticos y reflejan el hecho de que las predicciones obtenidas por la

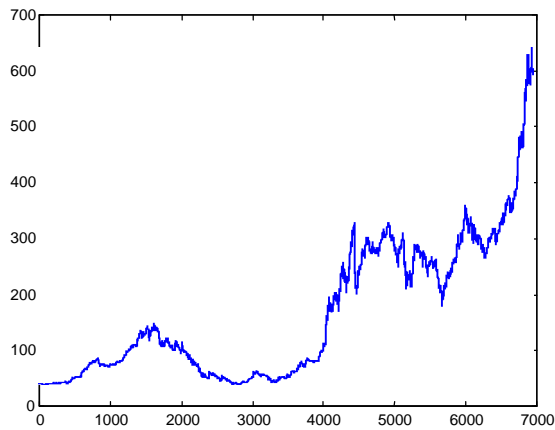


Fig. 3: Tramo de predicción alcista

Fig. 4: Tramo de predicción bajista

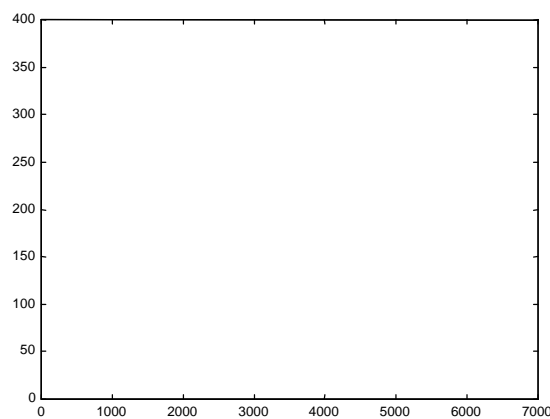


Fig. 2: Tramo de predicción estable

red sean significativamente mejores que las realizadas por algún otro procedimiento standard de predicción (en finanzas suele ser el modelo de camino aleatorio). En este terreno se sitúa el contraste de Pesarán y Timmermann (1992) que nosotros emplearemos para contrastar si la mejora en la predicción de los signos de los rendimientos es estadísticamente significativa.

El objetivo de cualquier inversión es maximizar el beneficio y minimizar el riesgo. Por ello, el criterio más adecuado para la valoración de la bondad de unas predicciones bursátiles no puede ser otro que la evaluación de los beneficios obtenidos por tales predicciones al ser convertidas en una regla técnica. Con este objetivo vamos a prestar especial atención a la "posición recomendada" que nos ofrece la red en cada instante. Esta posición recomendada será +1 si la red predice, para ese instante, rendimientos positivos. Será -1 cuando prediga rendimientos negativos.

La predicción de rendimientos positivos indicará, entonces, posiciones compradoras, mientras que las de rendimientos negativos posiciones vendedoras. Para ello las salidas de la red inferiores a 0 se igualarán a -1 y las salidas superiores a 0 se igualarán a 1. Cada día se realizarán dos operaciones. Una de apertura y otra de cierre. En este trabajo no consideraremos la existencia de costes de transacción.

Definimos, pues, el rendimiento total estimado como

$$\hat{R}_t = \sum_{t=n+1}^{n+r+1} \hat{y}_t r_t$$

donde p es el horizonte de predicción, $r_t = \log\left(\frac{P_t}{P_{t-1}}\right)$ es el rendimiento de la acción en el periodo t e \hat{y}_t es la posición recomendada en el periodo t . Esta variable se modelizará como función de los rendimientos pasados; n representa el número de observaciones.

Hemos comparado los beneficios obtenidos por la regla técnica que proporciona nuestra red con los beneficios R_b obtenidos por la sencilla estrategia de "comprar y mantener" que usaremos como referencia, donde

$$R_b = \log\left(\frac{P_{t+r}}{P_t}\right)$$

siendo p el periodo en que se mantiene en bolsa la posición.

3 **Resultados empíricos**

Los beneficios proporcionados por una regla técnica han sido evaluados atendiendo a diversos criterios. El ratio de Sharpe mide la media de los rendimiento por unidad de riesgo, es decir, a mayor ratio mayor es el rendimiento y menor la volatilidad.

El ratio de beneficio ideal que mide el consiente entre el rendimiento total estimado por la red y un predictor perfecto, es decir, un predictor en el que se ha acertado correctamente todas las posiciones.

$$R_i = \frac{\sum_{t=n+1}^{n+r+1} \hat{y}_t r_t}{\sum_{t=n+1}^{n+r+1} |r_t|}$$

De acuerdo con esto, si $R_i = 1$ las predicciones han sido perfectas, si $R_i = -1$ todas las predicciones han sido calculadas de forma incorrecta. Por lo que $R_i = 0$ se utilizará como referencia para evaluar la bondad de la estrategia de inversión.

Desde el punto de vista estadístico hemos empleado el test direccional de Pesaran y Timmerman. Dicho contraste evalúa la significatividad en la predicción de los signos. Bajo la hipótesis nula de independencia entre los valores reales y los predichos, el estadístico de dicho contraste se distribuye,

asintóticamente, como una $N(0,1)$, por lo que los valores críticos al 1%, 5% y 10% como niveles de significación serán 2.33, 1.645 y 1.82, respectivamente.

El porcentaje en los signos mide el porcentaje de veces que la predicción acierta con la posición correcta de compra o venta

	<i>Tramo bajista</i>	<i>Tramo estable</i>	<i>Tramo alcista</i>
Rendimiento total estimado	0.48	0.27	0.29
Ratio de beneficio ideal	0.25	0.17	0.14
Ratio de Sharpe	0.19	0.13	0.11
Predicción en signos	0.54	0.57	0.58
Test de Pesaran y Timmerman	0.97	2.24	2.26
"Comprar y mantener"	-0.40	0.0019	0.44

Se puede observar como el tramo en donde se ha conseguido un mayor rendimiento es el tramo bajista que coincide con el peor resultado que se podría obtener con la estrategia "comprar y mantener". En este tramo vemos que hemos obtenido casi la mitad de lo invertido mientras que siguiendo la estrategia "comprar y mantener" se hubiese perdido el 40% de la inversión. Aún así, se estaría muy lejos de la rentabilidad ideal ya que hemos conseguido un 25% de lo que se podría haber obtenido si se hubiese acertado todas las posiciones correctamente.

En los otros dos tramos el rendimiento estimado está en torno al 30%, aunque en el tramo alto se hubiese conseguido un 44% de la inversión si se hubiese optado por la estrategia "comprar y mantener". Aunque en ambos tramos se ha conseguido acertar entre el 57% y el 58% de los signos, es decir, se ha conseguido predecir correctamente la posición del inversor, y además significativamente como muestra el test de Pesaran y Timmerman, el beneficio ideal es inferior al 20% en dichos tramos por lo que se encuentran lejos de ese rendimiento ideal.

El tramo donde mejor se ha comportado la red es en este tramo bajista ya que aparte de ser donde más rentabilidad se ha obtenido y más cerca se ha estado de la rentabilidad ideal es el que tiene un más alto Ratio Sharpe. El tramo bajista es por tanto, entre los tres, donde se ha obtenido una mayor rentabilidad por unidad de volatilidad (riesgo). El tramo de peor comportamiento de las predicciones de la red ha sido el alcista; no sólo por los altos beneficios aportados por la estrategia de "comprar y mantener", sino también porque se ha conseguido los peores ratios de beneficio ideal y de Sharpe.

4 Conclusiones

En este trabajo estudiamos la posibilidad de implementar reglas técnicas basadas en modelos no paramétricos de redes neuronales. Nuestras conclusiones apuntan a que resulta imposible batir, con

nuestro modelo, a la estrategia de "comprar y mantener" en un periodo de alzas generalizadas tan fuertes como el examinado. No obstante, nuestras reglas técnicas son capaces de producir rendimientos superiores a la estrategia de "comprar y mantener" en tramos bajistas o estables del IGBM. Los horizontes de predicción empleados en todos los casos son de 250 observaciones, un periodo de tiempo estimable y que evita la posibilidad de "fisgoneo".

Sin embargo, estas conclusiones deben tomarse con cautela. Para examinar tanto la capacidad de generar beneficios por estas reglas, como las consiguientes implicaciones en el terreno de la eficiencia del mercado, deben tenerse en cuenta los costes de transacción. Esto es algo que realizaremos en una segunda versión de este trabajo.

5 Bibliografía

- Alexander, S. S. (1961): "Price movements in speculative markets: Trends or random walks?", *Industrial Management Review* 2, pp. 7-26
- Brock, W. A.; Lakonishock, J. and LeBaron, B. (1992): "Simple technical trading rules and the stochastic of stock returns", *Journal of Finance* 47, pp. 1731-1764
- Fama, E. F. (1970): "Efficient capital markets: A review of theory and empirical work", *Journal of Finance* 25, pp. 383-417
- Fama, E. F. (1991): "Efficient capital markets: II", *Journal of Finance* 46, pp 1575-1617
- Fama, E. F. and Blume, M. E. (1966): "Filter rules and stock market trading", *Journal of Business* 39, pp. 226-241
- Fama, E. F. and French, K. R. (1988): "Dividend yields and expected stock returns", *Journal of Financial Economics* 22, pp. 3-2
- Fama, E.F. (1965): "The behaviour of Stock-Market Prices", *Journal of Bussines* 38, pp. 34-105
- Fernández-Rodríguez, F.; Sosvilla-Rivero, S. and García Artiles, M. D. (1999): "Nearest Neighbour forecast for the nikkey index", aceptado en *Japan and the World Economy*.
- Fernández-Rodríguez, F.; Sosvilla-Rivero, S. and García Artiles, M. D. (1997): "Using nearest-neighbour predictors to forecast the spanish stock markets", *Investigaciones Económicas XXI*, pp. 75-91
- Gençay, R. (1988): "Optimization of technical trading strategies and the profitability in security markets", *Economics Letters* 59, pp. 249-254

- Kuan, C. M. and White, H. (1994): "Artificial neural networks: An econometric perspective", *Econometric Reviews* 13, pp.1-91
- Lo, A. and MacKinlay, A. C. (1990): "Data snooping biases in test of financial asset pricing models", *Review of Financial Studies* 3, pp. 431-468
- Olmeda, I. and Pérez, J. (1995): "Non-linear dynamics and chaos in the spanish stock market", *Investigaciones Económicas XIX*, pp. 217-248
- Pesaran, M. H. and Timmermann, A. (1992): "A simple nonparametric test of predictive performance", *Journal of Business and Economic Statistics* 10, pp. 461-465