

Predicción del índice IBEX-35 aplicando Máquinas de Soporte Vectorial y Redes Neuronales.

Forecasting IBEX-35 index using Support Vector Machines and Neural Networks.

Rosillo R¹, Dunis CL, De la Fuente D, Pino R.

Abstract The aim of this research is to examine the application of Support Vector Machines (SVMs) to forecast the weekly change in the IBEX-35 index. The data covers the period between 18/10/1990 and 29/10/2010. The inputs of the SVM are the Relative Strength Index (RSI) and the Moving Average Convergence Divergence (MACD). The SVM is used in order to determine the best situations to buy or sell the market. The two outputs of the SVM are both the direction of the market and the probability attached to each forecast market move. Better results are obtained analyzing recent periods rather than using all the database information. The SVM results are compared with other techniques.

Resumen El objetivo de este artículo es examinar la aplicación de las Máquinas de Soporte Vectorial (SVM) para predecir el movimiento del IBEX-35 semanalmente. La base de datos del estudio está comprendida entre el 18/10/1990 y el 29/10/2010. Las entradas de la SVM son: Relative Strength Index (RSI) y Moving Average Convergence Divergence (MACD). La SVM encuentra la mejor situación de compra o venta en el Mercado. Las salidas de la SVM son la dirección que va a tomar el Mercado y la probabilidad de que ese movimiento se produzca. Se obtienen mejores resultados analizando periodos recientes que usando toda la base de datos, además los resultados de la SVM son comparados con otras técnicas.

Keywords: RSI, MACD, SVM, quantitative finance; **Palabras clave:** RSI, MACD, SVM, finanzas cuantitativas.

¹ Rafael Rosillo Cambor (✉)

Grupo de Ingeniería de Organización (GIO). Escuela Politécnica de Ingeniería de Gijón, Campus de Viesques s/n, 33204 Gijón, Spain
e-mail: rosillo@uniovi.es

1.1 Introducción

El propósito de este trabajo es presentar un método para predecir el movimiento semanal del IBEX-35. Para realizar esta predicción, nuestro modelo necesita los indicadores del Análisis Técnico, entre los que hemos escogido: Relative Strength Index (RSI) y el Moving Average Convergence Divergence (MACD) ya que son los más utilizados por los inversores. Nuestro estudio también muestra la probabilidad² asociada a cada movimiento, esta salida puede ayudar a los inversores en su toma de decisión porque no todos están dispuestos a asumir el mismo riesgo.

La metodología utilizada por el algoritmo son las Máquinas de Soporte Vectorial (SVMs).

El resto del artículo está estructurado como sigue; en la sección 2, se presenta el estado del arte de las SVMs La sección 3 se explica la regla de trading aplicada. La sección 4 muestra los resultados empíricos de la regla de trading y las conclusiones del trabajo son expuestas en la sección 5.

1.2 Estado del Arte de las Máquinas de Soporte Vectorial

Las SVMs fueron desarrolladas originalmente por Vapnik (1998). Para una introducción más detallada se recomiendan Burges (1998) y Evgeniou et al. (2000).

El objetivo de las SVMs es minimizar el riesgo empírico (ERM) y construir modelos más confiables. Esto quiere decir, que tratan de dar prioridad a la construcción del modelo y que por lo tanto pueden existir muchos errores para un periodo concreto del tiempo, así minimizan el riesgo estructural.

La principal idea de las SVMs es construir un hiperplano como superficie de decisión tal que el margen de separación entre los ejemplos positivos y negativos sea máximo (Xu et al., 2009), esto se conoce como hiperplano óptimo de separación, como se muestra en la Figura 1.

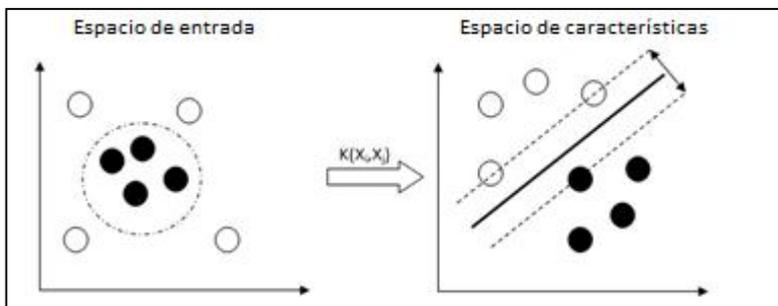


Fig 1.1 Ejemplo de cómo funciona una función kernel

² Se entiende por probabilidad el grado de pertenencia a un conjunto.

Las SVMs pueden ser usadas de diferentes maneras: regresión o clasificación.

La aplicación de las SMVs para resolver problemas financieros se comienzan a aplicar a principios del Siglo XX, y las más relevantes relacionadas con nuestra investigación son descritas a continuación.

En 2001, Cao y Tay (2001) compararon SVMs con otras técnicas como las redes neuronales back-propagation. Los Algoritmos Genéticos son utilizados en Tay y Cao (2001a, 2001b) para el ajuste de los parámetros de las SVMs. Dos aplicaciones para la predicción de las series financieras con SVMs fueron desarrolladas en 2003: en Cao y Tay (2003) las SVMs se aplican al problema de la predicción de varios contratos de futuros del Mercado Mercantil de Chicago, las cuales muestran la superioridad de las SVMs sobre las NN back-propagation, en Kim (2003), las SVMs se utilizan para predecir la dirección del cambio en el índice de Corea y son comparadas con las NN back-propagation. Los resultados experimentales muestran que las SVMs superan a los otros métodos y que deben ser consideradas como una metodología prometedora para el análisis de las series financieras. En Huang et al. (2005), utilizaron las SVMs en modo clasificación para predecir la dirección del movimiento del índice Nikkei225 con resultados muy interesantes. Las SVMs mejoran los resultados de las NN En este estudio se emplea una SVM en modo Clasificación.

1.3 Base de Datos y Diseño de la Regla de Trading

En la siguiente sección se explica el diseño de la regla de trading.

1.3.1 Base de Datos

El conjunto de datos ha sido obtenido de <http://finance.yahoo.com/>. Los datos son almacenados de forma diaria entre el 18/10/1990 y el 29/12/2010 y la SVM es entrenada en diferentes periodos de tiempo con el conjunto de datos para conseguir diferentes resultados y luego compararlos.

Aunque nuestro conjunto de datos es diario, hay que recordar que lo que se busca es una predicción semanal del movimiento del IBEX-35.

En la Tabla 1 se muestran algunos ejemplos de entrenamiento y test, además del porcentaje de acierto en el periodo de test.

Tabla 1.1 Diversos ejemplos realizados con distintos periodos de entrenamiento y test

Color	Periodo de Entrenamiento	Periodo de Test	% de acierto con C = 1
Rojo	29/12/2009-16/09/2010	17/09/2010-22/10/2010	87.5%
Verde	03/11/2008-12/08/2010	13/08/2010-22/10/2010	100%
Azul	26/11/2002-12/01/2010	13/01/2010-22/10/2010	47.15%
Negro	26/11/1998-19/08/2009	20/08/2009-22/10/2010	47.33%

1.3.2 Datos de entrada

Los datos de entrada de la SVM son los indicadores RSI y MACD. Estos indicadores han sido elegidos porque son los más utilizados en el Análisis Cuantitativo. En Rosillo et al. (2013), se explica como el RSI genera unas mayores ganancias de capital sobre blue chips, mientras que el Momento genera mayores rentabilidades para Small Caps, también han sido analizados el MACD y el Estocástico sobre las compañías del Mercado Continuo Español.

Los datos de entrada muestran un caso no separable, no es posible separar las recomendaciones de compra y las recomendaciones de venta sin errores.

1.3.2.1 RSI

El RSI fue diseñado por J. Welles Wilder Jr. (1978), es un indicador del Análisis Técnico que muestra la fuerza del precio mediante la comparación de los movimientos individuales al alza o a la baja de los sucesivos precios de cierre.

A continuación se muestra una breve explicación del cálculo de su valor mediante las ecuaciones (1) y (2). Si se necesitase una explicación más detallada, se recomienda J. Welles Wilder Jr. (1978).

Para cada día se calcula un cambio al alza (U) o un cambio a la baja (D).

Los días a la baja se caracterizan porque el precio de cierre actual es menor que el precio de cierre del día anterior.

$$U = cierre_t - cierre_{t-1}; D = 0$$

Los días al alza se caracterizan porque el precio del cierre del día actual es mayor que el precio del cierre del día anterior.

$$U = 0; D = cierre_t - cierre_{t-1}$$

$$RS = \frac{EMA[N]of U}{EMA[N]of D} \quad (1.1)$$

$$RSI = 100 - 100 \frac{1}{1+RS} \quad (1.2)$$

El periodo más utilizado para el cálculo del RSI es de 14 días y es el empleado en este estudio, el RSI oscila entre 0 y 100, sin embargo ha sido normalizado entre 1 y -1 para introducirlo en la SVM.

1.3.2.2 MACD

El MACD está diseñado principalmente para identificar los cambios de tendencia, se basa en el cálculo de dos medias móviles exponenciales (EMA) de diferente longitud. A continuación se explica su cálculo en la ecuación:

$$MACD(n) = \sum_{i=1}^n EMA_k(i) - \sum_{i=1}^n EMA_d(i) \quad (1.3)$$

donde $k=12$ and $d=26$

$$EMA_n(i) = \alpha * p(i) + (1 - \alpha) * EMA_n(i - 1) ; \alpha = \frac{2}{1+n}$$

y n es el número de días y $p(i)$ es el precio del activo en el día i .

En este estudio se han seleccionados unas EMAs de 12 y 26 días porque son las más utilizadas para calcular el MACD, si se necesitase una mayor explicación sobre el MACD se recomienda consultar (Murphy, 1999).

El rango del MACD ha sido normalizado entre -1 y +1, al igual que el del RSI, para ser usado como entrada en la SVM.

1.3.3 Diseño de la Regla de Trading

Un esquema de la misma, aparece en la Figura 2.

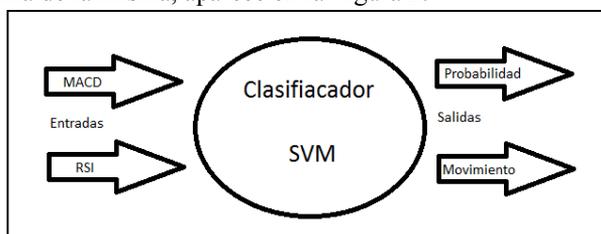


Fig 1.2 Diseño de la regla de trading

La regla de trading se genera con la utilización del Clasificador de las SVM (SVMC). A continuación se detallan los pasos que han sido llevados a cabo para la realización de la misma.

Primero, el SVMC analiza el RSI y el MACD.

Luego, el SVMC clasifica los diferentes valores del IBEX-35 en dos clases: situación de compra y situación de venta. Como se mencionó anteriormente, la predicción se hace semanal.

Posteriormente se ajusta el SVMC en base a la aversión al riesgo que tenga el inversor. Además se ajustan los propios parámetros de la Máquina de Soporte Vectorial como el Kernel y el parámetro C.

Por último se calcula el ratio de aciertos dependiendo de la probabilidad seleccionada por el inversor.

Finalmente, dado un valor de RSI y un valor del MACD, el SVMC es capaz de predecir el movimiento que va a ocurrir y estimar con qué probabilidad ese movimiento ocurrirá.

1.3.5 Salidas

Las salidas de la SVM son el movimiento(alcista o bajista) que tomará el Mercado la próxima semana y la probabilidad de que ese movimiento suceda, se entiende probabilidad como grado de pertenencia a un conjunto en este estudio.

1.4 Resultados

Se han realizado muchas pruebas y las más relevantes se reflejan a continuación:

En la Figura 3 se muestra una comparación de los resultados obtenidos en las 4 pruebas más relevantes. Los diferentes periodos estudiados en la Figura 3, son mostrados en la Tabla 1.

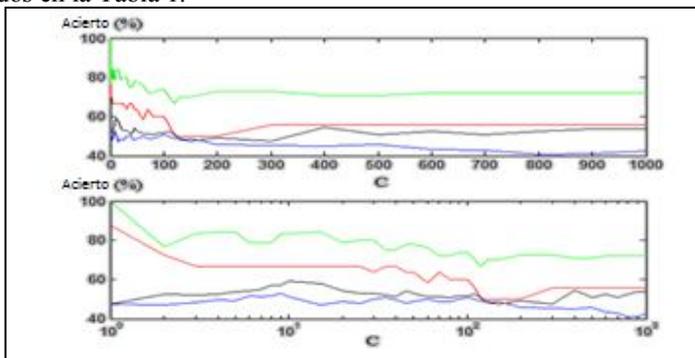


Fig 1.3 Comparación de los resultados obtenidos con periodos largos y cortos de entrenamiento.

Como se puede apreciar en la Figura 3, los mejores resultados son obtenidos con cortos periodos de entrenamiento. Por el contrario, los peores resultados se obtienen con largos periodos de entrenamiento. El eje vertical muestra la proporción de aciertos en % conseguidos por la SVM en el periodo de test, y el eje horizontal muestra el valor del parámetro C de la SVM. Como se puede apreciar, se consiguen mejores resultados con un valor bajo de C.

La Figura 4 muestra una de las pruebas realizadas en la que se asume que el inversor es adverso al riesgo, y solo toma posiciones en el mercado cuando el grado de pertenencia a un conjunto sea superior a 0.9 o inferior a -0.9. Los resultados obtenidos han sido generados con un periodo de entrenamiento entre el 03/11/2008 y el 12/08/2010. El valor del parámetro C es $C=1$.

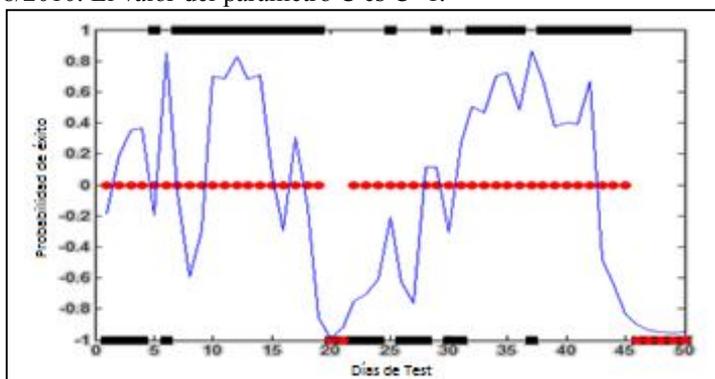


Fig 1.4 Resultados generados con el periodo de entrenamiento entre el 03/11/2008 y el 12/08/2010

Los cuadrados negros son la salida deseada, la línea azul muestra para cada día de test el grado de pertenencia al conjunto (probabilidad de éxito), y los puntos rojos indican la acción sugerida por la SVM. La forma de interpretar este gráfico es la siguiente: si coincide el punto rojo con el cuadrado negro quiere decir que la SVM ha acertado y si el punto tiene un valor de probabilidad de éxito 0, la SVM no hace ninguna recomendación. Para esta prueba en particular, la SVM ha generado 7 recomendaciones de las que no ha fallado ninguna.

A continuación se comparan los resultados de la SVM con otras estrategias, como se puede observar en la Tabla 2.

La Tabla 2 muestra los puntos del IBEX-35 que cada técnica ha conseguido, el mejor resultado se logra utilizando la SVM. Cada valor de la misma explica los puntos que un inversor hubiese conseguido usando la técnica analizada en un periodo concreto del tiempo. En el contrato de futuros del IBEX-35, 1 punto = 10 euros, entonces 757 puntos = 7570€. NN han sido entrenadas con 10 entradas ($RSI_n, RSI_{n-1} \dots RSI_{n-4}, MACD_n, MACD_{n-1}, \dots, MACD_{n-4}$), porque los resultados con tan solo dos entradas era muy malos.

Tabla 1.2 Ganancias acumuladas en los periodos de test

Técnica	17/09/2010-29/10/2010	13/08/2010-29/10/2010
SVM	757	695
NN	486	-2583
Compra y Espera	224	537

1.5 Conclusiones

Este estudio muestra que las SVMs producen mejores resultados cuando se utilizan en períodos cortos de entrenamiento, mientras que si se utiliza la base de datos completa el porcentaje de acierto disminuye. Esto puede ser causado por el overfitting y el cambio en las condiciones del mercado.

El mejor resultado conseguido por la SVM ha sido un 100% de aciertos seleccionando aquellas situaciones que cumplan un 90% o más de grado de pertenencia a un conjunto. En nuestro estudio, para el periodo y el mercado seleccionado, las SVMs producen mejores resultados que las redes neuronales y la estrategia de Compra y Espera. Por otra parte, el protocolo de optimización SVM desarrollado permite al usuario determinar su propio nivel de aversión al riesgo, lo que hace este trabajo aplicable a una amplia gama de inversores.

Agradecimientos

Soporte Financiero otorgado por el Gobierno del Principado de Asturias.

1.6 Referencias

- Burges, C. (1998) A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2, 121-167.
- Cao, L. and Tay, F. (2001) Financial forecasting using support vector machines. *Neural Computing Applications*, 10, 184-192.
- Cao, L. and Tay, F. (2003) Support vector machine with adaptive parameters in financial time series forecasting. *IEEE Transactions on Neural Networks*, 14, 1506-1518.
- Evgeniou, T., Pontil, M. and Poggio, T. (2000) Regularization networks and support vector machines. *Advances in Computational Mathematics*, 13, 1-50.
- Huang, W., Nakamori, Y. and Wang, S. Y. (2005) Forecasting stock market movement direction with support vector machine. *Computers & Operations Research*, 32, 2513-2522.
- Kim, K. (2003) Financial time series forecasting using support vector machines. *Neurocomputing*, 55, 307-319.
- Murphy, J. J. (1999) *Technical Analysis of the Financial Markets*. Institute of Finance, New York.
- Rosillo, R., De la Fuente, D., Brugos, JAL. (2013): Technical analysis and the Spanish stock exchange: testing the RSI, MACD, momentum and stochastic rules using Spanish market companies. *Applied Economics*, 45, 1541-1550.
- Tay, F. and Cao, L. (2001a) Application of support vector machines in financial time series forecasting. *Omega*, 29, 309-317.
- Tay, F. and Cao, L. (2001b) A comparative study of saliency analysis and genetic algorithm for features selection in support vector machines. *Intelligent Data Analysis*, 5, 191-209.
- Vapnik, V. N. (1998) *Statistical Learning Theory*. New York. Wiley.
- Welles Wilder, J. Jr. (1978) *New Concepts in Technical Trading Systems*, Greensboro, N.C.: Hunter Publishing Company.
- Xu, X., Zhou, C. and Wang, Z. (2009) Credit scoring algorithm based on link analysis ranking with support vector machine. *Expert Systems with Applications*, 36, 2625-2632.