

Predicción del ciclo solar 23 con un algoritmo Adaline basado en funciones potenciales de base radial

H.D. NAVONE*, P.F. VERDES, P.M. GRANITTO, H.A. CECCATTO

IFIR (UNIVERSIDAD NACIONAL DE ROSARIO - CONICET)
BV. 27 DE FEBRERO 210 BIS - (2000) ROSARIO - ARGENTINA
e-mail: hnavone@ifir.edu.ar

El uso de funciones potenciales permite la generalización a problemas no lineales de algoritmos de aprendizaje de naturaleza lineal. Dichas funciones, denominadas 'kernels', permiten mapear los datos originales a un espacio de alta dimensión (espacio de linearización) en el cual es posible aplicarles técnicas lineales ya sea de regresión o clasificación. En este trabajo se utiliza la versión generalizada basada en kernels del algoritmo ADALINE de Widrow y Hoff para predecir el ciclo solar 23. Los kernels se definen usando funciones de base radial y el ajuste del parámetro que define a estas funciones se realiza utilizando un intervalo de validación. Se concluye que los resultados que proporciona este algoritmo, conceptualmente simple y numéricamente robusto, compiten favorablemente con los ya reportados en la literatura sobre el mismo problema obtenidos con otras metodologías.

The use of potential functions allows the generalisation to non-linear problems of learning algorithms of linear nature. Such functions, called 'kernels', allow mapping the original data to a high dimensional space (linearisation space) in which is possible to apply linear techniques to them, either of regression or classification. In this work the generalised version based on kernels of the Widrow and Hoff's Adaline algorithm is used to predict the 23 solar cycle. The kernels are defined using radial basis functions and the adjustment of its internal parameter is done by using a validation interval. In conclusion, the results given by this algorithm, conceptually simple and numerically robust, compete favourably with the already reported on the same problem obtained with other methodologies.

PACS: 07.05.Mh

Introducción

La predicción de la actividad solar y de la ocurrencia de los valores máximos de cada ciclo solar es de gran importancia debido a su relación con tormentas solares y sus posibles efectos sobre los sistemas de telecomunicaciones y el clima. Es por ello que este fenómeno es muy estudiado en la literatura, siendo motivo del desarrollo y de la aplicación de numerosos métodos de predicción.

En este trabajo se predice el máximo correspondiente al ciclo solar 23 (ciclo actual) y su comportamiento subsecuente mediante un nuevo algoritmo que constituye la generalización no lineal del Adaline de Widrow y Hoff^{1,2}.

La idea de extender algoritmos lineales simples al dominio no lineal ya ha sido explorada en el método de funciones

potenciales³ y actualmente constituye una de las operaciones matemáticas necesarias que intervienen en la construcción de las Support Vector Machines⁴ (SVM).

Las funciones potenciales o kernels son utilizadas para mapear los vectores de entrada (inputs) en un espacio de alta dimensión, denominado espacio de atributos (feature space) o espacio de linearización, en donde se computan implícitamente algoritmos de naturaleza lineal. El Adaline es uno de estos métodos, el cual trata de construir el modelo

$$f(\vec{x}) = \vec{w} \cdot \vec{x} + b \quad \vec{w}, \vec{x} \in \mathcal{R}^d, b \in \mathcal{R}$$

basándose en el conjunto de datos de entrada-salida $\{\vec{x}_i, t_i\}$, $i = 1 \dots L$. Este ajuste se realiza por minimización del error cuadrático medio entre

* Autor a quien debe dirigirse la correspondencia.

el modelo $f(\cdot)$ y los valores de salida medidos t correspondientes a cada vector de entrada \vec{x} .

Aplicando descenso según el gradiente resultan las siguientes reglas de actualización de los parámetros del modelo:

$$\begin{aligned}\bar{w} &\leftarrow \bar{w} + \eta (t - f(\vec{x})) \bar{x} \\ b &\leftarrow b + \eta (t - f(\vec{x}))\end{aligned}$$

siendo η una constante positiva.

Los pesos \bar{w} pueden ser expresados como una combinación lineal de los vectores de entrada presentes en el conjunto de ajuste²

$$\bar{w} = \sum_{i=1}^L \alpha_i \bar{x}_i$$

Usando esta representación basada en los datos, el Adaline toma la siguiente forma^{6,7}:

$$f(\vec{x}_j) = \sum_{i=1}^L \alpha_i (\bar{x}_i \cdot \vec{x}_j) + b$$

Las reglas de actualización se aplican ahora sobre los parámetros α , resultando:

$$\begin{aligned}\alpha &\leftarrow \alpha + \eta (t - f(\vec{x})) \\ b &\leftarrow b + \eta (t - f(\vec{x}))\end{aligned}$$

La representación dependiente de los datos del Adaline permite introducir la extensión no lineal del modelo. Esto se logra aplicando una transformación fija de los datos de entrada $\vec{z}_j = \varphi(\vec{x}_j)$, lo suficientemente flexible como para capturar la forma funcional subyacente a los mismos, y entrenando el Adaline en el espacio de los \vec{z} . De esta manera el Adaline no lineal (K-Adaline) expresado en la representación dependiente de los datos resulta:

$$f(\vec{x}_j) = \sum_{i=1}^L \alpha_i K(\bar{x}_i, \vec{x}_j) + b$$

donde

$$K(\bar{x}_i, \vec{x}_j) = \varphi(\bar{x}_i) \cdot \varphi(\vec{x}_j)$$

es el producto escalar definido en el espacio de los atributos. Por lo tanto, no es necesario conocer $\varphi(\vec{x})$ si es posible definir funciones $K(\cdot)$ que correspondan a productos escalares en algún espacio de atributos. La solución a este problema viene dada por el teorema de Mercer, ya aplicado por Vapnik en el desarrollo de las SVM⁴. En particular, las funciones de base radial (RBF) cumplen con las condiciones del teorema de Mercer y por lo tanto pueden ser utilizadas como funciones potenciales que representan productos escalares en el espacio de los atributos. El kernel RBF se define de la siguiente manera:

$$K(\vec{x}_i, \vec{x}_j) = \exp\left(\frac{-|\vec{x}_i - \vec{x}_j|^2}{2\sigma^2}\right)$$

donde σ es un parámetro a ajustar en el proceso de modelado.

Predicción de series temporales con el Adaline no-lineal

El problema a resolver consiste en reconstruir la dinámica $t = F(\vec{x})$ subyacente en un registro histórico a partir de los propios datos. Dicha reconstrucción se lleva a cabo en el espacio de pseudofases en donde residen los vectores $\vec{x} = (x_{t-1}, x_{t-2}, \dots, x_{t-d})$ conformados por los valores de la serie temporal. De esta manera el modelo predictivo se obtiene ajustando el Adaline no-lineal sobre la base de datos generada por los pares de valores $\{\vec{x}, t\}$, siendo $t = x_t$.

A los efectos de determinar los parámetros del modelo y evaluar los resultados obtenidos se divide el registro histórico disponible en tres subconjuntos: ajuste $\{\vec{x}, t\}_{t=1 \dots L}$, validación $\{\vec{x}, t\}_{t=L+1 \dots M}$ y test $\{\vec{x}, t\}_{t=M+1 \dots N}$.

Para cada valor del parámetro σ que define al kernel, el monitoreo de la evolución del error en el conjunto de validación permite detener el proceso de ajuste antes de que se produzca sobre-entrenamiento. Por otra parte, el registro de los valores mínimos de este error para cada kernel nos permite seleccionar el valor óptimo

del parámetro σ y establecer el modelo predictivo final.

La medida de error utilizada es la varianza relativa promedio (ARV) definida por la siguiente expresión:

$$ARV(S) = \frac{\sum_{i \in S} (t_i - p_i)^2}{\sum_{i \in S} (t_i - \bar{t}_i)^2}$$

donde S indica alternativamente el conjunto de entrenamiento, validación o test usado, p_i es la predicción del Adaline no-lineal para t_i , y \bar{p}_i es el promedio de p_i en S . Esta cantidad es 1 para una predicción constante igual al promedio de los registros y 0 para una predicción perfecta.

Predicción del Ciclo Solar 23

La serie temporal estudiada corresponde al promedio del número de manchas solares anuales. Los datos utilizados abarcan el período 1700-1998, se define como intervalo de test el ciclo solar 21 (1976-1986) y como conjunto de validación los 35 patrones restantes más cercanos al ciclo 23. La dimensión del espacio de pseudofases usado para realizar la reconstrucción de la dinámica se fijó en $d = 12$ de acuerdo a resultados ya reportados en la literatura y a los efectos de permitir la comparación directa con otros métodos⁸.

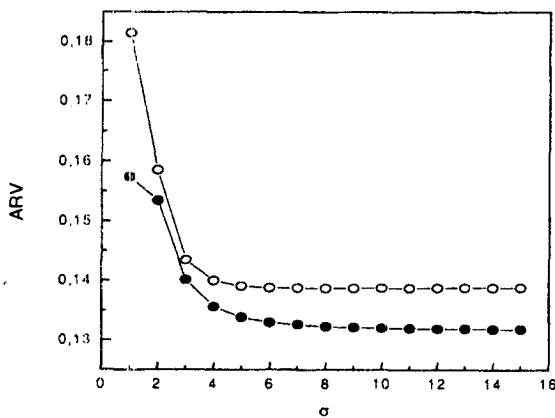


Fig.1: Mínimo ARV de validación (círculos vacíos) obtenido para cada kernel RBF y el correspondiente ARV de test (círculos sólidos).

En la Fig.1 se muestra la evolución de los mínimos errores de validación y los correspondientes errores en el intervalo de test obtenidos para cada uno de los kernels RBF. La

curva de validación registra un mínimo en $\sigma = 8$ y luego tiende a crecer muy lentamente. La evolución de los errores para el conjunto de test muestra un comportamiento similar, no registrándose grandes variaciones a partir del σ óptimo. En la Tabla 1 se muestran los valores de ARV de test correspondientes al K-Adaline y a un modelo basado en redes neuronales artificiales⁸ (ANN) con arquitectura 12:3:1. A partir de estos resultados se verifica que el K-Adaline logra una mejor generalización respecto del modelo ANN, sobre todo teniendo en cuenta que en este último caso, el ciclo solar 21 se usó como intervalo de validación.

TABLA 1: ARV CALCULADOS SOBRE EL CICLO SOLAR 21

Métodos	ARV
ANN 12:3:1	0.1402
K-ADALINE	0.1323

En la Fig.2 se presenta la predicción del ciclo solar 23 obtenida con el K-Adaline. Como puede apreciarse el máximo se registra en el año 2000, alcanzando un valor de 120 ± 32 manchas solares promedio. El error fue estimado calculando el error promedio de las predicciones a 2-pasos sobre todo el intervalo de test, ya que ésta fue la cantidad de iteraciones necesarias para alcanzar el máximo en el ciclo 23.

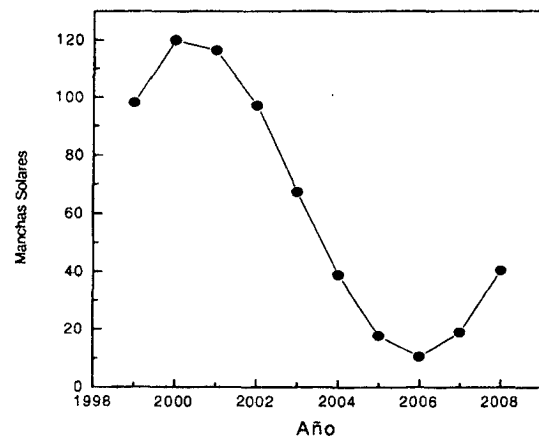


Fig.2: Predicción del ciclo solar 23 con el modelo construido por el algoritmo K-Adaline.

En la Tabla 2 se muestran los valores obtenidos con distintos métodos para el máximo del ciclo solar 23 reportados por Kane⁹, además de otros reportados recientemente en la literatura.

TABLA 2: PREDICCIONES DEL MÁXIMO DEL CICLO SOLAR 23

Referencia	Valor
Bouvar et al. ¹⁰ (1997)	158 ± 18
Calvo et al. ¹¹ (1995)	166.5
Jain ¹² (1997)	166.2
Joselyn et al. ¹³ (1997)	160 ± 30
Kane ⁹ (1998)	152 ± 29
Kulkarni et al. ¹⁴ (1998)	144 ± 36
Li ¹⁵ (1997)	149.3 ± 19.9
Obridko ¹⁶ (1997)	203.2 ± 10.7
New Index ¹⁶	74.7 ± 6.9
Schatten et al. ¹⁷ (1998)	153 ± 30
Shastri ¹⁸ (1998)	≈152
Sofia et al. ¹⁹ (1998)	143 ± 30
Wilson et al. ²⁰ (1998a)	168 ± 15
Wilson et al. ²¹ (1998b)	144.0 ± 28.8
Thompson ²² (1999)	156 ± 29
Hathaway ²³ (1999)	140 ± 25
Verdes et al. ⁸ (2000)	115 ± 18
Verdes et al. ⁸ (2000)-ANN	135 ± 17
K-ADALINE	120 ± 32

Conclusiones

En este trabajo se presenta la extensión de algoritmos de naturaleza lineal al dominio no lineal mediante el uso de funciones potenciales o kernels, dando lugar al desarrollo de metodologías computacionalmente simples y numéricamente robustas. Se muestra que es posible aplicar este tipo de algoritmos en la predicción de series temporales usando para ello la extensión no lineal del Adaline de Widrow y Hoff. Este estudio cobra relevancia al analizar los resultados obtenidos en la predicción de la actividad solar, ya que los mismos son altamente competitivos en relación a otras metodologías ya establecidas en la literatura, tal como redes neuronales artificiales. En particular, la predicción del máximo valor del ciclo solar 23 parece mostrar un resultado acorde con el desarrollo actual del ciclo.

Finalmente, es necesario mencionar que es posible modelizar este fenómeno usando otro tipo de funciones potenciales y que dicho estudio será desarrollado en trabajos posteriores.

Referencias

1- Widrow, B. & Hoff, M. *Adaptive Switching Circuits*, 96-104 (1960).

2- Frieb T-T & Harrison, R.F., *Intelligent Data Analysis* 3, 307-313 (1999).

3- Aizerman, M., Braverman, E. & Rozonoer, L. *Automations and Remote Control* 25, 821-837 (1964).

4- Cortes, C., Vapnik V., *Machine Learning*, 20, 273-297 (1995).

5- Courant A., Hilbert D., *Methods of Mathematical Physics*, Wiley, New York, 1951.

6- Frieb, T-T, Harrison R.F., *Research Report No. 728*, The University of Sheffield, Sheffield, 1998.

7- McClelland J., Rumelhart D., *Explorations in Parallel Distributed Processing*, MIT Press, Cambridge, MA, 1988.

8- Verdes, P.F., Parodi, M.A., Granitto, P.M., Navone, H.D., Piacentini, R.D. & Ceccatto, H.A., *Solar Phys.* 191, 421 (2000).

9- Kane, R.P., *Geophys. Res. Lett.* 25, 3121 (1998).

10- Bouvar, K.H., Cliver, E.W., and Boriakoff, V., *Solar Phys.* 176, 211 (1997).

11- Calvo, R.A., Ceccatto, H.A., and Piacentini, R.D., *Astrophys. J.* 444, 916 (1995).

12- Jain, R., *Solar Phys.* 176, 431 (1997).

13- Joselyn, J. A. et al., *EOS Trans. American Geophys. Union* 78, 205 (1997).

14- Kulkarni, D.R., Pandya, A.S. and Parikh, J.C., *Geophys. Res. Lett.* 25, 457 (1998).

15- Li, Y., *Solar Phys.* 170, 437 (1997).

16- Obridko, V.N., *Solar Phys.* 156, 179 (1997).

17- Schatten, K.H. and Hoyt, D., *Geophys. Res. Lett.* 25, 599 (1998).

18- Shastri, S., *Solar Phys.* 180, 499 (1998).

19- Sofia, S., Fox, P., and Schatten, K., *Geophys. Res. Lett.* 25, 4149 (1998).

20- Wilson, R.M., Hathaway, D.H., and Reichmann, E.J., *J. Geophys. Res.* 103, 6595 (1998a).

21- Wilson, R.M., Hathaway, D.H., and Reichmann, E.J., *J. Geophys. Res.* 103, 17411 (1998b).

22- Radio & Space Services, Department of Industry Science and Resources, Australian Government (<http://www.ips.oz.au>).

23- National Aeronautics and Space Administration (<http://www.ssl.msfc.nasa.gov>).

CEILAP
CITEFA CONICET
ZUFRIATEQUI Y VARELA
1603
REPUBLICA ARGENTINA