
ENTRENAMIENTO MUESTRAL DE MODELOS DINÁMICOS CON SVM

SAMPLE TRAINING OF DYNAMIC MODELS WITH SVM

César Andrés Burgos Molina¹

Max Chacón Pacheco²

RESUMEN

Las máquinas de vectores soporte con recurrencia externa (Support Vector Regression, SVR) son utilizadas en el estudio de sistemas dinámicos. Un inconveniente que ellas presentan, es que requieren un gran número de iteraciones durante el entrenamiento consumiendo así una gran cantidad de tiempo y recursos computacionales. La disminución de los tiempos se logra aplicando la técnica muestral, consistente en la disminución de la cantidad de iteraciones al elegir (estratégicamente) una menor cantidad de valores, entre los posibles, de los parámetros participantes. Se logró disminuir este tiempo en 17 veces con respecto al entrenamiento clásico y a la mitad con respecto al que utiliza el algoritmo genético de Vasconcelos. La aplicación de la técnica permitirá reducir los tiempos de estudios y aplicar la SVR en áreas que necesiten resultados rápidos.

Palabras clave: Support Vector Machine, SVR, Técnica de reducción tiempo entrenamiento, Algoritmo Vasconcelos, Entrenamiento tradicional de SVR.

ABSTRACT

The support vector machine with external recurrence (Support Vector Regression, SVR) are used in the study of dynamical systems. One drawback they have is that they require a large number of iteration during training and consuming a large amount of time and computational resources. The time reduction is achieved by sampling the technique consisting in reducing the number of iterations to choose (conveniently) a smaller number of possible values of the parameters involved. It was possible to reduce this time to 17-fold relative to normal training and half with respect to using the Vasconcelos genetic algorithm. The application of the technique will reduce study time and apply the SVR in areas that need fast results.

Keywords: Support Vector Machine, SVR, Training time reduction technique, algorithm Vasconcelos, traditional training.

INTRODUCCIÓN

La Máquina de Vectores Soporte (SVM) es un sistema de aprendizaje supervisado basado en la teoría de aprendizaje estadístico introducida por Vapnik y colaboradores [1], donde su ventaja principal, por sobre otros sistemas es que logra encontrar un mínimo global para obtener los valores óptimos.

Dentro de la familia de las SVM se encuentra la máquina de vectores soporte de regresión (SVR, "Support Vector Regresión"), la cual conlleva la idea básica, presente en cualquier problema de regresión, de encontrar la función que mejor interpole un conjunto de datos dados. Para realizar el modelamiento de fenómenos dinámicos en el tiempo (caso estudiado en este artículo) estos se pueden representar por una SVR estática añadiéndole recurrencia externa, es decir, la SVR por si misma sólo es estática pero para poder tratar fenómenos dinámicos en el tiempo se le añade recurrencia externa.

La SVR con recurrencia externa utiliza de una serie de parámetros específicos para encontrar el mayor valor del coeficiente de correlación dentro de la etapa de validación, donde cada uno de estos parámetros posee a su vez cierto rango de valores posibles. La búsqueda de la combinación correcta de la tupla conformada, consiste en realizar una gran cantidad de iteraciones para revisar todas las combinaciones posibles (para llegar al óptimo), lo cual consume una gran cantidad de tiempo y recursos computacionales.

Variados han sido los intentos de reducir los tiempos del proceso de entrenamiento dentro del área de las Máquinas de vectores soporte. Por ejemplo han existido intentos donde el entrenamiento de máquinas de vectores soporte permiten discriminar puntos en un conjunto, usando el algoritmo de chunking suave el cual inicia con un subconjunto aleatorio de vectores para hallar los vectores soporte que entreguen una separación para dos subconjuntos del conjunto dado. Utiliza una fase de pre-procesamiento de los datos. Con el algoritmo de

¹ Departamento de Ingeniería Informática. Universidad de Santiago de Chile. Alameda 3363, Santiago, Chile. E-mail: cesar.burgos@usach.cl

² Departamento de Ingeniería Informática. Universidad de Santiago de Chile. Alameda 3363, Santiago, Chile. E-mail: max.chacon@usach.cl

GetBorder logra reducciones hasta del 30% pero el trabajo no indica con que calidad se generan dichos resultados [2]. Por otro lado se han realizado estudios que indican que a mayor cantidad de datos, mayor será el tiempo empleado para el entrenamiento. Los resultados de reducción del tiempo (50% a 90%) varían dependiendo del largo de los datos y se enfocan en la eliminación de los datos menos significativos. El estudio abarca los algoritmos SMO, LibSVM, FCM-SVM2, MEB-SVM2, RS-SVM [3]

Otro enfoque es utilizar algoritmos para la búsqueda de los parámetros necesarios y así acotar el rango de los valores posibles [4]. Así también, se tiene la utilización de algoritmos genéticos para encontrar la tupla de parámetros que arroje la mejor correlación [5].

El propósito de este artículo es la presentación de una técnica de reducción de tiempos de entrenamiento para máquinas de vectores soporte de regresión con recurrencia externa. A esta técnica se le ha dado el nombre de técnica muestral de entrenamiento.

Planteamiento Del Problema

La SVR [6] con recurrencia externa requiere determinar una serie de valores (elegidos por el investigador) de los parámetros involucrados para encontrar el óptimo, el cual corresponde al mayor valor de coeficiente de correlación obtenido en la etapa de validación. La búsqueda de la combinación correcta de los valores de los parámetros (la cantidad de parámetros es fija) consiste en realizar una serie de iteraciones del entrenamiento, con cada una de las tuplas posibles, lo cual consume una gran cantidad de tiempo y recursos computacionales.

Por ejemplo, en el caso de un entrenamiento univariado lineal, los parámetros del SVR son ν (nu) y C , más los 2 retardos (n) correspondientes a la variables de entrada más el retardo de la variable de salida. Por tanto, una tupla sería (ν, C, n_x, n_y) donde los rangos de número de valores podrían ser (1-9, -1-5, 1-8, 1-8) dando como resultado 4032 ($9 \times 7 \times 8 \times 8$) tuplas a ser evaluadas.

Por otro lado, para un caso univariado de entrenamiento no lineal, (el que es abordado en este estudio) los parámetros del SVR son ν (nu), σ (sigma) y C , más los 2 retardos correspondientes a una variable de entrada más el retardo de la variable de salida. Por tanto, una tupla sería (ν, σ, C, n_x, n_y) donde los rangos de número de valores podrían ser (1-9, 1-11, -1-5, 1-8, 1-8) dando como resultado 44352 ($9 \times 11 \times 7 \times 8 \times 8$) tuplas a ser evaluadas. El tiempo promedio consumido por el entrenamiento es de

alrededor de 24 horas para un solo sujeto en un estado específico.

El reducir los tiempos del proceso de entrenamiento manteniendo la calidad de los resultados, es una característica deseable para facilitar la utilización de SVR por parte de investigadores en el ámbito de modelos complejos (señales dinámicas no lineales), puesto que reduce el tiempo utilizado en investigaciones y permite aplicar esta herramienta en áreas que requieran resultados en tiempos menores que los actuales.

ANÁLISIS TEÓRICO

SVR

El algoritmo SVM utilizado fue el ν -SVM, introducido por Schölkopf [7] a finales de 1990. Este algoritmo se basa en la teoría estadística de aprendizaje el cual introdujo la regresión como el ajuste de los datos a un tubo de ϵ radio [8]. La frontera para determinar el radio del tubo está dado por un pequeño subconjunto de ejemplos de entrenamiento llamados vectores de soporte (SV) como se ilustra en la figura 1.

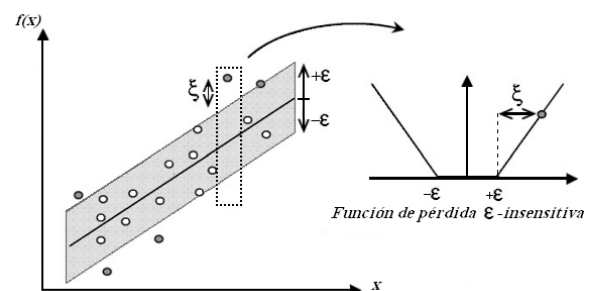


Figura 1. Interpretación geométrica de la regresión SV en el que la ecuación de regresión se estima por medio de un tubo de ϵ radio. El compromiso entre la complejidad del modelo y los puntos de la izquierda fuera del tubo de regresión se controla por las variables de holgura. Los puntos oscuros corresponden a los vectores de soporte.

Asumiendo que \vec{x} ($\vec{x} \in \mathbb{R}^N$) representa el vector de los datos de entrada, el valor de salida $f(\vec{x})$ está dado por la SVM de regresión usando un vector de peso \vec{w}

$$f(\vec{x}) = (\vec{w} \cdot \vec{x}) + b \quad (1)$$

donde b ($b \in \mathbb{R}$) es una constante obtenida de \vec{w} .

Una característica clave de la SVM es la determinación de la distancia de los datos a la línea de regresión

utilizando funciones de pérdida como la ε -insensitiva ilustrada en la figura. 1.

$$|f(\vec{x}) - y|_{\varepsilon} = \begin{cases} 0 & \text{si } |f(\vec{x}) - y| < \varepsilon \\ |f(\vec{x}) - y| - \varepsilon & \text{otro caso} \end{cases} \quad (2)$$

Donde $\varepsilon > 0$ corresponde a un valor elegido a priori.

El algoritmo se ejecuta minimizando el riesgo funcional $\|\vec{w}\|^2$, añadiendo penalización dada por variables de holgura ξ (figura 1) dependiendo de la distancia de los puntos de datos a la línea de regresión. C es un parámetro del modelo que determina el compromiso entre la complejidad del modelo expresada por \vec{w} y los puntos que se mantienen fuera del tubo.

$$\text{Minimizar } \theta(\vec{w}, \xi) = \frac{1}{2} \|\vec{w}\|^2 + C \sum_{i=1}^l \xi_i \quad (3)$$

La variación del v-SVM introducido por Schölkopf [7], consiste en agregar ε en la minimización del problema, ponderado por una variable v que ajusta la contribución de ε entre 0 y 1:

$$\text{Minimizar } \theta(\vec{w}, \xi) = \frac{1}{2} \|\vec{w}\|^2 + C \left(l v \varepsilon + \sum_{i=1}^l \xi_i \right) \quad (4)$$

En la ecuación anterior, l representa la dimensión total de los datos (número de casos). La solución de este problema de minimización para la obtención de los vectores de ponderación \vec{w} es encontrado por el procedimiento de optimización estándar para un problema con las restricciones de desigualdad al aplicar las condiciones de Kuhn-Tucker del problema dual. La principal ventaja de la introducción del parámetro $v \in [0-1]$ es hacer posible el control de la fracción del error y el número (o fracción) de SV con sólo un parámetro normalizado.

Para resolver un problema de regresión no lineal es suficiente sustituir el producto interno entre dos variables originales independientes $\vec{x}_i \cdot \vec{x}_j$ (ec. (1)) por una función kernel $k(\phi(\vec{x}_i) \cdot \phi(\vec{x}_j))$. Varias funciones se pueden utilizar como kernel, tales como la función de base radial gaussiano (RBF)

$$k(\vec{x}_i, \vec{x}_j) = \exp \left(-\frac{\|\vec{x}_i - \vec{x}_j\|^2}{2\sigma^2} \right) \quad (5)$$

o la función polinómica $k(\vec{x}_i, \vec{x}_j) = (\vec{x}_i \cdot \vec{x}_j)^p$.

Muestreo aleatorio simple

El muestro aleatorio simple consiste en que cada individuo tiene la misma probabilidad de ser extraído dentro de un conjunto [9].

El tamaño de muestra en el muestreo aleatorio simple se calcula de la siguiente manera [10]:

$$n = \frac{n_0}{1 + \frac{n_0}{N}} \quad (6)$$

$$n_0 = \frac{Z^2 \alpha \sigma^2}{E^2} \quad (7)$$

Donde, n es el tamaño de la muestra, n_0 es el tamaño de muestra aproximado, N es el tamaño de la población bajo estudio, $Z\alpha$ es el valores de nivel de significancia, E es el error de tolerancia de la estimación, $1-\alpha$ es el nivel de significancia y σ^2 es el varianza de la Variable.

La varianza es igual a PQ (el valor de σ^2 de la ecuación 7 es reemplaza por PQ produciendo la ecuación 8), donde P denota la proporción estimada o esperada de la variable. Si no se conoce tal valor se reemplaza por 0.5, es decir, $P=1/2$ y $Q=1-P$.

$$n_0 = \frac{Z^2 P Q}{E^2} \quad (8)$$

El error muestral o de estimación es el generado a causa de observar una muestra en lugar de la población completa. Este valor se encontrará generalmente entre 0% y 10%.

El nivel de confianza en la probabilidad a priori de que el intervalo de confianza contenga el valor del parámetro. Se indica por $1-\alpha$ y sus valores usuales son el 95%, 99% y 99,9%. En este estudio se utiliza α del 5%.

En el caso de este estudio, para modelamiento no lineal se tiene que la cantidad de iteraciones totales son las siguientes 44352.

Recordando, en el caso No Lineal una tupla contendrá los siguientes parámetros (v , σ , C , np , nv) donde los rangos de número de valores son 1-9, 1-11, 1-5, 1-8 y 1-8 respectivamente dando como resultado 44352 es decir $(9 \times 11 \times 7 \times 8 \times 8)$ tuplas a ser evaluadas.

El cálculo es el siguiente con el nivel de significancia del 95%.

$N = 44352$

$\alpha = 5\%$

$Z\alpha = 1.96$ (Tabla Normal)

$E = 5\%$.

$\sigma^2 = PQ$

$P = Q = 0.5$

Resultados:

$n_0 = 348.16$

$n = 345,45$

En este caso de entrenamiento no lineal este número debe ser de al menos 345. Estas cantidades de muestras (tuplas) mínimas requeridas para una significancia del 95% en la técnica muestral de entrenamiento se satisface, puesto que en el caso no lineal se utilizan 1024. Los cálculos utilizados para obtener estos últimos valores se desarrolla más adelante.

Equidistancia entre valores de parámetros

La elección de los valores a utilizar según parámetros participantes, se realizó según las siguientes condiciones:

- Determinación de la cantidad de valores de parámetros a utilizar. Si por ejemplo el parámetro contempla 8 valores sucesivos, se necesita determinar cuántos de estos se utilizarán, es decir, 2, 3 o 4. Este número lo determina el investigador recordando el número de muestras mínimas necesarias para la toma de muestras y que permita un índice de confianza aceptable.
- Las cotas de los rangos deben estar presentes: Es decir, si un rango fuese desde el 1 al 8, el 1 y el 8 deben estar presentes dentro de los valores que serán elegidos.
- Valores intermedios: Los valores intermedios deben tener pasos semejantes entre sus intervalos. Siguiendo con el ejemplo anterior, si se desean 4 valores para el rango entre el 1 y el 8, podrían elegirse el 1, 3, 6, 8. En este caso

los pasos fueron de 1 con excepción del valor entre el 3 y el 6 que fueron 2 unidades.

$$\text{Paso} = \frac{\text{Cantidad total de valores}}{\text{Cantidad de valores a utilizar}} \quad (9)$$

donde:

- Cantidad total de valores: Es la cantidad total de valores distintos y sucesivos dentro del rango (número discretos). Si el rango es del 1 al 8, serán 8 valores distintos.
- Cantidad de valores a utilizar: Es la cantidad de valores que se utilizan realmente dentro del rango. Si el rango es del 1 al 8, el investigador podría elegir utilizar solo 4 valores.
- Paso: Distancia aproximada entre valores sucesivos.

Ausencia de sesgo

Los cálculos para la determinación del número de muestras necesarias al utilizar el muestreo aleatorio simple son factibles si cada uno de los individuos que se elijan desde una población determinada posea la misma probabilidad de ser elegidos.

En el presente estudio, los individuos se identifican por el coeficiente de correlación obtenido en la etapa de validación dentro del entrenamiento de una máquina de vectores soporte, el cual representa un modelo específico.

Antes del entrenamiento y validación de la SVR se desconoce por completo que coeficiente de correlación obtendrá cierto modelo al entrenar con cierta tupla. A cada modelo le corresponde solo una tupla (compuesta por los valores de los parámetros e hiperparámetros) los cuales son conocidos porque serán asignados para realizar el entrenamiento, pero que no poseen importancia por ellos mismos hasta después del entrenamiento.

En la figura 2, se presenta un gráfico que muestra la relación entre el valor del coeficiente de correlación y la correspondiente iteración dentro de la validación del entrenamiento que a su vez corresponde a cierta tupla (o combinación de parámetros única). Se aprecia la no existencia de relación entre el valor del coeficiente de correlación y las iteraciones (tuplas) consecutivas.

Adicionalmente se calculó la regresión lineal múltiple (con respecto al coeficiente de correlación), donde el coeficiente de determinación R^2 entregó un valor de 0,1965 lo que indica que las variables tienen muy poca relación entre sí.

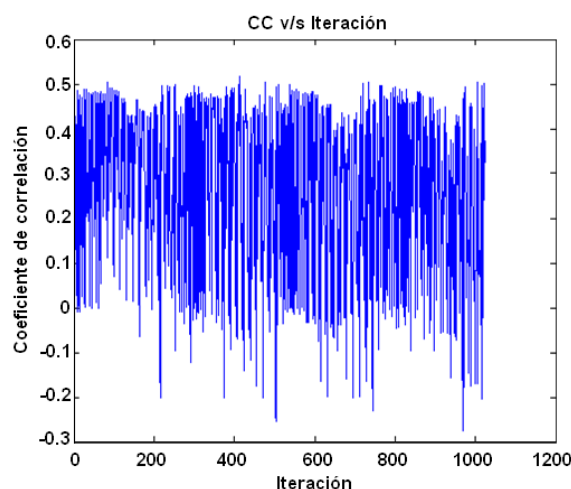


Figura 2. Coeficiente de correlación versus número de iteración dentro de la etapa de validación.

MATERIALES Y MÉTODOS

Las señales utilizadas para demostrar la reducción del tiempo corresponden a las de 31 pacientes con traumatismo encefalo craneano severo, es decir, con una medición en la escala de coma de Glasgow (GCS: Glasgow Come Scale) ≤ 8 , los cuales fueron admitidos en el Adult Intensive Care Unit, Queen's Medical Centre, Nottingham, Inglaterra.

La velocidad de flujo sanguíneo cerebral, (VFSC) fue registrada en la arteria cerebral media utilizando un sistema Doppler Transcraneal Scimed con un transductor de 2 MHz. La presión arterial media (PAM) fue medida en mmHg con un catéter intravascular insertado en la arteria radial (Baxter PX-600F). Por último, la presión intracraneal (PIC) se midió con un transductor subdural (Codman MicroSensor).

Se usó la transformada rápida de Fourier para extraer la máxima frecuencia en la señal de VFSC, con una resolución temporal de 5 ms. Las señales de PAM y PIC fueron digitalizadas y muestreadas a 200 muestras/s. El comienzo y final de cada ciclo cardiaco fue detectado en la señal de PAM, y luego se calcularon los valores medios de PAM, PIC y de VFSC para cada uno de los ciclos. Finalmente se les aplicó interpolación polinomial, generando nuevas señales continuas muestreadas cada 0,2 segundos.

Como plataforma de trabajo se utilizó Matlab 6.5.2 y la librería Spider 1.71.

El objetivo principal de utilizar las señales descritas fue el hecho de que el investigador que trabajó con ellas [11] registró los tiempos de entrenamiento correspondientes a cada sujeto y de esta manera se pueden comparar con los tiempos obtenidos con la presente técnica.

Modelamiento utilizado

La estructura de modelamiento utilizado corresponde al tipo univariado no lineal teniendo como entrada la señal de PIC, mientras que la salida correspondió a la señal VSFC. Se realizó el entrenamiento solo con kernel no lineal y bajo solo un escenario. El kernel no lineal en este caso corresponde al kernel de base radial puesto que en los trabajos anteriores [12] y [13] este se utilizó y se usará el mismo para realizar la evaluación en idénticas condiciones.

En el entrenamiento de las SVM se utilizó la estrategia llamada validación cruzada balanceada, la cual proporciona confiabilidad a los modelos respecto de su comportamiento con señales desconocidas puesto que a partir de dos conjuntos de datos, uno de ellos es empleado para entrenar el modelo y el otro para probar su eficiencia, para luego realizar el mismo procedimiento intercambiando ambos conjuntos. Se utiliza esta estrategia puesto que al usar el parámetro C (en el entrenamiento y validación) se realiza un proceso de regularización y por tanto solo se requiere un conjunto de entrenamiento y uno de prueba.

Selección de valores de parámetros

Como se explicó dentro del planteamiento del problema, para el entrenamiento clásico no lineal univariado la cantidad de tuplas (combinación de parámetros) utilizadas para ser evaluadas fue de 44352 ($9 \times 11 \times 7 \times 8 \times 8$). Esto significa lo siguiente:

- ν (Nu): Toma valores desde el 0.1 al 0.9 en pasos de 0.1. Corresponde a 9 valores distintos.
- σ (Sigma): Toma valores desde el 2 al 12 en pasos de 1. Corresponde a 11 valores distintos.
- C: Toma valores desde el -1 al 5 en pasos de 1. Corresponde a 7 valores distintos.
- Retardo Entrada (np): Retardo PIC: Toma valores desde el 1 al 8 en pasos de 1. Corresponde a 8 valores distintos.
- Retardo Salida: Retardo VFSC (nv): Toma valores desde el 1 al 8 en pasos de 1. Corresponde a 8 valores distintos.

Luego de la elección equidistante de los valores de los parámetros, el entrenamiento no lineal univariado requirió de una cantidad de 1024(4x4x4x4) tuplas, lo que significa lo siguiente:

- ν (Nu): Toma valores desde el 0.1 al 0.9 en pasos de 0.2. Tomó por tanto 4 valores, a saber, 0.1, 0.3, 0.6, 0.9.
- σ (Sigma): Toma valores desde el 2 al 12 en pasos de 4. Corresponde a 4 valores distintos, a saber, 2, 6, 10, inf.
- C: Toma valores desde el -1 al 5 en pasos de 2. Corresponde a 4 valores distintos, a saber, -1, 1, 3, 5.
- Retardo Entrada: Retardo PIC (np): Toma valores desde el 1 al 8 en pasos de 2. Corresponde a 4 valores distintos, a saber, 1, 3, 5, 8.
- Retardo Salida: Retardo VFSC (nv): Toma valores desde el 1 al 8 en pasos de 2. Corresponde a 4 valores distintos, a saber, 1, 3, 5, 8.

Con la selección espaciada de los valores de los parámetros se reduce la cantidad de iteraciones desde 44352 a 1024 lo que reduce el tiempo total de entrenamiento.

Comparación estadística

Para realizar la comparación estadística necesaria en este estudio, se aplica el contraste de adherencia de normalidad el cual permite determinar si la secuencia en estudio está normalmente distribuida (Shapiro-Wilk) y luego se aplicaron dósimas de hipótesis (Test Wilcoxon (signed Rank)), donde estas, permiten conocer la similitud o diferencia estadística entre dos modelos diferentes, a través de valores obtenidos mediante una determinada métrica. En ambos test se sostiene que existen diferencias estadísticas significativas si el valor-p entregado es menor a 0,05 ($p < 0,05$).

RESULTADOS Y ANÁLISIS

Para determinar la calidad de la Técnica de muestreo, se utilizaron los coeficientes de correlación (CC) obtenidos de los entrenamientos de cada uno de los 31 modelos cuyos datos pertenecen a los 31 pacientes (un modelo por paciente) con traumatismo encéfalo craneano. Estos CC se comparan con los obtenidos por Ruz [11] y con los resultados de la utilización del algoritmo genético de Vasconcelos [14]. La estructura del modelo utilizado en estos entrenamientos corresponde al tipo univariado no

lineal donde la señal de entrada corresponde a la PIC y la señal de respuesta corresponde a VFSC. Además por tratarse de un modelo auto regresivo, se incorpora al menos una retroalimentación de VFSC en cada sujeto.

Para determinar la eficiencia en la disminución en los tiempos de entrenamiento se registró el tiempo de cada uno de los entrenamientos de los 31 sujetos. Estos tiempos se comparan con los tiempos registrados por Ruz [11] y con los tiempos registrados utilizando Vasconcelos.

En la tabla 1, se presenta la calidad (por medio del CC) de las 3 técnicas de entrenamiento comparadas y también se presentan los tiempos promedios requeridos según técnica. Se utilizó el test de Wilcoxon para determinar diferencias significativas ($p < 0,05$) entre los coeficientes de correlación de las etapas de validación entre las técnicas de entrenamiento clásico y muestral como también entre Vasconcelos y entrenamiento muestral.

Se determinó que existe diferencia significativa con respecto al coeficiente de correlación de la etapa de validación entre la técnica de entrenamiento clásico (barrido completo) y muestral, mientras que entre las técnicas Vasconcelos y muestral no existe esta diferencia significativa por lo cual serían equivalentes.

En este estudio se logra que la técnica muestral obtenga mejores resultados con respecto a los coeficientes de correlación debido a que la técnica de entrenamiento clásico utilizado en el trabajo de Ruz [11] restringió el rango de valores de uno de sus parámetros, en particular, se restringió el número de retardos de la VFSC tomando valores desde el 1 al 4, en vez de que fuesen desde el 1 al 8.

Técnica	Entrenamiento CC	Validación CC	Tiempo utilizado
Entrenamiento clásico	0,9348 $\pm 0,092$	0,6789 $\pm 0,164$ †*	23:59:07 $\pm 8:05:59$
Vasconcelos	0,9406 $\pm 0,090$	0,7121 $\pm 0,137$ §*	2:56:27 $\pm 1:23:15$
Muestral	0,9417 $\pm 0,094$	0,7141 $\pm 0,131$ *	1:20:54 $\pm 0:27:08$

†* $p=0.0034$

§* $p=0.6311$

Tabla 1. Resultados promedio de comparación de técnicas

Con respecto del tiempo ocupado entre las técnicas, la técnica muestral es notoriamente inferior a la técnica de entrenamiento clásico (17,8 veces). En cambio, entre las técnicas de Vasconcelos y entrenamiento muestral, el tiempo que ocupa la técnica de Vasconcelos es aproximadamente el doble (2,2 veces).

CONCLUSIONES

El estudio demuestra que reduciendo la cantidad de valores de los parámetros participantes, es posible reducir los tiempos de entrenamiento y obtener resultados significativamente semejantes.

Su relevancia y ventaja se aprecia en que con respecto al entrenamiento clásico se reduce el tiempo de entrenamiento 17,8 veces en promedio. La importancia se ilustra en el siguiente ejemplo: En cierto estudio del sistema autorregulatorio cerebral, se utilizaron datos de 16 sujetos en dos estados de respiración distintos. En el fondo se realizaron 32 entrenamientos los cuales requirieron de 768 horas al utilizar el entrenamiento clásico contra las 42 horas necesarias en el entrenamiento muestral.

Su limitación es que no se realizan todas las combinaciones posibles para todo el rango de valores de los parámetros posibles. Sin embargo esta limitación será relevante solo en casos determinados donde se requiera lograr los máximos absolutos en términos de coeficiente de correlación de la etapa de validación.

Estos resultados se aplican tanto a estudios donde es requerido hacer gran cantidad de entrenamientos, como también para la realización de una exploración preliminar de rangos de valores de los parámetros participantes.

Como trabajo futuro se procederá, luego de aplicar la técnica de entrenamiento muestral, a buscar tempranamente patrones dentro de los parámetros e hiperparámetros participantes con el fin de disminuir aún mas estos tiempos de entrenamiento.

AGRADECIMIENTOS

Se agradece a R.B. Parenai del Departamento de Ciencias Cardiovasculares de la Universidad de Leicester en Inglaterra, por facilitar los datos utilizados en este trabajo.

NOTAS

Los datos utilizados en el presente artículo, han sido usados en el estudio del problema de autorregulación cerebral dentro de otros artículos como [12] y [13]. En esos casos los tiempos de entrenamiento son excesivos, pues se evalúan los parámetros en todos su rangos.

REFERENCIAS

- [1] C. Nello. "An Introduction to Support Vector Machines and Other Kernel-based Learning Methods". Cambridge University Press . 2000.
- [2] J. Flores. "Entrenamiento de maquinas de soporte vectorial". CONAIS. Universidad Autónoma del Carmen CA Tecnologías de la Información, CA Ciencias de la Computación. Campeche, México. 2007.
- [3] Cervantes J. "Clasificación de grandes conjuntos de datos vía Máquinas de Vectores Soporte y aplicaciones en sistemas biológicos". Departamento de computación, Centro de investigación y de estudios avanzados del Instituto Politécnico Nacional, México D.F. , México. 2009.
- [4] D. García. "Boosting Support Vector Machines". Revista de Ingeniería, Número 24, pp. 62-70. Universidad de los Andes, Colombia. Noviembre, 2006.
- [5] C. Sanhueza. "Evaluación de las redes neuronales recurrentes y maquinas de vectores soporte para la estimación no invasiva de la presión intracraneal". Departamento de Ingeniería Informática, Universidad de Santiago de Chile. Santiago, Chile. 2011.
- [6] A. Smola, B. Scholkopf. "A tutorial on support vector regression". Statistics and Computing 14: 199–222, 2004.
- [7] B. Schölkopf, A. Smola, RC. Williamson, PL Bartlett. "New support vector algorithms". Neural Computing, 12:1083–121. 1998.
- [8] Vapnik V, Golowich S, Smola A. Support vector method for function approximation. Regression estimation, and signal processing. In: Mozer M, Jordan M Petsche TT, editors. Neural

-
- information processing systems, 9. Cambridge, MA: MIT Press; 1997. p. 281–7. 1997.
- [9] D. Raj. “Teoría del Muestreo”. Fondo de cultura económica. ISBN 9681603559, 9789681603557. 1980.
- [10] M. Spiegel. “Estadística”. Serie Schaum. Editorial McGraw-Hill. 1978.
- [11] J. Ruz. “Modelamiento no lineal del sistema de autorregulación cerebral y su asociación con la predicción de mortalidad en pacientes con traumatismo encéfalo craneano”. Tesis de Magister en Ingeniería Informática, Universidad de Santiago de Chile. Santiago, Chile. 2009.
- [12] M. Chacón, C. Araya, M. Muñoz, R. Panerai. “Comparison between SVM and ANN for Modeling the Cerebral Autoregulation Blood Flow System”. IJCCI. 522-525. 2009.
- [13] M. Chacón, C. Araya, RB. Panerai. “Non-linear multivariate modeling of cerebral hemodynamics with autoregressive Support Vector Machines”, Medical Engineering & Physics. 33:2, pp 180-187, 2011.
- [14] Kuri-Morales A. "Pattern Recognition via Vasconcelos Genetic Algorithm". Progress in Pattern Recognition, Image Analysis and Applications. Alberto Sanfeliu, José Francisco Martínez Trinidad, y Jesús Ariel Carrasco Ochoa, editors. Springer Berlin Heidelberg. 3287:328-335. 2004.