

**Fornari, Javier ; Luccini, Eduardo ; Grieco, Sebastián ;  
Parodi, Miguel ; Vidali, Esteban**

*Identificación de variables en modelización de  
líneas de producción industrial y su análisis  
mediante redes neuronales artificiales*

Energeia, Año 9, N° 9, 2011

Este documento está disponible en la Biblioteca Digital de la Universidad Católica Argentina, repositorio institucional desarrollado por la Biblioteca Central "San Benito Abad". Su objetivo es difundir y preservar la producción intelectual de la institución.

La Biblioteca posee la autorización del autor para su divulgación en línea.

Cómo citar el documento:

Fornari, J., Luccini, E., Grieco, S. et al. Identificación de variables en modelización de líneas de producción industrial y su análisis mediante redes neuronales artificiales [en línea]. *Energeia*, 9(9), 2011. Disponible en:  
<http://bibliotecadigital.uca.edu.ar/repositorio/revistas/identificacion-variables-modelizacion-lineas-produccion.pdf>

(Se recomienda indicar fecha de consulta al final de la cita. Ej: [Fecha de consulta: 19 de agosto de 2010]).

## Identificación de Variables en Modelización de Líneas de Producción Industrial y su Análisis Mediante Redes Neuronales Artificiales

Javier Fornari<sup>1</sup>, Eduardo Luccini<sup>1,2</sup>, Sebastián Grieco<sup>1</sup>, Miguel Parodi<sup>1,2</sup> y Esteban Vidali<sup>1</sup>

<sup>1</sup> Facultad de Química e Ingeniería "Fray Roger Bacon", Pontificia Universidad Católica Argentina, Rosario, Argentina.

<sup>2</sup> Grupo de Energía Solar, Instituto de Física de Rosario (CONICET-UNRosario), Rosario, Argentina.

**Palabras Clave:** Procesos Industriales, optimización, Redes Neuronales Artificiales.

**Resumen.** En el marco de la integración de etapas en procesos de producción industrial en serie, se identifican, clasifican y caracterizan las variables que permiten modelizar dicho proceso, como estrategia con vistas a su optimización. El estudio se aplica en el ámbito de la industria metalúrgica, a partir de los datos de un conjunto de empresas de la región centro y sur de Santa Fe. La estructura secuencial de un proceso industrial hace que cualquier falla en una etapa cause demoras o reducción de calidad en el producto final obtenido. Como herramienta integradora de análisis, se plantea la implementación de un sistema de redes neuronales artificiales.

### 1 INTRODUCCIÓN

El concepto de proceso proviene del latín *processus*, y se define como "el conjunto de las fases sucesivas de un fenómeno natural o de una operación artificial". Aplicado a la industria, un proceso industrial es "el conjunto de operaciones necesarias para modificar las características de las materias primas hasta la obtención de un determinado producto", y requiere en general múltiples operaciones individuales.

La modelización de un sistema de producción con fines de mejorar el proceso se enmarca dentro de la Teoría de Optimización (e.g. Forst y Hoffmann, 2010). Es una disciplina que ha ganado enorme desarrollo en las últimas décadas, aplicada a la planificación y el estudio del desenvolvimiento de líneas de producción industrial (e.g. Rao, 2009). Entre otros aspectos, la optimización implica el estudio de eventos críticos, o interrupciones en el sistema de producción por diversas causas, y sus consecuencias. Para ello se han desarrollado diversos modelos que simulan las etapas de la línea de producción y los efectos de un evento crítico (Fornari et al., 2010). Más generalmente, la simulación numérica permite integrar las distintas etapas de un sistema de producción a fin de optimizar su funcionamiento. Para lograrlo, se deben determinar las variables que rigen el comportamiento del sistema y las leyes que las gobiernan. Los algoritmos y las técnicas de análisis disponibles son muy variados, y su elección depende fuertemente de las características del problema a resolver (e.g. Rao, 2009).

Uno de los algoritmos de análisis que ha cobrado más desarrollo y número de aplicaciones en las últimas décadas es el de redes neuronales artificiales (RNA). Una RNA es una arquitectura de procesamiento en paralelo distribuido, integrada por unidades de procesamiento masivamente interconectadas, también denominadas neuronas o nodos. Existe una amplia variedad de tipos de RNA, entre ellas (en inglés sólo a modo de mención) Probabilistic Neural Networks, General Regression Neural Networks, Radial Basis Function Networks, Cascade Correlation, Functional Link Networks, Kohonen networks, Gram-Charlier networks, Learning Vector Quantization, Hebb networks, Adaline networks, Heteroassociative networks, Recurrent Networks, Hybrid Networks. En particular, uno de los tipos de RNA más empleados y que utilizaremos en este trabajo es el Perceptrón Multicapa (Multilayer Perceptron Network). La RNA se compone de neuronas organizadas en capas. Usualmente existen tres capas: de entrada, oculta y de salida. El número de neuronas se determina según el problema a resolver y la información fluye sólo en una dirección, desde la entrada hacia la salida, en una variedad restrictiva denominada *feed-forward*. Las RNA han demostrado su potencial en control, optimización y detección de errores en procesos industriales (Zarenezhad y Aminian, 2011; Govindhasamy et al., 2005; Alves and Nascimento, 2007).

En este trabajo analizaremos en detalle el proceso industrial correspondiente a una empresa metalúrgica de la región de Rafaela, provincia de Santa Fe, identificando cada una de las variables que intervienen y permiten simular numéricamente el proceso de modo integral, enfatizando el abordaje del análisis mediante RNA.

### 2 VARIABLES Y DATOS DE PROCESOS INDUSTRIALES

Los valores de cada una de las variables en estudio provienen de 38 empresas del rubro metalúrgico de la zona de Rafaela, provincia de Santa Fe, y comprenden dos meses del año 2011. En este estudio inicial, reflejaremos la información provista por una de estas empresas, PyME autopartista perteneciente al sector industrial metalmeccánico, especializada en diseño, fabricación, comercialización y servicio de bombas de vacío para los mercados de

terminales y de reposición local e internacional. Las bombas de vacío son utilizadas en automóviles y vehículos comerciales livianos, equipados con motores diesel y sistemas de freno servo-asistidos. Un dato importante a tener en cuenta es el bajo recambio de estas piezas. La empresa tiene una organización de ingeniería altamente especializada, cuenta con máquinas de Control Numérico por Computadora (CNC), una sala de montaje presurizada con bancos de prueba informatizados para una completa trazabilidad del producto, un laboratorio de ensayos de validación y realiza innovaciones tecnológicas algunas de las cuales fueron patentadas. La figura 1 muestra un esquema de las etapas en el proceso de producción en serie de las bombas de vacío, y la figura 2 un diagrama más detallado con las principales operaciones del proceso.

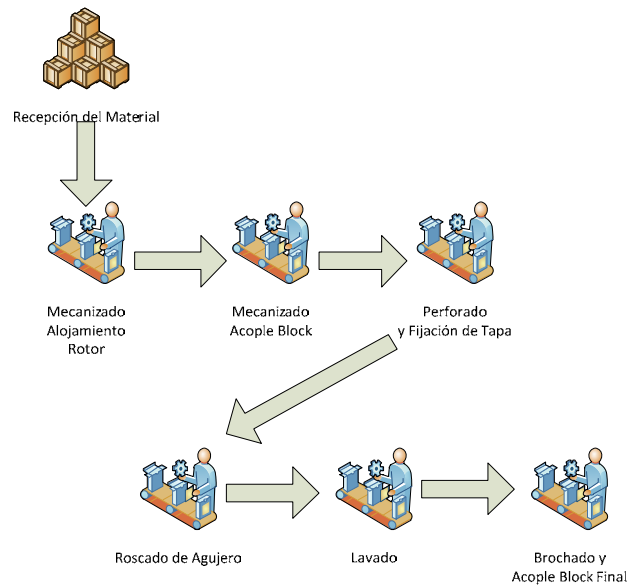


Figura 1. Etapas del proceso de producción de la bomba de vacío bajo estudio.

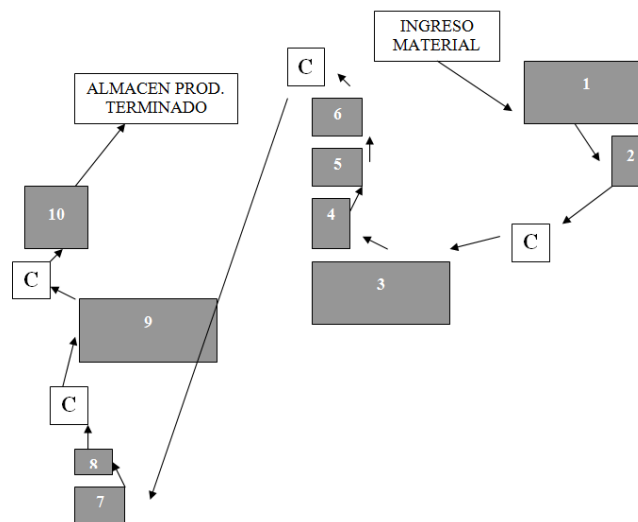


Figura 2. Diagrama de las principales operaciones en el proceso de producción de bomba de vacío bajo estudio.

**Referencias:**

1. Torno 1 (Maq1)
2. Dispositivos de Control 1 (DC1)
3. Torno 2 (Maq2)
4. Dispositivos de Control 2 (DC2)

5. Perforadora (Maq3)
6. Avellanadora (Maq4)
7. Roscadora (Maq5)
8. Dispositivos de Control 3 (DC3)
9. Lavadora (Lav)
10. Prensa (Maq6)

### C. Capachos

La tabla 1 muestra la planilla de control de cada etapa del proceso. Esta planilla es muy importante, ya que permite dictaminar la marcha del proceso y relacionarla con los tiempos de estadía de cada pieza en cada operación, relación que es crucial a los fines de modelizar fielmente los márgenes de los valores de cada variable dentro de los cuales el proceso obtiene resultados positivos.

Tabla 1. Cursograma del proceso de producción de bomba de vacío bajo estudio.

PROCESO	TRANSPORTE	CONTROL	DEMORA	ALMACENAJE	Pieza N°:	3323	Responsable:	Xxxxxxxx
					Denominación:	Cuerpo B.V. VW	Realizó:	Xxxxxxxx
					Proveedores:	XXXXXX	Fecha:	XX/XX/XXXX
					DESCRIPCIÓN DE LA ACTIVIDAD	MÉTODO DE CONTROL		
○	◡	■	D	△	Recepción del material	S/ Hoja de Instrucción		
○	◡	□	●	△	A la espera de mecanizado			
●	◡	□	D	△	Mecanizado alojamiento rotor	S/ Hoja de Proceso		
○	◡	□	●	△	A la espera de mecanizado			
●	◡	□	D	△	Mecanizado acople block	S/ Hoja de Proceso		
●	◡	□	D	△	Perforado agujero fijación tapa	S/ Hoja de Proceso		
●	◡	□	D	△	Avellanado agujero fijación tapa	S/ Hoja de Proceso		
○	◡	□	●	△	A la espera de roscado			
●	◡	□	D	△	Roscado agujero fijación tapa	S/ Hoja de Proceso		
○	◡	□	●	△	A la espera de lavado			
●	◡	□	D	△	Lavado	S/ Hoja de Proceso		
○	◡	□	●	△	A la espera de brochado			
●	◡	□	D	△	Brochado acople block	S/ Hoja de Proceso		
○	◡	■	D	△	Control del proceso	S/ Hoja de Instrucción		
○	◡	□	D	▲	A la espera de armado			

La tabla 2 detalla concretamente todas las operaciones correspondientes al proceso de producción de bomba de vacío. Cuando se trata de una operación de traslado o almacenamiento se detallan las características correspondientes, mientras que en operaciones de maquinado se detalla el tiempo estimado para cada pieza. Cuando se trata de simular numéricamente el proceso, precisamente, cada operación es una variable. Los datos recabados durante cada operación constituyen una serie temporal de cada variable, que puede ser estudiada mediante diferentes métodos, en este trabajo se lo plantea a través de las RNA. La disposición secuencial e interrelacionada de las variables en el proceso ocasiona que todas tengan un rango limitado de flexibilidad a fines de que el proceso no se interrumpa o detenga.

Tabla 2. Operaciones en el proceso de producción de bomba de vacío bajo estudio.

Número de operación	Tipo de operación	Características	Observaciones
1	Llegada del material	Llegan recipientes con 410 productos a granel	
2	Torno 1	Tiempo de ciclo: 130 seg.	
3	Dispositivo de control 1	Tiempo de ciclo: 40 seg.	
4	Colocación en capacho	Capacidad: 33 piezas	
5	Traslado del capacho al Torno 2		
6	Torno 2	Tiempo de ciclo: 160 seg.	c/ 170 piezas puesta a punto de herramienta 300 seg.
7	Dispositivo de Control 2	Tiempo de ciclo: 80 seg.	c/ 5 piezas se lleva una a la oficina de calidad para controlar rugosidad, cuyo tiempo extra es de 34 seg.
8	Perforadora	Tiempo de ciclo: 20 seg.	
9	Avellanadora	Tiempo de ciclo: 5 seg.	
10	Colocación en capacho	Capacidad: 33 piezas	
11	Traslado del capacho a Roscado		
12	Roscado	Tiempo de ciclo: 35 seg.	
13	Dispositivo de Control 3	Tiempo de ciclo: 10 seg.	
14	Colocación en capacho	Capacidad: 33 piezas	
15	Traslado del capacho a Lavadora		
16	Lavadora	Tiempo de lavado: 600 seg.	
17	Colocación en capacho	Capacidad: 33 piezas	
18	Traslado del capacho a Prensa		

19	Prensa (Brochado)	Tiempo de ciclo: 10 seg.	
20	Colocación en capacho	Capacidad: 33 piezas	
21	Traslado del capacho a la Zona de Almacenamiento		
22	Almacenamiento final		

Como se mencionó, la acumulación de datos de una determinada variable en el proceso establece una serie temporal de datos de dicha variable, apta para ser analizada. En este caso, del total de 22 operaciones o variables se grafican en las figuras 3 y 4 respectivamente, a modo de ejemplo, la evolución en el tiempo de la llegada de la materia prima y el tiempo de actividad del torno 1, recabados durante 2 meses del año 2011. En el eje de las abscisas se representan 1000 iteraciones o actividades realizadas, mientras que en el eje de las ordenadas se representa el tiempo empleado, en segundos. Se observa que la serie temporal es marcadamente "ruidosa" en términos de fluctuaciones estadísticas. Sin embargo es posible, aun a simple vista, observar periodicidades y modulaciones suaves que son precisamente aquellas útiles desde el punto de vista del proceso y que pueden ser detectadas por diferentes algoritmos de análisis.

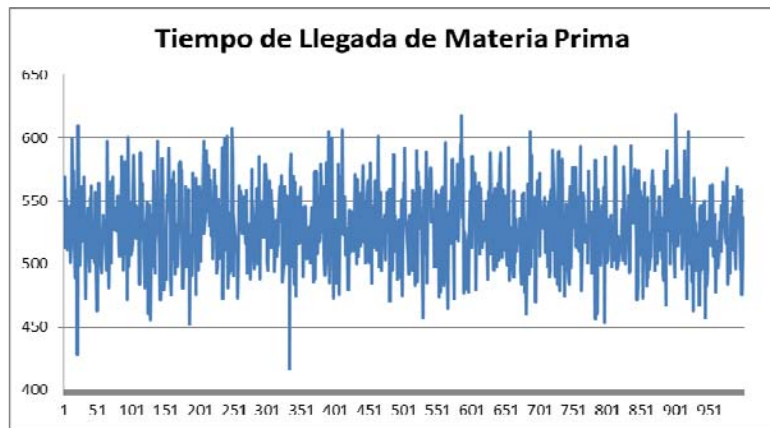


Figura 3. Variación del tiempo (seg.) de llegada de la materia prima.

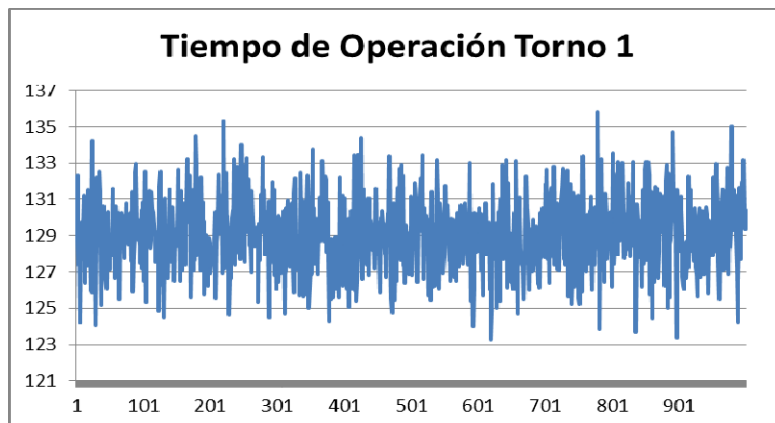


Figura 4. Variación del tiempo (seg.) de uso de la primera actividad torno 1.

Los datos del resto de las operaciones muestran un comportamiento similar, con fluctuaciones alrededor de un valor promedio aptas para su estudio en el marco de la optimización del proceso.

Los diagramas de dispersión de las figuras 5 y 6 representan la forma en que se distribuyen los tiempos para las operaciones anteriormente mencionadas. Este tipo de diagrama es extremadamente útil. En el caso del estudio de tiempos en dos turnos de trabajo donde el proceso es el mismo en función de las actividades realizadas, se pueden detectar desviaciones y optimizar un turno en función del otro. En este caso no se ha dado esta situación, ya que solamente se realiza un turno de trabajo, pero en otras empresas en donde se realizan varios turnos o que la llegada del material se realiza por diversos proveedores, este tipo de gráficos permite detectar las inconsistencias y tomar las acciones correctivas correspondientes. El problema podría surgir luego en el entrenamiento de la RNA, ya que al tener datos del mismo proceso industrial pero con diferentes valores en función del turno de trabajo, los patrones de entrenamiento podrían generar inconsistencias en la predicción.

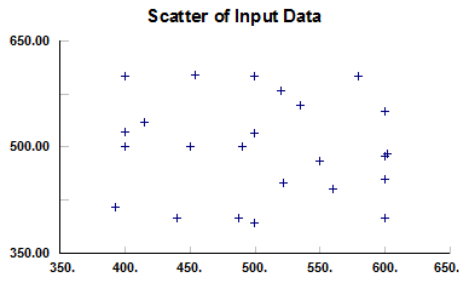


Figura 5. Diagrama de dispersión de los tiempos (seg.) de llegada de la materia prima

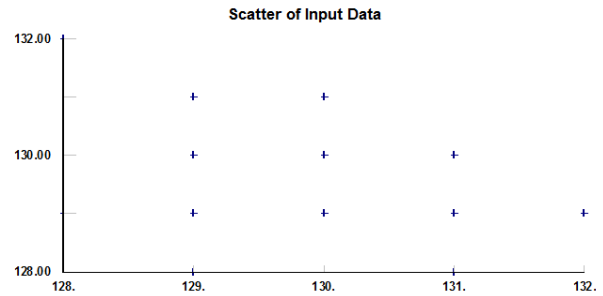


Figura 6. Diagrama de dispersión de los tiempos (seg.) de operación del torno 1

### 3 ANALISIS MEDIANTE REDES NEURONALES ARTIFICIALES

El tipo de RNA que se intenta aplicar es el perceptrón multicapa, entrenado mediante el método de backpropagation estándar. En cada operación dentro del proceso productivo se aplica una RNA, como se ilustra en la figura 7, donde  $s_i$  representa el vector o conjunto de variables de entrada estáticas (es decir que no cambian con el tiempo, como tipo de máquina, proceso industrial referenciado, etc.) y  $a_i$  representa el vector o conjunto de variables de entrada dinámicas (es decir que cambian con el tiempo, como cantidad de operarios, tiempo de proceso del recurso utilizado, etc.) en la etapa  $i$ -ésima, y  $Q$  representa la función de salida de la RNA.

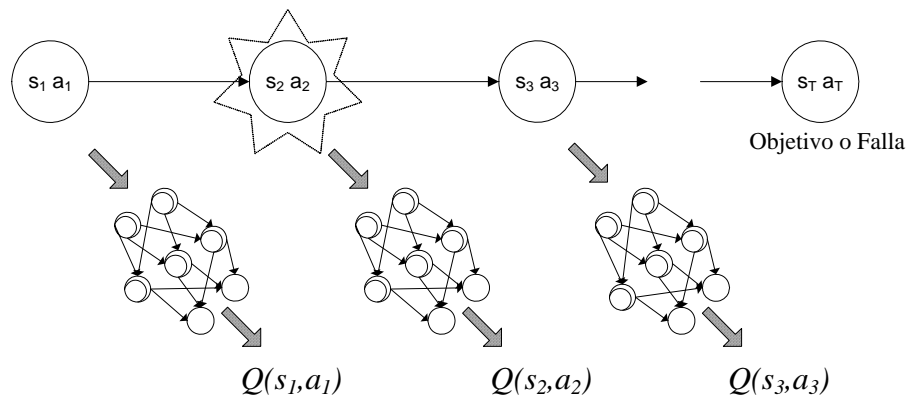


Figura 7. Esquema del análisis mediante RNA de un proceso industrial, como por ejemplo clasificación de equipos por arranque de viruta.

En cada RNA, el número de neuronas  $n$  de la capa de entrada es igual la suma de los parámetros de los vectores  $s_i$  y  $a_i$ . A la capa intermedia se le asignan  $(n/2)+1$  neuronas. Finalmente, la capa de salida posee sólo una neurona, cuyos valores  $Q$  pueden ser:

- +100 si el proceso es satisfactorio,
- -100 si el proceso indica una falla o demora en la producción, u
- otro valor, indicando en qué nivel de eficiencia se encuentra el proceso industrial.

La cantidad de patrones de datos disponibles (número de secuencias temporales de entrada en pasos de 1) depende

de cada tarea en particular, ya que al estar asociados a un proceso productivo existen tareas que son más rápidas que otras, generando mayor cantidad de patrones. Entre los datos típicos obtenidos a partir de este tipo de industrias metalmeccánicas, la cantidad mínima de patrones para una tarea es de 75 (tareas que se realizan un máximo de 2 veces al día) y la máxima es de 1850 (tareas que se realizan hasta 50 veces por día). A los fines de evitar el sobreentrenamiento, todas las RNA se entrenan sobre un total de 2600 épocas (o iteraciones). La cantidad de patrones utilizados como intervalo de validación es del 10% del total de patrones para cada RNA, aplicando siempre la técnica de *bootstrapping* (Efron and Tibshirani, 1993) para la selección de los patrones que quedan en el intervalo de validación. Por cada época realizada el vector de datos que se utiliza para validación varía en forma aleatoria, seleccionando hasta un máximo del 10% de la cantidad de patrones a entrenar. La cantidad de patrones utilizados como intervalo de predicción pura es de sólo 1 patrón o vector, seleccionado aleatoriamente para cada RNA.

En el Apéndice se explica en detalle la técnica de RNA y el procedimiento para su implementación.

#### 4 CONCLUSIONES

El presente trabajo realiza un relevamiento detallado de un proceso de producción industrial, basado en datos de una empresa del rubro metalmeccánico de autopartes de la región de Rafaela, provincia de Santa Fe, durante dos meses del año 2011. Se determinaron las operaciones o variables intervinientes en el proceso, las series temporales de datos de algunas de ellas y se analizó la técnica de RNA como potencial herramienta de análisis a fines de optimizar el proceso productivo. La interacción con las empresas relevadas confirma que valoran seriamente la importancia de una buena gestión del proceso industrial y la optimización de sus recursos productivos.

La posibilidad de incorporar nuevos datos de otras empresas y diferentes procesos productivos aportaría aún más información a las RNA, permitiendo mejorar su predicción (Bigus, 1996). Diferentes números de neuronas en la capa intermedia de las RNA permitirán optimizar su estructura en función del problema estudiado. Se continuará con el relevamiento de datos y se analizará también el uso de redes neuronales de mapas autoorganizativos para proceder a la clasificación de las diversas actividades y procesos industriales.

#### APÉNDICE – Técnica de Redes Neuronales Artificiales (Perceptrón Multicapa Feed-Forward)

##### A.1 Arquitectura y Entrenamiento de las RNA multicapas

En la aplicación de RNA multicapas en problemas de *clasificación*, por ejemplo en estudios de Ingeniería o de Clínica Médica, muchas veces se desconocen exactamente cuáles y cuántas son las variables de entrada necesarias para que la red pueda ajustar su conjunto de pesos a fin de relacionar correctamente las entradas presentadas con los valores conocidos de salida, durante su entrenamiento. Tampoco se conocen *a priori* el número mínimo de capas ocultas necesarias. Se recurre entonces al método de prueba de diversas configuraciones para compararlas y determinar las mejores, de acuerdo a criterios de evaluación de sus performances (Al-Zubaidi et al., 2011). Parte de este proceso de optimización consiste en cambiar la cantidad y/o las variables de entrada y de salida. La optimización se completa con la definición del número de neuronas de la o las capas ocultas. La mejor estructura es la que logra satisfacer un criterio de calidad para la función costo elegida. Más adelante se justificará la capacidad de las RNA multicapas, suficientemente *densas*, para relacionar o “mapear” correctamente los valores de entrada con los de salida.

Dos ejemplos tipo de aplicación de RNA multicapas en problemas de clasificación en los cuales se practicó este proceso de optimización son: el pronóstico del flujo del Nilo Azul (Shamseldin, 2010) y la predicción de metástasis esquelética en cáncer de próstata (Chiu et al., 2008). En el primer caso, después de ensayar diversas variables de entrada, las mejores matrices obtenidas correspondieron a las RNA que tuvieron como variables de entrada las expectativas estacionales del índice de precipitación y de la descarga observada. En el segundo caso, con el antígeno específico prostático (PSA) y la edad del paciente como las dos variables ya definidas de antemano en la entrada de la red, resultaron las RNA con cuatro capas ocultas las de mejor aprendizaje después del entrenamiento.

Otra aplicación muy extendida de las RNA multicapas es en predicción *por regresión no lineal* de series de tiempo. Para que las redes puedan aprender la dinámica del proceso que genera la serie, el número de unidades de entrada debe ser lo suficientemente grande como para que la dimensión del espacio de entrada (subespacio de “incrustamiento”) permita reproducir la geometría fractal del atractor en el verdadero espacio de estados del sistema dinámico (Kantz y Schreiber, 1996). Varios teoremas establecen el valor *mínimo* de la dimensión  $m$  del espacio de entrada en relación con la dimensión efectiva (box counting dimension),  $D_E$ , del atractor de la serie de tiempo muestreada (Whitney, Taskens, Sauer). El más general de éstos, el Teorema de Sauer establece que en el caso que las mediciones se hagan a través de una función de medición “suave”, entonces debe ser  $m > 2 D_E$  (Sauer, 1991). En casos donde la dimensión del atractor es conocida puede aplicarse esta relación para dimensionar el número de neuronas de entrada en el diseño de la arquitectura de la RNA. Un ejemplo para el caso de un láser de Resonancia Magnética Nuclear modulado periódicamente puede encontrarse en Kantz y Schreiber (1996).

Sin embargo, las incertezas en la estimación de  $D_E$  hacen que, en la mayoría de los casos, sea inviable la aplicación de las tesis de los teoremas mencionados para estimar el número  $m$  de neuronas de entrada. Una alternativa bien conocida es el seguimiento de la derivada de la función de correlación acumulada en función del número de



elementos de entrada. Si los valores de la derivada en la zona lineal de la función saturan después de cierto valor, se toma a éste como la mínima dimensión  $m$  de los vectores de entrada de la RNA. La aplicación de este proceso permitió dimensionar el número de neuronas de entrada de las RNA entrenadas para el análisis de la serie de la actividad solar, medida a través del número relativo de Wolf (Calvo et al., 1995). Trabajos posteriores aplicando RNA a la predicción de la actividad solar han empleado la arquitectura óptima 12:3:1 hallada por Calvo et al. (Verdes et al., 2000). Las predicciones del Ciclo Solar 23 obtenidas con estas redes han sido exitosas, superando las de algunos modelos predictores reconocidos (Parodi et al., 2010).

Con el único supuesto que la dinámica del proceso que genera la serie de tiempo bajo análisis es *esencialmente determinística*, es posible encontrar la estructura óptima de la RNA multicapas aún sin el conocimiento de la dimensión fractal del atractor que permite definir el número de unidades de entrada. El proceso determinístico se representa como una curva continua en el espacio de estados y también, si el número  $m$  de unidades de entrada es lo suficientemente grande, en el pseudo espacio de dimensión  $m$  donde están los vectores de entrada, ya que la trayectoria definida por éstos es *equivalente* a la verdadera trayectoria en el espacio de fases, en el sentido que existe una relación uno a uno entre los puntos de ambos espacios (Kantz y Schreiber, 1996). El objetivo del entrenamiento es que la RNA aprenda la función  $F$  que relaciona los vectores actuales de entrada con los que representan los estados futuros (para ello se presentan a la RNA pares de entrada - salida durante el entrenamiento). Ahora bien,

para cualquier RNA en capas, con una *única* capa oculta y con la *suficiente* cantidad de neuronas de entrada y de unidades en la capa oculta, se puede demostrar su capacidad para aprender *cualquier función continua* con el grado de precisión que se desee, después de la presentación de un número *suficiente* de patrones de entrada - salida (Rojas, 1995). Esta propiedad justifica el método de prueba de configuraciones de RNA con una sola capa en aplicaciones a la predicción de las realizaciones de una serie temporal por regresión, haciendo crecer "paso a paso" la complejidad de la red, tanto en la cantidad de unidades de entrada como en el número de neuronas ocultas. Más adelante se encuentra una descripción detallada de este proceso de crecimiento, ejemplificado con la aplicación de una regla empírica que relaciona la cantidad de conexiones con el tamaño de la muestra de patrones disponible.

La pregunta que cabe hacerse es si una red neuronal en capas con alimentación feed-forward y con un número suficiente de unidades en las capas ocultas, después de ser entrenada con ejemplos, puede relacionar ("mapear") cualquier conjunto  $m$ -dimensional de entradas con su correspondiente conjunto  $n$ -dimensional de salida, *ya sea en aplicaciones de predicción o clasificación*. Esto efectivamente es así y en Cybenko (1989) puede encontrarse la demostración de este resultado general.

## A.2 Crecimiento de la conectividad de las RNA multicapas

De lo anterior resulta teóricamente viable obtener RNA en capas que aprendan la dinámica del proceso de una serie temporal, sin un análisis previo de la dimensión del atractor de la serie de tiempo bajo estudio. En esta alternativa empírica se comienza con redes poco densas (es decir con pocas unidades o neuronas), de menor a mayor complejidad. Los parámetros que definen las conexiones entre las neuronas en la RNA son los pesos y los bias (hay un bias por cada unidad oculta). Los primeros son los factores, aleatorios al inicio del entrenamiento, que multiplican el valor de la entrada antes de ser computada por cada unidad de la capa oculta. La función de activación sigmoidea de cada neurona de esta capa tiene un valor umbral a partir del cual esta función tiene valores crecientes ante el aumento del valor de entrada pesado. En la práctica, el umbral o bias se implementa con una entrada extra por neurona que tiene el valor fijo de cada bias. Igual que los pesos iniciales, los bias resultan números aleatorios al comienzo del entrenamiento. Por último, la salida de cada unidad oculta es multiplicada por otro peso y resulta una entrada para la unidad colectora en la segunda capa. La salida de esta unidad es sencillamente la suma de todas estas entradas pesadas, pues su función de activación es la función identidad. En muchas aplicaciones la capa de salida sólo tiene una neurona.

El número total de parámetros de la RNA (número de pesos + número de bias) debe estar dentro de cierto intervalo, de forma que sea lo suficientemente alto para poder aprender la dinámica del sistema y a su vez acotado para no producir sobreaprendizaje ("overfitting" o aprendizaje de ruido). En el caso de aplicaciones de regresión, una regla empírica muy empleada propone que  $p$  debe pertenecer al intervalo  $(5N-10N)$  donde  $p$  es el número de patrones de entrada y  $N$  es el número de parámetros (Calvo et al., 1995). Por ejemplo: suponiendo que se disponga de 100 patrones de datos, se comienza considerando una estructura simple, por ejemplo 2:2:1 (unidades de entrada: unidades ocultas: unidad de salida); entonces  $N = 2 \times 2 + 2 \times 1 + 2 = 8$  y el intervalo es (40-80),  $p$  no está incluido y por lo tanto la RNA es demasiado elemental como para lograr la generalización. Luego se agrega una unidad de entrada para obtener una estructura 3:2:1. Ahora  $N = 3 \times 2 + 2 \times 1 + 2 = 10$ , el intervalo es (50-100) y esta RNA ya puede aprender. Si aún no aprende la dinámica, es decir, no se alcanza un mínimo de la función costo en el intervalo de validación elegido, se incrementa la densidad de la red. Esto tiene un límite. Por ejemplo, la red 5:3:1 está sobredimensionada, según la regla empírica antedicha, pues como  $N = 5 \times 3 + 3 \times 1 + 3 = 21$  esto implica que  $p$  no pertenece al intervalo correspondiente, (105-210). Así, paso a paso, se hace "crecer" la RNA hasta obtener aprendizaje en el sentido explicado más arriba. Entre cada paso hay que entrenar y observar la performance de la función costo en el intervalo de validación, donde la red entrenada hace predicciones a partir de valores de entrada

no utilizados en su entrenamiento para evaluar su capacidad de generalización, es decir de aprendizaje de la dinámica intrínseca del problema en estudio. (Parodi et al., 2006).

### A.3 Criterios de detención (“Stopping”) en el entrenamiento y de calidad de las RNA

Las mejores RNA resultantes del entrenamiento son aquellas que aprenden la dinámica del proceso bajo estudio y no la componente estocástica introducida durante la medición de los datos. Las RNA son entrenadas hasta que la función costo, que mide la calidad predictiva en el intervalo de validación, comienza a *crecer*. Se aplica este criterio de “stopping” porque el aumento de la función costo significa aprendizaje de *ruido* en el entrenamiento. Durante éste son presentados todos los  $p$  patrones generados a partir de los datos, con  $p = N - m$ , donde  $m$  es el número de unidades de entrada. Cada secuencia de presentación del conjunto completo de patrones se denomina época. Habitualmente se elige como función costo el error cuadrático medio, RMSE (root mean square error). Las aplicaciones informáticas actuales de RNA como, por ejemplo la Toolbox de RNA de Matlab (MathWorks), permiten el seguimiento gráfico *online* del RMSE época a época. El criterio estándar de “stopping” del entrenamiento establece que cuando el RMSE comienza a crecer, se detiene el aprendizaje y se archiva la matriz de pesos y bias de la RNA. Las mejores RNA obtenidas son aquellas para las cuales el valor del RMSE alcanzado en el intervalo de validación es uno de los más bajos al momento de detener el entrenamiento.

En general, por supuesto, el número de épocas aumentará con la complejidad de la dinámica: una baja relación señal - ruido y/o valores altos del primer coeficiente de Lyapunov, es decir, fuerte comportamiento caótico (Kantz y Schreiber, Tsonis). Por otro lado, muchas veces ocurre que no se llega a observar el aumento del RMSE en el seguimiento on line en validación, pues la RNA queda “atrapada” en un mínimo local de la función costo.

En aplicaciones de RNA multicapas a problemas de clasificación, diversos criterios de calidad son empleados para controlar el crecimiento de la red por prueba y error.

El método de optimización de la arquitectura por crecimiento controlado de la conectividad de la RNA se ha empleado en los más diversos campos, tanto para predicción como para clasificación, por ejemplo en Parodi et al. (2006).

### REFERENCIAS

- Al-Zubaidi S., J.A. Ghani, and C. Hassan Haron, Application of ANN in Milling Process: A Review. Modelling and Simulation in Engineering. Article ID 696275, 2011:1-7. doi:10.1155/2011/696275, 2011.
- Alves R.M.B., and Nascimento, C.A.O., Analysis and detection of outliers and systematic errors in industrial plant data. Chemical Engineering Communications, 194:382-397, DOI:10.1080/00986440600899955, 2007.
- Bigus J., Data Mining with Neural Networks, McGraw-Hill, 1996.
- Calvo RA, Ceccatto, HA, Piacentini, RD., Neural Network Prediction of Solar Activity. The Astrophysical Journal, 444(2):916-921, 1995.
- Chiu J.S., Wang, Y.F., Su, Y.C., Wei, L.H., Liao, J.G., Li Y.C., Artificial neural network to predict skeletal metastasis in patients with prostate cancer. J. Med. Syst. 33(2):91-100, 2009.
- Cybenko G. "Approximation by Superpositions of a Sigmoidal Function," Math. Control Signals Systems, 2, 303-314, 1989.
- Efron B., and Tibshirani R.J., An Introduction to the Bootstrap, Chapman & Hall, New York, 1993.
- Fornari J., Grieco, S., Parodi M., y Luccini E., Análisis de líneas de producción utilizando técnicas tradicionales. Energeia, 8:41-52, 2010. <http://www.uca.edu.ar/index.php/site/index/es/universidad/facultades/rosario/quimica-ingenieria/investigacion/revista-energeia/>
- Forst W., y Hoffmann, D., Optimization-Theory and Practice. DOI: 10.1007/978-0-387-78977-4\_8. Springer, 2010.
- Govindhasamy J.J., McLoone, S.F., Irwin, G.W., French, J.J., and Doyle, R.P., Neural modelling, control and optimisation of an industrial grinding process. Control Engineering Practice, 13:1243-1258, 2005.
- Kantz H. and T. Schreiber, Nonlinear time series analysis, Cambridge University Press, Cambridge, UK, 1997.
- Parodi M., A. Ceccatto, D. Piacentini. Neural network analysis of the effect of Mount Pinatubo eruption on the Southern Hemisphere High Latitude ozone WSEAS Transactions on Information Science & Applications 3(3):588-594, 2006.
- Parodi M., E. Luccini, R. Piacentini, S. Grieco, J. Fornari y L. Herrera. Análisis de la capacidad predictiva de una red neuronal artificial aplicada al número de manchas solares en el periodo 1997-2008 (ciclo solar 23). Avances en Energías Renovables y Medio Ambiente. 14, 2010.
- Rao S.S., Engineering Optimization: Theory and Practice, Fourth Edition, John Wiley & Sons, 2009.
- Rojas R., Neural Networks-A Systematic Introduction, Springer-Verlag, New York, 502 pp., 1995.

Sauer T., J. A. Yorke, and M. Casdagli. Embedology. *Journal of Statistical Physics*, 65(3/4):579-616, 1991.

Shamseldin Y., Artificial neural network model for river flow forecasting in a developing country, *J. Hydroinformatics*, 12:22-35, 2010.

Verdes P. F., M. A. Parodi, P. M. Granitto, H. D. Navone, R. D. Piacentini, H. A. Ceccatto. Predictions of the maximum amplitude for solar cycle 23 and its subsequent behavior using nonlinear methods. *Solar Physics*, 191(2):419-425, 2000.

Zarenezhad B., y Aminian, A., Application of the neural network-based model predictive controllers in nonlinear industrial systems. Case study. *Journal of the University of Chemical Technology and Metallurgy*, 46:67-74, 2011.