



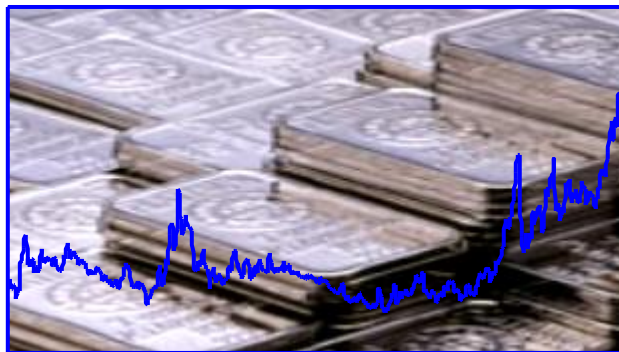
UNIVERSIDAD DE CHILE

FACULTAD DE CIENCIAS ECONOMICAS Y ADMINISTRATIVAS

Escuela de Graduados

**Prediciendo cambio de signo en el
precio de la Plata:
Modelos de Redes Ward con
Algoritmos Genéticos y Optimización Aleatoria.**

Evidencias para el periodo septiembre 1995 - marzo 2006



Tesis para optar al Grado de : **MAGÍSTER EN FINANZAS**

Alumno : **Renato Alejandro Bueno Bórquez**

Profesor Guía : Ph.D. Antonio Parisi Fernández

Santiago de Chile, Invierno 2006

A ti Renatito, que siempre has sido mi inspiración y me ha dado Fuerza.

**A ti Jenny, mi compañera incondicional que con su consejo,
Belleza y ternura, permitió culminar esta importante etapa en mi Vida.**

**A mis Padres, mi hermano Claudio y tía Bruni, forjadores de Amor
que con su Sabiduría, me instaron a seguir siempre adelante.**

A mis queridos amigos Antonio y Manuel, que gracias a su critica, disposición y

conocimiento, pude imprimirle un sello de calidad a esta tesis.

A mis queridos amigos Ivonne, Michelle, Iván, Camilo y Christian,

por su constante compromiso y preocupación.

TABLA DE CONTENIDO

<u>1</u>	<u>INTRODUCCION</u>	<u>4</u>
<u>2</u>	<u>MARCO TEORICO</u>	<u>9</u>
	2.1 <i>MERCADOS DE LA PLATA Y SU RELACION CON OTROS ACTIVOS FINANCIEROS.</i>	9
	2.1.1 <i>OFERTA Y DEMANDA DE LA PLATA.</i>	9
	2.1.2 <i>MERCADOS, PRECIO DE LA PLATA Y FACTORES RELEVANTES LIGADOS A SU COMPORTAMIENTO.</i>	11
	2.1.2.1 <i>Precio de la Plata relativo al Oro.</i>	12
	2.1.2.2 <i>Precio de la Plata relativo al Cobre y al Petróleo.</i>	13
	2.1.2.3 <i>Precio de la Plata relativo a GBP/USD y USD/YJP.</i>	15
	2.2 <i>MÉTODOS NO-LINEALES DE PREDICCIÓN "REDES NEURONALES"</i>	16
	2.3 <i>MÉTODOS NO-LINEALES DE PREDICCIÓN "ALGORITMOS GENÉTICOS"</i>	21
<u>3</u>	<u>COMENTARIOS Y REGULARIDADES EMPÍRICAS OBSERVADAS EN LOS DATOS</u>	<u>24</u>
<u>4</u>	<u>METODOLOGÍA Y DATOS</u>	<u>26</u>
	4.1 <i>BÚSQUEDA DE ARQUITECTURA ÓPTIMA</i>	27
	4.2 <i>MODELOS MULTIVARIADOS ARIMAX</i>	29
	4.3 <i>EVALUACIÓN DE LA CAPACIDAD PREDICTIVA Y SIGNIFICANCIA ESTADÍSTICA</i>	31
	4.4 <i>ANÁLISIS DE LOS RESULTADOS</i>	32
	4.4.1 <i>RESUMEN DE LOS MEJORES RESULTADOS OBTENIDOS POR METODOLOGÍA.</i>	35
<u>5</u>	<u>CONCLUSIONES</u>	<u>38</u>
<u>6</u>	<u>BIBLIOGRAFÍA Y REFERENCIAS</u>	<u>39</u>
<u>7</u>	<u>ANEXOS</u>	<u>42</u>
	7.1 <i>OBTENCIÓN DE VARIABLES SOBRE LAS CUALES SE APLICO LA RED WARD ALEATORIA Y RED WARD ALEATORIA CON ALGORITMO GENETICO .</i>	42
	7.2 <i>ANEXO 1: RESULTADOS RED NEURONAL - ARQUITECTURA WARD – MODELO ALEATORIO</i>	44
	7.3 <i>ANEXO 2: RESULTADOS RED NEURONAL POTENCIADA CON ALGORITMO GENETICO - ARQUITECTURA WARD – MODELO ALEATORIO</i>	45
	7.4 <i>ANEXO 3: CARACTERÍSTICAS Y ANTECEDENTES GENERALES DE UNA RED NEURONAL ARTIFICIAL (RNA).</i>	47
	7.4.1 <i>FUNCIONES DE ACTIVACIÓN O TRANSFERENCIA.</i>	47
	7.4.2 <i>ESTRUCTURA BÁSICA DE UNA RED NEURONAL ARTIFICIAL</i>	50
	7.4.3 <i>CARACTERÍSTICAS GENERALES DE LAS REDES NEURONALES ARTIFICIALES</i>	50
	7.4.4 <i>CLASIFICACIÓN DE LAS REDES NEURONALES ARTIFICIALES</i>	51
	7.4.5 <i>VENTAJAS DE LAS REDES NEURONALES ARTIFICIALES.</i>	52
	7.4.6 <i>DESVENTAJAS DE LAS REDES NEURONALES ARTIFICIALES.</i>	53
	7.5 <i>ANEXO 4: FUNCIONAMIENTO ROLLING Y RECURSIVO EN MODELOS DE PREDICCIÓN.</i>	54
	7.5.1 <i>FUNCIONAMIENTO RECURSIVO</i>	54
	7.5.2 <i>FUNCIONAMIENTO ROLLING</i>	55
	7.6 <i>TEST DE ACIERTO DIRECCIONAL DE PESARAN & TIMMERMANN (1992).</i>	57

1 INTRODUCCION

El propósito de la presente investigación es estudiar y desarrollar modelos flexibles, versátiles y eficientes, estadísticamente y económicamente significativos, para proyectar la variación futura del precio de la Plata. Se consideró a éste activo para su estudio, principalmente por el evento de no existir evidencias de trabajos predictivos con estas técnicas para este activo, además de la liquidez y profundidad del mercado, en que éste metal se cotiza. El porqué, de utilizar modelos de redes neuronales para la predicción, surge básicamente, por estudios de proyección sobre series temporales previos, cuyos resultados han sido un importante aporte para las Finanzas Modernas.

Otro factor de relevancia que se debe considerar en la definición del modelo de redes neuronales, son las variables explicativas, de las cuales depende en gran medida el éxito de ésta y otras investigaciones. Respecto a este ítem, después de un acabado estudio de 14 variables de relevancia mundial (índices accionarios, bolsas de metales, precios de commodities y monedas), se efectuó pruebas de análisis técnico (cruce de medias móviles con 5, 10 y 30 rezagos) y econometría básica (análisis de regresiones – coeficientes de correlación múltiples), donde finalmente se definieron 7 variables y sobre las cuales se desarrollaron los modelos predictivos.

Durante el desarrollo, surge la pregunta de si es posible mejorar los resultados de porcentaje predicción de signo de una red neuronal, ante lo cual, considerando la teoría darwiniana sobre la “evolución de las especies”, se incorporó un Algoritmo Genético.

Para cumplir con el propósito enunciado, se desarrolló un proceso de análisis con series de datos semanales del precio de la Plata de los últimos 10 años (precio cierre en USD-onza), fijando como principal objetivo, que a través de metodologías sustentadas en Algoritmos Genéticos y Redes Neuronales, se pueda anticipar la dinámica que experimenta el precio de la Plata (en términos de cambio de signo) y comprobar si los resultados obtenidos en ambos modelos, demuestren un comportamiento estable y sinérgico. Adicionalmente, se comprobará si la fusión de ambos modelos, logra generar mejoras considerables en el porcentaje de predicción de signo, depurando la información existente, de tal manera, de asegurar retornos económicamente más atractivos que en los modelos en particular mencionados anteriores.

Finalmente, el desempeño de cada uno de los métodos a utilizar, será medido por el número de predicciones correctas (hits) del signo de la variación del precio, aplicando para ello el test de acierto direccional (DAT) ¹ de Pesaran & Timmermann (1992), comprobando de esta forma que las proyecciones obtenidas por los modelos versus los resultados alcanzados con una estrategia “*buy and hold*” ² (independiente de la significancia estadística), son económicamente superiores o si estos permiten minimizar las pérdidas.

Respecto al análisis predictivo en otros activos financieros, es conocido el aporte realizado por, Parisi y Parisi³ a la literatura financiera con el uso de sistemas predictivos no paramétricos tales como las redes neuronales, lógica difusa y los algoritmos genéticos para proyectar variables financieras y que sirven de base para conformar portafolios de inversión. Estas técnicas no basan sus capacidades predictivas en la utilización de supuestos de normalidad en los retornos de los activos financieros, demostrando que es posible obtener altas capacidades predictivas, estadísticamente significativa y económicamente rentables, punto que pretendemos demostrar y reafirmar en el presente trabajo.

Con la presencia de estas dinámicas no lineales, se podría dar la posibilidad de explorar este campo para realizar predicciones más exactas que aquellas otras proporcionadas por un modelo estocástico lineal y, en concreto, por el modelo referente en la literatura: el paseo aleatorio (random walk).

¹ *Directional Accuracy Test*

² La estrategia “*buy and hold*” es una estrategia de inversión pasiva que no aplica gestión alguna. consiste en comprar el activo en cuestión y mantenerlo durante el periodo de la inversión, para liquidarlo al final de éste. La rentabilidad de esta estrategia está dada por la diferencia entre el valor de venta y el valor de compra del activo en estudio, que para este caso es el de la Plata.

³ Vease por ejemplo: Parisi, A., Parisi, F. & Guerrero, J.L.(2003).“Modelos Predictivos de Redes Neuronales en Índices Bursátiles ” , El Trimestre Económico, México.

Los trabajos de Takens (1981) y Casdagli (1989), entre otros, han establecido la metodología necesaria para la modelización no-lineal de series temporales. En concreto, el Teorema de Takens establece que, dada una serie temporal determinista, existe una función tal que donde τ es el factor de retardo y m la dimensión de incrustación. Lo que permite la posibilidad de anticipar la dinámica futura de una serie temporal determinista considerando únicamente sus valores pasados. El problema a resolver consiste en encontrar una buena representación o aproximación funcional. Gracias a los avances en el área de la informática, se han desarrollado potentes y sofisticadas técnicas no-lineales de predicción para la aproximación funcional como son, las redes neuronales y los algoritmos genéticos, lógica difusa, fractales, autómatas celulares entre otras. En un primer momento, las aproximaciones más empleadas fueron aquellas basadas en generalizaciones del método de ocurrencias análogas. Por ejemplo, Diebold y Nason (1992) aplicaron el método de regresión localmente ponderada, y Lisi y Medio (1997) una regresión local. En los últimos años se observa una mayor intensidad en el uso de redes neuronales (Kuan and Liu, 1995; Tenti, 1996; Yao et al., 1997; Zang and Hu, 1998; Hu et al., 1999; Yao and Tan, 2000; Walzack, 2001).

En esta materia, Leung, Daouk & Chen (2000) compararon la capacidad predictiva de los modelos de clasificación con los de estimación de nivel⁴ y concluyeron que los primeros⁵ se desempeñan mejor que los segundos en términos de su tasa de acierto⁶ y son capaces de generar beneficios. Mientras que Lo & MacKinlay (1988), quienes registran una correlación serial positiva entre los retornos semanales, se añaden los resultados de DeBondt & Thaler (1985), Fama & French (1988), Poterba & Summers (1988) y Chopra, Lakonishok & Ritter (1992), quienes encuentran una correlación serial negativa en los retornos de los activos individuales y varios portafolios, sobre intervalos de tres a diez años.

En este contexto, Jegadeesh (1990) examina la predictibilidad de los retornos mensuales sobre activos individuales y encuentra una correlación serial negativa de primer orden altamente significativa para rezagos de dos meses y una correlación serial positiva para rezagos mayores.

⁴ Adaptive exponential smoothing, vector autoregression, multivariate transfer function y multilayered feedforward neural network.

⁵ Los cuales son propuestos para predecir el signo o la dirección del retorno de los índices bursátiles S&P 500.

⁶ Medida por el número de veces en que la dirección pronosticada es correcta.

Un aspecto interesante y escasamente considerado en el ejercicio predictivo, consiste en aprovechar los efectos sinérgicos que pudieran existir entre diferentes métodos predictivos. Es así que, en esta tesis, se busca potenciar la red neuronal con un algoritmo genético, lo que permitiría obtener una aproximación a la dinámica subyacente en los datos, a partir de una combinación de las predicciones obtenidas con métodos individuales. Se han presentado dos razones básicas por las cuales se recomienda el uso de la combinación de datos. En primer lugar, diferentes métodos de predicción poseen distintas ventajas y reflejan desigual información. Por lo tanto se puede esperar que una amalgama de predicciones permita aprovechar las ventajas de cada método y mejorar los resultados en términos predictivos. Como segundo punto importante, la composición permite reducir los sesgos que pudieran existir empleando sólo un método predictivo.

Con ello se pretende que los sistemas de predicción estudiados cumplan las siguientes características:

- Que los sistemas cuyas variables endógenas y exógenas puedan presentar relaciones no lineales entre ellas.
- Que se adapten a la nueva información en la medida que ésta disponible y, que refleje el cambio de las expectativas de los agentes presentes en el mercado de financiero.

Al momento de llevar a cabo un ejercicio predictivo basado en la combinación de modelos, como son las redes neuronales supervisadas y algoritmo genético, es necesario contar con sistemas de predicción flexibles, que logren capturar relaciones no-lineales presentes entre las variables. Por lo tanto es preciso definir una forma de combinar predicciones. El procedimiento habitual consiste en emplear una regresión para asignar pesos relativos a cada una de las predicciones obtenidas por cada método individual (Kwok and Lubecke, 1990). Recientemente, asumiendo la limitación de una perspectiva lineal, se ha empleado un método de combinación basado en redes neuronales (Lubecke et al., 1998). Sin embargo, la adopción de una red neuronal requiere el largo y tedioso proceso de diseñar y especificar una arquitectura previa.

Se empleó un algoritmo genético para combinar las predicciones obtenidas por 2 métodos individuales (en particular, un algoritmo genético y una red neuronal).

Las conclusiones del estudio nos indican, que se puede lograr capacidad predictiva estadísticamente y económicamente significativa con la Plata (en términos semanales), acciones sustentables a través de Redes Neuronales fusionados con Algoritmo Genéticos, por medio de las cuales se logró alcanzar un 62,55% de acierto extra muestral, en la predicción de cambio de signo para el periodo comprendido entre el 22 de septiembre de 1995 y el 31 de marzo de 2006.

Finalmente, esta tesis se estructura en 7 secciones. Posterior a ésta introductoria, se presentan los métodos predictivos empleados y su Marco Teórico. En la sección 3, se comentan ciertos aspectos del ejercicio predictivo desarrollado, describiéndose brevemente algunas características de las series analizadas y las regularidades empíricas observadas. En la sección 4, se muestran los resultados obtenidos por cada método en términos de predicción puntual y del porcentaje de aciertos de signo (tanto a uno como a diferentes horizontes temporales). Finalmente en el punto 5, se presentan las conclusiones. En los puntos 6 detallamos la bibliografía de referencia y el material de apoyo utilizado, culminando en la sección 7 con anexos y cuadros de los resultados obtenidos, además de otras informaciones relacionadas de las materias desarrolladas.

2 MARCO TEORICO

2.1 MERCADOS DE LA PLATA Y SU RELACION CON OTROS ACTIVOS FINANCIEROS.

2.1.1 OFERTA Y DEMANDA DE LA PLATA.

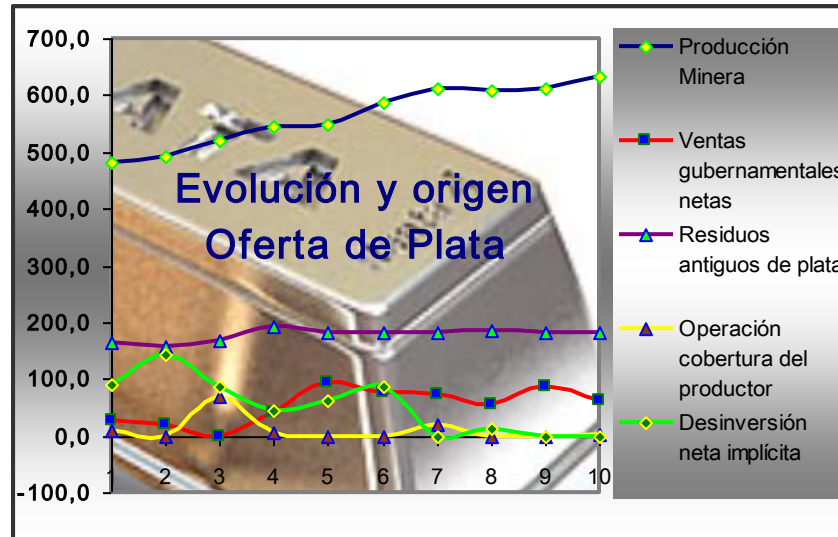
La plata como metal precioso, determina su precio tanto por su rol como producto como también de activo financiero. En efecto, al uso industrial, fotográfico y de joyería como principales sectores demandantes, se le une el requerimiento que existe por parte de los inversionistas como una alternativa al oro en lo referente a reserva de valor ante la inflación, depreciación de la moneda e inestabilidad política, entre otros. Sus inventarios son numerosos, pues se utiliza como una forma de atesoramiento, principalmente por el sector privado, y en menor medida por sector oficial.

Respecto de la fuerza oferente, ésta proviene principalmente de 3 fuentes. La primera y la más importante, se genera de la producción directa de minas, cuyo volumen de producción representa aproximadamente el 72% de la oferta total, concentrando la mitad de dicha producción en México, Perú y Estados Unidos, quienes se ubican en el primer, segundo y séptimos lugar de producción respectivamente, siendo su tercer productor Australia. Del total producido en minas, la extracción primaria de Plata representa sólo un 30%, la porción restante corresponde a sub productos del Plomo y Zinc con un 30%, del Cobre con el 25% y Oro con un 15% (cifras aproximadas).

La segunda fuente de stock esta dada por el reciclaje de chatarra con un 20,5% del total de la oferta, seguida por las ventas gubernamentales con 7% y finalmente por operaciones de cobertura de riesgo de los productores con una cifra cercana al 0,5%.

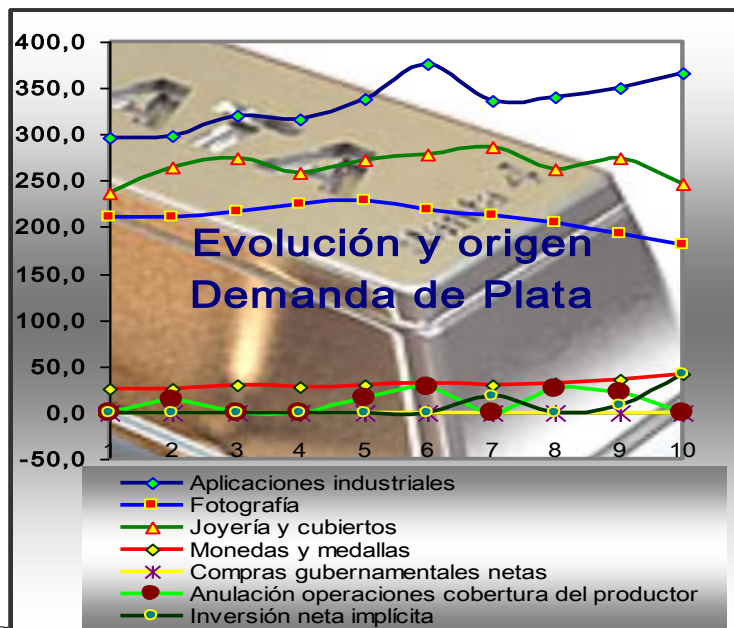
Evolución de la oferta de Plata periodo 1994 – 2004, cifras expresadas en MM/onzas

(fuente: The Silver Institute)



En relación a la demanda de éste activo, ésta se genera principalmente por usos industriales con 42%, seguido por la joyería y fotografía con el 27% y 21%, respectivamente. Otros usos del metal dicen relación con la confección de monedas e inversiones implícitas (gubernamentales) con el 5% cada una .

Evolución de la demanda de Plata periodo 1994 – 2004, cifras expresadas en MM/onzas



(fuente: The Silver Institute)

2.1.2 MERCADOS, PRECIO DE LA PLATA Y FACTORES RELEVANTES LIGADOS A SU COMPORTAMIENTO.

La Plata como activo financiero, se transa predominantemente en el “**London Bullion Market**” (LBM) y en el “**Commodity Exchange**” (filial Comex del NYMEX de Nueva York). El primero, corresponde al principal centro de transacciones extra bursátiles y principal mercado físico del mundo, siendo el más importante centro OTC u over-the-counter (principal a principal), y es aquí donde el proceso de negociación genera un precio referencia diario, el que es conocido como “**fix**”. El segundo mercado de importancia en que se transa este producto y que fue mencionado anteriormente, corresponde a una bolsa de futuros y opciones, la que de acuerdo a sus cifras transaccionales, es la que concentra la mayor parte de la actividad de los fondos de inversión de este metal.

Variaciones de precio periodo comprendido entre el



22 de septiembre de 1995 y 31 de marzo de 2006 (fuente bloomberg)

Si se considera que los inventarios disponibles de plata superan el año de consumo, se tiene un primer indicio de la importancia de los inversionistas en la determinación del precio de la plata. El precio de la plata como producto determina que su precio también tenga relación con el ciclo económico, el cual a su vez explica el precio de los metales base y del crudo.

De acuerdo a lo anterior, se podría sostener que este es el principal argumento para justificar que el mercado físico tiene una importancia menor en la determinación del precio de este metal.

Es por lo anterior que, al analizar el precio de la Plata se hace necesario considerar la evolución de factores económicos como por ejemplo la fortaleza o debilidad del dólar tras el aumento de expectativas de inflación en EEUU.; la amenaza de epidemia de gripe aviar en Europa y; el alza sostenida del precio del petróleo tanto por tensiones geopolíticas y factores climáticos (situación nuclear de Irán, dificultades en Nigeria, Irak y Arabia Saudita y desastres naturales en EEUU. y en el Golfo de México). En razón a este punto, se consideró necesario incorporar como variables dependientes y que puedan explicar el comportamiento del precio de la Plata a los siguientes activos: Oro, Cobre, Petróleo y monedas como el Dólar Americano, Yen Japonés y Libra Esterlina, activos financieros sobre los cuales en los siguientes puntos, se presentan algunas referencias.

Según antecedentes recogidos de la Comisión Chilena del Cobre (Cochilco), la mayor independencia alcanzada por la plata, se sustentaría en los efectos que ha tenido el fondo Zares de Barclay's Capital, que ha influido positivamente en el precio de esta. Recientemente, la Security and Exchange Comition (SEC) aprobó que este fondo se listara en la American Exchange (AMEX), lo cual generó efectos alcistas en la cotización, toda vez que este tipo de fondos requieren respaldar la emisión de cuotas con el material físico. Además, se debe destacar que los buenos fundamentos de demanda que están actuando sobre el mercado de los metales básicos, también influyen en la Plata, pues ésta tiene usos industriales que explican una relación de segundo orden.

2.1.2.1 Precio de la Plata relativo al Oro.

Con respecto a los precios del Oro y Plata al primer trimestre del 2006, podemos señalar que su valor ha aumentado 14,1% y 19,4% (respecto a mismo periodo del año 2005) con promedios que oscilan en 554,03 y 9,62 US\$/onza, respectivamente, alcanzando máximos históricos que no se habían visto, desde los años 80.

La relación existente entre el Oro y la Plata dista de ser siempre fuerte y directa, existiendo en el corto plazo periodos en que la relación entre ambos se hace difusa, aunque en un horizonte prolongado de

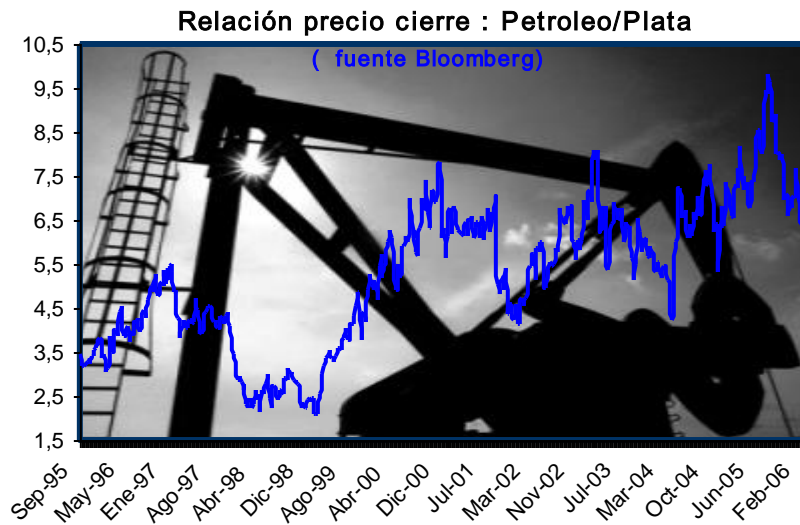
tiempo, se observa una relación más estable, que permite proyectar su movimiento futuro en conjunto. En cuanto a la condición de sustituto con el oro, existe una correlación positiva entre ambos precios en el largo plazo, que para el periodo analizado (septiembre 1995 marzo 2006), alcanza un indicador de 57%. Esto nos indica que la evolución del precio de la plata esta altamente ligada al del oro, razón por la cual esta variable se consideró en los modelos ARIMAX utilizados en el método de predicción de Algoritmo Genético.

Por lo anterior, es que generalmente en las investigaciones sobre el comportamiento de este activo, se analiza la evolución de la relación precios Oro/Plata, que en el presente estudio, nos demuestra una relación promedio de 63,35 veces con una desviación de 8,28 y valores mínimos y máximos de 41,18 y 80,45 al cierre del 06 febrero 1998 y 30 de mayo 2003, respectivamente. Al cierre del periodo en estudio, esta relación asciende a 70,46 veces, cifra superior al promedio del periodo, lo que implica que la plata se ha debilitado respecto al Oro, sin embargo, demuestra una tendencia fortalecida respecto a su máximo alcanzado en el año 2003.

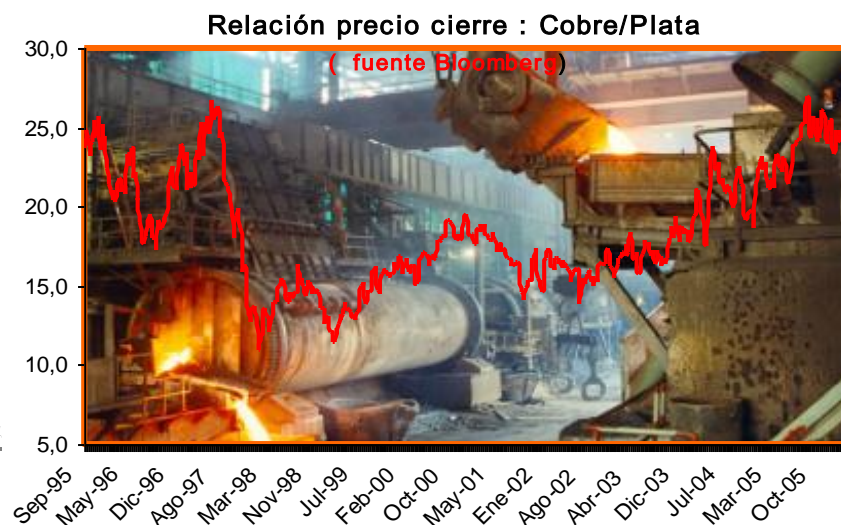


2.1.2.2 Precio de la Plata relativo al Cobre y al Petróleo.

Respecto al Petróleo, existe una correlación 10,9%, con un promedio de en su relación de precio en función de la Plata que asciende a 5,28 veces con una desviación de 1,63. El índice mínimo se alcanzó el 19 de febrero de 1999 con 2,09 veces y su máximo de 9,83 veces el 26 de agosto de 2005.

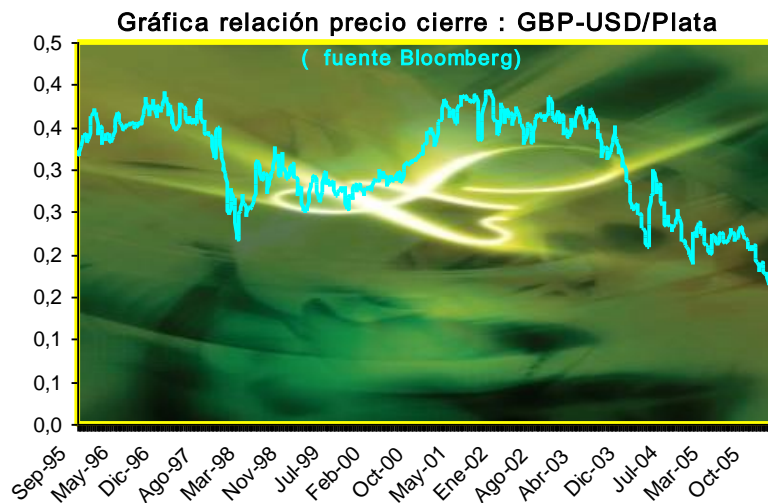


En relación al Cobre, su precio respecto a de la plata debería registrar una relación paralela en el largo plazo, lo cual no se cumple, sin embargo, por los comportamientos demostrados en el mediano y corto plazo, el precio de la plata se explica en un 25,4% por el comportamiento del cobre. El precio registra un promedio de 18,53 veces y una desviación de 3,66, con una relación mínima de 10,92 el 06 de febrero 1998 y un máximo de 27,03 veces el 26 de agosto 2005.



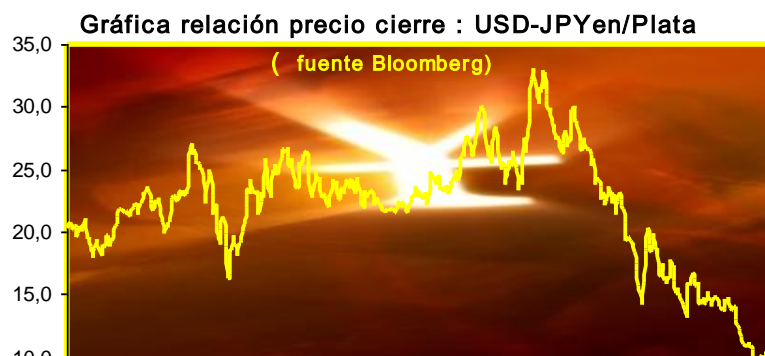
2.1.2.3 Precio de la Plata relativo a GBP/USD y USD/YJP.

Variaciones relación tipo de cambio Libra Esterlina/ Dólar EEUU - precio cierre Plata



Valor promedio de la relación asciende a 0,305, con una desviación de 0,059 y valores mínimos y máximos de 0,135 y 0,395, relaciones alcanzadas el 31 de marzo de 2006 y 23 de noviembre de 2001, respectivamente.

Variaciones relación tipo de cambio Dólar EEUU / Yen Japonés - precio cierre Plata



Valor promedio de la relación asciende a 21,775, con una desviación de 4,799 y valores mínimos y máximos de 8,807 y 33,043 alcanzados 31 de marzo de 2006 y 23 de agosto de 2002, respectivamente.

2.2 MÉTODOS NO-LINEALES DE PREDICCIÓN “ REDES NEURONALES”

Este método no-lineal, fue inspirado en los estudios sobre el funcionamiento del cerebro y del sistema nervioso y ha sido empleado para resolver numerosos problemas económicos y financieros destacando, en gran medida, su aplicación en la predicción de series temporales.

Una Red Neuronal Artificial (RNA) es un modelo de procesamiento de la información inspirado en el modo en que el sistema nervioso biológico, como el cerebro, procesa la información. La pieza clave de este paradigma es la estructura original del sistema de procesamiento, compuesta por un gran número de elementos interconectados (neuronas), trabajando en armonía para resolver problemas específicos.

Las RNA funcionan al igual que en el caso de los seres humanos, a través de ejemplos. Una Red Neuronal se configura para una tarea específica, como el reconocimiento de patrones o la clasificación de datos, a través de un proceso de aprendizaje. Un sistema biológico aprende ajustando las conexiones sinápticas entre las neuronas. Una Red Neuronal simula este proceso.

En la literatura especializada se distinguen diferentes tipos de redes neuronales (Gatley, 1996), de las cuales, la feedforward multilayer network con un algoritmo de aprendizaje basado en la técnica backpropagation (Rumelhart y McClelland, 1986), corresponde a la red más popular en economía y finanzas (Wong, 1995; Yoano, Li and Tan, 1997). La principal ventaja de las redes neuronales consiste en su elevada capacidad para detectar y explotar la no-linealidad existente en los datos, aun en condiciones donde existen datos incompletos o cuando la presencia de ruido es importante. En nuestra aplicación empírica empleamos una feedforward network con 3 niveles compuesta por 5 capas cuya formulación estadística puede ser expresada como:

$$\hat{X}_t = \Phi \left(\beta_0 + \sum_{h=1}^H \beta_h * \Psi_h * \left(a_0 + \sum_{j=1}^J a_{hj} * X_t - 1 \right) \right)$$

donde \hat{X}_t es el output del modelo. Las funciones $(\cdot)\Psi$ y $(\cdot)\Phi$ se denominan función de transferencia del *nivel oculto* y del *nivel output*, respectivamente. La red presenta J inputs (retardados de la variable X_t), H unidades de proceso (neuronas) en el nivel oculto y 1 output. Inicialmente, los pesos a_{hj} y β_h son determinados aleatoriamente adoptando valores dentro de un determinado rango. Por medio de un proceso iterativo de aprendizaje basado en la técnica backpropagation, los valores de estos pesos son modificados de forma que la diferencia entre el valor real y el valor estimado (output de la red neuronal) sea mínimo. Ha sido demostrado por numerosos trabajos teóricos que este tipo de red, con un único nivel oculto y con un número suficientemente elevado de unidades, es capaz de aproximar cualquier función no-lineal con un determinado grado de precisión (Cybenko, 1989; White, 1990). Su empleo ha sido mayoritario en las aplicaciones empíricas sobre predicción de series financieras (Huetal., 1999).

Además de la complejidad existente en los datos, el éxito predictivo de una red depende en gran medida de la correcta determinación de su arquitectura. Por tanto, es preciso especificar de forma óptima el número de inputs (J) y el número de unidades de proceso en el nivel oculto (H) así como seleccionar la estructura de las funciones de transferencia. Por ejemplo, un número demasiado elevado de H puede originar problemas de overfitting y, en consecuencia, una ausencia de generalización. Por otro lado, con un número insuficiente de unidades de proceso la red puede perder capacidad predictiva al no explotar plenamente la no-linealidad existente en los datos.

En definitiva, siguiendo las recomendaciones existentes en la literatura y la evidencia empírica presentada por los papers de Parisi y Parisi se aplicó una *Red Neuronal feedforward backpropagation network con 3 layers compuesta por 7 neuronas cada una*, con el objetivo de predecir la dinámica del precio de la Plata respecto a otras 7 variables (incluida la Plata). En cuanto a su arquitectura, el número de inputs y de unidades de proceso fueron determinados por un proceso de prueba y error. Por su parte, se consideró una función de transferencia tangente hiperbólica para el nivel oculto y lineal para el nivel output.

Es posible distinguir al menos dos aplicaciones de las redes neuronales en las economías y finanzas: primero, la clasificación de agentes económicos tales como compañías, para obtener una estimación de la probabilidad de quiebra (por ejemplo, Odom&Sharda, 1990, Wilson&Sharda, 1994, Rahimian,1993), compañías de seguros (Brockett, Cooper, Golden&Pitaktong, 1994), la capacidad acreedora de clientes bancarios (Marose, 1990), tenedores de tarjetas de créditos (Jagielska&Jaworski,1996), segundo, la predicción de series de tiempo (Tang,1991, et,all). Con respecto a esto último, cabe señalar que el propósito de un modelo de predicción es capturar patrones de comportamiento en datos multivariados que distinguen varios resultados, cosa que es bien realizada por los modelos no paramétricos de redes neuronales (Gorr,1994). Los modelos de redes neuronales han sido desarrollados para predecir valores de índices bursátiles y de activos individuales, situándose la mayoría de las primeras investigaciones y aplicaciones en mercados establecidos en USA (Bosarge, 1993; Tsibouris&Zeidenberg, 1995, White, 1993), Gran Bretaña (Tsibouris&Zeidenberg, 1995) y Japón (Yoda, 1994). Dichos modelos han sido empleados para predecir el nivel o el signo de los retornos de índices bursátiles, entre otras aplicaciones relacionadas a la toma de decisiones en las áreas de finanzas e inversión (Hawley, Jonson&Raina, 1990; Refenes, 1995). En el estudio se empleo el modelo de redes neuronales que usan el algoritmo de aprendizaje hacia atrás, con el objeto de predecir el signo de la variación del precio de la Plata. El modelo se caracteriza porque la red se entrena a sí misma hasta que es capaz de predecir el resultado asociado con los datos de entrada (Dasgupta, Dispensa&Ghose, 1994).

La herramienta de construcción de las redes Neuronales fue desarrollada en Visual Basic ejecutable en Excel 2003 Profesional, sobre la cual se programó una base de datos que se divida en dos bloques: **(a)** conjunto de entrenamiento⁷ (conjunto de datos empleados para que la red aprenda el problema), **(b)** conjunto de prueba (conjunto de datos utilizados para evitar el sobre aprendizaje⁸ de la red). Es necesario mencionar que algunos autores igualmente dividen el período de datos en dos segmentos, donde el primero es utilizado para el entrenamiento de la red, mientras que el segundo es usado para determinar y validar su arquitectura y la especificación del modelo (Chen&Leung, 1998; Leung, Daouk&Chen,2000).

⁷ En una red Neuronal hay pares de inputs y outputs que son utilizados para entrenar la red. Pueden haber distintos inputs (variables explicativas) y multiples outputs (proyecciones de distintas variables).Entre los inputs y los outputs hay una capa (o multiples capas) de procesamiento que imitan el trabajo del cerebro humano. Luego, dado un nuevo conjunto de inputs, la red puede producir un nuevo outputs (proyección) sobre la base de lo que aprendió de los pares de inputs y outputs que le fueron provistos. El analista puede controlar algunos aspectos del proceso, tales como la tasa de aprendizaje y la precision deseada del output (Wilson&Keating).

⁸ El sobreajuste a los datos de la muestra o sobreaprendizaje de la red se produce cuando el sistema se ajusta demasiado a los datos de entrenamiento, aprendiendo incluso el ruido presente en ellos, por lo que crece el error ante patrones diferentes a los empleados en el entrenamiento y disminuye la precisión de la proyección.

Con el objetivo de obtener resultados en términos de capacidad predictiva se utilizó un Red Ward de 5 capas: una capa de entrada, una capa oculta y otra de salida. La capa de entrada (**Slab 1**) posee una neurona por cada i -ésima variable de entrada, su función es escalar el dominio que va de -1 a 1; (**Slab 2**) que mantiene una función Gaussiana que captura los datos que se encuentran cerca de las colas, la tercera capa, (**Slab 3**) Tangente Hiperbólica, que rescata el signo (+ o -); (**Slab 4**) cuarta capa cuya función es la Gaussiana Complemento y por ultimo un (**Slab 5**) que es una función Logística que reescala el rango desde -1 a 1 llevando los datos al rango original.

$$i_{pj}^h = f_j^h \left(Neta_{pj}^h \right), \dots, \text{conj} = 1, 2, \dots, n.$$

donde la expresión se utiliza para los Slabs ocultos respectivamente:

$$Neta_{nj}^h = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1i} \\ w_{21} & w_{22} & \dots & w_{2i} \\ \dots & \dots & \dots & \dots \\ w_{\left[\frac{j}{2}\right]_1} & w_{\left[\frac{j}{2}\right]_2} & \dots & w_{\left[\frac{j}{2}\right]_i} \end{bmatrix}^h \begin{matrix} * \\ \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ x_i \end{bmatrix} \end{matrix} \begin{matrix} \\ + \\ \begin{bmatrix} \theta_1 \\ \theta_2 \\ \cdot \\ \theta_{\left[\frac{j}{2}\right]_1} \end{bmatrix} \end{matrix} = \begin{bmatrix} Neta_1 \\ Neta_2 \\ \cdot \\ Neta_{\left[\frac{j}{2}\right]_1} \end{bmatrix}^h$$

La salida neta de la capa de salida queda expresada de acuerdo a la ecuación anterior, para los Slabs 2 y 3:

$$o_{pk} = f_k^0 \left(Neta_{pk}^0 \right)$$

donde:

$$Neta_{pk}^0 = \begin{bmatrix} w_{11}, w_{12}, \dots, w_{\left[\frac{j}{2}\right]_1} \end{bmatrix}^0 \begin{matrix} * \\ \begin{bmatrix} i_1 \\ i_2 \\ \cdot \\ i_{\left[\frac{j}{2}\right]_1} \end{bmatrix} \end{matrix} \begin{matrix} \\ + \\ \begin{bmatrix} \theta_1 \\ \theta_{\left[\frac{j}{2}\right]_1} \end{bmatrix} \end{matrix} = Neta_{\left[\frac{j}{2}\right]_1}^0$$

Por medio de un proceso iterativo de aprendizaje, los valores de los pesos son modificados de forma que la diferencia entre el valor real y el valor estimado (output de la red neuronal) sea mínimo. Es importante destacar que numerosos trabajos teóricos han demostrado que es capaz de aproximar cualquier función no lineal con un determinado grado de precisión (Cybenko, 1989; White, 1990), empleándose en aplicaciones empíricas sobre predicción de series financieras (Huetal., 1999)

2.3 MÉTODOS NO-LINEALES DE PREDICCIÓN “ ALGORITMOS GENÉTICOS”

En los últimos años se ha incorporado el análisis basado en las teorías darwinianas de selección natural y supervivencia. Estos procedimientos, denominados algoritmos genéticos o evolutivos, han sido desarrollados por Holland (1975) y difundidos por Goldberg (1989) y Koza (1992). Los algoritmos genéticos (AG) ya han demostrado su robustez para aproximar explícitamente la forma funcional en el análisis no-lineal de series temporales (Álvarez et al., 2001). El interés en este tipo de modelo ha crecido intensamente en los últimos años entre los investigadores de múltiples disciplinas incluida la economía (Koza, 1995; Szpiro, 1997; Beenstock and Szpiro, 2002) y las finanzas (Neely et al., 1997; Allen and Karjalainen, 1999; Fyfe et al., 1999; Kaboudan, 2000).

Un AG permite encontrar una solución óptima a partir de la evolución de una población inicial de soluciones alternativas. Supone un proceso de búsqueda eficiente a través de un amplio número de posibilidades con el objetivo de hallar una expresión matemática que, dentro del espacio de posibles soluciones, mejor describa la dinámica presentada por una serie temporal. Esta técnica presenta toda una serie de ventajas frente a las tradicionales técnicas no-lineales de predicción. En primer lugar, a diferencia de las redes neuronales y de los métodos basados en ocurrencias análogas, los AG obtienen explícitamente una ecuación matemática como aproximación a la verdadera dinámica de la serie. Además, se presenta como un método predictivo más flexible que las redes neuronales ya que no requieren la costosa especificación de una arquitectura previa.

Los algoritmos genéticos trabajan con códigos que representan a cada una de las posibles soluciones al problema. Por ello, es necesario establecer una codificación para todo el rango de soluciones, antes de comenzar a trabajar con el algoritmo. Al respecto, Davis (1994) señala que la codificación mas utilizada es la representación de las soluciones por medio de cadenas binarias (conjuntos ceros y unos).

En este trabajo se ha empleado un algoritmo genético para la predicción de series. Su funcionamiento se basa en la simulación en un ordenador del proceso evolutivo observado en la Naturaleza y que puede ser explicado por medio de una serie de pasos iterativos. Para empezar, en el primer paso se generan de forma aleatoria una población inicial de N ecuaciones matemáticas por medio de una sencilla combinación aleatoria de operadores y operandos de la siguiente forma:

$$S_j = \left((A \otimes B) \otimes (C \otimes D) \right) \quad \forall 1 \leq j \leq N$$

en donde A, B, C y D son los argumentos (genes operandos) y el símbolo \otimes representa a los operadores matemáticos (genes operadores). Los argumentos considerados pueden ser números reales pertenecientes a un intervalo (coeficientes de las ecuaciones) o bien valores retardados de la variable $(x_{t-\tau}, x_{t-2\tau}, \dots, x_{t-m\tau})$. Por su parte, los operadores matemáticos (\otimes) empleados serán la suma (+), resta (-), multiplicación (*) y división (\div) este último operador estará protegido para evitar cocientes entre 0 o números muy pequeños. También cabe la posibilidad de incluir otros operadores matemáticos (como el logaritmo o los trigonométricos, por ejemplo) pero a costa de incrementar la complejidad en el proceso de optimización funcional. Además, trabajos previos llevados a cabo por (Parisi y Parisi)⁹, (Álvarez et al., 2001) así como con otros AG (Szpiro, 1997; Yadavalli et al., 1999) han demostrado la posibilidad de describir dinámicas complejas por medio de expresiones matemáticas construidas simplemente con los operadores aritméticos.

En el segundo paso, una vez determinada la población inicial, comienza el proceso evolutivo seleccionando aquellas ecuaciones que presentan una mayor fortaleza según algún criterio de ajuste. Para el caso de predicción puntual se ha determinado como criterio de fortaleza el *Error Cuadrático Medio Normalizado* definido como se aprecia en la siguiente ecuación:

$$ECMN_j = \frac{1}{Var(x_t)} * \frac{\sum_{t=m+1}^M \left[x_t - \hat{x}_t \right]^2}{M}$$

⁹ Paper. Algoritmo Genético y Modelos Multivariados Recursivos en la Predicción de Índices Bursátiles de America del Norte: IPC, TSE, NASDAQ y DJI

Donde: $ECMN_j$ es el Error Cuadrático Medio Normalizado presentado por la ecuación j -ésima ($\forall 1 \leq j \leq N$), $Var(x_t)$ es la varianza de la serie temporal, \hat{x}_t el valor predicho y M el número total de observaciones destinadas a entrenar en el AG. Una generalización de esta expresión ha sido empleada como criterio de fortaleza en los AG de Szpiro (1997) y Yadavalli et al. (1999).

A continuación, todas las ecuaciones de la población inicial son clasificadas en orden decreciente según su $ECMN_j$. Las ecuaciones con valores muy elevados del ECMN son aniquiladas mientras que, por el contrario, las que poseen un valor más pequeño tienen una mayor probabilidad de supervivencia constituyendo la base de la siguiente generación.

Las ecuaciones supervivientes al proceso de selección son empleadas para generar a los individuos de una nueva generación (proceso de reproducción). Para ello se les aplican los denominados operadores genéticos: Clonación, Cruzamiento y Mutación. Con la clonación las mejores ecuaciones son copiadas exactamente a la siguiente generación.

Por su parte, con el operador cruzamiento se seleccionan parejas de ecuaciones con pequeños valores de $ECMN_j$ para intercambiar partes de sus argumentos y operadores matemáticos. Por último, la mutación implica el reemplazamiento aleatorio de algún operador o argumento en un porcentaje reducido de ecuaciones.

En definitiva, la nueva población generada a partir de la población inicial estará constituida por individuos clonados, mutados y cruzados. A partir de este momento el proceso evolucionará repitiendo los pasos de selección y reproducción de forma iterativa en busca de aquella expresión matemática que menor $ECMN$ presente. Después de un número determinado de generaciones del proceso evolutivo, la iteración terminará y el AG ofrecerá explícitamente como resultado una ecuación matemática que considera óptima para representar la verdadera dinámica de la serie temporal¹⁰.

¹⁰ Para nuestro problema específico de predicción, se configuró el Algoritmo Genético de tal manera que el número máximo de operadores eran 7 y cada población estaba constituida por una población máxima de 313 caracteres

3 COMENTARIOS Y REGULARIDADES EMPÍRICAS OBSERVADAS EN LOS DATOS

La base de datos empleada en el desarrollo del presente estudio, fue obtenida desde terminales Bloomberg bajo la modalidad semanal de precios de cierre de Plata, Oro, Cobre, Petróleo y los tipos de cambio UDS/YJP y GBP/USD. La muestra finalmente seleccionada abarca el periodo comprendido entre el 22 de septiembre de 1995 y el 31 de marzo de 2006, correspondiente a 550 observaciones. La elección de una periodicidad semanal se justifica para minimizar los sesgos originados por el efecto día de la semana (Lo and Mackinlay, 1988; LeBaron, 1993) y el efecto fin de semana (Zhang and Hu, 1998). Además, se asume que la periodicidad semanal contiene suficiente información como para capturar la dinámica mostrada por los precios de los diversos activos financieros (Yao and Tan, 2000).

Como es habitual en la mayoría de las aplicaciones empíricas, se considera para el análisis la diferencia logarítmica de los precios de los activos ya mencionados. Al contrario de un análisis univariante, con el Algoritmo Genético se permite buscar relaciones funcionales entre 2 o más series temporales, donde es el precio de la Plata el objeto de análisis, considerando su rendimiento. Esta transformación ha llegado a ser estándar en el análisis financiero ya que permite obtener una serie estacionaria, puede ser interpretada como una rentabilidad y, además, se presenta como una variable de mayor interés para los operadores financieros (Brooks, 1996).

Siguiendo las recomendaciones existentes en la literatura (Yao and Tan, 2000), se ha dividido la muestra disponible en 2 subconjuntos: Entrenamiento y Producción. El subconjunto de entrenamiento, compuesto por las 275 primeras observaciones, está reservado para el entrenamiento de la red neuronal y la evolución del algoritmo genético¹¹. El subconjunto de producción, integrado por las 275 siguientes observaciones, permite seleccionar la dimensión de incrustación para cada uno de los métodos y, en particular, determinar la arquitectura de la red neuronal.

¹¹ En el caso del Algoritmo Genético, el conjunto de entrenamiento se reserva para su evolución tanto en el caso de la predicción individual como en la fusión de predicciones.

Este criterio para evaluar la capacidad predictiva de los métodos considerados, ha sido empleado en multitud de trabajos dedicados a la predicción de series financieras (Elms, 1994; Yao, Tan and Poh, 1999; Yao and Tan, 2000; Tenti, 1996). Se basa en la comparación de los errores del método no-lineal respecto a los errores que se obtendrían si se considerase la media de la serie como predictor. De esta forma, un valor del ECMN mayor/igual/menor que uno implicaría una capacidad predictiva peor/igual/mejor que emplear la media como predictor. Por otro lado, para la predicción de un alza o baja se empleó como criterio el porcentaje de signos correctamente predichos.

4 METODOLOGÍA Y DATOS

La importancia de esta investigación radica en el hecho de que a través de la aplicación de redes neuronales y algoritmo genético, obtengamos modelos que cumplan las propiedades enunciadas en el apartado 1 (introducción), de manera de modelar el proceso de adaptación de expectativas de alza o baja en el precio de la Plata que establecen los agentes que integran el mercado y como éstas van cambiando con el transcurso del tiempo ante el surgimiento de nuevos eventos, cuestión que no es tomada en consideración por el funcionamiento estándar de la red neuronal y el algoritmo genético. Para ello, la red es constantemente reentrenada y probada ante datos que nunca ha visto. Mientras que los algoritmos genéticos, van mejorando y optimizando el modelo a medida que va transcurriendo la información.

Como fue enunciado anteriormente, el estudio se basó en 550 observaciones semanales correspondientes al periodo comprendido entre el 22 de septiembre de 1995 y el 31 de marzo del 2006. Los modelos utilizados consideran como variable o *input* de salida la primera diferencia del precio de cierre de la Plata, correspondiendo esta al período “*t*” ($\Delta\text{Plata } t$), el cual representa nuestro futuro inmediato, *one step a head*.

Con respecto a las variables de entrada consideradas para la estructura de los modelos, son las siguientes:

- Las variables consideradas para la estructura de los modelos predictivos de cambio de signo en el precio de la Plata, corresponden a los precios de cierre de Plata, Oro, Cobre, Petróleo y a los tipos de cambios USD/JPY (dólar EEUU - yen japonés) y GBP/USD (Libra Esterlina – dólar de EEUU). Las variables de entrada o *inputs* corresponden a la primera diferencia del precio de cierre de la **Plata** hasta con uno y cinco periodos ($\Delta\text{Plata } t-1$, $\Delta\text{Plata } t-5$) respectivamente – en tercer orden, la primera diferencia del precio del **Oro** con tres rezagos ($\Delta\text{Oro } t-3$) – en cuarto y quinto lugar, la primera diferencia del precio de cierre del **Cobre** con 3 rezagos ($\Delta\text{Cobre } t-3$) y el precio de cierre del **Petróleo** con 1 rezago ($\Delta\text{Petróleo } t-1$), respectivamente – culminando nuestra serie, con los tipos de cambios **USD/JPY** y **GBP/USD**, ambos con 2 periodos de rezagos ($\Delta\text{USD/JPY } t-2$) y ($\Delta\text{GBP/USD } t-2$) respectivamente.

La forma funcional del modelo para determinar cambio de signo en el precio de la Plata en las redes neuronales, se expresa de la siguiente manera:

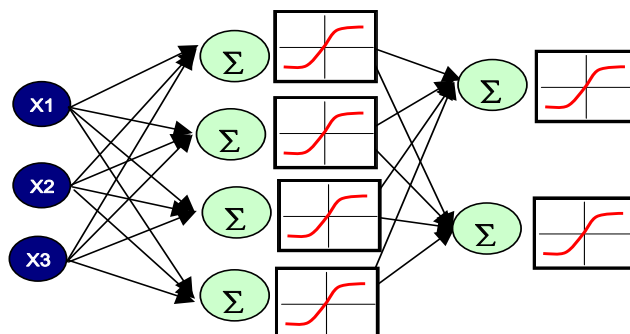
$$\Delta Plata_t = f(\Delta Plata_{t-1}, \dots, \Delta Plata_{t-5}, \Delta Oro_{t-3}, \dots, \Delta Cobre_{t-3}, \Delta Petroleo_{t-1}, \dots, \Delta USD/JPY_{t-2}, \Delta GBP/USD_{t-2})$$

4.1 BÚSQUEDA DE ARQUITECTURA ÓPTIMA

En base a los trabajos efectuados por Parisi y Parisi (2003), Gonzáles y Jiménez (2003) y Díaz (2005) donde probaron diferentes arquitecturas para las Redes Neuronales Artificiales y los estudios de Parisi, Parisi y Cornejo (2003) donde investigan modelos multivariados ARIMAX dinámicos a partir de arquitecturas basadas en algoritmos genéticos, la presente investigación se desarrolla bajo el mismo esquema de trabajo.

Para el caso de la **redes neuronales** se utilizó la **Red Ward**, la cual es capaz de detectar diferentes formas y patrones lineales y no lineales en los datos, a través de diferentes grupos de neuronas y diferentes funciones de activación en sus capas ocultas. Esta red posee 2 o más capas ocultas cada una con una función de activación distinta. Sea f la función teórica, dinámica y desconocida que explica la tendencia alcista o bajista del precio de la Plata, la cual cuenta con un vector de variables explicativas $X = (\Delta Plata_{t-1}, \dots, \Delta Plata_{t-5}, \Delta Oro_{t-3}, \dots, \Delta Cobre_{t-3}, \Delta Petroleo_{t-1}, \dots, \Delta USD/JPY_{t-2}, \Delta GBP/USD_{t-2})$ relacionadas a la variable dependiente $\Delta Plata_t = f(X)$. Así, f puede ser aproximada por una red neuronal multicapa.

ARQUITECTURA BASICA DE UNA RED WARD



La red neuronal Ward puede ser escrita como una función de n (número de neuronas de la capa oculta) variables explicativas x_i ($i = 1,2,3,\dots,7$) definidas en X , con funciones de activación Lineal, Gaussiana, Tangente Hiperbólica, Gaussiana Complementaria y Logística.

		Capa 1	Capa 2	Capa 3	Capa 4	Capa 5	
Red Ward	Arquitectura	Lineal	Gaussiana	Tangente Hiperbólica	Gaussiana Complemento	Logística	Total
	Neuronas	7	7	7	7	1	29

Con respecto a los objetivos del **algoritmo genético** es encontrar en primer orden, un mecanismo que mejore los resultados de una Red Ward (activando o desactivando variables al interior de los Slabs) y en segundo lugar, un modelo multivariado dinámico ARIMAX¹² que maximice el porcentaje de predicción de signo de las variaciones semanales del precio de la Plata. Los modelos multivariados dinámicos utilizados son modelos de series de tiempo que expresan el comportamiento de una variable en función de sus valores rezagados, de rezagos de variables exógenas y de rezagos de los residuos del modelo. Como variable exógena en esta metodología se ha considerado el precio del Oro. La variable elegida corresponde a un indicador de relación directa en el comportamiento del precio del activo en estudio a largo plazo. Los modelos se basaran en el *algoritmo genético* simple, el cual trabaja en cadenas binarias de largo fijo en representación de las posibles soluciones al problema.

Los modelos multivariados dinámicos usados para predecir el signo de variación de la Plata se presentan a continuación:

$$\text{ARIMA (3,1,1)} : \Delta \text{Plata } t = \alpha_1 \Delta \text{Plata } t-1 + \alpha_2 \Delta \text{Oro } t-1 + \alpha_3 \Delta \text{Plata } t-2 + \alpha_4 \Delta \text{Oro } t-2 + \alpha_5 \Delta \text{Plata } t-3 + \alpha_6 \varepsilon_t$$

$$\text{ARIMA (2,1,2)} : \Delta \text{Plata } t = \alpha_1 \Delta \text{Plata } t-1 + \alpha_2 \Delta \text{Oro } t-1 + \alpha_3 \Delta \text{Plata } t-2 + \alpha_4 \Delta \text{Oro } t-3 + \alpha_5 \Delta \text{Plata } t-4 + \varepsilon_t$$

Donde ε_t corresponde al término de error del modelo y los subíndices AR, MA y X representan el máximo orden de rezagos de las variables independientes.

¹² ARIMAX, Modelo Auto regresivo integrado de Medias Móviles

En este sentido se trabajó con máximo de 4 rezagos¹³ para los términos AR, MA y X, por lo que el máximo número de variables fue de 6 variables para ambos modelos y entre los cuales el *algoritmo genético* realizará la búsqueda de la mejor solución.

4.2 MODELOS MULTIVARIADOS ARIMAX

En 1970, los estadísticos británicos Box y Jenkins desarrollaron una metodología destinada a identificar, estimar y diagnosticar modelos dinámicos de series temporales en los que la variable tiempo desempeña un rol fundamental. Parte importante de esta nueva metodología se orientaba a eliminar la tediosa tarea de especificar un modelo, dejando que sea los propios datos temporales de la variable a estudiar indiquen la estructura probabilística subyacente, contraponiéndose a la forma estándar o tradicional de identificación y especificación de modelos. Estos nuevos modelos desarrollados se caracterizan porque una variable explicativa utiliza exclusivamente una variable exógena que es su propio pasado.

Dentro de los modelos univariantes más conocidos son los modelos ARIMA. Los modelos ARIMA o Autoregresivos Integrados de Medias Móviles son modelos econométricos en los cuales uno se aproxima a una nube de puntos a través de una regresión lineal, ajustando los ponderadores (W_i) de las variables explicativas para obtener el menor error cuadrático medio posible, en este caso, la regresión de las variables explicativas son los rezagos de la misma variable, los errores de predicción de los periodos anteriores y variables exógenas.

Se define a un modelo como autoregresivo si la variable endógena de un período t es explicada por las observaciones de ella misma correspondientes a períodos anteriores añadiéndose, como en los modelos estructurales, un término de error. En el caso de procesos estacionarios con distribución normal, la teoría estadística de los procesos estocásticos indica que, bajo determinadas condiciones previas, toda Y_t puede expresarse como una combinación lineal de sus valores pasados (parte sistemática) más un término de error (innovación).

¹³ Se trabajo con un máximo de 4 rezagos por cada variable debido a las restricciones que se tuvo en recursos computacionales, capacidad de procesamiento de datos y de tiempo.

La expresión genérica de un modelo autorregresivo, AR(p) sería la siguiente:

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_p Y_{t-p} + \varepsilon$$

Un modelo de los denominados de medias móviles es aquel que explica el valor de una determinada variable en un período t en función de un término independiente y una sucesión de errores correspondientes a períodos precedentes, ponderados convenientemente. Estos modelos se denotan normalmente con las siglas **MA**, seguidos, como en el caso de los modelos autorregresivos, del orden entre paréntesis. Así, un modelo con q términos de error MA(q) respondería a la siguiente expresión:

$$Y_t = \mu + \alpha_1 \varepsilon_{t-1} + \alpha_2 \varepsilon_{t-2} + \dots + \alpha_p \varepsilon_{t-p} + \varepsilon$$

4.3 EVALUACIÓN DE LA CAPACIDAD PREDICTIVA Y SIGNIFICANCIA ESTADÍSTICA

El proceso de evaluación empírica se realizó sobre la base de los datos del conjunto extra muestral (275 observaciones). Se procedió a medir el número de observaciones correctas (*hits*) del signo de la variación del precio de la Plata, para posteriormente, aplicar el *directional accuracy test* (DAT) de Pesaran & Timmermann (1992), con el objetivo de medir la significancia estadística de la capacidad predictiva de los modelos de proyección. Los autores demostraron que bajo la hipótesis nula de que ΔY_{n+i} y Y_{n+i} están independientemente distribuidos, dicho test sigue una distribución normal estándar.

La capacidad de los modelos de Algoritmo Genético para predecir el signo de las variaciones del precio de la Plata de acuerdo al tamaño de muestra y la técnica recursiva, fue contrastada con la capacidad predictiva de un tamaño de muestra similar para la técnica rolling. De esta forma se pretende encontrar si existen combinaciones óptimas de técnica y tamaño muestral que permitan obtener mejores rendimientos en términos de aciertos en el signo de la predicción.

En esta distribución se agregó el cálculo de la rentabilidad obtenida de seguir las recomendaciones de compra cuando el valor proyectado para el futuro próximo t indicaba un alza en el precio y una recomendación de venta en el caso contrario. El cálculo de las rentabilidades de la técnica empleada (Pesaran & Timmerman, 1992), consideró un monto de inversión de US\$100.000 de manera referencial, valorándose la estrategia como el valor de la cartera, es decir, la suma del inventario de Plata más el dinero disponible. Las acciones de compra fueron ejecutadas en la medida que la cartera contaba con disponible al momento de generarse la señal, por el contrario al ejercer una acción de venta, esta sólo se materializaba al registrar stock de plata disponible para su venta. Este ejercicio no consideró ventas cortas, costos de transacción ni de almacenaje del activo.

4.4 ANÁLISIS DE LOS RESULTADOS

Con el propósito de dar mayor validez y robustez a las variables utilizadas y por las cuales se sustentan los modelos encontrados, se efectuaron múltiples ejercicios por cada uno de los métodos en estudio ¹⁴, cuyos resultados y estadísticas se presentan en el punto (4.4.1).

1. Con respecto a las **Redes Neuronales**, después de analizar variadas combinaciones de modelos aleatorios, los mejores resultados se obtuvieron con una Red Neuronal Ward con 7 neuronas de entradas, 3 capas ocultas y 29 neuronas a través de un modelo aleatorio con 500,000 iteraciones, en que para el precio de la Plata, la mejor combinación de funciones de activación y escalamiento acertó 59,64% intra muestral y 66,91% extra muestral, con un acierto direccional de 5,03 estadísticamente significativo en un 5%. En dicho test, sobre un total de 275 observaciones (extra muestrales), la proyección de cambio de signos acertó 184 veces, lo que implica obtener un retorno equivalente al 450,61% sobre la inversión inicial en el evento de haber seguido las recomendaciones, cifra que triplica la rentabilidad esperada de una estrategia “*buy and hold*”, la que alcanzó un retorno de 149,13%.

¹⁴ Se ejecutó 50 Redes Ward Aleatorias y otras 50 fusionadas con Algoritmo Genético con un máximo de 500.000 iteraciones y un promedio de 200.000 (aproximado), proceso que demandó 15 horas en promedio por cada ejercicio. Se probó resultados con distintas variables, sobre las cuales, los mejores resultados en cuanto a número de aciertos intra y extra muestrales, estabilidad, significancia estadística (DAT) y beneficios al final del periodo, se definió que las variables más representativas y mejor predicen el comportamiento del precio de la Plata, correspondían a: (Δ Plata t-1 ,,,, Δ Plata t-5, Δ Oro t-3, ,,,, Δ Cobre t-3 , Δ Petroleo t-1, ,,,, Δ USD/JPYt-2, Δ GBP/USD t-2) y sobre las cuales se presentaron detalles en los puntos anteriores.

2. En relación a los **modelos multivariados dinámicos** construidos a partir de **Algoritmos Genéticos *rolling* y *recursivos***, se observa que estos modelos no alcanzan mayores PPS extra muestral en comparación con las redes neuronales supervisadas, obteniéndose en porcentaje de predicción de signo para el modelo 1¹⁵, equivalente al 61,45% (Rolling de 22) y un 54,91% para proceso recursivo, cuyo test de acierto direccional asciende a 2,71 y 1,96 significativos al 5%, respectivamente. En relación al modelo 2¹⁶, el mayor PPS se registró con un Rolling de 55, cuyo porcentaje de predicción de signo alcanza 57,45% con test de acierto direccional de 2,56 y una rentabilidad de 172,35%. Respecto al modelo 2, existen evidencias en algunos procesos de Rolling, que indican que el comportamiento de la serie representa una caminata aleatoria, de acuerdo al test de Pesaran & Timmermann, lo que da baja confiabilidad a este modelo respecto al anterior, que es significativo al 5 y 10%. En promedio, podemos señalar que las rentabilidades de los modelos multivariados 1 y 2 al seguir las recomendaciones, generan riquezas menores que en el evento de haber efectuado una estrategia “*buy and hold*” (143,19%), alcanzando sólo 138,71% y 135,65% respectivamente. Sin embargo, hay que notar que la eficiencia en porcentaje de predicción de signo intra muestral de estos modelos presenta un grado de estabilidad ya que los resultados tanto intra como extra muestral, tienden a ser parecidos. La capacidad de procesamiento en términos computacionales de este método, fue más eficiente que el de las redes neuronales, debido a la menor complejidad de su arquitectura, lo que se traduce en la utilización de menos recursos (ordenadores).
3. Respecto a la tercera técnica utilizada, **Redes Ward Aleatorias** fusionadas con **Algoritmos Genéticos**, el mejor porcentaje de predicción de signo intra y extra muestral obtenido alcanza a 61,82% y 62,55% respectivamente con un beneficio de 171,46% y acierto direccional de 3,36. En este caso, el retorno de la estrategia trading (171,46%) fue notablemente inferior respecto al mejor beneficio obtenido por el modelo de Red Ward Aleatorio sin fusión (450,61%), cuyo porcentaje de predicción de signo intra muestral fue de 59,64% y extra muestral de 66,91%, con un test de acierto direccional de 5,03. La explicación a esta diferencia, creemos se debe a que el calce de los aciertos proyectados, corresponden a días con bajas variaciones de precio.

¹⁵ ARIMA (3,1,1) : $\Delta \text{Plata } t = \alpha_1 \Delta \text{Plata } t-1 + \alpha_2 \Delta \text{Oro } t-1 + \alpha_3 \Delta \text{Plata } t-2 + \alpha_4 \Delta \text{Oro } t-2 + \alpha_5 \Delta \text{Plata } t-3 + \alpha_6 \varepsilon t-1$

¹⁶ ARIMA (2,1,2) : $\Delta \text{Plata } t = \alpha_1 \Delta \text{Plata } t-1 + \alpha_2 \Delta \text{Oro } t-1 + \alpha_3 \Delta \text{Plata } t-2 + \alpha_4 \Delta \text{Oro } t-2 + \alpha_5 \Delta \text{Plata } t-3 + \alpha_6 \varepsilon t-2$

En base a las pruebas efectuadas, podemos señalar que los modelos de Redes Ward Aleatorios fusionados con Algoritmos Genéticos, no alcanzan un mayor porcentaje de predicción de signo extra muestral en comparación con el modelo de Redes Ward Aleatorias (sin fusión), sin embargo, podemos concluir que la fusión:

- Ajusta la diferencia entre los resultados intra y extra muestrales.
- Aumenta el promedio del DAT disminuyendo su desviación estándar.
- Ajusta el beneficio de la estrategia de Trading (no significativo, disminuyendo su desviación estándar) mejoras que fueron medidas en 15 pruebas para cada Red Ward, aleatoria y aleatoria con algoritmo genético, lo que no implica restar validez a la fusión. Es necesario destacar que el proceso ejecutado con algoritmo genético, ayuda a limpiar los modelos de variables (alfas) que más que estabilizarlos, los desmejoran.

Es necesario mencionar que en función del problema planteado, en anteriores estudios se indica que menos inputs o neuronas, parecen ser preferibles a más, concluyendo que a mayor información se produce “ruido” en el modelo predictivo, sin embargo, las evidencias encontradas en los resultados obtenidos, satisfacen ampliamente las expectativas y objetivos enunciados en la sección de introducción.

4.4.1 RESUMEN DE LOS MEJORES RESULTADOS OBTENIDOS POR METODOLOGÍA.

RED WARD ALEATORIAS Y RED WARD ALEATORIAS CON ALGORITMO GENETICO.

Plata t f(Plata t-1 ; Plata t-5 ; Oro t-3 ; Cobre t-3 ; Petroleo t-1 ; USD/JPY t-2 ; GBP/USD t-2)							
Cuadro Resumen Porcentaje Predicción de Signo PPS Plata (t)				Retorno por Estrategia (**)			
Metodología Empleada	Intra Muestral	Extra Muestral	Extra-Intra	DAT (*)	Trading	Buy and Hold	Dif. Retorno
Red Ward Aleatorio	59,64%	66,91%	7,27%	5,03	450,61%	149,13%	301,48%
Red Ward Aleatorio	61,82%	60,00%	-1,82%	3,90	416,51%	149,13%	267,38%
Red Ward Aleatorio	61,82%	62,18%	0,36%	4,15	390,77%	149,13%	241,64%
Red Ward Aleatorio	61,45%	61,82%	0,37%	2,90	333,87%	149,13%	184,74%
Red Ward Aleatorio	62,55%	61,45%	-1,10%	3,41	315,21%	149,13%	166,08%
Red Ward Aleatorio	60,73%	63,27%	2,54%	4,40	301,63%	149,13%	152,50%
Red Ward Aleatorio	60,36%	61,45%	1,09%	2,78	299,52%	149,13%	150,39%
Red Ward Aleatorio	61,45%	61,45%	0,00%	3,36	274,04%	149,13%	124,91%
Red Ward Aleatorio	58,91%	62,18%	3,27%	3,51	252,99%	149,13%	103,86%
Red Ward Aleatorio	63,64%	59,64%	-4,00%	3,88	251,61%	149,13%	102,48%
Red Ward Aleatorio	60,73%	62,18%	1,45%	3,07	227,91%	149,13%	78,78%
Red Ward Aleatorio	62,91%	61,45%	-1,46%	2,11	205,15%	149,13%	56,02%
Red Ward Aleatorio	61,82%	61,82%	0,00%	2,93	198,05%	149,13%	48,92%
Red Ward Aleatorio	62,91%	60,36%	-2,55%	3,48	189,60%	149,13%	40,47%
Red Ward Aleatorio	64,73%	58,55%	-6,18%	2,01	185,85%	149,13%	36,72%
Promedios	61,70%	61,65%	-0,05%	3,39	286,22%	149,13%	137,09%
Desviación estándar	1,52%	1,88%	3,17%	0,81	83,69%	0,00%	83,69%
Máximos	64,73%	66,91%	7,27%	5,03	450,61%	149,13%	301,48%
Mínimos	58,91%	58,55%	-6,18%	2,01	185,85%	149,13%	36,72%

(*) Resultados con 95% de confianza, el modelo tiene capacidad predictiva (significancia estadística 5%, valor crítico 1,96).

(**) No considera ventas cortas - costos de transacción ni de almacenaje en el caso de tranzar en el London Bullion Market.

Plata t f(Plata t-1 ; Plata t-5 ; Oro t-3 ; Cobre t-3 ; Petroleo t-1 ; USD/JPY t-2 ; GBP/USD t-2)							
Cuadro Resumen Porcentaje Predicción de Signo PPS Plata (t)				Retorno por Estrategia (**)			
Metodología Empleada	Intra Muestral	Extra Muestral	Extra-Intra	DAT (*)	Trading	Buy and Hold	Dif. Retorno
Red Ward Aleatorio - Algoritmo Genético	60,73%	61,09%	0,36%	3,45	398,37%	149,13%	249,24%
Red Ward Aleatorio - Algoritmo Genético	60,45%	60,73%	0,28%	4,12	398,37%	149,13%	249,24%
Red Ward Aleatorio - Algoritmo Genético	60,36%	61,09%	0,73%	3,71	334,64%	149,13%	185,51%
Red Ward Aleatorio - Algoritmo Genético	62,18%	61,45%	-0,73%	3,95	326,08%	149,13%	176,95%
Red Ward Aleatorio - Algoritmo Genético	61,45%	60,73%	-0,72%	4,12	312,89%	149,13%	163,76%
Red Ward Aleatorio - Algoritmo Genético	62,18%	61,82%	-0,36%	3,47	303,51%	149,13%	154,38%
Red Ward Aleatorio - Algoritmo Genético	60,73%	61,82%	1,09%	2,93	284,05%	149,13%	134,92%
Red Ward Aleatorio - Algoritmo Genético	61,45%	61,09%	-0,36%	3,11	235,47%	149,13%	86,34%
Red Ward Aleatorio - Algoritmo Genético	61,09%	62,18%	1,09%	3,04	248,45%	149,13%	99,32%
Red Ward Aleatorio - Algoritmo Genético	61,09%	61,45%	0,36%	3,39	266,36%	149,13%	117,23%
Red Ward Aleatorio - Algoritmo Genético	63,27%	61,45%	-1,82%	3,42	250,12%	149,13%	100,99%
Red Ward Aleatorio - Algoritmo Genético	64,73%	60,73%	-4,00%	3,26	235,52%	149,13%	86,39%
Red Ward Aleatorio - Algoritmo Genético	61,45%	62,18%	0,73%	3,13	230,98%	149,13%	81,85%
Red Ward Aleatorio - Algoritmo Genético	59,64%	61,82%	2,18%	3,25	229,28%	149,13%	80,15%
Red Ward Aleatorio - Algoritmo Genético	61,82%	62,55%	0,73%	3,36	171,46%	149,13%	22,33%
Promedios	61,51%	61,48%	-0,03%	3,45	281,70%	149,13%	132,57%
Desviación estándar	1,25%	0,57%	1,46%	0,37	64,10%	0,00%	64,10%
Máximos	64,73%	62,55%	2,18%	4,12	398,37%	149,13%	249,24%
Mínimos	59,64%	60,73%	-4,00%	2,93	171,46%	149,13%	22,33%

(*) Resultados con 95% de confianza, el modelo tiene capacidad predictiva (significancia estadística 5%, valor crítico 1,96).

(**) No considera ventas cortas - costos de transacción ni de almacenaje en el caso de tranzar en el London Bullion Market.

MODELOS MULTIVARIADOS ROLLING Y RECURSIVOS (MODELO 1)

ARIMA (3,1,2) : $\Delta Plata t = \alpha 1 \Delta Plata t-1 + \alpha 2 \Delta Oro t-1 + \alpha 3 \Delta Plata t-2 + \alpha 4 \Delta Oro t-2 + \alpha 5 \Delta Plata t-3 + \varepsilon t-1$							
Porcentaje Predicción de Signo PPS Plata (t) y otros valores					Retorno por Estrategia (***)		
Metodología Empleada Modelo 1	Intra Muestral	Extra Muestral	Resultado Generaciones	DAT (*)	Trading	Buy and Hold	Dif. Retorno
Algoritmos Genéticos Recursivo	52,57%	52,00%	54,91%	1,96	6,40%	149,13%	-142,73%
Algoritmos Genéticos Rolling 10	47,45%	50,15%	51,36%	1,98	50,13%	149,13%	-99,00%
Algoritmos Genéticos Rolling 20	48,90%	50,91%	56,00%	1,96	154,71%	149,13%	5,58%
Algoritmos Genéticos Rolling 21	53,31%	57,09%	60,73%	2,35	166,66%	149,13%	17,53%
Algoritmos Genéticos Rolling 22	51,47%	54,18%	61,45%	2,71	192,08%	149,13%	42,95%
Algoritmos Genéticos Rolling 23	50,74%	49,09%	56,36%	2,05	111,53%	149,13%	-37,60%
Algoritmos Genéticos Rolling 24	48,16%	50,91%	56,36%	1,98	139,80%	149,13%	-9,33%
Algoritmos Genéticos Rolling 25	50,74%	54,18%	56,36%	2,09	148,17%	149,13%	-0,96%
Algoritmos Genéticos Rolling 26	52,57%	49,09%	61,45%	2,71	181,95%	149,13%	32,82%
Algoritmos Genéticos Rolling 27 (**)	49,63%	51,27%	55,64%	1,67	105,75%	149,13%	-43,38%
Algoritmos Genéticos Rolling 28	49,26%	52,00%	58,18%	2,50	115,44%	149,13%	-33,69%
Algoritmos Genéticos Rolling 29	51,47%	51,64%	56,36%	2,00	97,47%	149,13%	-51,66%
Algoritmos Genéticos Rolling 30	52,57%	52,00%	60,73%	2,35	173,08%	149,13%	23,95%
Algoritmos Genéticos Rolling 31	52,21%	54,18%	60,36%	2,17	153,22%	149,13%	4,09%
Algoritmos Genéticos Rolling 32 (**)	50,74%	51,64%	55,64%	1,89	128,57%	149,13%	-20,56%
Algoritmos Genéticos Rolling 33	52,94%	59,64%	59,64%	2,19	163,66%	149,13%	14,53%
Algoritmos Genéticos Rolling 34	46,69%	49,82%	56,73%	2,06	169,60%	149,13%	20,47%
Algoritmos Genéticos Rolling 35 (**)	48,09%	54,91%	56,36%	1,90	134,21%	149,13%	-14,92%
Algoritmos Genéticos Rolling 40	51,10%	54,18%	60,36%	2,30	160,36%	149,13%	11,23%
Algoritmos Genéticos Rolling 45	51,10%	56,00%	61,82%	2,88	178,22%	149,13%	29,09%
Algoritmos Genéticos Rolling 50	53,31%	57,82%	60,73%	2,14	155,81%	149,13%	6,68%
Algoritmos Genéticos Rolling 55	53,31%	57,82%	60,73%	2,14	155,81%	149,13%	6,68%
Algoritmos Genéticos Rolling 60	51,10%	49,45%	60,36%	2,49	151,69%	149,13%	2,56%
Algoritmos Genéticos Rolling 65	50,74%	52,36%	58,55%	2,93	168,57%	149,13%	19,44%
Algoritmos Genéticos Rolling 70	55,15%	56,00%	56,00%	2,58	104,74%	149,13%	-44,39%
Promedios	51,01%	53,13%	58,13%	2,24	138,71%	149,13%	-10,42%
Desviación estándar	2,08%	2,99%	2,70%	0,33	42,42%	0,00%	42,42%
Máximos	55,15%	59,64%	61,82%	2,93	192,08%	149,13%	42,95%
Mínimos	46,69%	49,09%	51,36%	1,67	6,40%	149,13%	-142,73%

(*) Resultados con 95% de confianza, el modelo tiene capacidad predictiva (significancia estadística 5%, valor crítico 1,96).
 (**) Resultados con 90% de confianza, el modelo tiene capacidad predictiva (significancia estadística 10%, valor crítico 1,645).
 (***) No considera ventas cortas - costos de transacción ni de almacenaje en el caso de tranzar en el London Bullion Market.

MODELOS MULTIVARIADOS ROLLING Y RECURSIVOS (MODELO 2)

ARIMA (2,1,2): $\Delta \text{Plata } t-2 = \alpha_1 \Delta \text{Plata } t-1 + \alpha_2 \Delta \text{Oro } t-1 + \alpha_3 \Delta \text{Plata } t-2 + \alpha_4 \Delta \text{Oro } t-3 + \alpha_5 \Delta \text{Plata } t-4 + \varepsilon t-2$							
Porcentaje Predicción de Signo PPS Plata (t) y otros valores					Retorno por Estrategia (****)		
Metodología Empleada Modelo 2	Intra Muestral	Extra Muestral	Resultado Generaciones	DAT (*)	Trading	Buy and Hold	Dif. Retorno
Algoritmos Genéticos Recursivo (***)	60,66%	53,09%	53,09%	0,89	36,97%	149,13%	-112,16%
Algoritmos Genéticos Rolling 10	46,15%	47,31%	48,39%	1,98	89,51%	149,13%	-59,62%
Algoritmos Genéticos Rolling 20	48,53%	44,36%	57,09%	2,30	170,67%	149,13%	21,54%
Algoritmos Genéticos Rolling 21	55,88%	53,45%	54,91%	2,37	107,55%	149,13%	-41,58%
Algoritmos Genéticos Rolling 22	49,26%	53,45%	54,18%	0,11	48,25%	149,13%	-100,88%
Algoritmos Genéticos Rolling 23 (***)	50,00%	51,27%	54,91%	1,22	88,45%	149,13%	-60,68%
Algoritmos Genéticos Rolling 24	50,74%	52,36%	56,36%	2,09	148,17%	149,13%	-0,96%
Algoritmos Genéticos Rolling 25 (**)	50,00%	55,27%	55,27%	1,71	147,65%	149,13%	-1,48%
Algoritmos Genéticos Rolling 26	46,26%	50,18%	54,55%	1,38	138,39%	149,13%	-10,74%
Algoritmos Genéticos Rolling 27	49,63%	55,64%	57,09%	2,44	165,95%	149,13%	16,82%
Algoritmos Genéticos Rolling 28	48,01%	53,14%	53,14%	2,11	149,60%	149,13%	0,47%
Algoritmos Genéticos Rolling 29	47,30%	52,56%	53,01%	2,04	138,00%	149,13%	-11,13%
Algoritmos Genéticos Rolling 30	50,00%	51,27%	55,27%	1,63	119,77%	149,13%	-29,36%
Algoritmos Genéticos Rolling 31	47,13%	49,56%	51,36%	1,97	139,23%	149,13%	-9,90%
Algoritmos Genéticos Rolling 32	48,85%	51,96%	52,18%	2,00	150,11%	149,13%	0,98%
Algoritmos Genéticos Rolling 33	50,00%	51,06%	53,25%	2,01	160,67%	149,13%	11,54%
Algoritmos Genéticos Rolling 34 (***)	50,05%	52,15%	53,25%	1,50	138,45%	149,13%	-10,68%
Algoritmos Genéticos Rolling 35 (***)	49,26%	54,55%	54,55%	1,38	141,77%	149,13%	-7,36%
Algoritmos Genéticos Rolling 40 (***)	53,68%	52,73%	53,09%	1,37	159,07%	149,13%	9,94%
Algoritmos Genéticos Rolling 45 (***)	53,68%	50,18%	53,09%	1,49	170,62%	149,13%	21,49%
Algoritmos Genéticos Rolling 50	50,74%	50,18%	56,00%	1,99	127,20%	149,13%	-21,93%
Algoritmos Genéticos Rolling 55	48,90%	48,73%	57,45%	2,56	172,35%	149,13%	23,22%
Algoritmos Genéticos Rolling 60	50,30%	49,05%	52,00%	2,30	160,15%	149,13%	11,02%
Algoritmos Genéticos Rolling 65	52,21%	50,55%	54,18%	2,21	154,39%	149,13%	5,26%
Algoritmos Genéticos Rolling 70	55,71%	49,82%	53,82%	1,90	168,28%	149,13%	19,15%
Promedios	50,52%	51,35%	54,06%	1,80	135,65%	149,13%	-13,48%
Desviación estándar	3,30%	2,52%	2,01%	0,55	36,24%	0,00%	36,24%
Máximos	60,66%	55,64%	57,45%	2,56	172,35%	149,13%	23,22%
Mínimos	46,15%	44,36%	48,39%	0,11	36,97%	149,13%	-112,16%

(*) Resultados con 95% de confianza, el modelo tiene capacidad predictiva (significancia estadística 5%, valor crítico 1,96).
(**) Resultados con 90% de confianza, el modelo tiene capacidad predictiva (significancia estadística 10%, valor crítico 1,645).
(***) No se tiene capacidad predictiva, presenta características de caminata aleatoria.
(****) No considera ventas cortas - costos de transacción ni de almacenaje en el caso de tranzar en el London Bullion Market.

5 CONCLUSIONES

Las redes neuronales y los algoritmos genético busca replicar la lógica del mercado, de esta forma los modelos no paramétricos si entregan tendencias en el porcentaje de predicción de signo.

De acuerdo a los resultados obtenidos, podemos concluir en primer lugar, que las variables definidas y con las cuales se sometió a prueba las Redes Neuronales Ward fusionadas con Algoritmo Genético (optimización aleatoria), con un proceso de reentrenamiento continuo y sin necesidad de levantar supuestos sobre el comportamiento de los retornos¹⁷, explican en gran medida los resultados en la predicción de signo en el precio de la Plata.

En segundo lugar, bajo las pruebas de acierto direccional (DAT) de Pesaran & Timmermann (1992), estos resultados son estadísticamente significativo con un nivel de confianza del 95% (5% de significancia), eliminado con ello la hipótesis de independencia entre las variables independientes y las proyectadas, o sea, "caminata aleatoria". Asimismo, se pudo constatar que las redes neuronales potenciadas con *algoritmos genéticos* más que ayudar a elevar o potenciar la eficiencia de predicción en términos de mejorar los porcentajes de predicción de signo extra muestral, logra ajustar los resultados en términos de reducir la divergencia entre ambos espacio muestrales, disminuyendo la volatilidad de los PPS, mejorando el DAT y manteniendo estables los retornos promedios, respecto los resultados de una Red Ward Aleatoria.

En tercer lugar, la rentabilidad promedio de las pruebas alcanzadas, de seguir las recomendaciones de compra o venta de una estrategia Trading, logra duplicar los beneficios respecto de una estrategia "Buy and Hold".

¹⁷ El mercado asimila en forma directa los factores exógenos (políticos, económicos, etc.), situación que se refleja instantáneamente en los precios de los activos.

6 BIBLIOGRAFÍA Y REFERENCIAS

Alvarez-Diaz, M. & A. Alvarez (2002) Forecasting exchange rates using genetic algorithms, *Applied Economic Letters*, forthcoming.

Baille R. & P.McMahon (1989) The foreign exchange market: Theory and evidence. Cambridge University Press.

Ball, R. & Kothari, S.P. (1989). Nonstationary expected returns: Implications for test of market efficiency and serial correlation in returns. *Journal of Financial Economics* 25, 51-74.

Blume, L., Easley D. & O'Hara M. (1994). Market statistics and technical analysis: The role of volume. *Journal of Finance* 49, 153-82.

Brock, W., Lakonishok J. & LeBaron B. (1992). Simple technical trading rules and the stochastic properties of stocks returns. *The Journal of Finance*, Vol. XLVII, N°5, 1731-1763 (December).

Chan, K.C. (1988). On the contrarian investment strategy. *Journal of Business* 61, 147-163.

Chopra, N., Lakonishok J. & Ritter J. R. (1992). Measuring abnormal returns: Do stocks overreact? *Journal of Financial Economics* 31, 235-268.

Conrad J. & Kaul G. (1988). Time-variation in expected returns. *Journal of Business* 61, 409-425.

Conrad J. & Kaul G. (1989). Mean reversion in short-horizon expected returns. *Review of Financial Studies* 2, 225-240.

Conrad J. & Kaul G. (1993). Long-term market overreaction or biases in computed return? *Journal of Finance* 48, 39-63.

DeBondt, W. F. M. & Thaler R. (1985). Does the stock market overreact? *Journal of Finance* 40, 793-805.

DeBondt, W. F. M. & Thaler R. (1987). Further evidence on investor overreaction and stock market seasonability. *Journal of Finance* 42, 557-581.

Diebold, F.X. & J.A Nason (1990) Nonparametric exchange rate prediction?, *Journal of International Economics*, 28, pp. 315-332

Diaz, D (2005) Prediciendo el Cambio de Signo del Precio del Oro: Modelos de Redes Neuronales *Rolling y Recursivos*. Tesis Facultad de Ciencias Económicas y Administrativas, Universidad de Chile.

Ferson, W. & Harvey, C. (1991). The variation of economic risk premiums. *Journal of Political Economy* 99, 385-415.

French, K. R. & Roll R. (1986). Stock return variances: The arrival of information and reaction of traders. *Journal of Financial Economics* 17, 5-26.

Friz, R (2005) Determinación de Optimo de *Rolling*, en Modelos ARIMAX. Tesis Facultad de Ciencias Económicas y Administrativas, Universidad de Chile.

Hu M.Y.,G.Zhang,C.X.Jiang & B.E Patuwo (1999) A cross – validation análisis of neural network out-of-sample performance in exchange rate forecasting. *Decision Sciences*, 30,1, 197-216.

Grundy, B. & McNichols M. (1989). Trade and the revelation of information through prices and direct disclosure. *Review of Financial Studies* 2, 495-526.

Goldberg D.E (1989) Genetic algorithms in search, optimization and machine learning. Reading, MA: Addison – Wesley.

González I. & Jiménez J.(2003) Redes neuronales aplicadas a la predicción del precio del tipo de cambio, *Rolling vs. Recursivo*. Tesis Facultad de Ciencias Económicas y Administrativas, Universidad de Chile.

Holland J.H (1975) Adaptation in natural and artificial systems, Ann Arbor. The University of Michigan Press.

Hsieh, D.A (1989) Testing for nolinear dependence in daily foreign exchange rates, *Journal of Business*, 62, pp. 329-368.

Inouye, G. (2005) Aplicación de redes neuronales y Algoritmos Genéticos para predecir la apreciación o depreciación de los tipos de cambio, evidencias para el Real Brasileiro y Nuevo Sol Peruano. Tesis Facultad de Ciencias Económicas y Administrativas, Universidad de Chile.

Jegadeesh, Narasimhan (1990). Evidence of predictable behavior of security returns. *The Journal of Finance*, Vol. 45(3), 881-898 (July).

Jensen, Michael C. & Benington, George A (May 1970). Random Walks and Technical Theories: Some Additional Evidence. *Journal of Finance*. Reprinted in *Security Evaluation and Portfolio Analysis*, E. Elton and M. Gruber, Editors (Prentice-Hall, 1972), and *Investment Management: Some Readings*, J. Lorie and R. Brealey, Editors (Praeger Publishers, 1972).

Kuan C.M. & T.Liu (1995) Forecasting exchange rates using feedforward and recurrent neural network. *Journal of Applied Econometrics*, 10, 4, 347 – 364.

Lehmann, B. (1990) Fads, martingales, and market efficiency. *Quarterly Journal of Economics* 105, 1-28.

Lisi F. & A. Medio (1997) Is a random walk the best exchange rate predictor?, *International Journal of Forecasting*, 13, 255-267

Lo, A. & MacKinley A. C. (1988). Stock market price do not follow random walk: Evidence from a simple specification test. *Review of Financial Studies* 1, 41-66.

Malkiel, B. (1981). A random walk down Wall Street, 2 ed. (Norton, New York).

Pesaran M.H. & Timmermann (1992) A simple no parametric test of predictive performance. *Journal of Business & Economic Statistics*, 10,4, 461- 465.

Parisi, F. & Vasquez, A. (2000). Simple technical trading rules of stock returns: evidence from 1987 to 1998 in Chile. *Emerging Markets Review*, Vol. 1, N°2, 152-164.

Parisi, A. (2002) "Evaluación de modelos de redes neuronales de predicción del signo de la variación del IPSA", *Estudios de Administración*, vol. 9, N° 1, Universidad de Chile, 2002, p. 76.

Parisi, Parisi (2003) Modelos de redes neuronales Rolling versus Recursivo: Precio del Oro, Facultad de Ciencias Económicas y Administrativas, Universidad de Chile.

Parisi A., Parisi F. & E.Cornejo (2003) "Algoritmos Genéticos: Predicción de Índices Bursátiles de América del Norte. Universidad de Chile

Poterba, J. M. & Summers L. H. (1988). Mean reversion in stock prices: Evidence and implications. *Journal of Financial Economics* 22, 27-59.

Shefrin, H. M. & Statman M. (1985). The disposition to ride winners too long and sell losers too soon: Theory and evidence. *Journal of Finance* 41, 774-790.

Urrutia, J. (1994). Time series properties of four Latin American equity markets: Argentina, Brazil, Chile and Mexico. *Estudios de Administración* 2, 1-8.

Vega, M. & Alvarez, J. (2001) Tipos de cambio flexibles y volatilidad: las regularidades empíricas de las observaciones diarias. Universidad de Navarra.

Vera, L. (2005) Predicción del cambio de signos en los tipos de cambio Peso chileno y Peso Mexicano. Tesis Facultad de Ciencias Económicas y Administrativas, Universidad de Chile.

Zarowin, P. (1990). Size, seasonality, and stock market overreaction. *Journal of Financial and Quantitative Analysis* 25, 113-125.

7 ANEXOS

7.1 OBTENCIÓN DE VARIABLES SOBRE LAS CUALES SE APLICÓ LA RED WARD ALEATORIA Y RED WARD ALEATORIA CON ALGORITMO GENÉTICO .

Después de numerosas pruebas y análisis que nos permitieran encontrar el mejor modelo o matriz con el cual aplicar la red neuronal, pasando por modelar medias móviles con 5, 10, 30 rezagos y regresiones lineales, ambas acciones ejecutadas inicialmente para un total de 14 variables (**Plata, Oro, Cobre, Petróleo WTI, DJComm, Dow Jones, Nikkei 225, S&P500, FTSE-100, USD/JPY, USD/SWF, GBP/USD, USD/DHK, USD/DNZ**), se definió que las variables dependientes más representativas con el metal en estudio en términos de correlación y que además sean correspondan a indicadores mundiales relevantes, correspondían a la **Plata, Oro, Cobre, Petróleo, USD/JPY, GBP/USD**, con las que a posterior se definieron los rezagos de manera estadística, tomando como criterio básico el mayor coeficiente de correlación múltiple de nos entregaba la regresión y que se describe a continuación:

Resumen

MODELO 27 (MODELO SELECCIONADO)

Estadísticas de la regresión	
Coefficiente de correlación múltiple	0,2088191
Coefficiente de determinación R ²	0,043605416
R ² ajustado	0,0311959
Error típico	0,03110954
Observaciones	550

ANÁLISIS DE VARIANZA

	Grados de libertad	Suma de cuadrados	F	Valor crítico de F
Regresión	7	0,023960195	0,003422885	3,536756163
Residuos	543	0,525517294	0,000967803	0,001000925
Total	550	0,549477489		

	Coefficientes	Error típico	Estadístico t	Probabilidad	Inferior 95%	Superior 95%	Inferior 95,0%	Superior 95,0%
Intercepción	0	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
Plata t-1	0,083907092	0,047473432	1,767453666	0,077713866	-0,0093471	0,177161283	-0,0093471	0,177161283
Plata t-5	-0,053147648	0,022460221	-2,366301199	0,018316885	-0,097267267	-0,009028029	-0,097267267	-0,009028029
Oro t-3	-0,043151095	0,044044175	-0,979723083	0,327659124	-0,129669043	0,043366853	-0,129669043	0,043366853
Cobre t-3	0,072087881	0,02770713	2,601780893	0,009527181	0,017661523	0,126514238	0,017661523	0,126514238
PETROLEO t-1	0,050155414	0,024889251	2,015143542	0,044381894	0,001264341	0,099046486	0,001264341	0,099046486
USD/JPY t-2	-0,038110353	0,061791347	-0,61675874	0,537652369	-0,159489868	0,083269161	-0,159489868	0,083269161
GBP/USD t-2	-0,198291626	0,08943931	-2,217052273	0,027032926	-0,373981273	-0,022601978	-0,373981273	-0,022601978

Resumen **MODELO 27B**

<i>Estadísticas de la regresión</i>	
Coefficiente de correlación múltiple	0,19052225
Coefficiente de determinación R²	0,036298728
R ² ajustado	0,023808474
Error típico	0,03122815
Observaciones	550

ANÁLISIS DE VARIANZA

	Grados de libertad	Suma de cuadrados	Media de los cuadrados	F	Valor crítico de F
Regresión	7	0,019945334	0,002849333	2,921801867	0,005217304
Residuos	543	0,529532155	0,000975197		
Total	550	0,549477489			

	Coefficientes	Error típico	Estadístico t	Probabilidad	Inferior 95%	Superior 95%	Inferior 95,0%	Superior 95,0%
Intercepción	0	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
Plata t-1	0,069057749	0,049347507	1,399417185	0,162258911	-0,027877772	0,16599327	-0,027877772	0,16599327
Plata t-4	-0,031565095	0,026304659	-1,199981126	0,230669956	-0,083236516	0,020106327	-0,083236516	0,020106327
Oro t-3	-0,055799556	0,044823973	-1,244859665	0,213720326	-0,143849296	0,032250184	-0,143849296	0,032250184
Cobre t-3	0,07063369	0,027907408	2,531001416	0,01165493	0,015813916	0,125453464	0,015813916	0,125453464
PETROLEO t-1	0,052913764	0,024954193	2,120435818	0,034422055	0,003895124	0,101932404	0,003895124	0,101932404
USD/JPY t-2	-0,038419692	0,06207106	-0,61896304	0,536200288	-0,160348659	0,083509275	-0,160348659	0,083509275
GBP/USD t-2	-0,194865139	0,089781364	-2,170440851	0,030406125	-0,371226698	-0,01850358	-0,371226698	-0,01850358

Resumen **MODELO 25**

<i>Estadísticas de la regresión</i>	
Coefficiente de correlación múltiple	0,153049298
Coefficiente de determinación R²	0,023424088
R ² ajustado	0,010791573
Error típico	0,031436055
Observaciones	550

ANÁLISIS DE VARIANZA

	Grados de libertad	Suma de cuadrados	Media de los cuadrados	F	Valor crítico de F
Regresión	7	0,012871009	0,001838716	1,860623336	0,073844785
Residuos	543	0,53660648	0,000988226		
Total	550	0,549477489			

	Coefficientes	Error típico	Estadístico t	Probabilidad	Inferior 95%	Superior 95%	Inferior 95,0%	Superior 95,0%
Intercepción	0	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
Plata t-1	0,087777232	0,06159797	1,425002037	0,154731121	-0,033222422	0,208776887	-0,033222422	0,208776887
Plata t-2	-0,027956022	0,04498072	-0,621511208	0,534524154	-0,116313667	0,060401624	-0,116313667	0,060401624
Oro t-3	0,033280678	0,064802426	0,513571475	0,60776053	-0,094013634	0,160574989	-0,094013634	0,160574989
Oro t-5	-0,094467984	0,047795897	-1,976487306	0,048605188	-0,188355608	-0,00058036	-0,188355608	-0,00058036
Petroleo t-3	0,004971795	0,016137131	0,308096604	0,758126974	-0,026727095	0,036670686	-0,026727095	0,036670686
GBP/USD t-2	-0,202344014	0,090164433	-2,244166657	0,025223714	-0,379458053	-0,025229975	-0,379458053	-0,025229975
USD/JPY t-3	-0,063706524	0,051634037	-1,233808693	0,217807961	-0,165133579	0,037720531	-0,165133579	0,037720531

7.2 ANEXO 1: RESULTADOS RED NEURONAL - ARQUITECTURA WARD - MODELO ALEATORIO

Mejor resultado en Porcentaje de Predicción de Signo
Mejor resultado en Rentabilidad con Estrategia Trading :

Acierto Intra Muestral	59,64%
Acierto Extra Muestral	66,91%

DAT: 5,03 Retorno Estrategia Trading: 450,61% Retorno Estrategia Buy and Hold: 149,13%

	Ponderadores Encontrados							
	SESGO	Plata t-1	Plata t-5	Oro t-3	Cobre t-3	PETROLEO t-1	USD/JPY t-2	GBP/USD t-2
Ponderador Slab 2	-0,343319585	0,975384124	-0,329862632	0,841745245	-0,797295914	0,179483013	-0,874674862	0,240834337
	-0,357224715	0,623574192	0,54479091	0,814185327	0,536746182	-0,19309635	0,023402349	0,257124541
	0,431649075	0,416196765	-0,108918984	0,729315233	-0,972438191	0,107183421	-0,928964882	-0,941451075
	0,100206976	-0,405266478	0,300571385	0,334227854	0,874404594	-0,049825256	-0,911180606	-0,282081504
	0,44300473	0,681554968	-0,026000553	-0,928779547	0,878726318	0,393824468	0,172752002	-0,979938871
	-0,037831083	-0,310569593	0,318682449	0,202983233	-0,079013623	0,429867258	0,148992085	-0,326082933
	-0,807154355	0,156994447	0,2651215	0,180907821	-0,881640501	-0,168703216	-0,411628222	-0,178111171

	Ponderadores Encontrados							
	SESGO	Plata t-1	Plata t-5	Oro t-3	Cobre t-3	PETROLEO t-1	USD/JPY t-2	GBP/USD t-2
Ponderador Slab 3	-0,362100869	-0,74207826	0,472065173	0,574720766	0,755298471	0,208937759	0,400382054	0,574810635
	0,578160155	-0,560785294	0,950278965	-0,887633639	0,972681951	-0,867906798	0,709994797	-0,718444608
	0,77313611	0,065426406	0,314218843	0,365915064	0,074499052	0,07784857	0,973457238	0,746191883
	-0,887536881	-0,901452909	0,510909558	-0,746366021	0,361962694	-0,741727329	-0,081440954	0,591048754
	-0,62759551	-0,362324672	0,032047635	0,739648534	-0,83751866	-0,848110697	-0,872500976	-0,409433892
	-0,300124867	-0,524433674	-0,040460057	-0,935569865	0,191009471	0,326671778	0,666181926	0,995344387
	0,365648372	-0,634955385	0,52582193	0,519825831	-0,36786397	-0,369396173	0,582841488	0,508903675

	Ponderadores Encontrados							
	SESGO	Neurona 1	Neurona 2	Neurona 3	Neurona 4	Neurona 5	Neurona 6	Neurona 7
Slab -2	0,388843831	0,681010174	0,118724523	-0,351779648	0,555366959	-0,8278926	0,883589689	-0,640264109
Slab -3	-0,926921686	0,230018996	0,533539514	-0,607608605	-0,077057833	0,807153336	0,896338702	-0,544684816

7.3 ANEXO 2: RESULTADOS RED NEURONAL POTENCIADA CON ALGORITMO GENETICO - ARQUITECTURA WARD – MODELO ALEATORIO

Mejor resultado en Porcentaje de Predicción de Signo:

Acierto Intra Muestral	61,82%
Acierto Extra Muestral	62,55%

DAT: 3,36

Retorno Estrategia Trading: 171,46%

Retorno Estrategia Buy and Hold: 149,13%

		Ponderadores Encontrados							
		SESGO	Plata t-1	Plata t-5	Oro t-3	Cobre t-3	PETROLEO t-1	USD/JPY t-2	GBP/USD t-2
Ponderador Slab 2		0,129142844	0,08037952	-0,170081125	0,055929842	-0,290366819	0,730128421	-0,986126931	-0,329935191
		0,955934236	0,575417602	-0,4010737	-0,522156406	0,324591702	0,237763265	-0,504321006	-0,513946311
		-0,842596945	-0,543448137	-0,781498503	-0,674146861	-0,373664508	0,663519616	0,848805427	0,540752953
		0,246862814	0,178392039	0,41086493	-0,472868536	0,38076564	-0,07799326	0,450849733	0,217882518
		-0,824651265	-0,418635279	-0,994418468	0,238884078	0,503186656	0,218803595	-0,707236316	0,65479612
		-0,254891279	-0,338614264	0,891968184	0,44218421	-0,886221974	0,836648497	-0,852453081	0,480941091
		0,102854554	-0,78159009	0,426384179	-0,05832761	-0,412799794	0,315878693	0,667300372	-0,020397088

		Ponderadores Encontrados							
		SESGO	Plata t-1	Plata t-5	Oro t-3	Cobre t-3	PETROLEO t-1	USD/JPY t-2	GBP/USD t-2
Ponderador Slab 3		0,35881708	-0,482096981	-0,251796025	-0,46610591	0,796509196	0,93947265	0,983549524	-0,137467044
		-0,534970933	-0,645864521	-0,612807358	0,041591822	0,895937724	-0,572955246	-0,570814358	0,454845981
		-0,630731466	0,940585737	-0,084818334	-0,578200561	-0,085058576	-0,937624153	0,015844476	0,03125411
		-0,431927047	0,047782665	-0,303790662	0,894566783	-0,036966067	-0,62108859	0,711608729	-0,868023071
		0,644997074	0,076322282	-0,01621822	-0,856484089	0,892416979	0,849805854	0,365943519	0,353958084
		0,917880501	-0,420912733	0,63870739	-0,832066888	0,69374326	-0,324792913	0,631459708	-0,011557354
		-0,238992314	-0,983130345	-0,90046652	0,940960355	-0,405701283	0,030349269	0,482829518	0,291348982

		Ponderadores Encontrados							
		SESGO	Neurona 1	Neurona 2	Neurona 3	Neurona 4	Neurona 5	Neurona 6	Neurona 7
Slab -2		0,985347241	0,182888605	0,6527883	-0,721947283	0,862351971	-0,47742198	-0,864596705	0,557230699
Slab -3		0,868129289	0,978646725	0,473120653	0,008922221	0,467126939	0,938922353	0,927053682	-0,720858825

	Decimal	Binario	Neurona 1	Neurona 2	Neurona 3	Neurona 4	Neurona 5	Neurona 6	Neurona 7
Slab 1	94	1011110	1	0	1	1	1	1	0
Slab 2	110	1101110	1	1	0	1	1	1	0
Slab 3	86	1010110	1	0	1	0	1	1	0
Slab 4 -a	99	1100011	1	1	0	0	0	1	1
Slab 4 -b	45	0101101	0	1	0	1	1	0	1

Mejor resultado en Rentabilidad con Estrategia Trading:

Acierto Intra Muestral	60,73%
Acierto Extra Muestral	61,09%

DAT: 3,45

Retorno Estrategia Trading: 398,37%

Retorno Estrategia Buy and Hold: 149,13%

		Ponderadores Encontrados							
		SESGO	Plata t-1	Plata t-5	Oro t-3	Cobre t-3	PETROLEO t-1	USD/JPY t-2	GBP/USD t-2
Ponderador Slab 2		-0,102378259	-0,340813802	-0,709690645	0,550826365	0,088379849	0,401148936	0,106357258	0,95728286
		0,742751725	0,434946664	0,033838474	0,750309063	-0,792039673	-0,198971976	0,318876896	0,112648416
		0,799221546	-0,857236662	-0,498600429	-0,332163592	0,244013969	0,883842562	0,640453586	0,317317643
		0,397144687	0,64942724	0,447582551	0,130885456	-0,151287286	0,630222908	-0,158161681	-0,883212047
		-0,138014795	0,930092498	0,861072614	-0,983204801	0,368305772	-0,446361744	0,703964373	0,056759936
		0,729855253	0,633112595	0,221444976	-0,766237591	-0,796731972	-0,282045617	0,452649721	-0,104438733
		-0,770425752	0,072398103	-0,555579013	0,081169667	-0,410049776	-0,676194395	-0,482497183	-0,182180336

		Ponderadores Encontrados							
		SESGO	Plata t-1	Plata t-5	Oro t-3	Cobre t-3	PETROLEO t-1	USD/JPY t-2	GBP/USD t-2
Ponderador Slab 3		-0,923715634	-0,103846098	0,550127776	-0,772454995	0,142151982	0,497267271	0,084524533	0,538097305
		-0,758465518	0,706622644	0,163642337	-0,445954695	0,701592826	0,765801971	-0,636189047	0,410021269
		0,742260973	0,068228572	0,495457243	0,308242372	-0,329013897	-0,822830476	-0,595446607	0,541525263
		-0,058369245	-0,850254583	0,072396603	-0,570309993	-0,591791349	0,439819371	-0,111306565	-0,719120357
		-0,084116605	0,516145053	-0,516782699	-0,90023206	-0,756405381	-0,234594212	0,472893182	0,706597417
		-0,552065145	-0,945911878	0,622100802	0,074630149	-0,634648649	-0,461725397	-0,182469626	0,388455888
		-0,744834387	-0,699168414	-0,110336458	0,808304691	0,783024484	0,506113501	0,963343906	-0,576842225

		Ponderadores Encontrados							
		SESGO	Neurona 1	Neurona 2	Neurona 3	Neurona 4	Neurona 5	Neurona 6	Neurona 7
Slab -2		-0,877560992	-0,743687268	0,987345495	-0,422540803	0,780623912	0,979351437	0,331091512	0,071347091
Slab -3		0,456514125	-0,86488524	-0,467198292	0,167674075	-0,821699064	0,313474334	0,590864289	-0,595862554

	Decimal	Binario	Neurona 1	Neurona 2	Neurona 3	Neurona 4	Neurona 5	Neurona 6	Neurona 7
Slab 1	30	0011110	0	0	1	1	1	1	0
Slab 2	124	1111100	1	1	1	1	1	0	0
Slab 3	102	1100110	1	1	0	0	1	1	0
Slab4 -a	106	1101010	1	1	0	1	0	1	0
Slab 4 -b	87	1010111	1	0	1	0	1	1	1

7.4 ANEXO 3: CARACTERÍSTICAS Y ANTECEDENTES GENERALES DE UNA RED NEURONAL ARTIFICIAL (RNA).

7.4.1 FUNCIONES DE ACTIVACIÓN O TRANSFERENCIA.

Al igual que las neuronas biológicas, las neuronas pertenecientes a una RNA pueden estar activadas (excitadas) o inactivadas (no excitadas), es decir, poseen un determinado “estado de activación”. De esta forma, la “función de activación o transferencia” será aquel elemento de la topología de una red que permite calcular el estado de actividad de una neurona, transformando la entrada neta (descontando o añadiendo el parámetro bias conocido como sesgo o constante, θ_k en un valor (estado de activación), cuyo rango normalmente va de (0 a 1) o de (-1 a 1), dependiendo de su estado de activación o inactividad. así, una neurona podrá estar, totalmente inactiva (0 ó -1), o totalmente activa (1).

Las funciones de activación más comunes y que se encuentran contenidas en las 5 capas de nuestra Red Ward, se presentan a continuación:

- **Función Lineal:** Esta función es especialmente útil para aquellos casos en los cuales la salida es una variable continua o, para aquellos casos en que se desea que la red aprenda los eventos menos frecuentes, es decir, de aquellos valores alejados de cero en la escala en cuestión. Ello se debe a que, a diferencia de otras funciones, la función lineal no se hace menos sensible al alejarse de cero.

En términos matemáticos su expresión es:

$$f(ak) = ak$$

- **Función Lineal a Tramos:** Esta función corresponde a una variación de la función lineal, siendo igual a esta última cuando el valor de la entrada neta se encuentra de un rango de valores $(-1/\beta$ a $1/\beta)$. Al estar fuera del rango la función se torna constante (-1 o 1). En términos matemáticos su expresión es:

$$f(ak) = \begin{cases} -1 & ak \leq -1/\beta \\ \beta ak & -1/\beta < ak < 1/\beta \\ 1 & ak \geq 1/\beta \end{cases}$$

En la figura se aprecia como la función pasa a ser la función lineal a partir del momento en que el valor de la entrada neta, considerando el parámetro bias, se encuentra entre el rango $-1/\beta$ hasta el valor $1/\beta$. Fuera de rango la función se torna constante con un valor de -1 desde, $-\infty$ hasta -1 y, con un valor de 1 , desde 1 hasta $+\infty$.

- **Función Gaussiana:** La principal diferencia entre la función gaussiana y las descritas anteriormente, consta en que ésta función no es monótona creciente, utilizándose para extraer información central (distribución normal, los valores más probables). Produce salidas en el rango $(0,1)$, quedando representada de la siguiente manera:

$$f(ak) = \exp(-ak^2)$$

- **Función Gaussiana Complemento o Gaussiana -1:** Al igual que en el caso de la función Gaussiana, esta función pretende enfatizar información importante en valores extremos de los datos, con la diferencia que para estos valores toma el valor de uno en vez de cero como en el caso de la Gaussiana. Esta función se describe matemáticamente de la siguiente manera:

$$f(ak) = 1 - \exp(-ak^2)$$

- **Función Tangente Hiperbólica:** Esta función posee las mismas propiedades que la función logística, sin embargo, el rango de salida permite respuestas simétricas entre los valores -1 y 1 , manteniéndose centrada en 0 . Esto puede ser de especial interés, ya que en muchos casos puede ser necesario atenuar los valores exageradamente altos o bajos, pero diferenciándolos de un nivel donde la información sea irrelevante, donde se tendrá como valores a aquellos cercanos a 0 .

La función en términos matemáticos, es la siguiente:

$$f(ak) = \frac{\exp(\beta ak) - \exp(-\beta ak)}{\exp(\beta ak) + \exp(-\beta ak)}$$

- **Función Logística Simétrica:** Esta función es similar a la función logística, sin embargo, las ventajas de utilizar esta función frente a una logística residen en que, la segunda sólo ofrece resultados en el rango positivo entre 0 y 1 , en cambio la logística simétrica da resultados entre -1 y 1 , por lo que amplía a números negativos, los posibles resultados.

Esta función se describe de la siguiente manera:

$$f(ak) = \frac{2}{1 + \exp(-\beta ak)} - 1$$

- **Función Logística:** Esta función (sigmoide), es la función de activación más utilizada en las aplicaciones financieras y económicas de Redes Neuronales Artificiales. Su salida siempre será continua, abarcando el rango entre 0 y 1 , utilizándose para concentrar el aprendizaje en valores no extremos donde deberían encontrarse la mayor parte de los casos. En términos matemáticos su expresión es:

$$f(ak) = \frac{1}{1 + \exp(-\beta ak)}$$

Donde β es la pendiente de la función sigmoidea.

7.4.2 ESTRUCTURA BÁSICA DE UNA RED NEURONAL ARTIFICIAL

La estructura de RNA consiste en la organización y disposición de las neuronas dentro de la misma. En este sentido, y tal como fue comentado, las “neuronas” se estructuran formando capas, es decir, se agrupan en conjuntos o grupos de neuronas, las que conforman las llamadas Redes Neuronales. Adicionalmente, las neuronas pueden formar “Slabs”, los cuales son un subconjunto de neuronas dentro de una determinada capa de la red neuronal, las que poseen un número de características similares, como por ejemplo, poseer una misma función de activación. Dentro de la terminología neuronal existen tres tipos de capas, a saber:

- **Capas de Entrada:** Es la capa que agrupa a las neuronas que reciben los datos d entrada al problema (Inputs o variables independientes). en esta capa no se realiza proceso alguno, sin embargo, los datos de la capa de entrada pueden ser escalados con alguna de las funciones de activación analizadas.
- **Capas Ocultas:** Es la capa que no tiene conexión directa con el entorno, pero en ellas es donde se procesan los datos del problema bajo análisis.
- **Capas de Salida:** Esta formada por las neuronas que proporcionan la respuesta al problema en cuestión. en los modelos más simples hay una sola salida.

7.4.3 CARACTERÍSTICAS GENERALES DE LAS REDES NEURONALES ARTIFICIALES

Las RNA se caracterizan por actuar de forma:

- **Paralela:** Las RNA cuentan con una gran cantidad de neuronas, cada una de ellas trabajando simultáneamente, y con una parte de un problema mayor.
- **Distribuida:** Las RNA cuentan con una gran cantidad de neuronas a través de las cuales distribuyen su memoria. esto los diferencia de los sistemas computacionales tradicionales, los que sólo cuentan con un procesador y memoria fija.
- **Adaptativa:** Las RNA poseen la capacidad de adaptarse al entorno modificando sus pesos sinápticos, aprendiendo de las experiencias, consiguiendo generalizar conceptos a partir de casos particulares, permitiéndole encontrar una solución aceptable al problema.

7.4.4 CLASIFICACIÓN DE LAS REDES NEURONALES ARTIFICIALES

1) Clasificación de acuerdo a su aprendizaje.

- **Supervisados:** Consiste en presentarle un conjunto de patrones a la red, junto con la salida deseada u objeto. De esta forma, la red es capaz de ajustar los pesos sinápticos de las neuronas hasta que su salida tiende a ser deseada, utilizando para ello información detallada del error que comete a cada paso. Así, la red es capaz de estimar relaciones entrada/salida sin necesidad de proponer una cierta forma funcional a priori.
- **No Supervisadas:** Se le presentan a la red una multitud de patrones, sin adjuntar la respuesta deseada. Así durante el proceso de aprendizaje, la red debe descubrir por si sola los rangos comunes, regularidades, correlaciones o patrones según similitud y, de esta forma, incorporarlas a su estructura interna de conexiones (pesos sinápticos). Se dice, por lo tanto, que las neuronas deben autoorganizarse en función de estímulos (señales o datos) procedentes del exterior.

Para obtener resultados la calidad de la red requiere de un cierto nivel de redundancia en las entradas procedente del exterior, en definitiva, de un número de patrones de aprendizaje suficiente.

- **Reforzados:** Se encuentran entre los 2 anteriores, de forma que, como aprendizaje supervisado, emplea información sobre el error cometido, pero solamente indicando lo bien o lo mal que está actuando, mientras que, como aprendizaje no supervisado, no suministra explícitamente la salida deseada.
- **Híbridos:** Combinación de los tipos básicos de aprendizaje, supervisado y no supervisado, normalmente en distintas capas de neuronas.

2) Clasificación según se arquitectura

- **Redes Monocapa:** Son aquellas compuestas por una única capa de neuronas que procesa la información del problema bajo análisis. Este tipo de redes se utilizan típicamente en tareas relacionadas con lo que se conoce como auto asociación, es decir, para aquellos casos en los cuales es necesario regenerar informaciones de entrada que se presenta distorsionada o incompleta.
- **Redes Milticapas (Layered Networks):** son aquellas RNA en las que las neuronas se organizan o estructuran en distintas capas. Su utilización abarca una amplia gama de ramas, debido a que este tipo de redes poseen la capacidad de ser “Aproximadores Universales o Funciones” .

3) Clasificación según flujo de datos o información

- **Unidireccionales (FeedForwards):** la información fluye en un solo sentido, desde las neuronas de entrada hacia las de salida.
- **Realimentadas (Feedback):** La información puede circular entre las capas en cualquier sentido, incluido el de salida-entrada. En este caso, se tiene un conjunto de neuronas que actúa como memoria de las características de los patrones procesados previamente.

4) Clasificación según tipo de resultado.

- **Pronóstico o estimación en Nivel:** La estimación realizada por la RNA es un valor, una cantidad, es decir, corresponde a una variable continua.
- **Clasificación:** La estimación realizada por la RNA es una probabilidad, entrega categorías, es decir, corresponde a una variable discreta.

7.4.5 VENTAJAS DE LAS REDES NEURONALES ARTIFICIALES.



- **Adaptabilidad:** Esta se basa en la capacidad que poseen las redes neuronales artificiales de aprender y realizar ciertas tareas basadas en un determinado entrenamiento o en una experiencia inicial. En este sentido, una red neuronal ya entrenada y operando puede realizar un aprendizaje continuo presentándole constantemente nuevos ejemplos. de esta forma, las reden neuronales pueden mejorar su capacidad de segregar entre entradas similares, adaptándose a los cambios del ambiente.
- **Autoorganización:** A diferencia de las tecnologías actuales , en donde es necesario que toda la información sea especificada explícitamente por el programador, las redes neuronales, durante el proceso de aprendizaje o entrenamiento, organizan y crean su propia representación de la información que recibe. Así, las redes neoronales son adecuadas para problemas en los cuales es difícil o imposible definir un modelo explícito para obtener una determinada solución. Esta es una ventaja clara, ya que existe un gran número de problemas de este tipo.
- **Tolerancia a Fallas:** En el caso de que una neurona o una porción completa de una red neuronal son destruidas, esta continuará funcionando adecuadamente. Cuando el daño llega a ser tan extensivo que el comportamiento de la red comienza a ser afectado, el efecto será una degradación gradual de su rendimiento, y no una falla completa inmediata.

La razón de lo anterior radica en que las redes neuronales son capaces de almacenar la información de una manera distribuida. el conocimiento o información no es codificada ni manejada en una sola neurona, sino que a través de diversas neuronas y capas.

- **Procesamiento rápido:** Gracias a la capacidad de paralelizar el procesamiento de datos mediante la división de tareas computacionales en pequeñas sub tareas (las cuales son repartidas sobre varios procesadores que operan simultáneamente), las redes neuronales aventajan a otras técnicas de procesamiento de datos, logrando uno de sus intereses más críticos: “La velocidad de procesamiento”.
- **Compatibilidad con las tecnologías existentes:** Como ocurre con la mayoría de las nuevas tecnologías, si los elementos necesarios para su uso fueran incompatibles y difíciles de integrar con otras tecnologías existentes y consolidadas, lo más probable, es que tendría muy poca oportunidad de éxito y aceptación. Con esto en mente, las redes neuronales son sumamente prácticas, ya que pueden ser fácilmente integrables a casi cualquier tipo de sistema.

7.4.6 DESVENTAJAS DE LAS REDES NEURONALES ARTIFICIALES.

- Las redes neuronales no permiten conocer a priori la topología con la que se va a solucionar del modo más eficiente un determinado problema. Si bien, existen métodos incipientes que tratan de solucionar de una manera formal el inconveniente anterior, aún son muy complejos, poco difundidos y no exactos, por lo que los métodos heurísticos cobran gran importancia.
- Una vez entrenada una red neuronal, se hace difícil interpretar su funcionamiento, los datos entran y se obtienen predicciones, pero no se revela las relaciones entre las variables (Efecto “Black Box”). aún más, no es fácil asegurar con que grado de acierto responderá ante casos nunca vistos (generalización).
- Ante un determinado problema de error de predicción, o sobre ajuste puede resultar dificultoso saber el por que del no funcionamiento adecuado de una red neuronal, ya que existen una multitud de factores que pueden influir en su rendimiento (Tasa de aprendizaje, momentum, número de capas ocultas, número de neuronas, escalamiento de datos, funciones de activación, entre otros).
- Los Modelos Neuronales necesitan una herramienta de procesamiento poderosa, lo cual se manifiesta principalmente en el proceso de aprendizaje, pero este inconveniente se contrasta con la facilidad de implementación en dispositivos de *hardware* específicos.

7.5 ANEXO 4: FUNCIONAMIENTO *ROLLING* Y *RECURSIVO* EN MODELOS DE PREDICCIÓN.

7.5.1 FUNCIONAMIENTO RECURSIVO

Esta metodología ha sido empleada anteriormente para medir el desempeño de modelos de redes neuronales que buscan predecir periodos de recesión en los Estados Unidos (Qi, 2001; Estrella & Michkin, 1998) y para proyectar el signo de la variación de índices bursátiles internacionales (Parisi, Parisi & Guerrero, 2003). El funcionamiento recursivo consiste en agregar mediante un algoritmo iterativo, nueva información a la ya estudiada por la red. Desde el punto de vista financiero, esto nos permite reflejar que es necesario que los agentes ajusten sus expectativas frente a la nueva información obtenida, pero sin dejar de considerar la totalidad de la información pasada, pues si razonan utilizando toda la información disponible sus predicciones serán más complejas y por lo tanto, probablemente más acertadas. De esta forma, si consideramos que inicialmente el conjunto extra muestral tiene “ m ” datos, al evaluar el funcionamiento de la red se consideró sólo la predicción del primer valor. Posteriormente, el dato analizado sale del conjunto extra muestral (quedando con “ $m-1$ ” datos) y para formar parte del conjunto de entrenamiento, por lo que la muestra de “ n ” datos que contiene los valores de entrada se incrementa a “ $n+1$ ”. Luego se realizó una nueva iteración, lo que implicó reconstruir los pesos del modelo para cada una de las “ $m-1$ ” proyecciones, permitiéndole a la red aprender del error cometido de la predicción y ajustar nuevamente los pesos estimados para los valores de entrada de las capas de salida y oculta, w_{jk} y w_{ij} , respectivamente. Este proceso se repitió hasta que en el conjunto extra muestral quedó sólo un dato, por lo que la red recalculó los pesos w_{jk} y w_{ij} hasta el momento en que la última observación (correspondiente al periodo “ $t-1$ ”) es utilizada para proyectar el valor que la variable de salida podría alcanzar en el momento “ t ”, el cual representa el futuro inmediato.

7.5.2 FUNCIONAMIENTO *ROLLING*

El funcionamiento Rolling pretende capturar el proceso de adaptación de expectativas, de manera tal que se le otorgue una mayor importancia a aquella información recientemente agregada y descartando la más antigua, simulando el hecho de que los agentes ajustan sus expectativas dando una mayor ponderación a la información más reciente. Específicamente, éste mantiene constante el tamaño de la muestra “ n ”, pero varía desde el punto de vista de su composición, pues para cada iteración, éste se va desplazando a través del conjunto total de datos, de manera que incluya la próxima observación y a la vez elimina la última ($-1+n+1=n$). Como en el caso recursivo, el proceso se repite “ $m-1$ ” veces, permitiéndole a la red recalcular sus pesos de acuerdo a lo aprendido.

7.6 TEST DE ACIERTO DIRECCIONAL DE PESARAN & TIMMERMANN (1992).

El **test de acierto direccional**¹⁸ se usa con el objetivo de medir la significancia estadística de la capacidad predictiva de los modelos analizados. El test de acierto direccional prueba la hipótesis nula de que las variaciones observadas están independientemente distribuidas de las variaciones proyectadas. Por ello, si se rechaza la hipótesis nula, se dice que existe evidencia estadística de que el modelo tiene capacidad para predecir la evolución futura de la variable observada.

Este test compara el signo de la proyección $\hat{\Delta P}_{i,t+1}^j$, con el del valor observado $\Delta P_{i,t+1}$ para cada i -ésima observación del conjunto extra muestral ($i=1,2,\dots,n$); donde el signo indica la dirección en que se moverá el mercado accionario: al alza, si es positivo, o a la baja, si es negativo. Si los signos coinciden, aumentan la efectividad de la predicción, y en caso de no existir coincidencia, aumenta el error de predicción del modelo (igual que la metodología usada para calcular PPS).

Luego, se calcula un ratio de éxito (SR) que se define como:

$$SR = n^{-1} \sum_{i=1}^m I_i \left[\Delta P_{i,t+1}, \hat{\Delta P}_{i,t+1}^j > 0 \right]$$

Donde $I_i[\cdot]$ es una función indicador que toma el valor de 1 cuando su argumento es cierto y 0 en otro caso. Además:

$$P = n^{-1} \sum_{i=1}^m I_i \left[\Delta P_{i,t+1} > 0 \right]$$

$$\hat{P} = n^{-1} \sum_{i=1}^m I_i \left[\hat{\Delta P}_{i,t+1}^j > 0 \right]$$

El ratio de éxito en el caso de independencia de $\hat{\Delta P}_{i,t+1}^j$ y $\Delta P_{i,t+1}$, SRI , está dada por:

$$SRI = P \cdot \hat{P} + (1 - P)(1 - \hat{P})$$

¹⁸ Directional Accuracy Test

Donde P corresponde al porcentaje de positividad de las observaciones y \hat{P} corresponde al porcentaje de positividad de las proyecciones. La varianza del ratio SRI se define Como:

$$VAR \left[\overline{SRI} \right] = n^{-2} \left[(2 \cdot \hat{P} - 1)^2 P(1 - P) + n(2 \cdot P - 1)^2 \hat{P}(1 - \hat{P}) + 4 \cdot P \cdot \hat{P}(1 - P)(1 - \hat{P}) \right]$$

Por su parte, la varianza del ratio de éxito (SR) se define como:

$$VAR \left[\overline{SR} \right] = n^{-1} \cdot SRI(1 - SRI)$$

Finalmente, el *test DA* de Pesaran & Timmermann (1992) está dado por:

$$DA = (VAR \left[\overline{SR} \right] - VAR \left[\overline{SRI} \right])^{-1/2} (SR - SRI)$$

Como dicho *test* sigue una distribución normal estándar, los valores de rechazo de la hipótesis nula serán el **-1,96** y el **1,96**¹⁹. Es decir; si el valor DA está entre los valores de rechazo, no se rechaza la hipótesis nula de que las variaciones observadas están independientemente distribuidas de las variaciones proyectadas.

De esto último se entiende que se busca rechazar la hipótesis nula; es decir, que el valor DA no se encuentre entre los valores críticos mencionados y que por ende si exista capacidad predictiva.

¹⁹ Significativo al 95%