

**UNIVERSIDAD MAYOR DE SAN ANDRÉS  
FACULTAD DE CIENCIAS PURAS Y NATURALES  
CARRERA DE INFORMÁTICA**



**TESIS DE GRADO**

**MODELO PREDICTIVO BASADO EN REDES NEURONALES  
RECURRENTES APLICADO AL CASO DE LA INFLACIÓN EN BOLIVIA**

**PARA OPTAR AL TÍTULO DE LICENCIATURA EN INFORMÁTICA  
MENCIÓN: INGENIERÍA EN SISTEMAS INFORMÁTICOS**

**POSTULANTE:** FLORES QUISPE JOSÉ LUÍS

**TUTOR:** FREDDY MIGUEL TOLEDO PAZ

**REVISOR:** JAVIER H. REYES

**La Paz – Bolivia  
2009**

# AGRADECIMIENTOS

A Dios.

A mis padres, por el Esfuerzo y el Apoyo.

A mis hermanos.

Al Lic. Javier H. Reyes.

# INDICE

1. INTRODUCCIÓN.....	2
2. PROBLEMÁTICA.....	3
3. HIPÓTESIS.....	4
4. OBJETIVOS.....	4
4.2 Objetivo general:.....	4
4.3 Objetivos específicos:.....	4
5. JUSTIFICACIÓN.....	5
Justificación Tecnológica.....	5
Justificación Social.....	6
Justificación Científica.....	6
6. LIMITES Y ALCANCES.....	6

## Capítulo I. MARCO TEÓRICO

7. SERIES DE TIEMPO.....	9
7.1 Generalidades.....	9
7.2 Algunos conceptos de las series de tiempo.....	11
7.2.1 Proceso estocástico.....	11
7.2.2 Proceso estacionario.....	12
7.2.3 Proceso puramente aleatorio.....	12
7.2.4 Caminata aleatoria.....	12
7.2.5 Procesos estocásticos integrados.....	12
7.2.6 Prueba de raíz unitaria.....	13
8. METODOLOGÍA BOX-JENKINS.....	14
8.1 ARIMA: Modelo Univariable.....	14
8.1.1 Proceso Autorregresivo (AR).....	14
8.1.2 Proceso de media móvil (MA).....	15
8.1.3 Proceso Autorregresivo y de media móvil (ARMA).....	15
8.2 Caracterización de las series de tiempo: función de Autocorrelación.....	16
8.2.1 Función de Autocorrelación.....	16
8.2.2 Función de Autocorrelación parcial.....	17
8.2.3 Suavizamiento de series de tiempo.....	18

9. REDES NEURONALES.....	19
9.1 Teoría y operación.....	20
9.2 Modelo General de una Neurona Artificial.....	22
9.3 Neurodinámica.....	24
9.4 Regla de propagación.....	25
9.5 Función de activación o función de transferencia.....	26
9.6 Función de Salida.....	26
9.7 Modelo estándar de una neurona artificial.....	27
9.8 Neuronas todo-nada.....	28
9.9 Neurona continua sigmoidea.....	28
9.10 Redes neuronales aplicadas a las series de tiempo en economía.....	29
9.11 Sistemas tradicionales.....	30
9.12 RNA como modelo no lineal.....	31
9.13 Modelos Paramétricos y No Paramétricos.....	32
9.14 Modelos Autorregresivos No Lineales.....	33
10. ANÁLISIS DE LA INFLACIÓN.....	34

**Capítulo II.  
METODOLOGÍAS PARA CONSTRUCCIÓN DEL MODELO**

11. METODOLOGÍA PARA PRONÓSTICOS MEDIANTE MODELOS ARIMA Y RNA.....	43
11.1 Metodología para predicciones mediante ARIMA.....	43
11.2 Metodología para predicciones mediante RNA.....	44

**Capítulo III  
DESARROLLO Y ANÁLISIS DE RESULTADOS**

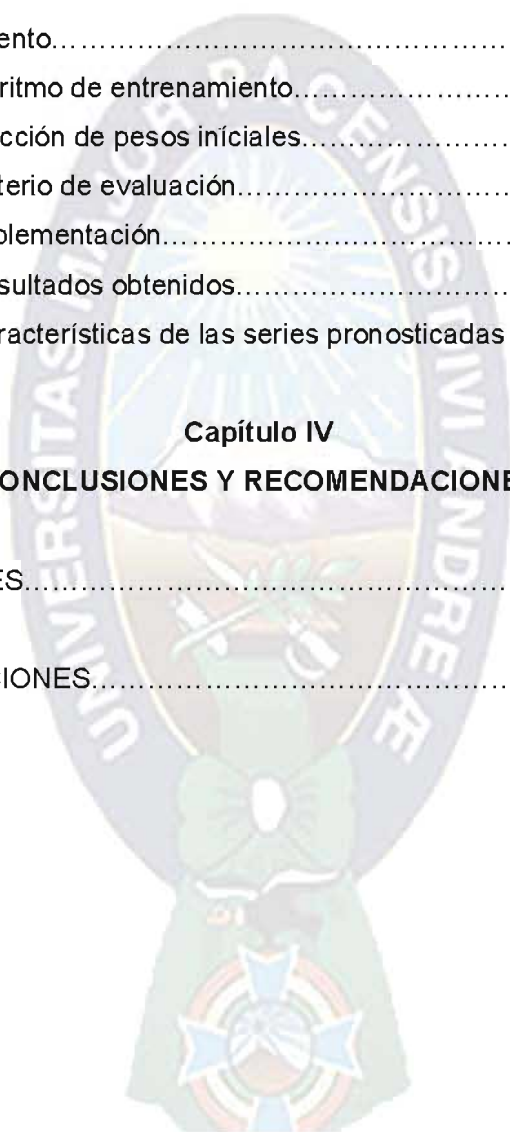
12. DATOS.....	53
12.1 Análisis exploratorio de la serie.....	54
12.2 Experimentación Modelo ARIMA con XLSTAT.....	55
12.3 Obtención del modelo RNA para la serie de de Variación del Índice de Precios al Consumidor.....	56
12.3.1 Selección de variables.....	56

12.3.2 Colección de datos.....	57
12.3.3 Pre-procesamiento de datos.....	58
12.3.4 Conjuntos de entrenamiento y prueba.....	59
12.3.5 Topología de la red.....	59
12.3.6 Determinación de las entradas.....	60
12.3.7 Determinación del número de neuronas en la capa oculta y entrenamiento.....	60
12.3.8 Algoritmo de entrenamiento.....	61
12.3.9 Selección de pesos iniciales.....	61
12.3.10 Criterio de evaluación.....	62
12.3.11 Implementación.....	62
12.3.12 Resultados obtenidos.....	65
12.3.13 Características de las series pronosticadas por ARIMA y RNA.....	65

#### Capítulo IV

### CONCLUSIONES Y RECOMENDACIONES

13. CONCLUSIONES.....	69
14. RECOMENDACIONES.....	70



## INDICE DE TABLAS

<b>Tabla 1</b> Patrones teóricos de FAC y FACP.....	17
<b>Tabla 2:</b> IPC, FUENTE: Instituto Nacional De Estadística.....	53
<b>Tabla 3:</b> Variación del IPC, FUENTE: Instituto Nacional De Estadística.....	53
<b>Tabla 4.</b> Pronóstico realizado mediante el modelo RNA, ARIMA, y El INE primer trimestre.....	65
<b>Tabla 5.</b> Comparación de estadísticos RNA y ARIMA.....	67

## INDICE DE FIGURAS

<b>Figura 1</b> Modelo de una neurona artificial Fuente (Rumelhart, 1986).....	23
<b>Figura 2</b> Curva de Phillips Inflación vs Desempleo.....	40
<b>Figura 3</b> Gráficas posibles del comportamiento de los errores en los conjuntos de entrenamiento y prueba de una RNA Fuente: Ivan Cruz Universidad Autónoma de México.....	50
<b>Figura 4</b> Bolivia: Índice de Precios al Consumidor por Año según Mes, Serie: 1998 – 2008 Fuente INE.....	54
<b>Figura 5</b> Gráfica de la Variación del Índice de Precios al Consumidor del 1 de enero de 1998 al 30 de diciembre del 2008 Fuente INE.....	55
<b>Figura 6</b> Gráfica de la Variación del Índice de Precios al Consumidor del 1 de enero de 1998 al 30 de diciembre del 2008 realizado por XLSTAT.....	56
<b>Figura 7.</b> Arquitectura de la RNA que proporciona los mejores resultados Fuente elaboración propia.....	62
<b>Figura 8.</b> Primera prueba con un error $e=0.012899$ en 300 épocas.....	63
<b>Figura 9.</b> Segunda prueba con un error $e=0.0049226$ en 300 épocas.....	63
<b>Figura 10.</b> Prueba X, con un error $e=0.004886$ en 300 épocas.....	64
<b>Figura 11.</b> Prueba $X+1$ , con un error $e=0.0029974$ en 300 épocas.....	64
<b>Figura 12.</b> Prueba $X+2$ , con un error $e=0.0012083$ en 89 épocas.....	64
<b>Figura 13.</b> Prueba que mejor ajusta con un error $e=0.0010339$ en 300 épocas.....	65
<b>Figura 14.</b> Grafica Obtenida por el modelo ARIMA con XLSTAT.....	66
<b>Figura 15.</b> Grafica Obtenida por el modelo RNA con Mat Lab.....	66

## RESUMEN

El objetivo de esta tesis es proponer un modelo que permita pronosticar la tendencia de la serie generada por la variación del IPC factor que mide proceso inflacionario, aplicando redes neuronales al análisis de series de tiempo, y comparar los pronósticos obtenidos de esta forma con los que se obtendría con un modelo Box-Jenkins tradicional, con el fin de conocer qué modelo ajusta mejor los datos dentro de muestra y qué modelo pronostica mejor los datos fuera de muestra.

El principal objetivo del trabajo de investigación es demostrar que el modelo propuesto basado en redes neuronales es capaz de obtener buenas aproximaciones en el pronóstico de series de tiempo, además de observar que, en este caso en particular, la aproximación resulta mejor que la generada por medio de la metodología Box-Jenkins. Así también, la tesis contiene una propuesta metodológica para la aplicación de las redes neuronales a las series de tiempo.

# 1. INTRODUCCIÓN

Mediante el uso de redes neuronales artificiales es posible realizar predicciones, estimaciones y/o proyecciones sobre el comportamiento en series de temporales, en la investigación presente se busca poder pronosticar la tendencia de la variación en la inflación. De acuerdo a la mayor parte de la literatura existente, se espera un mejor desempeño de las RNA en comparación con un modelo Box-Jenkins utilizado, ya que las redes neuronales son capaces de identificar relaciones no lineales, las cuales, como se ha demostrado en algunos estudios.

Las redes neuronales artificiales (RNA), son modelos computacionales que tratan de replicar, de manera simplificada, el complejo funcionamiento del cerebro humano. Su capacidad de aprendizaje a través de ensayos repetidos, las ha hecho muy populares en una amplia variedad de aplicaciones en todas las ciencias. Su reciente implementación en economía se debe al hecho de que en las series económicas y financieras es más probable que aparezcan relaciones no lineales que lineales (Granger, 1991) como las exigidas por los modelos econométricos tradicionales. Las RNA han demostrado ser una herramienta muy útil por su sorprendente habilidad para capturar relaciones no lineales entre variables. De hecho, pueden aproximar cualquier función no lineal si son correctamente especificadas (Tkacz y Hu, 1999).

La importancia de saber el comportamiento de determinada variable que influye en la variación de los niveles del índice de precios sugiere contar con un instrumento por el cual poder tomar decisiones adecuadas y en un tiempo pertinente para poder controlar de alguna manera los cambios que presenta el mercado.

Las técnicas de pronóstico o estimación de series de tiempo por los métodos estadísticos tradicionales permiten llevar a cabo esta labor como, por ejemplo, por medio de la metodología Box-Jenkins (ARIMA), la cual permite obtener buenas



aproximaciones en el caso de que dicho método sea bien aplicado, ya que ésta metodología tiene una parte de arte en el sentido de que el proceso es iterativo hasta alcanzar el mejor modelo de acuerdo a la experiencia del investigador.

Existen en general, dos grupos de modelos propuestos para pronosticar series de tiempo, por un lado, los modelos estadísticos tradicionales como los ARIMA, los modelos GARCH y mixtos, y por otro, los modelos basados en áreas nuevas como la inteligencia artificial, las Redes Neuronales Artificiales y los Algoritmos Genéticos.

## 2. PROBLEMÁTICA

La predicción de series de tiempo ha adquirido gran atención dentro de las investigaciones empíricas, tanto que los econométricos han empezado a dedicar mayores esfuerzos a este tipo de información.

Muchas propuestas de trabajos de investigación se han enfocado al aprendizaje de las RNA utilizando variables de entrada con las cuales se calcularía el valor de una opción, por ejemplo, por medio de la fórmula Black y Scholes (Pérez, 1998), por lo que el presente trabajo de investigación pretende determinar la factibilidad de emplear Redes Neuronales Artificiales en la predicción, estimación y/o proyección de la serie de tiempo de la variación del IPC en la inflación.

Formulación del problema:

¿Cuál es la capacidad predictiva de las Redes Neuronales Artificiales en comparación con las metodologías Box-Jenkins?

A partir de la cual se plantean las siguientes preguntas específicas:

- ¿El modelo propuesto basado en redes neuronales ajusta mejor los datos dentro de muestra que el modelo ARIMA usado como referencia?

- ¿El modelo propuesto basado en redes neuronales estima, pronostica o proyecta mejor los datos fuera de muestra que el modelo ARIMA usado como referencia?

### 3. HIPÓTESIS

#### **Hipótesis general:**

Un modelo basado en redes neuronales es capaz de ajustar y predecir la tendencia de los datos de la serie de tiempo de la variación del IPC en la inflación, así como también mejorar los resultados obtenidos por medio de un modelo Box-Jenkins (ARIMA) utilizado como referencia.

A partir de la cual se plantean las siguientes hipótesis específicas:

- El modelo propuesto basado en redes neuronales ajusta mejor los datos dentro de muestra que el modelo ARIMA usado como referencia.
- El modelo propuesto basado en redes neuronales pronostica mejor los datos fuera de muestra que el modelo ARIMA usado como referencia.

### 4. OBJETIVOS

#### **a) Objetivo general:**

Desarrollar un modelo basado en redes neuronales es capaz de ajustar o predecir la tendencia de la variación de la inflación, así como también realizar un análisis comparativo con un modelo Box-Jenkins (ARIMA) utilizado como referencia para determinar qué modelo proporciona los mejores resultados.

#### **b) Objetivos específicos:**

- Determinar si el modelo propuesto basado en redes neuronales ajusta mejor los datos dentro de muestra que el modelo ARIMA usado como referencia por medio de criterios estadísticos.

- Determinar si el modelo propuesto basado en redes neuronales pronostica mejor los datos fuera de muestra que el modelo ARIMA usado como referencia por medio de criterios estadísticos.

## 5. JUSTIFICACIÓN

La intención del presente trabajo de investigación es mostrar la capacidad de las RNA de realizar análisis de estimación, pronósticos en series de tiempo con variables económicas, como una herramienta alterna a las técnicas estadísticas tradicionales como la metodología ARIMA, además de explicar de manera clara y sencilla pero sin perder de vista la rigurosidad matemática que involucran este tipo de temas, el funcionamiento de una RNA (Red Neuronal Artificial).

Existen pocos estudios realizados por medio de redes neuronales relacionados al presente trabajo de investigación, y los que existen son aplicados sobre todo a mercados de países desarrollados, por lo que se pretende que este estudio pueda aplicarse a series de tiempo dentro la economía Bolivia, incrementando con ello el abanico de posibilidades para la mejor toma de decisiones en empresas, Instituciones e investigadores involucrados al estudio del comportamiento del índice de precios.

Por todo lo anterior, existe la necesidad de desarrollar un marco conceptual que permita fundamentar la propuesta metodológica del presente trabajo de investigación, y elaborar una investigación empírica con el fin de comprobar las hipótesis planteadas.

### a) Justificación Tecnológica.

Los modelos de predicción a la fecha sigue siendo estudiado por la econometría y la estadística, en la actualidad se cuenta con herramientas tecnológicas en hardware y software que pueden procesar grandes cantidades datos y realizar enormes cálculos, al

medio de software de econometría como ser el Econometric Eviews o XLSTAT 2009.

- Este trabajo de investigación no pretende establecer la superioridad de una metodología sobre otra, sino de evaluar, mediante ejemplos, la posibilidad de predecir utilizando Redes Neuronales Artificiales, y discutir la ventajas y desventajas de éstas comparadas con las metodologías de Box-Jenkins, utilizadas comúnmente en la econometría por medios estadísticos.





Capítulo I.  
**MARCO TEÓRICO**

## 7. SERIES DE TIEMPO

En este capítulo se establecen las bases teóricas para el desarrollo del trabajo de investigación, incluyendo también la revisión de la literatura relacionada con la aplicación de las RNA en la predicción de series de tiempo económicas.

### 7.1 Generalidades

De acuerdo a Bowerman y O'Connel, 1993. Una serie de tiempo es una secuencia cronológica de observaciones de una variable particular.

En términos formales, una serie de tiempo se define como una secuencia de vectores, dependientes del tiempo transcurrido  $t$ :

$$\bar{x}(t)$$

Donde  $t = 0, 1, 2, \dots$

Teóricamente,  $\bar{x}$  se puede ver como una función continua de la variable de tiempo  $t$ . Sin embargo, para cuestiones prácticas, el tiempo generalmente se ve en términos de pasos de tiempo discretos, conduciendo a una instancia de  $\bar{x}$  en cada punto de un intervalo de tiempo, generalmente de tamaño fijo (Dorffner, 1996). De acuerdo a Chatfield (1978), son varios los objetivos por los cuales se desea analizar una serie de tiempo:

- Descripción: Al tener una serie de tiempo, el primer paso en el análisis es graficar los datos y obtener medidas descriptivas simples de las propiedades principales de la serie.
- Explicación: Cuando las observaciones son tomadas sobre dos o más variables, es posible usar la variación en una serie para explicar la variación en las otras series.
- Predicción: Dada una serie de tiempo se intenta predecir los valores futuros de la serie. Este es el objetivo más frecuente en el análisis de series de tiempo.

- Control: Si una serie de tiempo se genera por mediciones de calidad de un proceso, el objetivo del análisis puede ser el control del proceso.

Una serie de tiempo generalmente se examina con la intención de descubrir patrones históricos que puedan ser útiles en la predicción. Para identificar esos patrones es conveniente pensar que una serie de tiempo consiste de varios componentes:

- I. Tendencia. Una serie de tiempo tiene tendencia cuando por largos periodos los valores crecen o disminuyen consistentemente. También puede definirse como cambios en la media.
- II. Ciclos. Se refiere a movimientos hacia arriba o hacia abajo alrededor del nivel de la tendencia. Estas fluctuaciones, medidas de pico a pico, pueden tener una duración larga.
- III. Variaciones estacionales. Son patrones históricos que ocurren y se repiten cada determinado tiempo, por ejemplo, anualmente. Estas variaciones son causadas típicamente por factores como el clima y las costumbres.
- IV. Fluctuaciones irregulares. Son movimientos irregulares en una serie de tiempo que no siguen un patrón regular, ni reconocible. Tales movimientos representan “lo que queda” en una serie de tiempo después de que la tendencia, ciclos y variaciones estacionales han sido explicados.

Pueden obtenerse valores futuros de una serie de tiempo observada mediante una gran cantidad de métodos de predicción. Estos métodos pueden clasificarse fundamentalmente en tres grupos:

- I. Subjetivos. Las predicciones se hacen sobre bases subjetivas usando el criterio, la intuición, el conocimiento en el área y otra información

relevante. Entre estos métodos se encuentran: Ajuste de una curva subjetiva, el método Delphi [1] y comparaciones tecnológicas en tiempo independiente.

- II. Univariados. Con este tipo de método se obtienen valores futuros de la serie basándose en el análisis de sus valores pasados, se intenta conseguir un patrón en estos datos, se asume que este patrón continuará en el futuro y se extrapola para conseguir predicciones. Son muchos los métodos que encajan en esta categoría, entre estos se encuentran: Extrapolación de curvas de tendencia, suavización exponencial, método de Holt-Winters[2] y método de Box-Jenkins (ARIMA). Este último es un método ampliamente utilizado.
- III. Causales o multivariados. Involucra la identificación de otras variables que están relacionadas con la variable a predecir. Una vez que esas variables han sido identificadas, se desarrolla un modelo estadístico que describe la relación entre esas variables y la variable a pronosticar. Entre estos métodos se pueden identificar: Regresión múltiple, modelos econométricos y métodos de Box y Jenkins (Modelo de función de transferencia).

## 7.2 Algunos conceptos de las series de tiempo

### 7.2.1 Proceso estocástico

Un proceso estocástico o aleatorio es una colección de variables aleatorias ordenadas en el tiempo. Si  $Y$  denota una variable aleatoria y continua, se denota como  $Y(t)$  pero si es discreta se expresa como  $Y_t$ .

---

1

El método Delphi pretende extraer y maximizar las ventajas que presentan los métodos basados en grupos de expertos y minimizar sus inconvenientes

2

El modelo de Holt-Winters es una ampliación perfeccionada del suavizamiento exponencial



### 7.2.2 Proceso estacionario

Un proceso estocástico es estacionario si su media y su varianza son constantes en el tiempo y si el valor de la covarianza entre dos periodos depende solamente de la distancia o rezago entre estos dos periodos de tiempo y no del tiempo en el cual se ha calculado la covarianza. En otras palabras, una serie de tiempo no estacionaria tendrá una media que varía con el tiempo o una varianza que cambia con el tiempo o ambas (Gujarati, 2004).

### 7.2.3 Proceso puramente aleatorio

Un proceso puramente aleatorio es un tipo especial de proceso estocástico que tiene una media igual a cero, una varianza constante  $\sigma^2$  y no está serialmente correlacionada.

### 7.2.4 Caminata aleatoria

Es una serie de tiempo estocástica en la que cada cambio sucesivo en  $Y_t$ , expresado como  $u_t$ , es extraído en forma independiente de una distribución de probabilidad con media 0 y varianza  $\sigma^2$  (Pindick y Rubinfeld, 2001). Por lo tanto,  $Y_t$  está determinada por:

$$Y_t = Y_{t-1} + u_t$$

### 7.2.5 Procesos estocásticos integrados

El modelo de caminata aleatoria es un caso específico de una clase más general de procesos estocásticos conocidos como procesos integrados. Se puede demostrar de manera sencilla que la primera diferencia del modelo de caminata aleatoria es estacionaria. Si expresamos la caminata aleatoria como:

$$(Y_t - Y_{t-1}) = \Delta Y_t = u_t$$

Resulta fácil probar que mientras  $Y_t$  no es estacionaria, su primera diferencia si lo es, ya que por definición  $u_t$  tiene media constante cero y una varianza de  $\sigma^2$ . Por tanto, se llama al modelo de caminata aleatoria sin variaciones proceso integrado de orden 1 y se denota como  $I(1)$ . De manera similar, si una serie de tiempo tiene que diferenciarse dos veces para hacerla estacionaria, se llamará a esa serie de tiempo integrada de orden 2 ó  $I(2)$ . En consecuencia, para convertir una serie no estacionaria en estacionaria se toman tantas diferencias como sea necesario, y se dice que el número de diferencias que se requiere para alcanzar la estacionariedad es el orden de integración de la serie.

#### 7.2.6 Prueba de raíz unitaria

Es una prueba para determinar la estacionariedad o no estacionariedad de una serie de tiempo.

Si se considera la siguiente serie:

$$Y_t = \rho Y_{t-1} + u_t$$

Donde:  $-1 \leq \rho \leq 1$

Si  $\rho = 1$ , como se demostró anteriormente, el proceso no es estacionario. De esta manera, si se realiza la regresión de  $Y_t$  sobre  $Y_{t-1}$  y se determina que la  $\rho$  es estadísticamente igual a 1, se concluiría que  $Y_t$  no es estacionaria.

Ahora, al diferenciar la serie se tiene la siguiente ecuación:

$$Y_t - Y_{t-1} = \rho Y_{t-1} - Y_{t-1} + u_t$$

O expresándola de otra manera:

$$\Delta Y_t = \delta Y_{t-1} + u_t$$

Donde  $\delta = \rho - 1$

Si  $\delta = 0$ , entonces  $\rho = 1$  y se dice que tiene una raíz unitaria, lo cual significa que la serie de tiempo bajo estudio no es estacionaria, pero al mismo tiempo  $\Delta Y_t = u_t$ . En consecuencia al tomar las primeras diferencias de la serie se obtiene una estacionaria.

Para saber si  $\delta = 0$ , la mayoría de los estudios aplican pruebas de raíces unitarias introducidas por Dickey y Fuller (1979) conocidas como *DF* o *DFA* (Dickey-Fuller aumentada), las cuales vienen integradas en la mayoría de los paquetes econométricos actuales.

## 8. METODOLOGÍA BOX-JENKINS

Box y Jenkins (1978) propusieron un modelo novedoso para la predicción de series de tiempo. Popularmente es conocida como metodología Box-Jenkins aunque técnicamente se conoce como metodología ARIMA, en la que se analizan las propiedades probabilísticas o estocásticas de las series de tiempo económicas bajo la filosofía de “permitir que la información hable por sí misma”.

### 8.1 ARIMA: Modelo Univariable

El proceso Auto Regresivo Integrado de Promedios Móvil: ARIMA (de sus siglas en inglés “Autoregressive Integrated Moving Average”) es denominado también Método Univariante de Box-Jenkins. La metodología de Box-Jenkins requiere que la serie sea estacionaria o sea estacionaria después de una o más diferenciaciones. Una serie es estacionaria si su media, su varianza, y su covarianza (en los distintos rezagos) permanecen constantes sin importar el momento en el cual se midan.

#### 8.1.1 Proceso Autorregresivo (AR)

En el proceso Autorregresivo de orden  $p$  la observación actual  $Y_t$  es generada por un promedio ponderado de observaciones pasadas que se

remontan  $p$  periodos, junto con una perturbación aleatoria en el periodo actual. Se denota este proceso como  $AR(p)$  y la ecuación que lo describe es:

$$Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \alpha_3 Y_{t-3} + \dots + \alpha_p Y_{t-p} + \delta + u_t$$

Donde  $\delta$  es un término constante que se relaciona con la media del proceso estocástico.

Es interesante observar que el modelo anterior solamente considera los valores actuales y anteriores de  $Y_t$ , no hay otros regresores.

### 8.1.2 Proceso de media móvil (MA)

En el proceso de media móvil de orden  $q$  cada observación  $Y_t$  es generada por un promedio ponderado de perturbaciones aleatorias que se remontan  $q$  periodos. Se denota este proceso como  $MA(q)$  y su ecuación respectiva es:

$$Y_t = \mu + \beta_0 u_t + \beta_1 u_{t-1} + \beta_2 u_{t-2} + \beta_3 u_{t-3} + \dots + \beta_q u_{t-q}$$

Donde  $\mu = E(Y_t)$

El modelo de media móvil supone que las perturbaciones aleatorias están distribuidas en forma independiente a lo largo del tiempo, es decir, son generadas por un proceso de ruido blanco.

### 8.1.3 Proceso Autorregresivo y de media móvil (ARMA)

Muchos procesos aleatorios estacionarios no pueden modelarse como autorregresivos puros o de medias móviles puras, ya que tienen cualidades de ambos tipos de procesos. De esta forma, se puede concebir un modelo mixto autorregresivo de media móvil de orden  $(p, q)$ . Se denota este proceso como  $ARMA(p, q)$  y se representa por medio de la ecuación:

$$Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_p Y_{t-p} + \delta + u_t + \beta_1 u_{t-1} + \beta_2 u_{t-2} + \dots + \beta_q u_{t-q}$$

Además, las series de tiempo económicas generalmente no son estacionarias, es decir, son integradas, razón por la cual si agregamos el

concepto de integración explicado anteriormente, de manera que se tenga que diferenciar una serie de tiempo  $d$  veces para hacerla estacionaria y luego aplicar a esta el modelo ARMA( $p, q$ ) se dice que la serie de tiempo original es del tipo ARIMA( $p, d, q$ ), es decir, es una serie de tiempo autorregresiva integrada de media móvil.

## 8.2 Caracterización de las series de tiempo: función de Autocorrelación

Existen dos funciones importantes en el análisis de series de tiempo a partir de las cuales es posible extraer los parámetros del proceso ARMA. De acuerdo con Box y Jenkins, éstas son la función de autocorrelación de la muestra y la función de autocorrelación parcial. La apariencia de cada una de estas funciones da una indicación de la naturaleza del proceso bajo estudio y, en teoría, la tarea de identificar el modelo y de obtener sus parámetros se puede llevar a cabo con base en ellas.

### 8.2.1 Función de Autocorrelación

Dada una muestra  $Y_0, Y_1, Y_2, \dots, Y_{n-1}$  de  $n$  observaciones, la función de Autocorrelación muestral (FAC) de la muestra al rezago  $k$ , denotada por  $\rho_k$ , se define como:

$$\hat{\rho}_k = \begin{cases} \frac{\gamma_k}{\gamma_0} & \text{para } k = 1, \dots, n-1 \\ 1 & \text{para } k = 0 \end{cases}$$

Donde:  $\gamma_0$  es la varianza de la muestra dada por:

$$\gamma_0 = \frac{1}{n} \sum_{t=1}^n (Y_t - \bar{Y})^2$$

$\gamma_k$  es la covarianza al rezago  $k$  definida como:

$$\gamma_k = \frac{1}{n} \sum_{t=1}^{n-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y})$$

La función de autocorrelación indica cuánta correlación existe entre datos individuales contiguos en la serie  $Y_t$ . Conforme el valor del retraso aumenta, el número de observaciones comprendidas en la autocovarianza disminuye hasta el elemento final. Al graficar  $\rho_k$  frente a  $k$ , la gráfica obtenida se conoce como correlograma. Cuando se grafica  $\rho_k$  se debe omitir el valor  $\rho_0$  el cual es invariablemente 1. Además, al interpretar la gráfica se debe tener cuidado de dar demasiada importancia a correlaciones en valores de retrasos que están significativamente altos en relación con el tamaño de la muestra.

### 8.2.2 Función de Autocorrelación parcial

La función de autocorrelación parcial (FACP) de la muestra  $\rho_{kk}$  en el retraso  $k$  es la correlación entre observaciones (series de tiempo) que están separadas  $k$  periodos de tiempo, manteniendo constantes las correlaciones en los rezagos intermedios (es decir rezagos menores de  $k$ ). En otras palabras, la autocorrelación parcial es la correlación entre  $Y_t$  y  $Y_{t-k}$  después de eliminar el efecto de las  $Y$  intermedias.

En la tabla 1. se dan algunos lineamientos generales acerca de los patrones típicos de las funciones de correlaciones muestral y parcial.

Tipo de Modelo	Patrón típico de FAC	Patrón Típico de FACP
AR(p)	Disminuye exponencialmente o con un patrón sinusoidal decreciente o ambos	Picos grandes a lo largo de los p rezagos
MA(p)	Picos grandes a lo largo de los q rezagos	Decrece exponencial mente
ARMA(p, q)	Decrece exponencial mente	Decrece exponencial mente

Tabla 1 Patrones teóricos de FAC y FACP

### 8.2.3 Suavizamiento de series de tiempo

Dependiendo de la serie de tiempo, algunas veces resulta necesario suavizarla para de esta manera eliminar algunas de las fluctuaciones a corto plazo más volátiles. Dicho suavizamiento puede realizarse antes de generar un pronóstico o también para contar con una serie de tiempo más fácil de analizar e interpretar. Una forma sencilla de suavizar una serie es tomar un promedio móvil de periodo  $n$ . La serie  $Y_t$  suavizada de esta manera queda expresada por:

$$\hat{Y}_t = \frac{1}{n}(Y_t + Y_{t+1} + \dots + Y_{t-n+1})$$

Mientras más grande sea la  $n$ , más suave resultará la serie.

Otro método de suavizamiento es el exponencial, el cual hace uso del modelo de promedio móvil con ponderaciones decrecientes en forma de progresión geométrica:

$$\hat{Y}_t = \alpha Y_t + \alpha(1-\alpha)Y_{t-1} + \alpha(1-\alpha)^2 Y_{t-2} + \dots$$

Donde el coeficiente  $\alpha$  puede tomar valores entre 0 y 1. La suma de los coeficientes siempre dará la unidad, ya que:

$$\alpha \sum_{s=0}^{\infty} (1-\alpha)^s = \frac{\alpha}{1-(1-\alpha)} = 1$$

Por lo tanto, se necesitan menos términos de la media móvil cuanto mayor sea el valor de  $\alpha$ . Para  $\alpha = 1$ , la media coincidirá con el valor del periodo. Con un  $\alpha$  cercano a cero, las ponderaciones de los valores serán todas muy pequeñas y el número de términos será mayor.

La elección del parámetro  $\alpha$  debe ajustarse para cada serie en particular, aunque muchos programas permiten el cálculo automático del valor óptimo de  $\alpha$ , en el sentido de seleccionar aquel que minimiza el error cuadrático medio.

Algunas veces se requiere hacer una mayor suavización de una serie pero no dar mucho peso a los datos individuales pasados. En este caso, el uso de la ecuación con un valor pequeño de  $\alpha$  no sería aceptable, pero en su lugar puede aplicarse un suavizamiento exponencial doble. Como su nombre lo indica, la serie suavizada inicialmente se suaviza otra vez.

La fórmula de suavizamiento exponencial simple también puede modificarse incorporando cambios promedio en la tendencia (incremento o disminución) a largo plazo de la serie. Esto fundamenta el método de suavizamiento exponencial lineal con dos parámetros o técnica de Holt-Winters.

## 9. REDES NEURONALES

Las Redes Neuronales Artificiales (RNA) forman parte de una tecnología de procesamiento de información que surge a raíz de los últimos desarrollos en software computacional aplicados al área de inteligencia artificial. Representan una herramienta muy innovadora para simular y analizar sistemas de relaciones complejas y cambiantes.

Las RNA son modelos matemáticos desarrollados con base en el funcionamiento de los sistemas nerviosos biológicos. De esta manera, se intenta contar con un sistema que imite el funcionamiento natural de las neuronas, es decir, que puedan “pensar” y “aprender” como el cerebro humano, el cual se puede ver como un conjunto interconectado de neuronas.

Sin embargo, cuando se pretende analizar un conjunto de datos con relaciones complejas y condiciones adicionales, la cantidad de procesamiento requerida se eleva considerablemente, lo cual ocasionó que los primeros intentos por utilizar este tipo de modelos fueran abandonados debido a la limitada capacidad de los sistemas de cómputo disponibles en la **época**.



Fue hasta los años 80 cuando se retoma el interés por su empleo de forma más generalizada. Esto debido a los avances en la potencia de cálculo de las nuevas computadoras, aunado a un mayor conocimiento del funcionamiento del propio cerebro humano. Así, los estudios realizados en este campo por Hopfield (1982, 1984, 1986), Rumelhart y McClelland (1986), Sejnowski y Rosemberg (1986), Feldman (1982), Grossberg (1986) y otros, permitieron el resurgimiento de esta nueva área del conocimiento.

Este nuevo interés se debió al desarrollo de algoritmos y nuevas topologías de red, nuevas técnicas de implementación de circuitos VLSI[1] (integración en escala muy grande) analógicos (Mead, 1986) y algunas demostraciones intrigantes por parte de Senowski y Hopfield, así como por la creciente fascinación por el funcionamiento del cerebro humano.

### **9.1 Teoría y operación**

Para poder entender el funcionamiento de un sistema nervioso biológico es necesario conocer las partes que lo componen, así como la interacción que existe entre sus propios elementos y su relación con el exterior. Ramón y Cajal (1899) proporciona un buen estudio e introducción en este sentido. También Berumen (1998) en su tesis de maestría realiza una completa síntesis de las redes neuronales y algunas aplicaciones a las finanzas. Lo importante del presente trabajo es el de tratar de imitar el extraordinario funcionamiento de los sistemas nerviosos por medio de redes neuronales artificiales, y aprovechar las ventajas que ofrecen en la resolución de problemas prácticos, gracias a la capacidad innata que tienen los sistemas biológicos de aprender a partir de ejemplos. Algo que la propia naturaleza ha logrado a lo largo de millones de años para adaptarse al medio que lo rodea y, resolver problemas que involucran el tratamiento de información masiva, imprecisa y distorsionada proveniente del entorno natural.

---

<sup>1</sup>

VLSI La integración en escala muy grande de sistemas de circuitos basados en transistores en circuitos integrados

Sin embargo, dada la complejidad de los sistemas neuronales biológicos, es necesario hacer una abstracción de las funciones y elementos que los componen para poder emplear modelos capaces de ser implementados con la tecnología disponible. De esta manera, y considerando a las neuronas desde un punto de vista práctico, éstas pueden ser vistas como pequeños procesadores de información. Las cuales haciendo una analogía con los sistemas informáticos tradicionales poseen un canal de entrada de información, las dendritas, un órgano de cómputo, el soma, y un canal de salida, el axón.

La conexión funcional entre dos neuronas para el transporte del impulso nervioso se denomina sinapsis. En relación a la sinapsis, se mencionan dos tipos de neuronas, presinápticas (las que envían señales) y postsinápticas (las que reciben señales). La intensidad de una sinapsis no es fija, sino que puede cambiar, dependiendo de determinadas condiciones. Esta variabilidad en la intensidad de la sinapsis constituye en buena medida el aprendizaje propiamente dicho, tal y como lo postuló Hebb[1]. Existen varias formas de modelar el sistema nervioso: por el establecimiento de nuevas conexiones, ruptura de otras, modelado de las intensidades sinápticas (plasticidad) o incluso mediante muerte neuronal. En especial, la modificación de las intensidades sinápticas es el mecanismo más empleado por los sistemas neuronales artificiales para llevar a cabo el aprendizaje.

Por medio de las redes neuronales artificiales se pretenden replicar tres características importantes de los sistemas nerviosos: el paralelismo de cálculo, la memoria distribuida y la adaptabilidad.

El paralelismo de cálculo está relacionado con la existencia de varios

---

<sup>1</sup> Hebb establece una teoría psicológica que siguió de cerca los avances en la investigación neurofisiológica. En su libro presenta una teoría de la conducta basada en la fisiología del sistema nervioso e intenta encontrar algo en común en las concepciones psicológicas y neurológicas

elementos de cálculo capaces de procesar información de manera simultánea o en paralelo, permitiendo al sistema un manejo de información mucho más rápido que un sistema que ejecuta instrucciones de manera secuencial.

Por último, la adaptabilidad se da por medio de la modificación de las sinapsis existentes en la red y a su capacidad de “aprender” por medio de ejemplos o patrones también llamada experiencia, para poder realizar generalizaciones a partir de ejemplos específicos y explotar esa característica esencial y hasta distintiva del sistema nervioso humano. De acuerdo a lo anterior, se puede decir entonces que las redes neuronales son sistemas paralelos, distribuidos y adaptables.

En un sistema neuronal biológico, las neuronas son los elementos básicos. Millones de neuronas se organizan en capas para formar sistemas con determinada funcionalidad, lo que en conjunto constituyen el sistema nervioso. En el caso de un sistema neuronal artificial se puede intentar copiar dicha estructura, en la cual, la neurona artificial se organiza en capas, varias capas forman una red y la red, con sus interfaces de entrada y salida constituyen un sistema neuronal completo.

## 9.2 Modelo General de una Neurona Artificial

El funcionamiento de una neurona o procesador elemental se puede modelar por medio de una función (Rumelhart y McClelland). Por lo que la operación de una neurona  $i$  se puede expresar de la siguiente forma:

$$y_i(t) = F_i\left(f_i\left[a_i(t-1), \sigma_i(W_{ij}, X_j(t))\right]\right)$$

Donde:  $X_j(t)$  representa el conjunto de entradas,  $W_{ij}$  son los pesos sinápticos de la neurona  $i$ , que representan la intensidad de la relación entre la neurona presináptica  $j$  y la neurona postsináptica  $i$ ,  $\sigma_i[W_{ij}, X_j(t)]$  es la regla de

propagación que proporciona el valor del potencial postsináptico  $h_i(t) = \sigma[W_{ij}, X_j(t)]$  de la neurona  $i$  en función de sus pesos y entradas,  $f_i[a_i(t-1), h_i(t)]$  es la función de activación que proporciona el estado de activación actual  $a_i(t) = f_i[a_i(t-1), h_i(t)]$  de la neurona  $i$ , en función de su estado anterior  $a_i(t-1)$  y de su potencial postsináptico actual.

$F_i[a_i(t)]$  Es la función de salida que proporciona la salida actual  $y_i(t) = F_i[a_i(t)]$  de la neurona  $i$  en función de su estado de activación. De acuerdo a Nelson e Illingworth (1991), existe una infinidad de formas de organizar una red neuronal aunque, tal vez sólo una docena de modelos son de uso común. Una red neuronal se puede describir en términos de su *neurodinámica* y *arquitectura*.

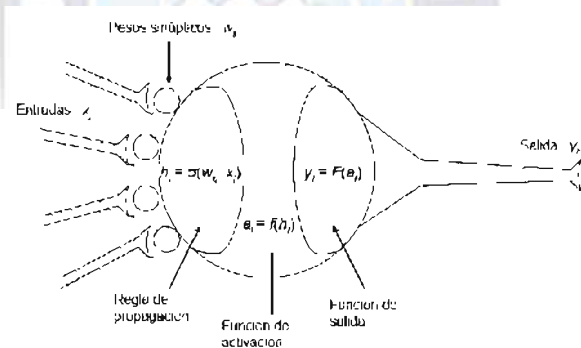


Figura 1 Modelo de una neurona artificial Fuente (Rumelhart, 1986)

La neurodinámica se refiere a las propiedades de una neurona artificial individual que consiste de lo siguiente:

- Combinación de entradas
- Producción de salidas
- Tipo de funciones de transferencia (o activación)
- Esquemas de pesos. inicialización de pesos, y algoritmos de entrenamiento.

Estas propiedades también pueden ser aplicadas a la red en su conjunto de un sistema básico. La arquitectura de red o topología define la estructura de la red e incluye las siguientes características básicas:

- I. Tipos de interconexiones entre neuronas
- II. Número de neuronas
- III. Número de capas

### 9.3 Neurodinámica

**Entradas.-** Las variables de entrada y salida pueden ser tanto analógicas como digitales, dependiendo del modelo como de las aplicaciones. La capa de entrada de una red neuronal comúnmente funciona como un regulador de las entradas, transfiriendo datos a la siguiente capa. Se puede requerir realizar un pre procesamiento para que la red sólo maneje datos numéricos. Esto puede involucrar el escalamiento de los datos de entrada y convertir o codificar los datos de entrada a una forma numérica que pueda ser usada por la red.

**Salidas.-** La capa de salida de una red neuronal funciona de una forma similar a la capa de entrada excepto que transfiere la información de la red al exterior. Frecuentemente se utiliza un postprocesamiento para convertir la información a una forma comprensible y útil hacia fuera de la red. Dependiendo del tipo de salida, las neuronas reciben nombres específicos. Por ejemplo, las neuronas estándar cuya salida sólo puede tomar el valor de 1 ó 0 se conocen como tipo McCulloch-Pitts<sup>1</sup>, mientras que las que toman a su salida los valores de -1 y +1 se llaman del tipo Ising, si puede adoptar diversos valores en la salida. -2, -1, 0, 1, 2, se dice que se trata de una neurona tipo Potts (Müller, 1990).

---

<sup>1</sup>

Fue uno de los primeros modelos matemáticos de una neurona que fue el propuesto

#### 9.4 Regla de propagación

La regla de propagación permite calcular el valor del potencial post-sináptico  $h_i$  de la neurona a partir de las entradas y pesos correspondientes. La función más común es la suma ponderada de las entradas y los pesos sinápticos.

$$h_i(t) = \sum_j W_{ij} X_j$$

Utilizando la notación de vectores, se puede representar como el producto escalar del vector de entrada y el de pesos.

$$h_i(t) = W_i^T \bullet X$$

El peso sináptico  $W_{ij}$  se interpreta como la intensidad de la relación entre la neurona presináptica  $j$  y la postsináptica  $i$ . Dependiendo de la entrada, si el peso es positivo y la entrada también positiva, se tenderá a excitar a la neurona postsináptica, pero si el peso es negativo se tenderá a inhibirla. Por lo que se habla entonces de sinapsis excitadoras (con peso positivo) o inhibitoras (con peso negativo). Una regla de propagación no lineal es la siguiente:

$$h_i(t) = \sum_{j_1 j_2 \dots j_p} W_{ij_1 j_2 \dots j_p} X_{j_1} X_{j_2} \dots X_{j_p}$$

que involucra la multiplicación de todas las entradas de la neurona, lo cual se acerca mucho a la funcionalidad de la neurona biológica, pero aumenta su complejidad tanto en su análisis como en el hardware necesario para implementarla. Este tipo de neurona se denomina de orden superior o neurona *sigma-pi*, debido al empleo de sumas y multiplicaciones (Rumelhart).

Otra regla de propagación empleada comúnmente en los modelos que emplean el cálculo de distancias entre vectores (como RBF, mapas de Kohonen o LVQ), es la distancia euclídea:

$$h_i^2(t) = \sum_j (X_j - W_{ij})^2$$

La cual determina la distancia entre el vector de entradas y el de pesos. Cuando los vectores de entrada y de pesos son similares, la distancia es muy pequeña, y cuando son diferentes la distancia se incrementa.

### 9.5 Función de activación o función de transferencia

Esta función determina el estado de activación actual  $a_i(t)$  a partir del potencial postsináptico  $h_i(t)$  y del estado de activación anterior  $a_i(t-1)$

$$a_i(t) = f_i(a_i(t-1), h_i(t))$$

Aunque en muchos modelos sólo se considera su estado actual, la función se expresa como sigue:

$$a_i(t) = f_i(h_i(t))$$

Existen varias funciones de activación, variable  $x$  representa el potencial postsináptico, mientras que la variable  $y = f(x)$  representa el estado de activación. La función identidad, la función escalón y la función lineal por intervalos son las más sencillas. Aunque dependiendo del algoritmo de aprendizaje, algunas veces se requiere que la función de activación sea derivable, las más empleadas en este caso son las funciones del tipo sigmoideo. La función gaussiana también es ampliamente utilizada junto con reglas de propagación que involucran el cálculo del cuadrado de las distancias entre los vectores de entrada y pesos, como se mencionó anteriormente. Las funciones sinusoidales también se emplean en los casos en los que es necesario expresar una periodicidad temporal.

### 9.6 Función de Salida

La función de salida determina la salida total de la neurona  $Y_i(t)$  en función de su estado de activación actual  $a_i(t)$ . Frecuentemente se usa la función identidad  $y = x$ , de tal forma que el estado de activación de la neurona se considera como la propia salida:

$$y_i(t) = F_i(a_i(t)) = a_i(t)$$

La función de salida también puede ser un escalón, para que la neurona dispare la salida hasta que se alcance cierto nivel de umbral.

### 9.7 Modelo estándar de una neurona artificial

El modelo anterior es muy general, por lo que de manera práctica se emplea un modelo estándar más sencillo.

Además, en algunas ocasiones se agrega al conjunto de pesos de la neurona un parámetro adicional  $\theta_i$  en las unidades de  $i$ , llamado umbral quedando el argumento de la función de la siguiente manera:

$$\sum_j W_{ij} X_j - \theta_i$$

Lo anterior implica agregar un grado de libertad a la neurona. Para el caso de las neuronas todo-nada mencionadas más adelante, el parámetro  $\theta$  representará el nivel mínimo o umbral que debe alcanzar el potencial postsináptico para que la neurona se active.

De esta forma, el modelo de la neurona queda:

$$y_i(t) = f_i \left( \sum_j W_{ij} X_j - \theta_i \right)$$

Para expresar el modelo de una forma más compacta, los índices  $i$  y  $j$  se hacen que comiencen en cero, definiendo  $W_{i0} = \theta_i$  y  $X_0 = -1$ .

Finalmente, el modelo se puede expresar de la siguiente manera:

$$y_i(t) = f_i \left( \sum_j W_{ij} X_j - \theta_i \right)$$

Donde:  $X_j(t)$  es el conjunto de entradas,  $W_{ij}$  son los pesos sinápticos,  $h_i(t) = \sum_j W_{ij} X_j$  es la regla de propagación y  $y_i(t) = f_i(h_i(t))$  es la función de



activación que representa tanto la salida de la neurona como su estado de activación.

Con el modelo de la neurona establecido solo es necesario especificar la forma de la función de activación para caracterizarla por completo. A continuación se muestran algunos de los modelos de neuronas más comunes.

### 9.8 Neuronas todo-nada

Tomando como referencia al modelo de la neurona estándar, considerando que las entradas son discretas, por ejemplo  $x_j = \{0, 1\}$ , y que la función de activación es el escalón  $H(x)$ , definido entre 0 y 1, se tiene:

$$y_i(t) = H\left(\sum_j W_{ij} X_j - \theta_i\right)$$

y como la función  $H(x) = 1$  cuando  $x \geq 0$ , y  $H(x) = 0$  cuando  $x < 0$ , se tiene lo siguiente:

$$y_i = \begin{cases} 1, & \text{si } \sum_j W_{ij} X_j \geq \theta_i \\ 0, & \text{si } \sum_j W_{ij} X_j < \theta_i \end{cases}$$

Por lo que, cuando el potencial de entrada supera el umbral  $\theta_i$  la neurona se activa, de lo contrario permanece inactiva. Debido a lo anterior este modelo se conoce también como dispositivo de umbral dada su forma de activación.

### 9.9 Neurona continua sigmoidea

Si el modelo de neurona estándar puede tener entradas tanto discretas como continuas pero la salida solamente continua, se puede emplear como función de activación a la función sigmoidea, que es una función continua y diferenciable en determinado intervalo. Las dos funciones más comunes son las siguientes:

$$y = \frac{1}{1 + e^{-x}}, \text{ con } y \in [0, 1]$$

$$y = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \text{tgh}(X), \text{ con } y \in [-1, 1]$$

El motivo de emplear funciones diferenciables es que algunas reglas de aprendizaje así lo requieren, tal como sucede con la retropropagación ó *BP*.

### 9.10 Redes neuronales aplicadas a las series de tiempo en economía

Las aplicaciones de las RNA en los negocios se dividen en dos principales categorías: clasificación y modelado. La clasificación se refiere a la forma de discriminar entre observaciones con características comunes en diferentes grupos. predicción de fallas corporativas, asignación de créditos y clasificación de bonos. Por otra parte, el modelado consiste en simular el comportamiento o las reacciones de una entidad o variable dado un estímulo externo, basado en observaciones pasadas, predecir los movimientos de los precios de las acciones o fluctuaciones en el tipo de cambio.

El primer artículo sobre redes neuronales que manejan información financiera fue realizado por White quien estudió la predicción de los precios de las acciones con un modelo de red neuronal. Posteriormente la predicción de la cotización de las acciones, del tipo de cambio y de otras variables económicas ha sido uno de los temas más atractivos para los investigadores y analistas financieros.

Hasta el momento se han presentado muchas ventajas de las redes neuronales con respecto a otras técnicas, pero en relación a lo que se conoce de la aplicación de las redes neuronales en el campo de la predicción bursátil, es interesante señalar que si fuera una herramienta infalible nadie lo publicaría sino más bien lo guardaría para sí y trataría de sacar el mayor provecho

posible. Lo anterior es difícil todavía de afirmar, pero en el ámbito académico se tiene la creencia de que es sólo un modelo más que realiza su función de predicción y optimización y, que realiza su tarea de una manera más eficiente puesto que considera a la información de una manera mejor organizada.

En la literatura existen varios modelos de redes neuronales diferentes aplicadas al procesamiento de series de tiempo de acuerdo al tipo de mecanismo a tratar con la información temporal. Como la mayoría de las redes neuronales han sido previamente definidas para reconocimiento de patrones con patrones estáticos, la dimensión temporal tiene que ser suministrada de una forma apropiada. Ulbricht (1996) distingue los siguientes mecanismos de manejo de series de tiempo por medio de redes neuronales:

- Capa de retraso sin realimentación (o ventanas de tiempo).
- Capa de retraso con realimentación.
- Unidad de retraso sin realimentación.
- Unidad de retraso con realimentación (bucles auto recurrentes).

El uso de las redes neuronales en este campo se puede ver en el contexto de la aproximación de funciones y clasificación, por lo que los principales tipos de redes neuronales se verán desde una forma más tradicional de procesamiento de secuencias de series de tiempo.

### **9.11 Sistemas tradicionales**

La forma tradicional de modelar los datos consiste en el análisis de series de tiempo. Pueden consistir ya sea de valores continuos o discretos. Hay dos tipos de modelos disponibles para modelar el comportamiento de las series de tiempo: los lineales y los no lineales.

- I. Sistemas lineales. Los métodos lineales clásicos están basados generalmente en realizar regresiones lineales de las series de tiempo

sobre una función lineal que posteriormente se utiliza para pronosticar valores futuros. La evidencia reciente muestra que las variaciones de los Índices de precios son no lineales (Bollerslev, 1992), Existen además, otros dos métodos para modelar series de tiempo: el análisis espectral de máxima entropía y la transformada rápida de Fourier. Sin embargo, para poderlos utilizar es necesario que exista cierta ciclicidad y poderla identificar.

- II. Modelos no lineales. Estas técnicas intentan reconstruir las series de tiempo con base en un muestreo de datos para pronosticar los valores futuros. Aunque estas técnicas son estadísticamente potentes, no han tenido mucho éxito en su aplicación en la predicción de los mercados financieros y otros. En el mundo real, los precios de mercado se generan por factores que incluyen las creencias, las relaciones atípicas entre los participantes, coberturas, arbitrajes, etc. Lo anterior crea un ambiente que contiene ruido estadístico, en donde aplicar las técnicas no lineales se vuelve complicado y donde estadísticamente se vuelve indeseable.

En estudios recientes, se comparan a las RNA con otros modelos de predicción tradicionales y concluyen que las estimaciones de las RNA son superiores cuando las entradas están altamente correlacionadas, cuando faltan datos o cuando los sistemas son no lineales.

### **9.12 RNA como modelo no lineal**

Una RNA sin capas ocultas es similar a un modelo lineal generalizado. Mientras que las RNA con alimentación hacia delante son un subconjunto de una clase más grande de modelos de regresión y discriminación no lineales. Generalmente se escoge un modelo no lineal sobre uno lineal cuando se sabe a priori que las relaciones entre las variables son no lineales o desconocidas.

Los modelos económicos lineales no son capaces de capturar patrones no lineales y tendencias en las relaciones de la mayoría de las variables de predicción econométricas usadas en contabilidad y finanzas. Por ejemplo, una serie de tiempo puede ser dividida en cuatro componentes, tendencia secular, variación cíclica, fluctuación estacional y fluctuación irregular. La fluctuación irregular puede a su vez ser dividida en comportamiento caótico determinista y ruido estocástico. Las técnicas lineales convencionales no pueden distinguir entre estos dos subcomponentes de ruido aleatorio y relaciones no lineales.

Cuando las RNA se aplican a tareas de clasificación, muestran la habilidad de inducir algoritmos para el reconocimiento de patrones. Así, la base del conocimiento es inferido del entrenamiento que se dé a la red. La desventaja, sin embargo, es que estas redes no tienen la facilidad de explicar cómo se llega a determinada conclusión. En general, una RNA no debe ser usada para modelar relaciones lineales inherentes. Las RNA deben ser aplicadas cuando exista alguna evidencia de no linealidad en las relaciones entre las variables dependientes y explicativas.

### **9.13 Modelos Paramétricos y No Paramétricos**

Los modelos paramétricos y en particular los modelos de regresión han llegado a abusar de los métodos estadísticos. Las pruebas se realizan de manera rutinaria y se hacen inferencias sin verificar la normalidad de los errores, independencia de los errores y constancia en la varianza del error.

Las RNA pueden ser usadas como modelos paramétricos, así que son necesarias las mismas consideraciones de distribución para los términos de error que para los modelos estadísticos.

Sin embargo, las RNA aplicadas como modelos no paramétricos pueden incorporar fácilmente múltiples fuentes sin simplificar consideraciones relativas

a la forma funcional de la relación entre las variables de salida y las predictivas. Se considera a las RNA como métodos estadísticos no lineales que ofrecen ciertas ventajas sobre sus contrapartes paramétricas. De esta forma, la primera decisión fundamental que se debe tomar es, si es apropiado usar una RNA para resolver el problema.

#### **9.14 Modelos Autorregresivos No Lineales**

Entre las redes neuronales más ampliamente utilizadas en clasificación y aproximación de funciones son las redes con alimentación hacia delante, como los perceptrones multicapa (MLP<sup>24</sup>; unidades ocultas con funciones de transferencia sigmoidales (Rumelhart) y las redes con función de base radial (RBFN<sup>25</sup>; unidades ocultas usando una regla de propagación de distancia y una función de transferencia gaussiana u otra (Broomhead y Lowe, 1988)). Ambos tipos de redes han sido probadas para ser aproximadores universales (Cybenko). y Hornik, 1989) para el MLP y (Kurkova, 1992 y Girosi, 1990) para aproximar cualquier función razonable por medio de las redes neuronales respectivas.

La aproximación de la no linealidad se realiza por medio del empleo de la función de transferencia no lineal, ya sea sigmoidea o gaussiana. Con un número fijo de neuronas ocultas (como en el caso de la mayoría de las aplicaciones) el método podría ser llamado aproximación semiparamétrica de funciones. No hace consideraciones específicas acerca de la forma de la función (como lo haría un método paramétrico), pero no puede aproximar cualquier función compleja arbitrariamente (como lo podría hacer una técnica no paramétrica)

De esta forma, los MLP y RBFN ofrecen una extensión a la forma clásica de modelar series de tiempo: los modelos autorregresivos lineales. El modelado autorregresivo lineal de series de tiempo (Box y Jenkins) considera a la

función objetivo como una combinación lineal de un número fijo de valores previos de la serie, incluyendo un término de error  $\varepsilon$ :

$$x(t) = \sum_{i=1}^p \alpha_i x(t-i) + \varepsilon(t)$$
$$x(t) = F^L(x(t-1), \dots, x(t-p)) + \varepsilon(t)$$

## 10. ANÁLISIS DE LA INFLACIÓN

El objetivo de este análisis es identificar la estructura su funcionamiento, e identificación de las variables que afectan a la inflación en sus variaciones por medio del IPC, puesto que estos datos serán quienes alimentaran el modelo neuronal.

### I. Introducción

Inflación es el crecimiento continuo y generalizado de los precios de los bienes y servicios y factores productivos de una economía a lo largo del tiempo. Otras definiciones la explican como el movimiento persistente al alza del nivel general de precios o disminución del poder adquisitivo del dinero.

Para comprender el proceso de desarrollo de la inflación y sus causas, resulta útil comprender que ante una inflación se suelen distinguir dos clases de fenómenos causales: las presiones inflacionarias y los mecanismos de propagación. Las presiones inflacionarias son un desequilibrio de oferta y demanda que provoca el ascenso en los precios en uno o varios mercados. Las presiones inflacionarias pueden ser de varios tipos: monetarias, fiscales, de costos, etc. Los mecanismos de propagación hacen que esos aumentos de precios se trasladen a otros mercados y se repitan a través del tiempo, posibilitando que la inflación se mantenga, aún luego de que hayan desaparecido los desequilibrios que la originaron. En la práctica, la evolución de la inflación se mide por la **variación del Índice de Precios al Consumidor (IPC)**. Para comprender el fenómeno inflacionario, se debe distinguir entre

aumentos generalizados de precios, que se producen de una vez y para siempre, de aquellos aumentos de precios que son persistentes en el tiempo.

Para determinar las causas de la inflación, ha sido el motor que impulsó a diversos teóricos a ensayar diferentes explicaciones sobre los procesos inflacionarios definidos en los siguientes tipos.

## II. Tipos de Inflación

### ➤ Inflación Monetaria

Cuando se está ante una inflación monetaria la oferta de dinero crece a una tasa superior a la tasa de crecimiento de la demanda de dinero. El principal fundamento teórico de quienes aplican esta teoría es la teoría cuantitativa del dinero. Los orígenes del exceso de oferta de dinero pueden ser una monetización del déficit del Gobierno u otra causa.

### ➤ Inflación de Demanda

La inflación de demanda corresponde con la explicación keynesiana[1] de la inflación: La demanda de bienes y servicios de una economía es mayor a la oferta de los mismos. La demanda puede provenir de diversos sectores y para diferentes finalidades:

- a) Por parte de las familias: productos y servicios finales, o sea, de bienes y servicios de consumo.
- b) Por parte de empresas para ampliar su capacidad productiva, es decir inversión.
- c) Por parte del gobierno, que puede ser inversión productiva o bien gasto que no aumentará directamente la oferta global del país.
- d) Por parte del sector externo, es decir, exportaciones.

---

<sup>1</sup> Keynes desarrolló todo un cuerpo teórico en donde muestra que las economías generalmente se encuentran produciendo por debajo de su nivel máximo potencial, fundamentalmente en razón de unas condiciones de demanda de bienes y servicios seriamente restringidas



Esta causa de la inflación se presenta cuando la capacidad productiva (oferta de bienes) no puede aumentar al mismo ritmo con el que crece la demanda de bienes. Puede estar relacionada con la política monetaria, por ejemplo, si la política monetaria mantiene baja o negativa la tasa de interés real, esto puede estimular el consumo pero no la inversión, resultando en una presión inflacionaria.

➤ **Inflación de Costos vacío**

Dado que el precio final de bienes y servicios está íntimamente relacionado con los costos incurridos en su producción, un aumento en los costos generará un aumento del precio final de los mismos.

Para analizar las causas del aumento de los costos de producción, resulta útil agrupar los insumos que se utilizan en categorías. Así tenemos que los insumos utilizados para producir bienes y servicios pueden ser:

- a) mano de obra;
- b) materias primas que pueden ser commodities internacionales como petróleo o granos, o materias primas cuyo precio no se fija en mercados internacionales;
- c) maquinaria, que también puede ser producida localmente o importada;
- d) servicios, por ejemplo, el transporte, que a su vez depende fuertemente del precio del petróleo, etc.

El aumento del precio de cualquiera de estos insumos, se puede trasladar al precio final de los bienes, lo que puede ser el gatillo que dispare los mecanismos de propagación para que este aumento de precios se transforme en inflación. Por ejemplo, el aumento de la demanda de commodities (petróleo, granos, metales, etc.) por parte de China e India provocó un aumento a nivel mundial en el precio de los mismos. Esto aumentó los costos de producción de numerosos bienes y servicios a nivel

mundial (costos de transporte, precios de insumos como cobre, acero, etc., precios de alimentos debido al aumento del precio de los granos, etc.).

➤ **Inflación estructural**

La inflación estructural se debe a alguna característica de la estructura productiva de un país que produce un aumento de precios, en particular hace referencia a "cuellos de botella" en un sector determinado, que se trasladan mediante aumentos de precios al resto de la economía.

➤ **Inflación Inercia**

El concepto de inflación inercial hace referencia a una situación en la que los mecanismos de propagación de la inflación ya se han instalado y son la principal causa de que la inflación perdure en el tiempo.

**III. ¿Qué variables influyen en la variación del IPC para una inflación?**

Para facilitar el análisis de las variables económicas que puedan proyectar la variación del IPC por el cual medir la inflación, se clasifican en tres grupos. El primer grupo incluye **variables financieras**, que son capaces de reflejar acciones de política monetaria, tales como, tasas de interés, dinero y tipo de cambio. El segundo incorpora **variables reales**, que presionan los precios vía excesos de demanda en la economía, como el crecimiento de la producción real, y la tasa de desempleo. El tercero considera **variables fiscales**, dada la hipótesis de que el déficit fiscal y su forma de financiamiento tienen un efecto sobre la inflación.

➤ **VARIABLES FINANCIERAS**

a) **Tasas de Interés**

Muchos bancos centrales utilizan las tasas de interés como variable instrumento para lograr el objetivo en precios y lo realizan a través de su política monetaria. En este sentido, incrementos en las tasas de

interés ocasionarían una reducción de la presión inflacionaria como consecuencia de la desaceleración de la actividad económica y de la demanda agregada asociada a una disminución de la inversión y del consumo.

Algunos bancos centrales realizan sus políticas basados en reglas de acción, tal como lo describen las denominadas “reglas tipo Taylor”, en lugar de tomar decisiones de tipo discrecional. Por ejemplo, una aplicación de este tipo de reglas, en el contexto de un modelo keynesiano convencional, se da cuando la tasa de interés de política se ajusta bajo el siguiente criterio

$$i = \pi + \lambda y + (1 - \lambda)(\pi - \pi^*) + \rho$$

Donde:

$i$  = es la tasa de interés nominal de política monetaria.

$\pi$  = es la tasa de inflación promedio de los últimos meses.

$\lambda$  = es la importancia relativa (ponderador) que la autoridad monetaria le otorga al objetivo de crecimiento económico versus objetivo de inflación.

$y$  = es la desviación del producto real con respecto a su nivel objetivo.

$\pi^*$  = es la tasa de inflación objetivo.

$\rho$  = es la estimación de la tasa real de interés de equilibrio.

Esta regla relaciona la actuación de los bancos centrales en cuanto a la tasa de interés nominal de corto plazo con los objetivos de inflación a largo plazo, teniendo en cuenta las desviaciones inflacionarias que se estén produciendo respecto del objetivo y atendiendo la brecha de producción que exista respecto del producto potencial de largo plazo.

## **b) Crecimiento del Dinero**

Para alcanzar su objetivo de estabilidad de precios, los bancos centrales suelen utilizar metas intermedias como los agregados

monetarios, y para ello parten de la ecuación de cambio de Fisher (1911), que en su forma logarítmica se expresa como:

$$m + v = p + y$$

Donde:

$m$  = logaritmo de oferta monetaria.

$v$  = logaritmo de la velocidad de circulación del dinero.

$p$  = logaritmo del nivel de precios.

$y$  = logaritmo del PIB real.

Dicha formalización indica que si el nivel de producto y la velocidad de circulación se mantienen relativamente estables, cambios en los agregados monetarios estarán directamente relacionados con la inflación en el largo plazo.

En general, la evidencia empírica ha respaldado esta conclusión mostrando que países con incrementos relativamente rápidos y sostenidos del *stock* del dinero terminarán con tasas de inflación relativamente altas, por lo que, como relación de largo plazo, se corrobora que el crecimiento del dinero es una de las causas de la inflación.

### c) Tipo de Cambio

El tipo de cambio real ( $TCR$ ) se define como el precio relativo de los bienes transables y no transables, y se aproxima por la relación entre el índice de precios internacionales y domésticos, expresados en la misma moneda por medio del tipo de cambio nominal.

$$TCR = \frac{P_T}{P_{NT}} \approx \frac{EP^*}{P}$$

Donde:

$P_T$  = precio de los transables.

$P_{NT}$  = precio de los no transables.

$E$  = tipo de cambio nominal (unidades monetarias del país local por unidad monetaria del extranjero).

$P^*$  = índice de precios internacionales.

$P$  = índice de precios del país local.

Según la ley de un solo precio (paridad del poder de compra) los precios domésticos se igualan a los precios internacionales expresados en la misma moneda ( $P = EP^*$ ). En este caso, una variación del tipo de cambio nominal estará asociada con un cambio de igual proporción en los precios del país local.

### ➤ Variables Reales

#### Tasa de Desempleo

En busca de un enfoque alternativo al de la inflación como fenómeno monetario, algunos análisis exploran la medición de la actividad económica real como indicador de presión inflacionaria en el corto plazo. Estos estudios se han desarrollado en torno a la aparente relación negativa entre desempleo e inflación señalada por la conocida curva de Phillips.

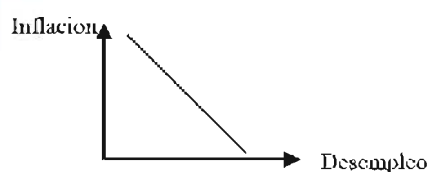


Figura 2 Curva de Phillips Inflación vs Desempleo  
Elaboración propia

La intuición inicial que estaba detrás de la curva de Phillips, es que a bajas tasas de desempleo usualmente las condiciones del mercado de trabajo propician incrementos en los salarios que pueden ser trasladados a precios, y por tanto, generen mayor inflación. Caso contrario es cuando hay altas tasas de desempleo que se reflejan en excesos de oferta de trabajadores, lo que

reduce la presión en salarios y en precios. Sin embargo, en los años setenta, las tasas de inflación y desempleo empezaron a crecer conjuntamente, con lo que la curva de Phillips ya no refleja la evidencia empírica. Por lo tanto, esos nuevos fenómenos económicos exigían otras explicaciones.

### ➤ Variables Fiscales

El gobierno puede utilizar la política fiscal, mediante variaciones discrecionales de sus ingresos y gastos, para influir en el nivel de actividad económica, con el objetivo de reducir las fluctuaciones y propiciar el crecimiento de la economía. La restricción presupuestaria del sector público, indica que éste no puede gastar más recursos que aquellos de los que puede disponer. Es decir, en cada período  $t$ , el sector público tiene un volumen nominal de gastos ( $G_t$ ) que debe financiar mediante impuestos ( $T_t$ ), endeudamiento con el sector privado ( $B_t - B_{t-1}$ ), endeudándose con el exterior ( $B_t^* - B_{t-1}^*$ ), o con crédito del sistema bancario ( $H_t - H_{t-1}$ ), por lo que en cada periodo tendríamos:

$$G_t = T_t + (B_t - B_{t-1}) + (B_t^* - B_{t-1}^*) + (H_t - H_{t-1})$$

En principio, si existe déficit en el presente ( $(G_t) > (T_t)$ ), será necesario aplicar una restricción fiscal en el futuro que conduzca a un superávit para poder repagar el endeudamiento incurrido por la insuficiencia de ingresos en el presente. No obstante, la posibilidad de acudir a la expansión monetaria (crédito del banco central) hace innecesario que el déficit presente conlleve necesariamente a superávit en el futuro. La restricción presupuestaria pone de manifiesto que el déficit público aparte de su posible efecto sobre la producción, puede también afectar la cantidad de dinero en circulación al disponer de crédito del ente emisor como forma de financiamiento, con lo que tendría un efecto inflacionario.



**Capítulo II.**  
**METODOLOGÍAS PARA CONSTRUCCIÓN DEL MODELO**

## 11. METODOLOGÍA PARA PRONÓSTICOS MEDIANTE MODELOS ARIMA Y RNA

Para construir un modelo ARIMA que aproxime aceptablemente las características de una serie de tiempo se hace uso de la metodología de Box-Jenkins, la cual puede ser estructurada en cinco etapas.

### 11.1 Metodología para predicciones mediante ARIMA

#### I. Análisis exploratorio de la serie

Se grafica la serie a través del tiempo, de manera que se puedan observar apriori sus componentes: tendencia, estacionalidad y ciclos. Podría notarse también la necesidad de aplicar diferencias, en la parte estacional o regular, para hacer que la media sea constante.

#### II. Identificación del modelo

Se debe sugerir un conjunto reducido de posibles modelos:

- Selección del conjunto de estimación: conjunto de datos que se usará para la estimación y adecuación del modelo y el conjunto de predicción: conjunto de datos para evaluar las predicciones.
- Determinación de la función de autocorrelación, la función de autocorrelación parcial y sus correspondientes correlogramas.
- Determinación del orden del componente autorregresivo  $p$  y promedio móvil  $q$  del modelo ARMA  $(p, q)$ , haciendo uso de los patrones que se observan en los correlogramas simple y parcial.
- Estudio de la estacionariedad. Si la serie no es estacionaria, se debe convertir en estacionaria antes de aplicar la metodología Box-Jenkins.



### **III. Estimación de parámetros**

Una vez identificado el modelo, se obtienen los parámetros mediante la minimización de la suma del cuadrado de los errores. Los parámetros se estimarán siguiendo los pasos descritos en libro de Pulido y López (1999).

### **IV. Verificación de diagnóstico**

Después de seleccionar un modelo ARIMA particular y de estimar sus parámetros, se trata entonces de ver si el modelo seleccionado se ajusta a los datos en forma razonablemente buena, ya que es posible que exista otro modelo ARIMA que también lo haga.

### **V. Pronóstico**

El pronóstico se basa en el modelo ARIMA seleccionado. Se predicen  $m$  periodos correspondientes al tamaño del conjunto de predicción con sus intervalos de confianza.

## **11.2 Metodología para predicciones mediante RNA**

Se seguirá la metodología propuesta por Kaastra y Boyd (1996) en el diseño de modelos de redes neuronales para predicción de series de tiempo, la cual está compuesta por ocho pasos descritos a continuación.

### **a) Verificación de diagnóstico**

Determinar qué variables de entrada son importantes en el pronóstico de series de tiempo es un asunto crítico. En este caso, el interés en la predicción involucra emplear como entradas datos técnicos en lugar de fundamentales por las características mismas de la investigación. La forma más simple de modelar redes neuronales involucra emplear valores retrasados de la variable dependiente o sus primeras diferencias como entradas. Generalmente, en las aplicaciones en mercados accionarios y de derivados es más probable de usar datos diarios en el diseño de una red

neuronal como un componente de todo un sistema completo. Sin embargo, también es posible emplear datos semanales o mensuales como entradas de la red en el caso de que se requiera formular una estrategia diferente que involucre intervalos de tiempo mayores.

#### **b) Colección de datos**

Al seleccionar las variables del paso anterior también se debe considerar su costo y disponibilidad. Típicamente los datos técnicos son más fácilmente disponibles por parte de varios proveedores a cierto costo, mientras que la información fundamental es más difícil de conseguir.

El tiempo dedicado a la recolección de datos no puede ser usado para el pre procesamiento, entrenamiento y la evaluación del desempeño de la red.

Es importante también mencionar que se deben calcular los retrasos de los datos apropiadamente de manera que se puedan introducir a la red.

#### **c) Pre-procesamiento de datos**

El pre-procesamiento de datos se refiere al hecho de analizar y transformar las variables de entrada y salida para minimizar el ruido, enfatizar relaciones importantes, detectar tendencias, y aplanar la distribución de la variable para ayudar a la red neuronal en el aprendizaje de patrones relevantes. Como las redes neuronales son buscadores de patrones, la representación de los datos es crítica en el diseño de una red exitosa. Las variables de entrada son raramente ingresadas a la red en su forma original. Al menos los datos deben ser escalados entre los límites inferior y superior de la función de transferencia (generalmente entre 0 y 1 ó -1 y 1).

Dos de las transformaciones de datos más comunes tanto en métodos tradicionales como en redes neuronales es tomar las primeras diferencias y calcular el logaritmo natural de la variable independiente.

La primera diferenciación es para quitar tendencias lineales en los datos. La transformación logarítmica es útil para datos que pueden tomar valores muy pequeños o muy grandes y que están caracterizados por una distribución de cola derecha extendida. Las transformaciones logarítmicas también convierten las relaciones multiplicativas o de razones en aditivas lo cual se cree que simplifica y mejora el entrenamiento de la red.

#### **d) Conjuntos de entrenamiento y prueba**

Comúnmente el total de datos de la muestra se divide en dos conjuntos, llamados de entrenamiento y prueba. El conjunto de entrenamiento es el más grande y es utilizado por la red neuronal para aprender los patrones existentes en los datos. El conjunto de prueba, que varía del 10% al 30% del conjunto de entrenamiento se utiliza para evaluar la capacidad de generalización de una red ya entrenada. El investigador selecciona la red que se desempeña mejor de acuerdo al conjunto de prueba. Algunas veces también se incluye un conjunto de validación el cual debe consistir de las observaciones más recientes.

El conjunto de prueba puede ser seleccionado ya sea aleatoriamente del conjunto de entrenamiento o consistir de un conjunto de observaciones ubicadas inmediatamente después del conjunto de entrenamiento. La ventaja de utilizar las observaciones más recientes como hechos de prueba es que dichas observaciones pueden ser más importantes que los datos anteriores.

#### **e) Paradigmas de las redes neuronales**

Existe un número infinito de maneras de construir una red neuronal; como ya se mencionó en Marco Teórico, la neurodinámica y la arquitectura son dos términos usados para describir la forma en que se organiza una red. La combinación de la neurodinámica y la arquitectura define el paradigma de la red neuronal.

De acuerdo a Caudill (1992), la red de retropropagación es la red multicapa más común utilizada casi en el 80% de las aplicaciones.

#### **f) Número de capas ocultas**

Las capas ocultas dan a la red la habilidad de generalizar, y en la práctica las redes neuronales con una ó dos capas ocultas son las más utilizadas y han tenido un buen desempeño. El incremento en el número de capas también incrementa el tiempo de procesamiento y el peligro de sobreajuste lo que conduce a un pobre desempeño en la predicción fuera de muestra. El sobreajuste ocurre cuando un modelo de predicción tiene muy pocos grados de libertad. En otras palabras, se tienen relativamente pocas observaciones en relación con sus parámetros y por lo tanto es capaz de memorizar datos individuales en lugar de aprender patrones generales.

#### **g) Número de neuronas ocultas**

Se traza una gráfica del criterio de evaluación (la suma del cuadrado de los errores) en el conjunto de entrenamiento como una función del número de neuronas ocultas para cada red neuronal. Dicha gráfica produce una forma de tazón, en la cual, la red que genere el menor error en el fondo del tazón será la red que sea capaz de generalizar mejor. Este método consume mucho tiempo pero generalmente da buenos resultados.

Para determinar el número de neuronas en la capa oculta no existen reglas generales o teorías, aunque existen algunas recomendaciones sugeridas por varios investigadores:

- Shih (1994), sugiere que la topología de la red debe tener una forma piramidal, esto es, tener el mayor número de neuronas en la capa de entrada y menos en las posteriores. Recomienda que el número de neuronas en cada capa este más o menos entre la mitad de la capa siguiente y el doble del número de la capa anterior.
- De acuerdo con Azoff (1994), son suficientes  $2N+1$  neuronas para  $N$  entradas.
- Baum y Haussler (1998) sugieren que el número de neuronas sea calculado por:  $j = \frac{me}{n+z}$  donde  $j$  es el número de neuronas en la capa oculta,  $m$  es el número de datos en el conjunto de entrenamiento,  $e$  es la tolerancia al error,  $n$  es el número de entradas y  $z$  es el número de salidas.

La mayoría de los investigadores no están convencidos de las recomendaciones anteriores ya que argumentan que no hay forma de determinar una buena topología de la red a partir sólo del número de entradas y salidas.

#### **h) Número de neuronas de salida**

Decidir el número de neuronas de salida es algo más sencillo porque hay muchas razones para emplear sólo una neurona de salida. Las redes neuronales con múltiples salidas, especialmente si éstas salidas están ampliamente espaciadas, producirán resultados inferiores en comparación con una red con una única salida. Lo recomendable es tener una red especializada para cada una de las salidas deseadas en cada predicción.

**i) Funciones de transferencia**

Siempre es recomendable ver los histogramas de las variables escaladas de entrada y salida, de manera que se pueda identificar la necesidad de realizar un escalamiento que produzca una distribución uniforme y poder emplear el número de neuronas disponibles de manera eficiente.

**j) Criterio de evaluación**

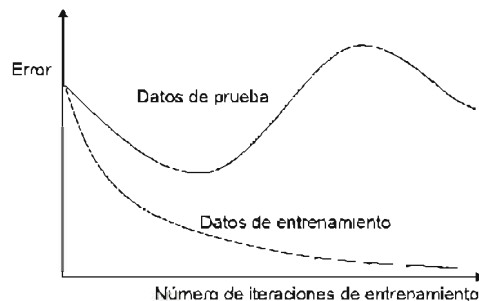
En redes neuronales la función de error más común a minimizar es la suma del cuadrado de los errores.

**k) Entrenamiento de la red**

Entrenar una red neuronal para aprender patrones involucra el presentarle ejemplos de manera iterativa de las respuestas correctas. El objetivo del entrenamiento es encontrar un conjunto de pesos entre las neuronas que determinan el mínimo global de la función de error. A menos que el modelo esté sobre ajustado, el conjunto de pesos debería proporcionar una buena generalización. Un término de momento y de cinco a diez conjuntos aleatorios de pesos iniciales pueden mejorar las oportunidades de alcanzar un mínimo global.

**l) Número de iteraciones de entrenamiento**

Muchas veces el procedimiento de agregar un número mayor de iteraciones en el entrenamiento producirá un menor error en el entrenamiento, pero éste a su vez no garantiza que se obtendrá el menor error en con el conjunto de prueba, tal y como se ilustra en la figura 3.



**Figura 3** Gráficas posibles del comportamiento de los errores en los conjuntos de entrenamiento y prueba de una RNA **Fuente:** Ivan Cruz Universidad Autónoma de México.

El objetivo de la convergencia en el entrenamiento es alcanzar un mínimo global. El punto en el cual la red ya no mejora se llama convergencia. Esto requiere entrenamiento para un número suficiente de operaciones utilizando un número razonable de pesos iniciales seleccionados aleatoriamente. Un método para determinar un valor razonable para el máximo número corridas es dibujar la gráfica de la correlación de la media, la suma de los cuadrados de los errores o cualquier otra medida de errores apropiada para cada iteración o en intervalos predeterminados hasta el punto en donde la mejora es imperceptible. Cada iteración se puede dibujar fácilmente si el software crea un archivo de estadísticas o, si la correlación de la media puede ser grabada a intervalos de 100 ó 200. Después de dibujar la correlación de la media para un número valores iniciales de los pesos seleccionados aleatoriamente, se debe escoger el máximo número de corridas basado en el punto donde la correlación de la media deja de incrementar y se estabiliza.

#### **m) Tasa de aprendizaje y momento**

La tasa de aprendizaje es una constante de proporcionalidad que determina el tamaño del cambio de los pesos. El cambio de los pesos de una neurona es proporcional al impacto del peso de las neuronas sobre el error. Un método de incrementar la tasa de aprendizaje y por lo tanto agilizar el

tiempo de entrenamiento sin caer en oscilaciones, es incluir un término de momento en la regla de aprendizaje de Backpropagation *BP*. El término de momento determina cómo los cambios pasados de los pesos afectan a los actuales cambios en los pesos. Este término suprime oscilaciones lado a lado filtrando variaciones de alta frecuencia. Cada nueva dirección de búsqueda es una suma ponderada de los gradientes actuales y previos.

La mayoría de los programas de software proporcionan valores por *default* para la tasa de aprendizaje y términos de momento que generalmente funcionan bien. Las tasas de aprendizaje usadas en los trabajos previos varían de 0.1 a 0.9. Es común empezar el entrenamiento con una tasa de aprendizaje como 0.7 y disminuirla conforme avanza el entrenamiento. Muchos de los programas disminuyen la tasa de aprendizaje y aumentan los valores de momento conforme se alcanza la convergencia.

#### **n) Implementación**

La mayoría de los vendedores de software proporcionan los medios para implementar las redes obtenidas ya sea por el programa mismo o como un programa ejecutable. También es posible replicar la respuesta de la red en una hoja de cálculo conociendo su arquitectura, funciones de transferencia y pesos. Así como también es importante tomar en cuenta todas las transformaciones de datos, escalamiento, y otros parámetros empleados en la etapa de pruebas.

En un mercado cambiante, es necesario re-entrenar la red periódicamente para que no pierda su habilidad de adaptación. Una vez implementada, el desempeño de la red se degradará con el tiempo, sin embargo, aún cuando se realice un entrenamiento periódico, no se garantiza el desempeño de la red porque se puede dar el caso de que las variables independientes seleccionadas cambien de importancia.





**Capítulo III**  
**DESARROLLO Y ANÁLISIS DE RESULTADOS**

## 12. DATOS

Los datos a utilizarse pertenecen a la serie de datos generados por la Variación del Índice de Precios al Consumidor por Año según Mes fuente de obtención INE (Instituto Nacional de Estadística) del 1 de enero de 1998 al 30 de diciembre de 2008. Representando 132 observaciones.

<b>Bolivia: Índice de Precios al Consumidor por Año según Mes</b>											
<b>Serie: 1998 - 2008</b>											
MES	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Enero	68,20	70,32	72,80	75,06	75,57	77,74	80,97	85,35	88,69	94,03	104,70
Febrero	68,77	70,59	73,10	75,04	75,73	77,57	81,13	85,40	88,96	94,81	107,44
Marzo	68,91	70,29	73,55	74,87	75,50	77,62	80,91	85,53	88,72	95,10	108,49
Abril	69,10	70,18	74,26	75,02	75,48	77,86	80,93	85,21	88,84	94,85	109,29
Mayo	69,32	70,24	73,33	74,89	75,51	77,87	81,26	85,79	89,59	95,28	111,33
Junio	69,48	70,50	73,47	75,44	75,59	78,03	81,87	87,11	90,14	96,09	112,73
Julio	69,64	70,64	74,00	76,34	75,90	78,50	82,30	86,70	90,65	98,66	113,25
Agosto	69,66	71,01	74,29	75,84	76,07	79,00	82,59	87,01	90,82	100,23	113,99
Septiembre	69,62	71,43	75,66	75,64	76,42	79,19	82,57	87,14	90,90	100,42	114,99
Octubre	70,31	71,94	76,67	75,70	76,83	80,18	83,24	87,46	91,32	101,67	115,20
Noviembre	70,33	71,99	74,72	75,53	77,25	79,76	83,69	87,84	92,01	102,92	115,35
Diciembre	70,22	72,42	74,89	75,58	77,43	80,48	84,20	88,33	92,70	103,57	115,84
Prom Anual	69,46	70,96	74,23	75,41	76,11	78,65	82,14	86,57	90,28	98,14	111,88

Tabla 2: IPC, FUENTE: Instituto Nacional De Estadística

<b>Bolivia: Variación del Índice de Precios al Consumidor por Año según Mes</b>											
<b>Serie: 1998 - 2008</b>											
MES	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Enero	1,40	0,15	0,53	0,22	-0,01	0,40	0,62	1,37	0,40	1,43	1,09
Febrero	0,83	0,38	0,42	-0,03	0,21	-0,21	0,19	0,05	0,31	0,83	2,62
Marzo	0,21	-0,42	0,61	-0,23	-0,31	0,06	-0,27	0,16	-0,28	0,30	0,97
Abril	0,28	-0,15	0,96	0,21	-0,03	0,32	0,02	-0,38	0,14	-0,26	0,74
Mayo	0,31	0,09	-1,25	-0,18	0,04	0,02	0,41	0,67	0,84	0,45	1,87
Junio	0,23	0,37	0,19	0,74	0,11	0,20	0,74	1,54	0,62	0,85	1,26
Julio	0,23	0,20	0,72	1,20	0,41	0,60	0,53	-0,46	0,56	2,68	0,46
Agosto	0,03	0,52	0,40	-0,66	0,22	0,64	0,35	0,35	0,19	1,59	0,65
Septiembre	-0,05	0,59	1,84	-0,26	0,46	0,23	-0,03	0,15	0,08	0,19	0,88
Octubre	0,99	0,71	1,34	0,07	0,54	1,25	0,81	0,37	0,47	1,25	0,18
Noviembre	0,03	0,07	-2,55	-0,23	0,54	-0,51	0,55	0,44	0,75	1,23	0,13
Diciembre	-0,16	0,59	0,22	0,07	0,23	0,90	0,60	0,56	0,75	0,63	0,43
PROM ANUAL	7,67	2,16	4,60	1,60	0,92	3,34	4,44	5,40	4,28	8,70	14,01
ACUMULADO	4,39	3,13	3,41	0,92	2,45	3,94	4,62	4,91	4,95	11,73	11,85

Tabla 3: Variación del IPC, FUENTE: Instituto Nacional De Estadística

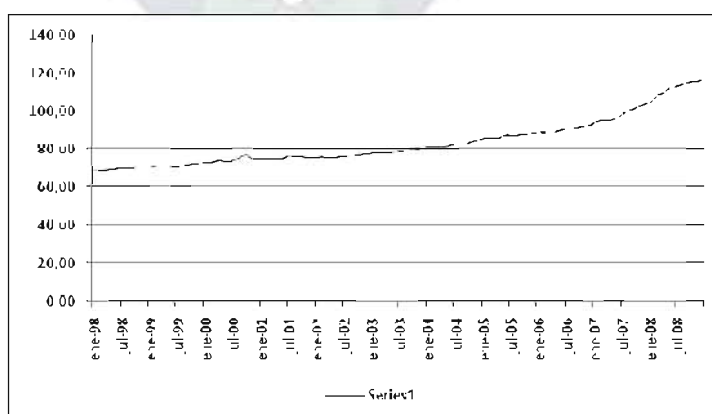
El análisis se realiza en la serie de variación del IPC, es obtenida por la siguiente fórmula:

$$VarIPC = \left\{ \left( \frac{\text{Índice del mes más actual}}{\text{Índice del mes anterior al mes más lejano}} \right) \times 100 \right\} - 100$$

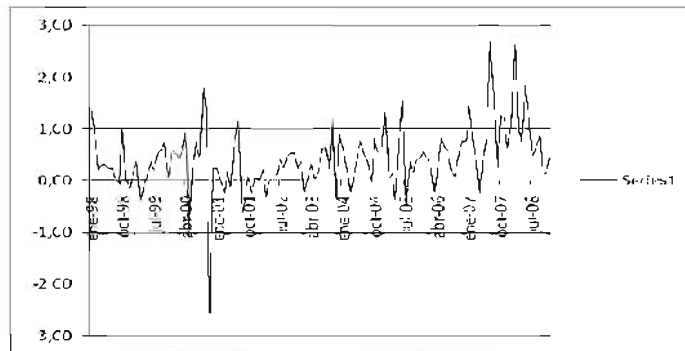
De acuerdo a Kaastra y Boyd (1996), suavizar tanto los datos de entrada como los de salida, empleando ya sea promedios móviles simples o exponenciales, puede ser una forma recomendable en el caso de que se intente predecir la tendencia de una serie, en lugar de intentar predecir los cambios en precios alrededor de la tendencia, en cuyo caso, lo mejor es modelar las series de tiempo por medio de modelos ARCH. En este trabajo, se realiza un suavizado exponencial de los datos, de manera que se obtengan pronósticos sobre la tendencia de la serie en lugar de los valores originales, tanto para el modelo ARIMA como para el modelo con RNA. **Obtención del modelo ARIMA para la serie de Variación del Índice de Precios al Consumidor** La muestra usada para ajustar el modelo de la serie del INE contiene el 80% del total de los datos y se usarán los componentes AR y MA como las variables explicativas. De esta manera, se obtienen 132 datos. Para poder realizar los análisis de una manera más ágil, se utiliza el software econométrico XLSTAT y Eviews

### 12.1 Análisis exploratorio de la serie

La gráfica de la serie original y la serie aplicando la formula de la variación para su análisis se presenta en las figuras 4 y 5.



**Figura 4** Bolivia Índice de Precios al Consumidor por Año según Mes, Serie 1998 – 2008 Fuente INE



**Figura 5** Gráfica de la Variación del Índice de Precios al Consumidor del 1 de enero de 1998 al 30 de diciembre del 2008 Fuente INE.

## 12.2 Experimentación Modelo ARIMA con XLSTAT

Aplicando y siguiendo las etapas y metodologías correspondientes con paquete econométrico XLSTAT se obtiene los resultados siguientes:

XLSTAT 2009.3.02 - ARIMA - 2009

Parámetros del modelo:  $p = 1 / d = 0 / q = 0 / P = 0 / D = 0 / Q = 0 / s = 0$

Optimizar Mínimos cuadrados (Convergencia = 0.00001 / Iteraciones = 5000)

Validación 1

Predicción 4

Intervalos de confianza (%): 95

Estadísticas descriptivas:

Variable	Observaciones	Obs. con datos perdidos	Obs. sin datos perdidos	Mínimo	Máximo	Media	Desviación típica
I39785307384561	131	0	131	-2,547	2,677	0,407	0,627

### Resultados de la modelización ARIMA de la serie I39785307384561:

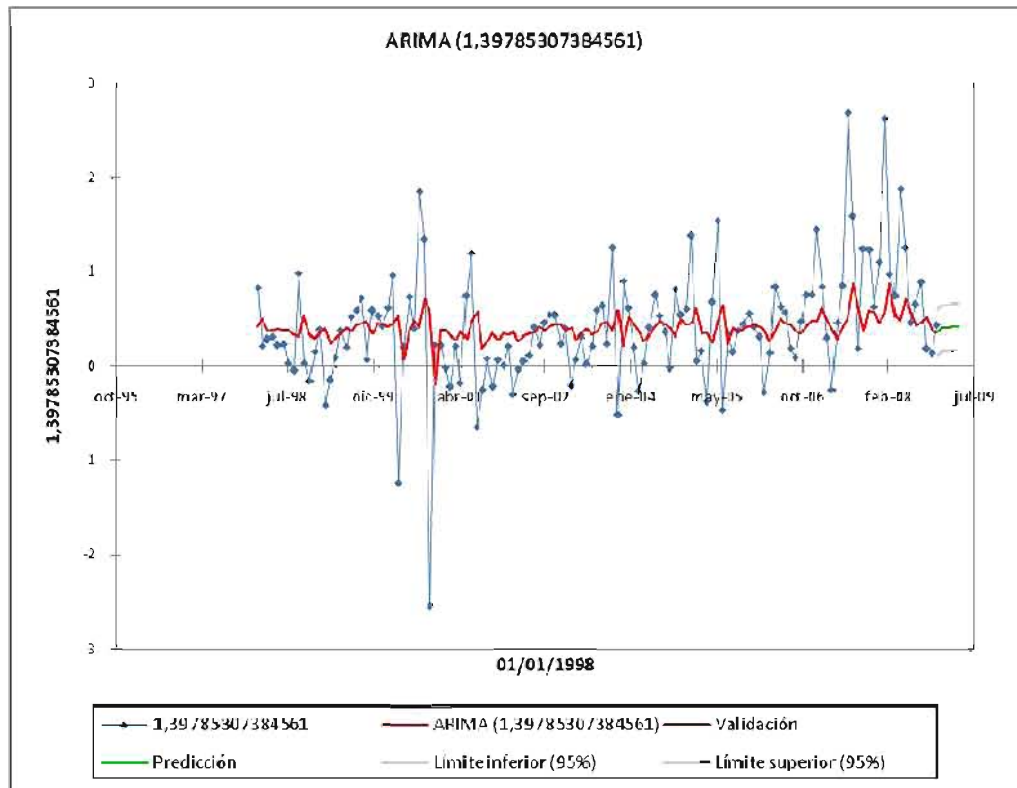
Estimación preliminar Yule-Walker (I39785307384561)

#### Coefficientes de ajuste

Observaciones	130.000
SEC	48.896
MAPE(Dif)	154.260
MAPE	230.532
Varianza RB	0.376
Varianza RB (estimada)	0.376
-2Log(Vero)	2418.49
FPE	0.382
AIC	2458.49
AICC	2459.43
SBC	25158.4

#### Parámetros del modelo

Constante	0.823
-----------	-------



**Figura 6** Gráfica de la **Variación del Índice de Precios al Consumidor** del 1 de enero de 1998 al 30 de diciembre del 2008 realizado por XLSTAT

### 12.3 Obtención del modelo RNA para la serie de de Variación del Índice de Precios al Consumidor

A continuación se especifican los pasos a seguir para construir, entrenar y probar una red neuronal para predecir valores futuros en el tiempo  $t+1$ , basándose únicamente en los valores históricos de la serie correspondiente a la Variación del Índice de Precios al Consumidor.

#### 12.3.1 Selección de variables

Los datos a utilizar son los obtenidos por medio de la fórmula de cálculo de la variación del IPC de la serie de tiempo.

Se propone utilizar como programa de desarrollo al *Neural Network Toolbox* de Matlab, ya que es una herramienta genérica de programación muy potente en la manipulación de matrices, redes neuronales y gran flexibilidad en la manipulación de parámetros.

### 12.3.2 Colección de datos

Para poder introducir la serie de tiempo a la red neuronal es necesario tener los datos de tal forma que se puedan manipular con facilidad en la RNA, es por eso que se propone copiar los datos a un vector columna A como el siguiente (Datos descritos en el Apéndice B):

$$A = \begin{bmatrix} a_{t-n} \\ \dots \\ a_{t-2} \\ a_{t-1} \\ a_t \end{bmatrix}$$

A partir del vector anterior se genera la matriz de retrasos  $B_r$ , que contiene los elementos de retraso que serán utilizados como entradas a la red neuronal artificial. La primera columna se considerará como el objetivo, mientras que las demás columnas serán las correspondientes a los retrasos hasta un máximo de  $r$ , el cual será fijado de acuerdo a los criterios mencionados en la determinación de las entradas.

$$B_r = \begin{bmatrix} a_{t-n} & 0 & 0 & \dots & 0 \\ a_{t-n+1} & a_{t-n} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{t-2} & a_{t-3} & a_{t-4} & \dots & a_{t-r-2} \\ a_{t-1} & a_{t-2} & a_{t-3} & \dots & a_{t-r-1} \\ a_t & a_{t-1} & a_{t-2} & \dots & a_{t-r} \end{bmatrix}$$

Donde:

$n$  es el número total de elementos de la serie

$r$  es el número de retrasos

### 12.3.3 Pre procesamiento de datos

El entrenamiento de la red neuronal se puede hacer más eficiente si se lleva a cabo un pre procesamiento en las entradas y objetivos de la red, que consiste en reescalar las entradas (Demuth y Beale, 1998).

En este caso teniendo como referencia que se aplicara la función de activación sigmoideal o sigmoidea, los datos de la serie se normalizan o transforman a valores comprendidos entre 0 y 1 empleando la siguiente ecuación:

$$z_t = \frac{y_t - y_{\min}}{y_{\max} - y_{\min}}$$

Donde:

$z_t$  = Serie de tiempo transformada en valores entre 0 y 1.

$y_t$  = Valores originales de la serie de tiempo.

$y_{\max}$  = Valor máximo de la serie de tiempo.

$y_{\min}$  = Valor mínimo de la serie de tiempo.

**Datos de ejemplo:**

$$y_t = 0,59$$

$$y_{\max} = 2,68$$

$$y_{\min} = -2,55$$

$$z_t = \frac{0,59 - (-2,55)}{2,68 - (-2,55)} = 0,600$$

La forma de la función tiene un pequeño efecto en la red aunque tiene un impacto significativo en la rapidez de entrenamiento. Al terminar de realizar la predicción, los datos obtenidos se re-escalan a su valor original invirtiendo la ecuación anterior.

#### 12.3.4 Conjuntos de entrenamiento y prueba

Los valores de la serie de tiempo se dividen en dos conjuntos de datos:

1. Conjunto de entrenamiento. Lo constituye el conjunto del 80% de los datos de la serie. Se seleccionan de forma consecutiva y ordenada. Este conjunto de datos es el que se utiliza para el entrenamiento de la RNA.
2. Conjunto de prueba. Está formado por el conjunto del 20% de los datos de la serie. Corresponde a los datos restantes, una vez seleccionados los patrones de entrenamiento. Este conjunto de datos se utiliza para evaluar la capacidad de generalización o predicción de la red.

Los conjuntos de entrenamiento y prueba se obtienen a partir de la matriz  $B_r$ . Primero se eliminan los primeros  $r$  renglones de la matriz. A los elementos que no tienen retrasos se les asigna cero (dado que ya no hay más datos anteriores). Posteriormente, la matriz restante se divide en dos partes, una que contiene el 80% de los renglones y otra con el 20% restante, correspondiente a los conjuntos de entrenamiento y prueba respectivamente.

De esta manera, para seleccionar un retraso específico y considerarlo en el entrenamiento basta con incluir la columna correspondiente a tal retraso. La primera columna será invariablemente el objetivo a perseguir, es decir, lo que la red deberá aprender con base en los retrasos proporcionados.

#### 12.3.5 Topología de la red

Se propone una topología de red con las siguientes características:

- Red multicapa unidireccional o perceptrón multicapa (*MLP*), es decir una red con conexiones que pasan las salidas en una única dirección a neuronas en la siguiente capa.



- Red totalmente interconectada en la que la salida de cada neurona es conectada a cada una de las otras neuronas de la siguiente capa.
- Número de entradas:  $p$
- Número de capas ocultas: 1
- Número de neuronas en la capa oculta:  $q$
- Función de transferencia o activación de las neuronas de la capa oculta: tangente hiperbólica.
- Función de transferencia o activación de la capa de salida: lineal.

### **12.3.6 Determinación de las entradas**

Como no existe una regla para determinar qué entradas son las mejores para realizar la predicción, se consideran dos opciones para determinar las entradas de la red neuronal:

- Se sugiere emplear 10 retrasos y como lo sugiere Faraway y Chatfield (1998), analizar los pesos correspondientes a cada entrada para seleccionar las entradas con los pesos de mayor magnitud, además de comparar el error generado tanto en el ajuste como en la predicción con las entradas seleccionadas.
- Una vez determinado el modelo Box Jenkins (ARIMA), se seleccionan como entradas los retrasos correspondientes de este modelo (coeficientes estadísticamente significativos).

### **12.3.7 Determinación del número de neuronas en la capa oculta y entrenamiento**

Al igual que en la determinación de las entradas, no existe una regla general para la obtención de los parámetros de la red que se desempeña mejor en cada situación. Se realiza entonces un procedimiento automático de búsqueda, por medio de un "barrido" de los parámetros que determinan el comportamiento de la red, por lo que, para obtener los parámetros de la red neuronal que arroja el menor error tanto en el ajuste como en la predicción, se

realizan tres ciclos anidados que evalúan diversas redes en donde se establecen como parámetros de variación, el porcentaje de error máximo permitido de 0.01 a 0.1, con incrementos de 0.01, el número de neuronas en la capa oculta de 1 a 10 con incrementos de 1, así como el número de retrasos también de 1 a 10 con incrementos de 1.

La red obtenida de este modo tiene 30 neuronas ocultas divididas en tres capas ocultas de 5, 10 y 15 respectivamente, con 3 retrasos y un porcentaje de error permitido de 0.001.

Gráficamente la red obtenida se puede observar en la figura 3.

#### **12.3.8 Algoritmo de entrenamiento**

Se propone emplear el algoritmo de retropropagación, el cual es por mucho el algoritmo más popular y más ampliamente usado por los investigadores en redes neuronales artificiales además de estar implementado en casi todos los paquetes de software comercial. Su popularidad se debe a la simplicidad en el diseño e implementación. La red neuronal se entrena en particular con el algoritmo Levenberg-Marquardt de Matlab, que es el método más rápido para entrenar redes de hasta varios cientos de pesos.

***newff***: Crea una red tipo Backpropagation (Neuronal Network - Mat Lab).

#### **12.3.9 Selección de pesos iniciales**

La elección de los pesos iniciales es importante y se recomienda probar con diferentes conjuntos de valores iniciales para obtener los mejores resultados. Los pesos iniciales se generan aleatoriamente 50 veces y se selecciona el modelo que tenga menor promedio entre la suma de cuadrados de los errores de ajuste y predicción.

### 12.3.10 Criterio de evaluación

En redes neuronales la función de error más común a minimizar es la suma de los cuadrados de los errores. El error tanto en el ajuste como en la predicción se calcula mediante:

$$S = \sum_t (\hat{X}_t - X_t)^2$$

Donde:

$X_t$  Es el valor de la serie original y  $\hat{X}_t$  es el valor de la serie ajustada

### 12.3.11 Implementación

Por medio de la ecuación de predicción definida por la red neuronal, se obtiene el valor de la predicción en el tiempo  $t+1$ . Se calcula entonces el error de generalización. A partir de los errores de entrenamiento y generalización se comparan las redes obtenidas y se selecciona aquella en la que ambos valores son mínimos. Como ya se explicó anteriormente, no es conveniente que el error de entrenamiento sea muy pequeño en comparación con el error de generalización, pues esto indica un sobreajuste o memorización. La correlación entre los valores originales de la serie y los estimados por la red neuronal puede usarse como una medida de la exactitud de la predicción.

La arquitectura del modelo de red estimada puede observar en la figura 7.

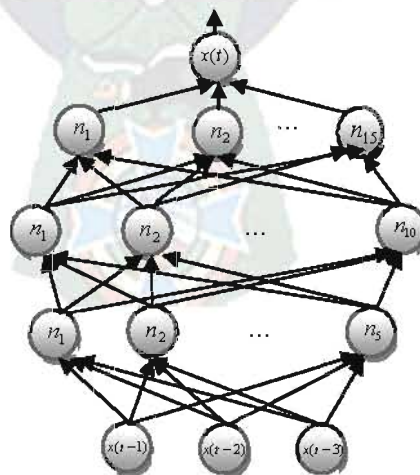
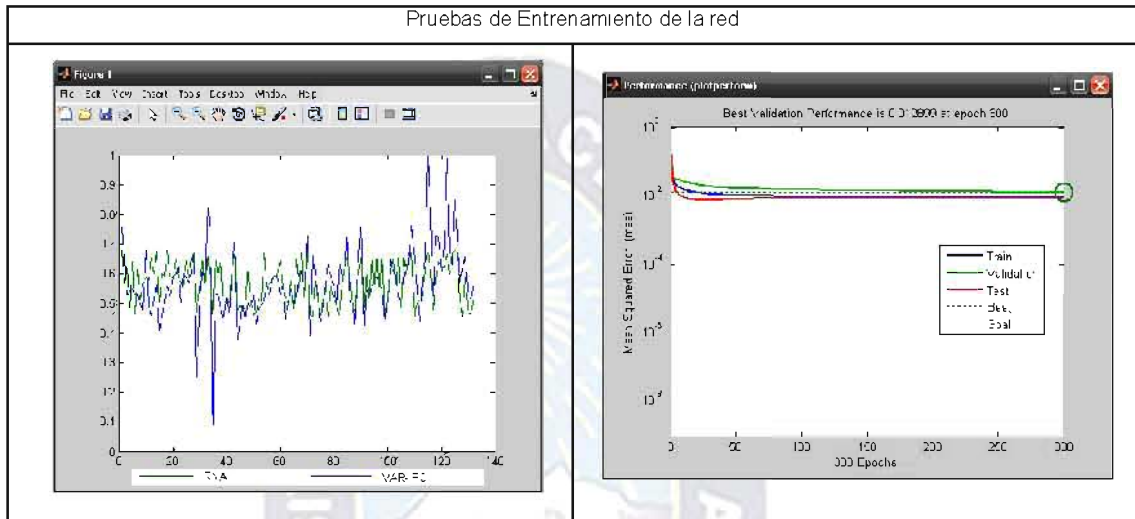
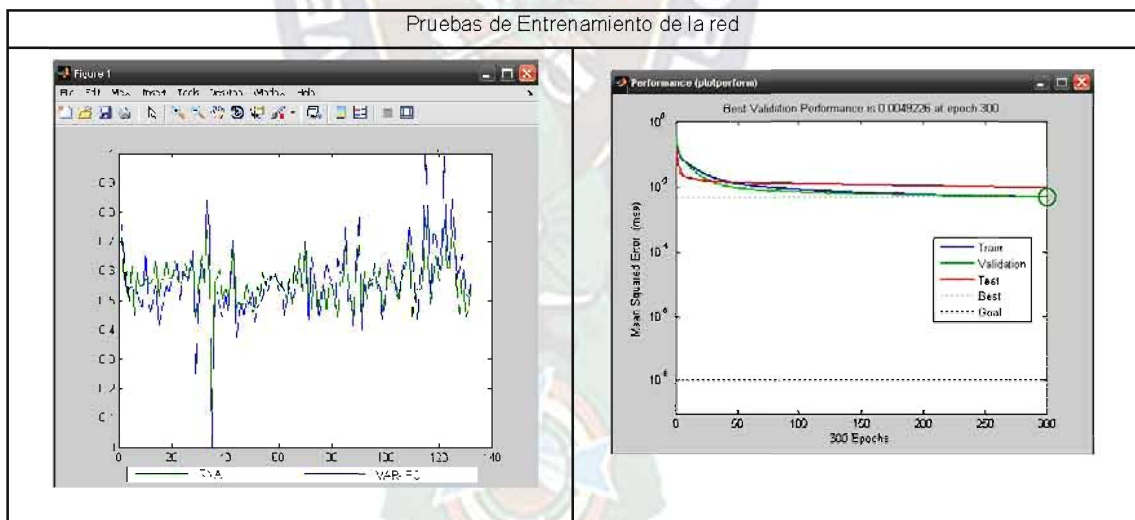


Figura 7. Arquitectura de la RNA que proporciona los mejores resultados Fuente elaboración propia

Una vez seleccionado el modelo, se puede pasar a la etapa de experimentación y predicción, las gráficas de la experimentación y pronóstico se muestra en las Figura 8, 9, 10, 11, 12.



**Figura 8.** Primera prueba con un error  $e=0.012899$  en 300 épocas



**Figura 9.** Segunda prueba con un error  $e=0.0049226$  en 300 épocas

Pruebas de Entrenamiento de la red

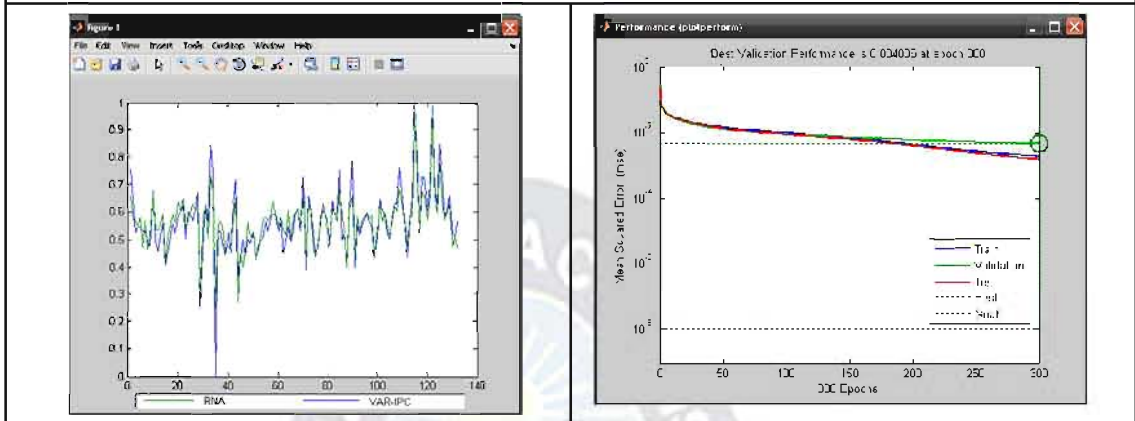


Figura 10. Prueba X, con un error  $e=0.004886$  en 300 épocas.

Pruebas de Entrenamiento de la red

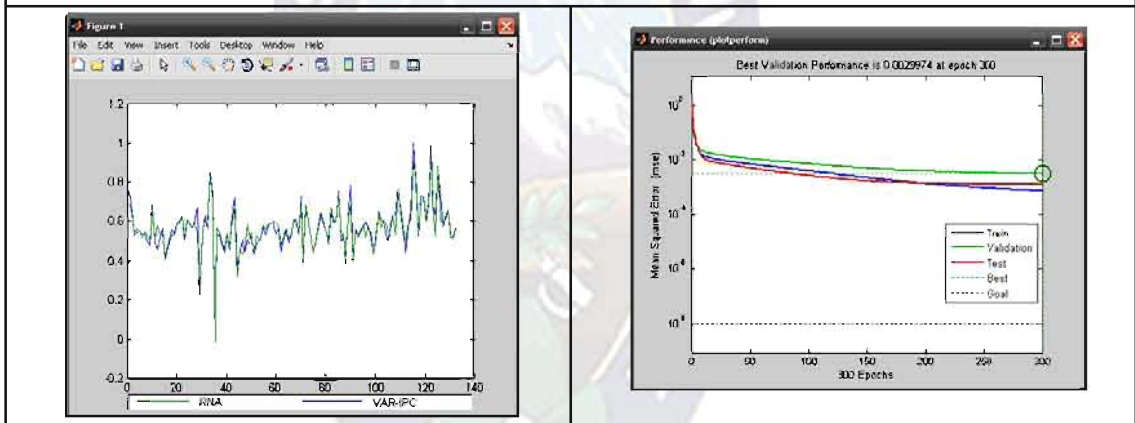


Figura 11. Prueba X+1, con un error  $e=0.002974$  en 300 épocas

Pruebas de Entrenamiento de la red

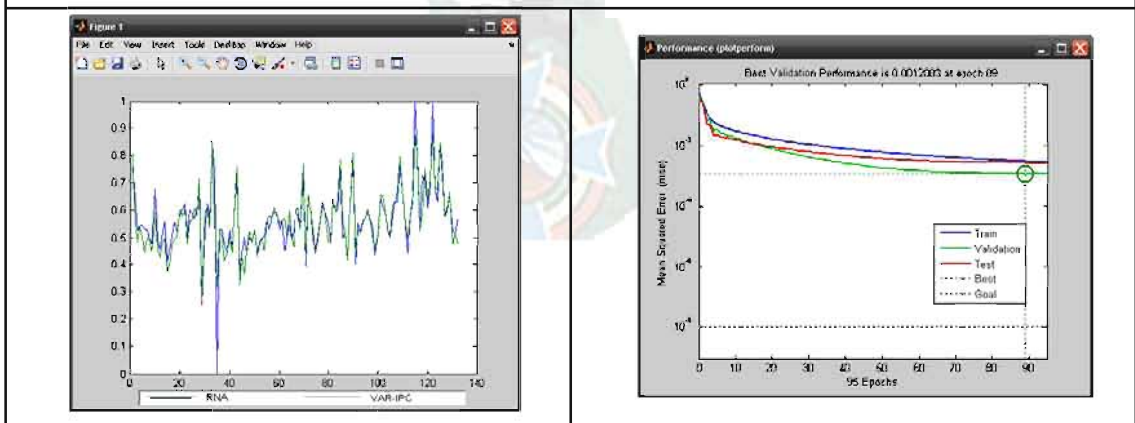


Figura 12. Prueba X+2, con un error  $e=0.0012083$  en 89 épocas

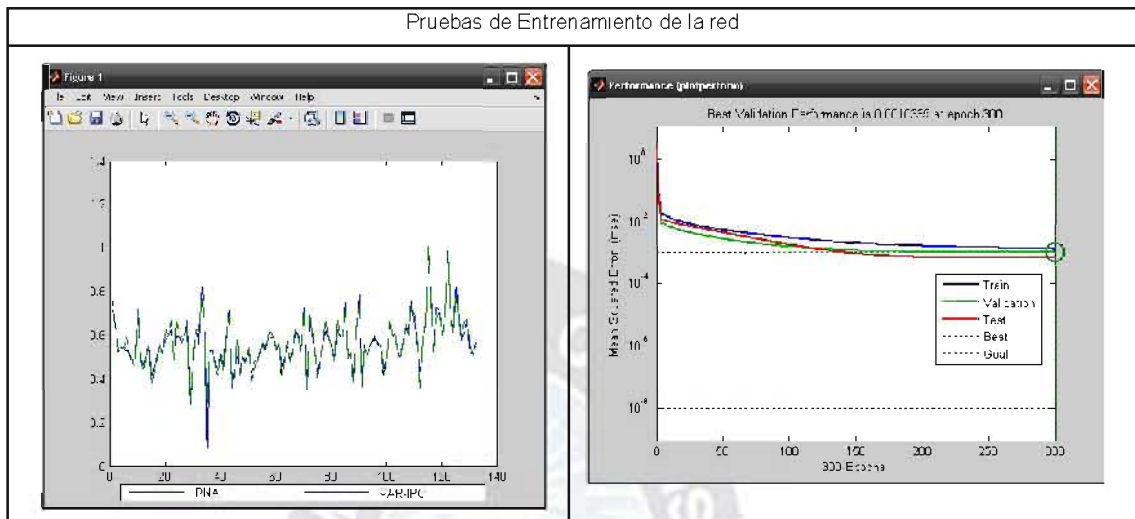


Figura 13. Prueba que mejor ajusta con un error  $e=0.0010339$  en 300 épocas

El menor error cuadrático está representado por la performance, se puede apreciar en el en la figura 13, en el último cuadro con resultados de 0.0010339 o un 0.01% un error mínimo cuadrático bastante aceptable con lo que revela al modelo neuronal con tres capas ocultas.

### 12.3.12 Resultados obtenidos

Resultados obtenidos por medio del modelo Neuronal toolbox de Mat Lab, XLSTAT, INE:

FECHA	ARIMA	RNA	INE
01/2009	0.395	0.37521	0.36
02/2009	-0.405	-0.52415	-0.07
03/2009	-0.407	-0.64571	-0.49
04/2009	-0.407	-0.45124	-0.43

Tabla 4. Pronóstico realizado mediante el modelo RNA, ARIMA, y El INE primer trimestre

### 12.3.13 Características de las series pronosticadas por ARIMA y RNA

Los pronósticos de la series pronosticadas por ARIMA y RNA en niveles se muestran en las figuras 14 y15, de la que se puede observar que el modelo obtenido por RNA pronostica mejor los datos en las primeras 300 observaciones, mientras que después, el modelo ARIMA sigue más de cerca

la tendencia original, sin embargo, la forma de la serie obtenida por RNA es más parecida a la original, aunque sobrevalúa los datos.

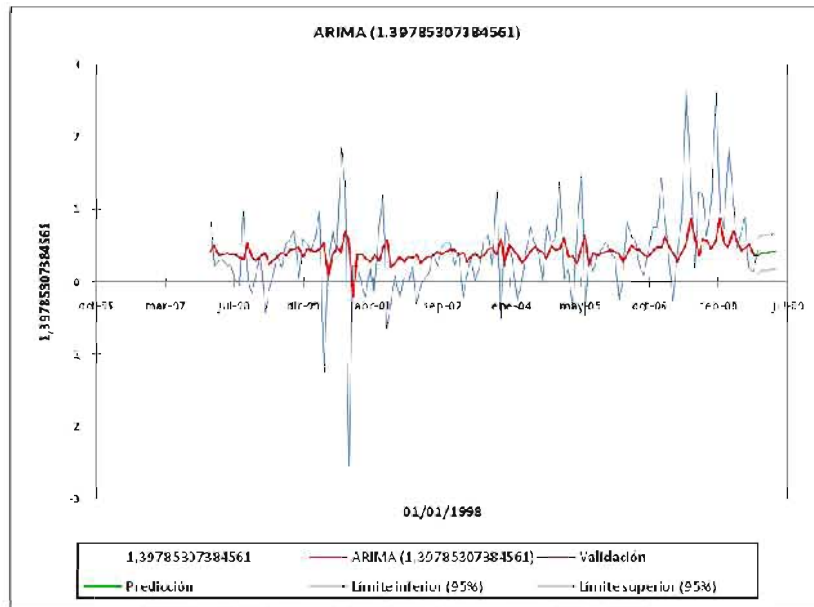


Figura 14. Grafica Obtenida por el modelo ARIMA con XLSTAT.

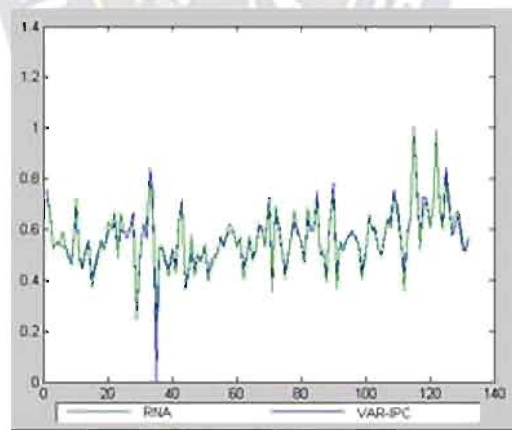


Figura 15. Grafica Obtenida por el modelo RNA con Mat Lab

Se evalúa también si los valores generados los ambos modelos son capaces de reproducir los primeros cuatro momentos de la distribución de los valores de la serie suavizada exponencialmente sobre los 112 datos pronosticados.

De la comparación entre los momentos de la distribución de los valores observados y los correspondientes pronósticos (tabla 4), se desprende que la media, la mediana y la desviación estándar de los valores del pronóstico realizado con el modelo ARIMA se encuentran más cerca de los valores correspondientes a la distribución de los valores observados para la serie original, sobre todo por la sobre valuación de los datos a partir del dato pronosticado.

	VARIACION	RNA	ARIMA
Media	0.414693	0.255925	0.277622
Mediana	0.369145	0.24117	0.260615
Máximo	2.676557	0.922450	1.139140
Mínimo	-2.547326	-0.857840	-0.861100
Desviación estándar	0.630234	0.21459722	0.24532489
Curtosis	4.996355	5.926982	4.529203

**Tabla 5.** Comparación de estadísticos RNA y ARIMA





#### Capítulo IV

### CONCLUSIONES Y RECOMENDACIONES

### 13. CONCLUSIONES

A partir del desarrollo de ambos modelos, de los resultados obtenidos es posible concluir lo siguiente:

Existen varios modelos de predicción de series de tiempo, algunos más sofisticados que los desarrollados en este trabajo de investigación, como la familia de modelos ARCH, sin embargo, para poder realizar un estudio comparativo entre distintas técnicas, es necesario hacerlo sobre las mismas bases, es decir, en este estudio se empleó una configuración de red neuronal equivalente a un modelo autorregresivo (AR) no lineal, razón por la cual se utilizó un modelo ARIMA como referencia.

Es bueno contar con un modelo que realice predicciones más precisas, ya que conocer el comportamiento futuro, o al menos contar con una aproximación más cercana a la real, nos brinda la posibilidad de tomar mejores decisiones en un ambiente incierto como lo es la economía.

Los resultados, las pruebas y los estadísticos indican que el modelo basado en RNA es mejor que el modelo ARIMA calculado y empleado como referencia para dentro de muestra.

Se puede decir entonces que el método cumple su objetivo de generar mejores pronósticos de la tendencia de la variación del IPC una medida de la inflación en comparación con el modelo ARIMA usado como referencia, pero sólo en la primeros datos del intervalo de predicción ya que posteriormente el modelo arroja valores que sobrevalúan a los originales e incluso el ARIMA proporciona mejores resultados en esta situación.

El modelo propuesto puede ser empleado como una herramienta adicional a las ya disponibles, más no como sustituto, pues el modelo lineal es más sencillo de generar y eventualmente genera buenas aproximaciones.

## 14. RECOMENDACIONES

Como trabajo futuro se puede recomendar el empleo de un modelo de redes neuronales que tenga la configuración de un modelo ARCH de manera que pueda captar de mejor manera la volatilidad existente en las series financieras en tiempos cortos, y aunque el modelo propuesto también es capaz de ser comparado de esta forma, no se efectuaría bajo las mismas consideraciones.

Debido a que se requiere bastante experiencia para determinar la arquitectura apropiada de una RNA, sería muy provechoso la creación de un sistema experto que ayude a identificar las entradas más adecuadas para la red, mediante el uso de todos los aspectos que involucra el diseño de una RNA para la predicción con series de tiempo.

Dado que el modelo de RNA diseñado es puramente no lineal, sería conveniente utilizar un modelo mixto que incluya la parte lineal de manera simultánea, para aprovechar las ventajas de ambos enfoques en un mismo modelo.

Otra aplicación interesante sería la evaluación de varias RNA aplicadas a mercados financieros como el de divisas, el bolsín de valores en el caso boliviano.

## BIBLIOGRAFÍA

**Título:** Inteligencia artificial Sistemas Expertos. Alianza editorial Madrid, 1986 (Biblioteca De Informática).

**Título:** Redes neuronales, algoritmos, aplicaciones y técnicas de propagación.

**Autor:** Jaime A Freeman, David M Skapura.

**Título:** Tesis de Grado Predicción de la demanda Eléctrica Utilizando redes Neuronales Artificiales.

**Autor:** Juan Pablo Quispe Poma (UMSA Informática).

**Título:** Pronósticos en el mercado de derivados utilizando redes neuronales

**Autor:** Ivan Cruz Torrez documento en pdf (UNAM).

**Título:** Pronóstico y series de tiempo: un acercamiento aplicado (3ª. Ed.), California: edición en español.

**Autor:** Bowerman, B. y O'Connell, R. (1993),

**Título:** "Análisis y predicción de la serie de tiempo del precio externo del café colombiano utilizando redes neuronales artificiales".

**Autor:** Garcia, I. (2003).

**Título:** Econometría, 3ª edición, Santa Fe de Bogotá: McGraw-Hill, 824 págs.

**Autor:** GUJARATI, D.N., (1997).

**Título:** Problemas de Econometría, 4ª edición, Madrid: Pirámide, 496 págs.

**Autor:** AZNAR, A. y GARCÍA, A., (1991).

## REFERENCIAS ELECTRÓNICAS

<http://grial.uc3m.es/~docweb/doctorado/rn/bibliografia.html>.

[http://fciencias.ens.uabc.mx/notas\\_cursos/inteligencia/neurona/biblio.htm](http://fciencias.ens.uabc.mx/notas_cursos/inteligencia/neurona/biblio.htm).

[http://www2.uca.es/dept/leng\\_sist\\_informaticos/preal/23041/default.htm](http://www2.uca.es/dept/leng_sist_informaticos/preal/23041/default.htm).

<http://www.eumed.net/libros/2007a/233/4d.htm>.

[http://www2.uca.es/dept/leng\\_sist\\_informaticos/preal/23041/transpas/E-Backpropagation/index.htm](http://www2.uca.es/dept/leng_sist_informaticos/preal/23041/transpas/E-Backpropagation/index.htm)

<http://www.xlstat.com/es/support/tutorials/hw.htm>.

## APÉNDICE A

### Resultados obtenidos por el software econométrico XLSTAT:

XLSTAT 2009 3 02 - ARIMA - 2009  
 Parámetros del modelo  $p = 1 / d = 0 / q = 0 / P = 0 / D = 0 / Q = 0 / s = 0$   
 Optimizar Mínimos cuadrados (Convergencia = 0,00001 / Iteraciones = 5000)

Validación 1  
 Predicción 4  
 Intervalos de confianza (%) 95

Estadísticas descriptivas										
Variable	Observaciones	Obs. con datos perdidos	Obs. sin datos perdidos	Mínimo	Máximo	Media	Desviación típica			
139 785 307 384 561	131	0	131	-2,547	2,677	0,407	0,627			

#### Resultados de la modelización ARIMA de la serie 1,39785307384561:

Estimación preliminar Yule-Walker (1,39785307384561)

##### Coefficientes de ajuste

Observaciones	130,000
SEC	48,896
MAPE(Dif)	154,260
MAPE	230,532
Varianza RB	0,376
Varianza RB (estimada)	0,376
-2Log(Vero)	241,849
FPE	0,382
AIC	245,849
AICC	245,943
SBC	251,584

##### Parámetros del modelo

Constante	0,323
-----------	-------

Parámetro	Valor	Desviación típica método	Límite inferior (95%)	Límite superior (95%)	Desviación típica asint	Límite inferior (95%)	Límite superior (95%)
AR(1)	0,205	0,086	0,037	0,373	0,053	0,102	0,308

##### Predicciones y residuos

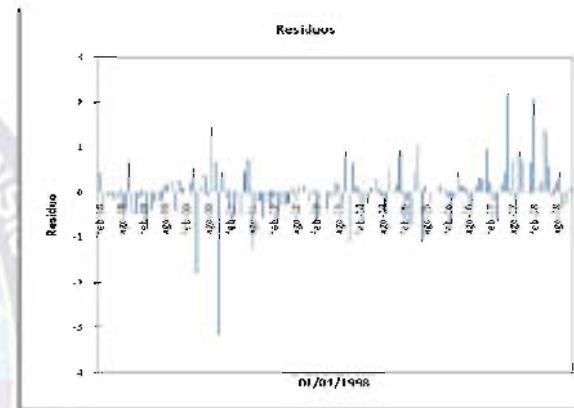
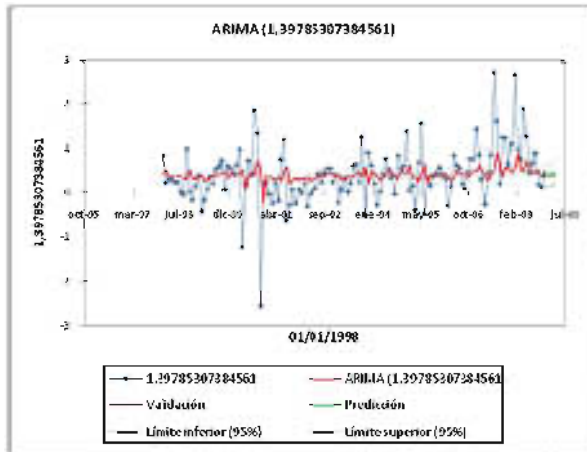
Observaciones	1,39785307384561	ARIMA(1,39785307384561)	Residuos	Residuos estandarizados	Desviación típica	Límite inferior (95%)	Límite superior (95%)
feb-98	0,830	0,416	0,414	1,099			
mar-98	0,211	0,494	-0,283	-0,753			
abr-98	0,278	0,367	-0,089	-0,237			
may-98	0,312	0,380	-0,068	-0,181			
jun-98	0,227	0,388	-0,161	-0,428			
jul-98	0,230	0,370	-0,139	-0,371			
ago-98	0,027	0,371	-0,344	-0,914			
sep-98	-0,051	0,329	-0,380	-1,011			
oct-98	0,987	0,313	0,674	1,793			
nov-98	0,028	0,526	-0,498	-1,324			
dic-98	-0,158	0,329	-0,487	-1,296			
ene-99	0,146	0,291	-0,145	-0,384			
feb-99	0,380	0,354	0,026	0,070			

mar-99	-0,420	0,401	-0,821	-2,184
abr-99	-0,152	0,237	-0,389	-1,035
may-99	0,089	0,292	-0,203	-0,540
jun-99	0,366	0,342	0,024	0,064
jul-99	0,199	0,399	-0,200	-0,530
ago-99	0,520	0,364	0,156	0,414
sep-99	0,594	0,430	0,164	0,436
oct-99	0,715	0,445	0,269	0,716
nov-99	0,067	0,470	-0,403	-1,071
dic-99	0,591	0,337	0,254	0,676
ene-00	0,525	0,445	0,080	0,213
feb-00	0,415	0,431	-0,016	-0,042
mar-00	0,612	0,409	0,203	0,540
abr-00	0,965	0,449	0,516	1,371
may-00	-1,245	0,522	-1,767	-4,696
jun-00	0,192	0,068	0,124	0,329
jul-00	0,720	0,363	0,357	0,949
ago-00	0,395	0,471	-0,076	-0,202
sep-00	1,843	0,405	1,439	3,825
oct-00	1,338	0,702	0,636	1,690
nov-00	-2,547	0,598	-3,145	-8,361
dic-00	0,221	-0,199	0,421	1,118
ene-01	0,225	0,369	-0,144	-0,383
feb-01	-0,026	0,370	-0,395	-1,051
mar-01	-0,226	0,318	-0,544	-1,446
abr-01	0,211	0,277	-0,066	-0,175
may-01	-0,180	0,367	-0,547	-1,455
jun-01	0,737	0,286	0,451	1,199
jul-01	1,195	0,475	0,720	1,915
ago-01	-0,655	0,569	-1,224	-3,254
sep-01	-0,262	0,189	-0,451	-1,200
oct-01	0,073	0,270	-0,196	-0,522
nov-01	-0,226	-0,339	-0,565	-1,501
dic-01	0,067	0,277	-0,210	-0,558
ene-02	-0,006	0,337	-0,344	-0,914
feb-02	0,212	0,322	-0,110	-0,292
mar-02	-0,306	0,367	-0,673	-1,790
abr-02	-0,031	0,261	-0,292	-0,775
may-02	0,044	0,317	-0,273	-0,726
jun-02	0,106	0,332	-0,226	-0,601
jul-02	0,410	0,345	0,064	0,171
ago-02	0,223	0,408	-0,184	-0,490
sep-02	0,460	0,369	0,090	0,240
oct-02	0,541	0,418	0,123	0,327
nov-02	0,536	0,435	0,101	0,269
dic-02	0,235	0,433	-0,199	-0,529
ene-03	0,397	0,372	0,025	0,066
feb-03	-0,213	0,405	-0,618	-1,642
mar-03	0,060	0,280	-0,219	-0,583
abr-03	0,316	0,336	-0,020	-0,053
may-03	0,016	0,388	-0,373	-0,991
jun-03	0,204	0,327	-0,123	-0,326
jul-03	0,595	0,365	0,230	0,611
ago-03	0,644	0,446	0,199	0,528
sep-03	0,232	0,456	-0,223	-0,594
oct-03	1,249	0,371	0,878	2,335
nov-03	-0,515	0,580	-1,095	-2,910
dic-03	0,896	0,218	0,678	1,802
ene-04	0,618	0,507	0,111	0,294

feb-04	0,191	0,450	-0,259	-0,689			
mar-04	-0,267	0,363	-0,629	-1,673			
abr-04	0,022	0,269	-0,246	-0,655			
may-04	0,410	0,328	0,081	0,217			
jun-04	0,744	0,408	0,337	0,895			
jul-04	0,526	0,476	0,050	0,132			
ago-04	-0,353	0,431	-0,078	-0,208			
sep-04	-0,026	0,396	-0,422	-1,121			
oct-04	0,812	0,318	0,494	1,314			
nov-04	0,546	0,490	0,056	0,148			
dic-04	0,605	0,435	0,169	0,450			
ene-05	1,372	0,448	0,924	2,456			
feb-05	0,051	0,605	-0,554	-1,472			
mar-05	0,158	0,334	-0,176	-0,467			
abr-05	-0,376	0,356	-0,732	-1,944			
may-05	0,675	0,246	0,428	1,138			
jun-05	1,539	0,462	1,077	2,864			
jul-05	-0,464	0,639	-1,104	-2,934			
ago-05	0,350	0,228	0,122	0,324			
sep-05	0,150	0,395	-0,245	-0,651			
oct-05	0,373	0,354	0,018	0,048			
nov-05	0,437	0,400	0,037	0,100			
dic-05	0,556	0,413	0,142	0,378			
ene-06	0,404	0,438	-0,034	-0,090			
feb-06	0,312	0,406	-0,095	-0,252			
mar-06	-0,277	0,387	-0,665	-1,767			
abr-06	0,140	0,267	-0,126	-0,336			
may-06	0,836	0,352	0,484	1,287			
jun-06	0,623	0,495	0,128	0,340			
jul-06	0,563	0,451	0,111	0,296			
ago-06	0,189	0,439	-0,250	-0,666			
sep-06	0,081	0,362	-0,281	-0,747			
oct-06	0,467	0,340	0,127	0,338			
nov-06	0,752	0,419	0,332	0,883			
dic-06	0,754	0,478	0,276	0,733			
ene-07	1,435	0,478	0,956	2,542			
feb-07	0,834	0,618	0,216	0,674			
mar-07	0,297	0,495	-0,197	-0,524			
abr-07	-0,256	0,385	-0,641	-1,703			
may-07	0,454	0,271	0,183	0,486			
jun-07	0,846	0,417	0,429	1,140			
jul-07	2,677	0,497	2,180	5,793			
ago-07	1,591	0,873	0,718	1,908			
sep-07	0,190	0,650	-0,460	-1,223			
oct-07	1,247	0,362	0,884	2,351			
nov-07	1,228	0,579	0,649	1,724			
dic-07	0,632	0,576	0,056	0,150			
ene-08	1,090	0,453	0,637	1,694			
feb-08	2,620	0,547	2,073	5,510			
mar-08	0,971	0,861	0,110	0,293			
abr-08	0,740	0,523	0,217	0,577			
may-08	1,870	0,475	1,395	3,707			
jun-08	1,260	0,707	0,553	1,469			
jul-08	0,460	0,582	-0,122	-0,325			
ago-08	0,650	0,418	0,232	0,617			
sep-08	0,880	0,457	0,423	1,125			
oct-08	0,180	0,504	-0,324	-0,861			
nov-08	0,130	0,360	-0,230	-0,612			
dic-08	0,430	0,350	0,080	0,212	0,126	0,103	0,597

dic-08		0,395
ene-09		0,405
mar-09		0,407
mar-09		0,407

	0,126	0,149	0,642
	0,126	0,158	0,651
	0,126	0,160	0,653
	0,126	0,160	0,654



Resultados tras la optimización (1,39785307384561)

**Coefficientes de ajuste**

Observaciones	130,000
SEC	48,896
MAPE(Dif)	154,603
MAPE	230,421
Varianza RB	0,379
Varianza RB (estimada)	0,376
-2Log(Vero)	242,853
FPE	0,385
AIC	246,853
AICC	246,947
SBC	252,588

**Parámetros del modelo**

Constante	0,323
-----------	-------

Parámetro	Valor	Desviación típica Hess	Límite inferior (95%)	Límite superior (95%)	Desviación típica asint	Límite inferior (95%)	Límite superior (95%)
AR(1)	0,206	0,012	0,182	0,230	0,053	0,103	0,309

**Predicciones y residuos**

Observaciones	1,39785307384561	ARIMA(1,39785307384561)	Residuos	Residuos estandarizados	Desviación típica	Límite inferior (95%)	Límite superior (95%)
feb-98	0,830	0,416	0,413	1,099			
mar-98	0,211	0,494	-0,284	-0,754			
abr-98	0,278	0,366	-0,089	-0,236			
may-98	0,312	0,380	-0,068	-0,181			
jun-98	0,227	0,387	-0,161	-0,428			



jul-98	0,230	0,370	-0,139	-0,370
ago-98	0,027	0,371	-0,344	-0,913
sep-98	-0,051	0,329	-0,380	-1,010
oct-98	0,987	0,312	0,675	1,794
nov-98	0,028	0,527	-0,499	-1,326
dic-98	-0,158	0,329	-0,487	-1,295
ene-99	0,146	0,290	-0,144	-0,383
feb-99	0,380	0,353	0,026	0,070
mar-99	-0,420	0,401	-0,821	-2,183
abr-99	-0,152	0,236	-0,388	-1,032
may-99	0,089	0,292	-0,203	-0,539
jun-99	0,366	0,341	0,024	0,065
jul-99	0,199	0,399	-0,200	-0,530
ago-99	0,520	0,364	0,156	0,414
sep-99	0,594	0,430	0,164	0,436
oct-99	0,715	0,446	0,269	0,715
nov-99	0,067	0,470	-0,403	-1,072
dic-99	0,591	0,337	0,255	0,677
ene-00	0,525	0,445	0,080	0,213
feb-00	0,415	0,431	-0,016	-0,043
mar-00	0,612	0,409	0,203	0,540
abr-00	0,965	0,449	0,516	1,371
may-00	-1,245	0,522	-1,767	-4,697
jun-00	0,192	0,066	0,126	0,334
jul-00	0,720	0,363	0,357	0,949
ago-00	0,395	0,472	-0,076	-0,203
sep-00	1,843	0,405	1,439	3,825
oct-00	1,338	0,703	0,634	1,686
nov-00	-2,547	0,599	-3,146	-8,363
dic-00	0,221	-0,202	0,424	1,126
ene-01	0,225	0,369	-0,144	-0,382
feb-01	-0,026	0,369	-0,395	-1,050
mar-01	-0,226	0,318	-0,543	-1,444
abr-01	0,211	0,277	-0,065	-0,173
may-01	-0,180	0,367	-0,547	-1,454
jun-01	0,737	0,286	0,452	1,200
jul-01	1,195	0,475	0,720	1,914
ago-01	-0,655	0,570	-1,225	-3,256
sep-01	-0,262	0,188	-0,450	-1,197
oct-01	0,073	0,269	-0,196	-0,520
nov-01	-0,226	0,338	-0,564	-1,500
dic-01	0,067	0,276	-0,209	-0,556
ene-02	-0,006	0,337	-0,343	-0,913
feb-02	0,212	0,322	-0,109	-0,291
mar-02	-0,306	0,367	-0,673	-1,790
abr-02	-0,031	0,260	-0,291	-0,773
may-02	0,044	0,317	-0,273	-0,725
jun-02	0,106	0,332	-0,226	-0,600
jul-02	0,410	0,345	0,065	0,172
ago-02	0,223	0,408	-0,184	-0,490
sep-02	0,460	0,369	0,090	0,240
oct-02	0,541	0,418	0,123	0,327
nov-02	0,536	0,435	0,101	0,269
dic-02	0,235	0,434	-0,199	-0,529
ene-03	0,397	0,371	0,025	0,067
feb-03	-0,213	0,405	-0,618	-1,642
mar-03	0,060	0,279	-0,219	-0,581
abr-03	0,316	0,336	-0,020	-0,052
may-03	0,016	0,388	-0,373	-0,990

jun-03	0,204	0,326	-0,122	-0,325
jul-03	0,595	0,365	0,230	0,612
ago-03	0,644	0,446	0,199	0,528
sep-03	0,232	0,456	-0,224	-0,594
oct-03	1,249	0,371	0,878	2,335
nov-03	-0,515	0,581	-1,096	-2,912
dic-03	0,896	0,217	0,679	1,804
ene-04	0,618	0,508	0,110	0,293
feb-04	0,191	0,451	-0,259	-0,689
mar-04	-0,267	0,363	-0,629	-1,672
abr-04	0,022	0,268	-0,246	-0,653
may-04	0,410	0,328	0,082	0,218
jun-04	0,744	0,408	0,337	0,895
jul-04	0,526	0,477	0,049	0,131
ago-04	0,353	0,432	-0,078	-0,209
sep-04	-0,026	0,396	-0,422	-1,121
oct-04	0,812	0,318	0,495	1,315
nov-04	0,546	0,491	0,055	0,147
dic-04	0,605	0,436	0,169	0,450
ene-05	1,372	0,448	0,924	2,456
feb-05	0,051	0,606	-0,555	-1,475
mar-05	0,158	0,334	-0,175	-0,466
abr-05	-0,376	0,358	-0,731	-1,944
may-05	0,675	0,246	0,429	1,140
jun-05	1,539	0,462	1,077	2,863
jul-05	-0,464	0,641	-1,105	-2,937
ago-05	0,350	0,227	0,123	0,327
sep-05	0,150	0,395	-0,245	-0,651
oct-05	0,373	0,354	0,018	0,049
nov-05	0,437	0,400	0,037	0,100
dic-05	0,556	0,413	0,142	0,378
ene-06	0,404	0,438	-0,034	-0,090
feb-06	0,312	0,406	-0,095	-0,252
mar-06	-0,277	0,387	-0,665	-1,767
abr-06	0,140	0,266	-0,126	-0,334
may-06	0,836	0,352	0,484	1,288
jun-06	0,623	0,496	0,127	0,339
jul-06	0,563	0,452	0,111	0,295
ago-06	0,189	0,439	-0,251	-0,666
sep-06	0,081	0,362	-0,281	-0,747
oct-06	0,467	0,340	0,128	0,339
nov-06	0,752	0,419	0,332	0,883
dic-06	0,754	0,478	0,276	0,732
ene-07	1,435	0,479	0,956	2,541
feb-07	0,834	0,619	0,215	0,571
mar-07	0,297	0,495	-0,198	-0,525
abr-07	-0,256	0,384	-0,641	-1,703
may-07	0,454	0,270	0,184	0,488
jun-07	0,846	0,417	0,429	1,140
jul-07	2,677	0,497	2,179	5,792
ago-07	1,591	0,875	0,715	1,902
sep-07	0,190	0,651	-0,461	-1,227
oct-07	1,247	0,362	0,885	2,351
nov-07	1,228	0,580	0,648	1,722
dic-07	0,632	0,576	0,056	0,148
ene-08	1,090	0,453	0,637	1,693
feb-08	2,620	0,548	2,072	5,508
mar-08	0,971	0,864	0,108	0,287
abr-08	0,740	0,523	0,217	0,576

may-08	1,870	0,476	1,394	3,706			
jun-08	1,260	0,709	0,551	1,465			
jul-08	0,460	0,583	-0,123	-0,327			
ago-08	0,650	0,418	0,232	0,617			
sep-08	0,880	0,457	0,423	1,124			
oct-08	0,180	0,505	-0,325	-0,863			
nov-08	0,130	0,360	-0,230	-0,612			
dic-08	0,430	0,350	0,080	0,213	0,127	0,102	0,598
dic-08		0,395			0,127	0,147	0,643
ene-09		0,405			0,127	0,157	0,653
mar-09		0,407			0,127	0,159	0,654
mar-09		0,407			0,127	0,159	0,655

