



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

FACULTAD DE CIENCIAS

**ANÁLISIS NO LINEAL DE UNA SERIE DE TIEMPO
BURSÁTIL**

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

ACTUARIO

P R E S E N T A:

GERARDO LAZARO AYALA SILVA



**DIRECTOR DE TESIS:
DR. PEDRO MIRAMONTES VIDAL
2014**



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Análisis no lineal de una serie de tiempo bursátil

Gerardo Ayala

9 de marzo de 2014

Índice general

Introducción	5
1. Las redes de neuronas artificiales	7
1.1. Aspectos fisiológicos de las neuronas	7
1.2. Modelo de la neurona elemental	10
1.3. Perceptrón	12
1.4. Perceptrón multicapa	21
1.5. Algoritmo de retropropagación	23
2. El sistema financiero	25
2.1. Los bancos	26
2.2. Mercados de deuda	27
2.3. Mercado cambiario	29
2.4. Mercados de acciones o valores	31
2.4.1. Los mercados eficientes y los mercados emergentes . . .	34
2.4.2. El Índice de Precios y Cotizaciones (IPC)	36
3. Antecedentes del uso de las redes neuronales artificiales en diversas áreas de negocio y en los índices de los mercados de valores	41
4. Metodología	49
4.1. Metodología	49
4.1.1. Selección de las variables	52
4.1.2. Recolección de los datos	52
4.1.3. Preparación de los datos	53
4.1.4. Conjuntos de entrenamiento, prueba y validación . . .	53
4.1.5. Aplicación de los paradigmas de las redes neuronales .	54

4.1.6. Criterios de evaluación	55
4.1.7. Entrenamiento de la red neuronal e implementación . .	56
4.2. Resultados	58
Conclusiones	59
Apéndices	61
A. Descripción detallada de los datos	61
A.1. Datos fuente IPC	61
A.2. Hoja para generación aleatoria	62
B. Código Matlab	65
Bibliografía	67
Índice de figuras	71
Índice de tablas	73

Introducción

El campo de aplicación de las redes neuronales artificiales (RNA), es muy extenso y, cada día se amplía más, gracias al crecimiento del poder de cómputo, la disponibilidad de las tecnologías de la información y los trabajos de los investigadores en cada rama de las ciencias. Particularmente, en las últimas décadas, se ha incrementado su aplicación en el estudio de las variables económicas que sirven para medir el comportamiento de los índices de las Bolsas de Valores en el mundo.

Estos índices se forman tomando las acciones o componentes más representativos y su comportamiento en los días de operación. Este valor numérico sirve como guía para las transacciones en los mercados y se busca que su generación sea lo menos sesgada posible para evitar especulaciones; además, se hace del conocimiento público para crear certeza en la situación económica del país. En México, desde el 30 de octubre de 1978, se establece como guía para las operaciones financieras de la Bolsa Mexicana de Valores (BMV) el Índice de Precios y Cotizaciones (IPC).

Uno de los indicadores que mide el comportamiento de la BMV es el comparativo del cierre del día contra el cierre del día hábil inmediato anterior representado por un número que puede ser positivo o negativo. Si es positivo, se puede decir en forma genérica, que las acciones participantes ese día tuvieron una ganancia y, por el contrario, si resulta negativo, entonces se habla de una pérdida. He aquí la importancia de poder lograr una predicción confiable del comportamiento del índice. Se ha encontrado que el uso de una RNA, es una herramienta útil para predecir el signo de la variación del retraso de la variable de tal manera que se pueda indicar una jornada favorable o desfavorable.

Éste es precisamente el objetivo del presente trabajo: lograr una predicción del signo de la variación a un día del IPC de la BMV con base en la aplicación de las RNA. Para este propósito, y con la ayuda del software Mat-

lab, se construyó una red neuronal con base en los resultados del día de la BMV comprendidos del 3 de enero de 2000 hasta el 6 de julio de 2012, con un total de 3,151 registros.

Por lo tanto, además de mostrar resultados recientes se conocerá el grado de confiabilidad en el uso de una herramienta como las RNA para la predicción del signo de la variación a un día del IPC de la BMV.

En el capítulo 1 se expone una descripción de los aspectos morfológicos y fisiológicos de las neuronas, así como el por qué surge la idea de usar el funcionamiento de ellas para la resolución de problemas. Se verá el surgimiento de los primeros modelos neuronales con la neurona elemental, su evolución al modelo del perceptrón y del perceptrón multicapa. Además, se presenta la descripción del algoritmo de retropropagación, mismo que se usará como técnica de la RNA usada en el capítulo 4.

El capítulo 2 se refiere al surgimiento del mercado financiero y sus componentes actuales: bancos, mercados cambiarios, de deuda y de acciones o valores. En él se expone, además, la definición de los mercados eficientes y los emergentes como teorías principales del comportamiento de los mercados financieros y como una ayuda para entender el comportamiento de la BMV.

Se hace también una presentación detallada del procedimiento para generar el IPC con lo que se pretende entender mejor la composición del mismo y cómo esta se relaciona directamente con el objetivo de la presente investigación.

Posteriormente, en el capítulo 3, se exponen algunos ejemplos de las diversas aplicaciones de las RNA con todas las referencias pertinentes del caso.

Asimismo, se verá la amplia gama de posibilidades de solución que ofrece el uso de una red neuronal a través del entendimiento del problema.

En el capítulo 4 se plantea la metodología para la aplicación de una RNA. En este sentido, se presenta desde la selección de las variables hasta la implementación del modelo, pasando por la recolección y preparación de los datos; la creación de los conjuntos de entrenamiento, prueba y validación y, por último, los paradigmas de las RNA junto con sus criterios de evaluación. Con esta metodología se formaliza la creación del modelo para éste caso de estudio y, hacia el final del capítulo, se realiza la presentación de los resultados obtenidos.

Finalmente, en las conclusiones se expone el análisis de la aplicación del modelo generado, en relación con los resultados obtenidos, para, por una parte, definir los alcances al mismo y, por otra, establecer el papel que el modelo propuesto podría jugar en investigaciones posteriores en este campo.

Capítulo 1

Las redes de neuronas artificiales

Desde los albores del siglo XX y, a lo largo de la historia, se ha estudiado el cerebro humano para poder entender la lógica de su funcionamiento y lograr modelos factibles de ser emulados en otras áreas de conocimiento.

Hoy en día se sabe que el componente básico del cerebro, y del sistema nervioso, es la neurona. La extraordinaria precisión en el funcionamiento de éstas ha sido fuente y motivo central de grandes investigaciones.

Las neuronas, que son las encargadas de recibir y transmitir la información que, posteriormente, se convertirá en estímulo para realizar una tarea específica, se presentan en una gran variedad de formas y tamaños. Existen neuronas que miden tan solo unas micras y otras, como las motoneuronas espinales, que por contar con axones que llegan hasta la punta del pie, pueden llegar a medir mas de un metro de largo.

La estructura típica de una neurona es: un cuerpo celular o soma; unos brazos o dendritas por donde entra la señal; un axón o terminación prolongada por donde corre la señal y unas terminaciones de tipo arbóreo llamada terminales del axón en donde se realiza la sinapsis o conexión con otras neuronas o con otras células. Un esquema sencillo de una neurona se puede apreciar en la Figura 1.1.

1.1. Aspectos fisiológicos de las neuronas

Las neuronas fueron vistas por primera vez por Santiago Ramón y Cajal, quién recibió el premio Nobel en Fisiología o Medicina en 1906 al descubrir los mecanismos que gobiernan la morfología y los procesos conectivos de las

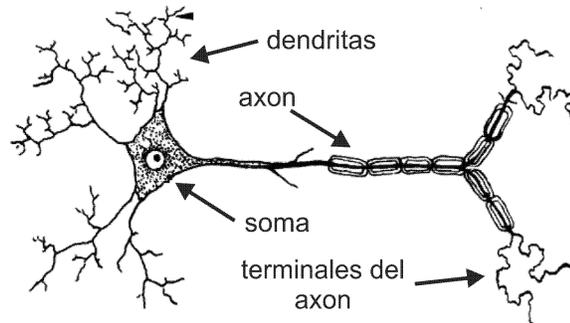


Figura 1.1: Esquema sencillo de una neurona donde se muestran sus principales partes.

celulas nerviosas. En el tiempo que Ramon y Cajal realizaba sus estudios no existía la posibilidad de poder tomar fotografías de lo que se observaba en el microscopio, por lo que sus observaciones quedaron registradas en dibujos realizados por él mismo. Un ejemplo de ello es la ilustración realizada para el tejido del neocortex de una rata que muestra las seis capas de tejido que lo componen. Figura 1.2.

Más tarde, en 1952, Alan Lloyd Hodgkin y Andrew Huxley desarrollaron un modelo para explicar los mecanismos iónicos subyacentes en la iniciación y propagación de los potenciales de acción en el axón gigante del calamar. En 1970, los investigadores Sir Bernard Katz, Ulf Von Euler y Julius Axelrod, recibieron el premio Nobel en Fisiología o Medicina por su trabajo en la identificación y descripción de la participación de los neurotransmisores acetilcolina y noradrenalina en la sinapsis de las neuronas.

Dichas investigaciones ayudaron a comprender la complejidad del tejido neuronal y la importancia de los diversos mecanismos y sustancias neurotransmisoras en las funciones del tejido neuronal.

El ser humano tiene un número de neuronas que está entre 10^{10} y 10^{11} [R. Beale y T. Jackson, 1990]. Se sabe también que cada una de ellas tiene cientos o miles de conexiones con otras neuronas, lo que nos lleva a un total de 10^{14} o 10^{15} conexiones; un número casi incalculable si se intenta imitar su comportamiento. A esto podemos añadir que los impulsos que llegan a la neurona están controlados por umbrales químicos que determinan qué conexiones se harán o no, o de qué intensidad serán en la sinapsis final y esto

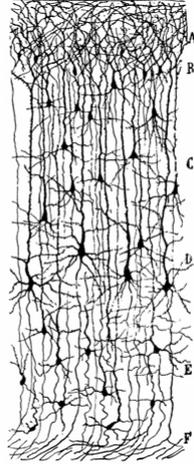


Figura 1.2: Dibujo del neocórtex de una rata realizado por Santiago Ramón y Cajal, publicado por primera vez en el volumen dos, parte dos de “Sistema nervioso del hombre y de los vertebrados” en 1904.

Recuperado el 15 de noviembre de 2012 de: <http://www.sfn.org/About/About-Neuroscience>

es sólo para un impulso. Ahora, si se multiplica esto por todas las funciones que realiza el sistema nervioso ya sean de movimiento, de procesamiento de ideas o de recopilación de recuerdos; resultaría una tarea casi imposible de realizar si este proceso se quisiera utilizar para estructurar la base de un modelamiento que permita realizar cálculos o interpretar datos de manera similar.

Este planteamiento permite poner en dimensión lo complejo y maravilloso del cerebro humano el cual realiza todas estas tareas de manera cotidiana, y a una velocidad asombrosa, a pesar de que tal número de conexiones e interacción entre las neuronas demande entre el 15 % y 20 % de la energía de la que dispone el cuerpo humano.

Si se observa desde el punto de vista de un modelo, los *cerebros*, como lo dice James Anderson: “son especialmente buenos para actividades tales como el reconocimiento de patrones, el control motor, la percepción, la inferencia flexible, la intuición y la realización de buenas suposiciones”. Pero, también admite que: “son lentos, imprecisos, hacen generalizaciones erróneas, están llenos de prejuicios y a menudo son incapaces de explicar sus acciones” [Anderson, 2007].

Umbral=1		
a	b	$i + 1$
0	0	0
1	0	1
0	1	1
1	1	1

Cuadro 1.1: Disyunción inclusiva

Fuente: Anderson, James A. *Redes Neuronales*, 2007.

Sin embargo, la experiencia y la investigación que se ha realizado a través de los años ha llevado a demostrar y usar modelos que son útiles para ciertos tipos de problemas. Los avances en la investigación del funcionamiento de las neuronas y sus procesos han continuado hasta la fecha y al mismo tiempo han evolucionado. Tampoco se han quedado atrás los intentos por desarrollar modelos copiando el comportamiento de ellas.

1.2. Modelo de la neurona elemental

El primer modelo de una neurona fue desarrollado por Walter Pitts y Warren McCulloch en 1943 [W. S. McCulloch y W. H. Pitts, 1943], llamado también “la neurona formal”. Siguiendo el comportamiento de las neuronas anteriormente señalado, en él se fija un umbral para la suma de las entradas que la neurona está recibiendo; si el umbral se excede, la neurona se activa, en caso contrario permanece inactiva. Por ejemplo, si el umbral se fija en 1 y existen dos entradas a y b cuyos valores pueden ser 0 o 1 y, si en un determinado tiempo i ambas entradas son cero, entonces al tiempo $i + 1$ la neurona no se activará puesto que no se ha rebasado el umbral fijado. Si a vale 1 y b vale cero, entonces se activará, al igual que cuando $a = 0$ y $b = 1$ y cuando ambas, tanto a como b valgan 1.

Esto se conoce como la tabla de valores para la *disyunción inclusiva* como se puede observar en el Cuadro 1.1. Cuando el valor del umbral es igual a 2, se conoce como *conjunción* y la tabla de valores se observa en el Cuadro 1.2.

Para generalizar, supongase que tenemos n entradas x que están ligadas a n pesos w . El modelo de la neurona considera la suma del producto de todos ellos:

Umbral=2		
a	b	$i + 1$
0	0	0
1	0	0
0	1	0
1	1	1

Cuadro 1.2: Conjunción

Fuente: Anderson, James A. *Redes Neuronales*, 2007.

$$\sum_{i=1}^n w_i x_i$$

Esta suma se compara con el valor fijado en el umbral, si es mayor se le asigna un valor de 1 y, si es menor, de 0. De igual forma, el umbral se puede restar de la suma ponderada y compararla con el cero, si el valor es positivo se le asigna el valor de 1, en caso contrario el valor de cero. Una manera de lograr el mismo efecto es quitando el valor del umbral del cuerpo de la neurona y conectándolo a una entrada extra que siempre estará activada, sólo que ahora en lugar de restarlo a la suma ponderada se multiplica por -1 y se añade a las otras entradas. Esto se conoce como *sesgar* la neurona, y a $-\theta$ que es el umbral, ahora negativo, es el *sesgo* o *desviación* de la neurona. Si la salida de la neurona es y , entonces:

$$y = f_h \left[\sum_{i=1}^n w_i x_i - \theta \right]$$

donde f_h es la función escalón:

$$\begin{aligned} f_x(z) &= 1 & z > 0 \\ f_x(z) &= 0 & z \leq 0 \end{aligned}$$

En la Figura 1.3 está la representación del modelo básico de esta neurona.

Para la época, el avance que se tenía en el estudio de la actividad neuronal sugería que el cerebro podría verse como la base para la construcción de una red lógica y de computación muy poderosa. Sin embargo, se sabe que,

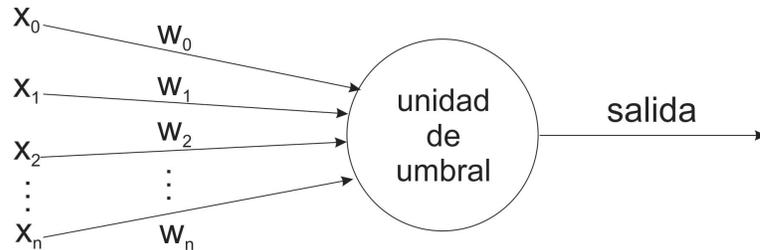


Figura 1.3: Modelo básico de la neurona de McCulloch-Pitts. Las n entradas se multiplican por sus correspondientes pesos y el resultado se compara en la unidad de umbral, donde se le asigna el valor de salida.

Fuente: Russell Beale and Tom Jackson, *Neural Computing: an introduction*, 1990.

particularmente, la función lógica relacionada con la toma de decisiones, aún no ha sido completamente emulada. A pesar de esto, la manera en la que se ha investigado el funcionamiento neuronal ha dado origen a muchos estudios posteriores que, de hecho, se aproximan cada vez más a la reproducción de la actividad neuronal humana.

1.3. Perceptrón

Uno de los modelos neuronales más usados y más conocidos es el perceptrón, el cual fue desarrollado por Frank Rosenblatt, en 1958 y presentando en un reporte técnico del Laboratorio Aeronáutico de la Universidad de Cornell. Años más tarde, en 1969, Marvin Minsky y Seymour Papert [M. Minsky y S. Papert, 1969] hicieron un fuerte señalamiento a las limitaciones teóricas de este modelo neuronal. En consecuencia, el estudio de las redes neuronales se estancó por casi una década. Fue hasta principios de los años ochenta que renació el interés por su estudio y el perceptrón resurge para continuar con su desarrollo.

Este modelo opera con un umbral θ , que en la literatura temprana se le

llamó la *unidad de lógica del umbral* o ULU y cuenta con n ponderaciones de entrada con w_i fuerzas y un sesgo b . La ULU es el producto interno entre los patrones de entrada y las ponderaciones, al cual se le asigna un valor de 0 ó 1, dependiendo de si el valor del umbral es rebasado o no. Es decir:

$$f(x) = \begin{cases} 1 & \text{si } \mathbf{w} \cdot \mathbf{x} + b \geq \theta, \\ 0 & \text{en otro caso} \end{cases} \quad (1.1)$$

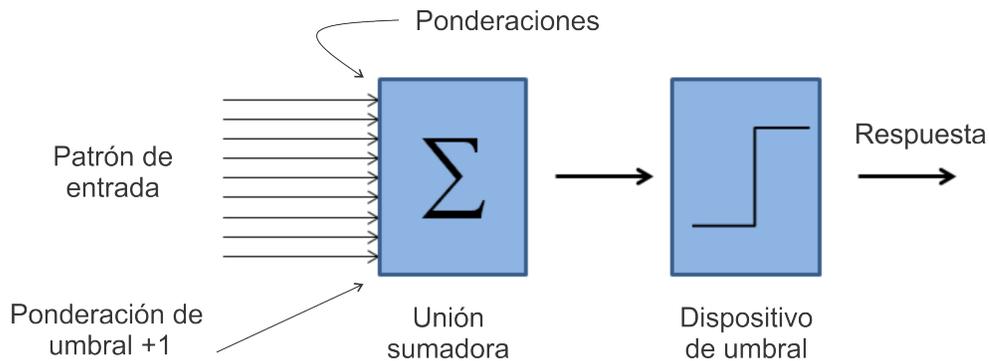


Figura 1.4: Esquema del perceptrón desarrollado por Frank Rosenblatt en 1958.

Fuente: Anderson, James A. *Redes Neuronales*, 2007.

Gráficamente se observa en la Figura 1.4; pero la arquitectura del perceptrón original es más complicada de lo que aquí se muestra. El perceptrón contaba con una superficie sensorial conocida como *retina* que tenía conexiones aleatorias con una *capa de asociación* (*capa A*). Las unidades de esta *capa A* se conectaban a todas las unidades de una tercera capa, la *capa de respuesta* (*R*), y viceversa; de esto, resulta un modelo complicado como puede observarse en la Figura 1.5. Asimismo, se observa que la retina no estaba totalmente conectada a la *capa A*; es decir, que cada unidad de la *capa A* estaba calculando una función diferente de la imagen en la *retina* del

modelo. Esto lleva a una fuerte suposición teórica, como lo señaló James Anderson: “Probablemente la mayoría de las arquitecturas de redes comienzan por asumir una conectividad total, lo que significa que todas las unidades en una capa totalmente conectada percibirán el mismo patrón de entrada” [Anderson, 2007].

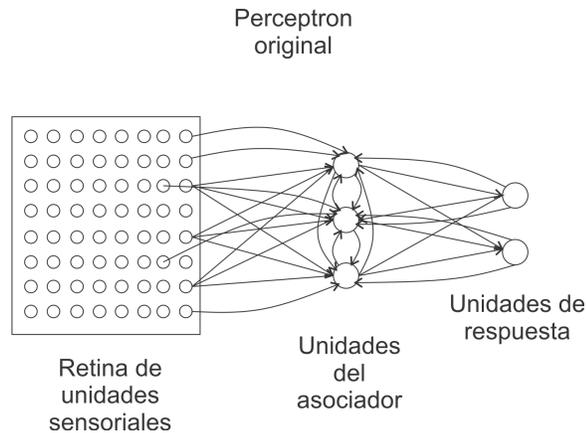


Figura 1.5: Perceptrón temprano. Todas las unidades de la capa de asociación se conectan entre si y estas con la capa de respuesta y viceversa.

Fuente: Anderson, James A. *Redes Neuronales*, 2007.

Un perceptrón diseñado de esta manera es muy complicado de analizar, por lo que se usó un modelo simplificado del mismo, es decir, uno que sólo ocupa conexiones progresivas como puede observarse en la Figura 1.6. Este modelo supone que todas las unidades de la *capa A* se conectan a cada unidad en la *capa R*. Las unidades en la *capa R* inician con pesos aleatorios y sólo se activan cuando detectan o “ven” un patrón específico; por ejemplo, pueden distinguir entre un triángulo y un cuadrado. Su utilidad es más la de un clasificador de patrones que la de un asociador de los mismos.

La pregunta sería entonces: ¿qué es lo que se está clasificando? Para responder esta interrogante se asigna al ULU un valor distinto, -1 o 1, para cada uno de los patrones que se presenten lo que nos lleva a tener dos tipos de ellos. Si se considera un punto \mathbf{x} en el espacio de estado los pesos de la ULU formarán un vector \mathbf{w} ; si se toma el producto interno $[\mathbf{x}, \mathbf{w}]$ cuando éste es igual a cero nos dará la ecuación de un hiperplano en el espacio de

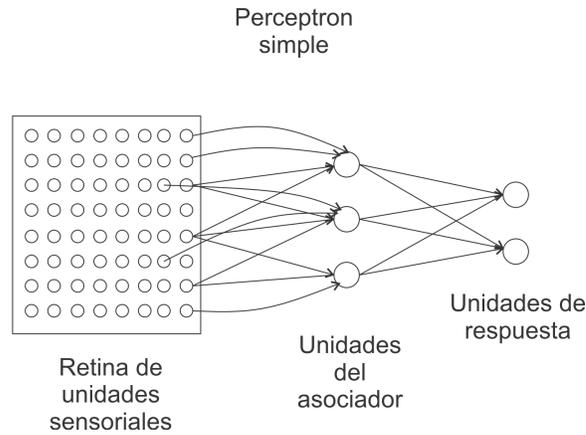


Figura 1.6: Perceptrón simplificado. A diferencia del modelo más complicado, este solo usa conexiones progresivas.

Fuente: Anderson, James A. *Redes Neuronales*, 2007.

estado:

$$[\mathbf{x}, \mathbf{w}] = \sum_i x[i]w[i] = 0 \quad (1.2)$$

Todos los puntos de un lado del plano reciben el valor de -1 y los del otro lado de +1. Cuando esto corresponde a dos grupos distintos de patrones $\{\mathbf{a}\}$ y $\{\mathbf{b}\}$ solamente es posible para la ULU si estos son linealmente separables; es decir, si pueden ser separados por un hiperplano. La separación lineal es intuitiva como se observa en la Figura 1.7. Sin embargo muchas clasificaciones interesantes no son linealmente separables y requerirán de varios hiperplanos para separarlas.

Si existen dos conjuntos de patrones linealmente separables, es posible encontrar un conjunto de pesos que haga la clasificación de todos ellos y, además, se puede demostrar que esto se puede encontrar en un tiempo finito. Esto es lo que se conoce como el *teorema de convergencia del perceptrón* o *teorema de aprendizaje del perceptrón*.

Si se tienen dos conjuntos de patrones linealmente separables $\{\mathbf{a}\}$ y $\{\mathbf{b}\}$ con ejemplos de cada clase y para *entrenar* a la red se presentan los patrones en una *secuencia de entrenamiento* $\{\mathbf{y}\}$. Al presentar un patrón a la red se observará el resultado: si el resultado esperado era +1 pero resulta -1,

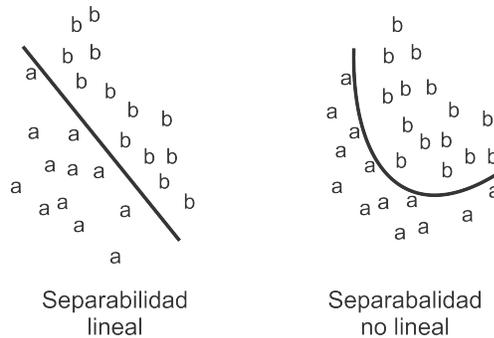


Figura 1.7: Separabilidad lineal. Como se observa en la gráfica de la izquierda, la separación lineal es intuitiva, por el contrario, patrones más complejos como el de la derecha requieren de varios hiperplanos para separarlos.

Fuente: Anderson, James A. *Redes Neuronales*, 2007.

entonces se sumará una cantidad c a los pesos de conexión con lo cual, se espera que la siguiente vez que se presente ese patrón, con los pesos ajustados, se estará más cerca del resultado esperado. Un proceso análogo, pero con signo contrario en c se realiza cuando el resultado esperado sea -1 y se obtenga lo contrario.

Esto se puede ver matemáticamente cuando se presenta el k -ésimo miembro de la secuencia de entrenamiento $\{\mathbf{y}_k\}$ con su conjunto de pesos \mathbf{w}_k . Si la clasificación es correcta, entonces se obtendrá $\mathbf{w}_{k+1} = \mathbf{w}_k$. Pero, si la clasificación es incorrecta, entonces se hace $\mathbf{w}_{k+1} = \mathbf{w}_k + c\mathbf{y}_k$ cuando el resultado debió haber sido $+1$, y se hace $\mathbf{w}_{k+1} = \mathbf{w}_k - c\mathbf{y}_k$, cuando el resultado esperado era -1 .

Para realizar la demostración de este teorema se podría suponer que, en efecto, los conjuntos de patrones $\{\mathbf{a}\}$ y $\{\mathbf{b}\}$ son linealmente separables y con las salidas deseadas de $+1$ y -1 respectivamente. Esto quiere decir que el producto interno de $\{\mathbf{a}\}$ con sus pesos w es positivo y los correspondientes de $\{\mathbf{b}\}$ negativos. De esto se deduce que muchos planos pueden ser la solución al problema de clasificación como se aprecia en la Figura 1.8.

Para entender mejor la prueba del teorema es útil conocer el concepto del *espacio de pesos*. Los pesos de la ULU forman un punto en un espacio con una dimensión igual al número de pesos. Dentro de ese espacio hay una región en la que todos los pesos serán soluciones correctas; si se considera

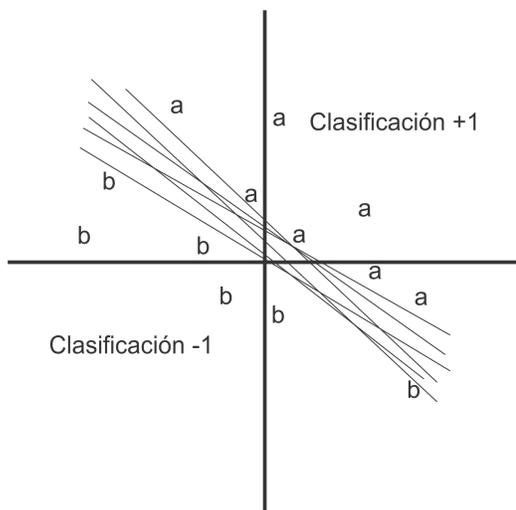


Figura 1.8: Para dos patrones linealmente separables existen múltiples planos de solución.

Fuente: Anderson, James A. *Redes Neuronales*, 2007.

clasificar únicamente con base en el signo del producto interno, se obtendrá una región como la figura de un cono que pasa a través del cero, como se puede observar en la Figura 1.9.

Con el fin de simplificar esta prueba se usará el producto interno negativo de $\{\mathbf{b}\}$, es decir:

$$[-\mathbf{b}_i, \mathbf{w}] > 0.$$

Ahora los conjuntos de entrenamiento de ambas clases $\{\mathbf{a}\}$ y $\{\mathbf{b}\}$ se encontrarán del lado positivo del hiperplano y se llamará al nuevo conjunto de entrenamiento $\{\mathbf{y}'\}$, para el cual también existirá un conjunto de pesos \mathbf{w} , tal que $[\mathbf{y}'_i, \mathbf{w}] > 0$. La prueba seguirá a la prueba 2 de Nils J. Nilsson [Nilsson Nils J., 1965].

Al estar dentro de la región de solución se sabrá que el producto interno es positivo y que este puede ser mayor que cualquier otro valor: ya que al ser mayor que cero implica que tenga un mínimo. Si se multiplica el producto interno por una constante, se creará lo que Nilsson llama una *región aislada*.

Sea m un valor más grande que la distancia al cuadrado en el conjunto de

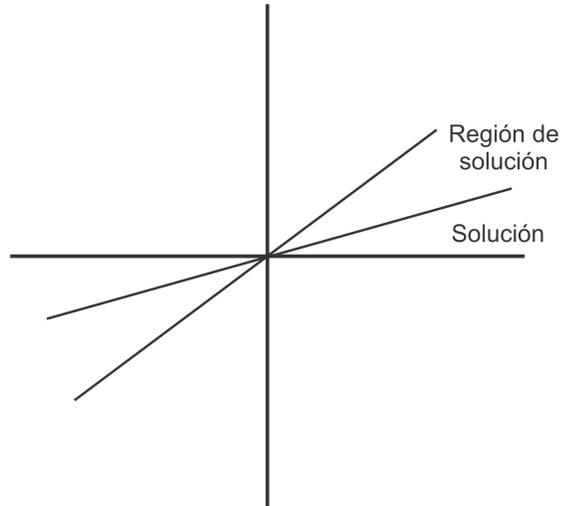


Figura 1.9: Al considerar la clasificación solamente por el signo del producto interno de los pesos, dentro del espacio de solución, se forma la región de solución.

Fuente: Anderson, James A. *Redes Neuronales*, 2007.

entrenamiento y b una constante arbitraria, la región aislada podría quedar como:

$$[\mathbf{y}'_i, \mathbf{w}] > \frac{(m + b)}{2}.$$

La prueba consistirá en demostrar que cada vez que *aprende* el perceptrón, el aprendizaje disminuirá en una cantidad mayor que una constante y, por ser mayor que cero, en algún momento deberá detenerse.

Sea \mathbf{w}_k un conjunto de pesos en la k -ésima prueba de aprendizaje y \mathbf{w} un vector en la región aislada de soluciones. La distancia entre los dos vectores estará dada por:

$$|\mathbf{w} - \mathbf{w}_k|^2 = [\mathbf{w}, \mathbf{w}] - 2[\mathbf{w}, \mathbf{w}_k] + [\mathbf{w}_k, \mathbf{w}_k].$$

Sea la constante $c = 1$, cada vez que el perceptrón aprende, se sumará el vector \mathbf{y}'_i a los pesos. El nuevo vector de pesos \mathbf{w}_{k+1} estará dado por:

$$\mathbf{w}_{k+1} = \mathbf{w}_k + \mathbf{y}'_i.$$

La distancia al cuadrado entre el vector \mathbf{w} y los dos conjuntos de pesos \mathbf{w}_k y \mathbf{w}_{k+1} forman una cantidad que se demostrará que es mayor que un mínimo; luego de las pruebas de aprendizaje:

$$\begin{aligned} d_{k+1} &= |\mathbf{w} - \mathbf{w}_k|^2 - |\mathbf{w} - \mathbf{w}_{k+1}|^2 \\ &= |\mathbf{w} - \mathbf{w}_k|^2 - |\mathbf{w} - (\mathbf{w}_k + \mathbf{y}'_i)|^2. \end{aligned}$$

$$d_{k+1} = -2[\mathbf{w}_k, \mathbf{y}'_i] + 2[\mathbf{w}, \mathbf{y}'_i] - [\mathbf{y}'_i, \mathbf{y}'_i].$$

Como se está analizando el aprendizaje del perceptrón, el primer término debe ser mayor o igual a cero por haber sido clasificado erróneamente; de manera que, si se elimina puede tomarse una desigualdad:

$$d_{k+1} \geq 2[\mathbf{w}_k, \mathbf{y}'_i] - [\mathbf{y}'_i, \mathbf{y}'_i].$$

Por suposición, los patrones en el conjunto de entrenamiento tienen un máximo de longitud, por lo que se tiene que:

$$d_{k+1} \geq 2[\mathbf{w}_k, \mathbf{y}'_i] - m.$$

Si se recuerda que:

$$[\mathbf{w}, \mathbf{y}'_i] > \frac{(m+b)}{2}$$

$$2[\mathbf{w}, \mathbf{y}'_i] - m > b.$$

Finalmente se tendrá:

$$d_{k+1} > b > 0.$$

Por lo tanto, la distancia entre el vector de pesos de solución y el real se reducirá siempre que el perceptrón aprenda en una cantidad finita.

Pero, el perceptrón tiene limitaciones: siempre trata de encontrar una línea que separe a las clases cuando muchos problemas no son linealmente separables.

Considerese el ejemplo de la *disyunción exclusiva* o XOR, de acuerdo con el Cuadro 1.3.

X	Y	Z
0	0	0
1	0	1
0	1	1
1	1	0

Cuadro 1.3: Disyunción exclusiva

Fuente: Anderson, James A. *Redes Neuronales*, 2007.

Aparentemente, es un problema sencillo de resolver; la única variante es que ahora cuando los valores de X y Y son 1, tenemos como resultado un cero. Gráficamente se puede observar en la Figura 1.10, donde se han identificado las regiones sombreadas para las entradas que producen como resultado un 1 y las claras para las que producen un 0.

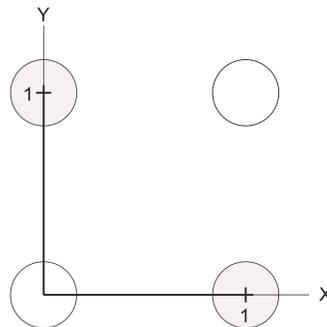


Figura 1.10: El problema XOR. No es posible trazar una recta que separe las clases, esto se solucionó añadiendo más capas de neuronas al modelo.

Fuente: Russell Beale and Tom Jackson, *Neural Computing: an introduction*, 1990.

Como se pudo observar, no se puede trazar una línea recta que separe a las clases y, por lo tanto, el perceptrón simple no sirve para resolver este tipo de problemas. Para solucionar esto surge, entonces, la idea de usar más de una capa de neuronas naciendo así el *perceptrón multicapa*.

1.4. Perceptrón multicapa

Con base en el problema XOR de la sección anterior, ahora se añade una capa adicional a la red neuronal planteada, los umbrales en cada nodo y los pesos en las conexiones, tal como lo plantea [R. Beale y T. Jackson, 1990] en la Figura 1.11.

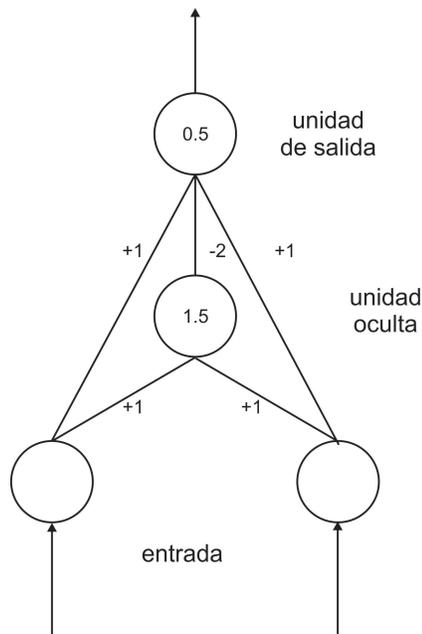


Figura 1.11: Una solución al problema XOR. Al añadir una capa más de neuronas y con los valores adecuados, el problema XOR de la sección anterior tiene solución ocupando una red neuronal.

Fuente: Russell Beale and Tom Jackson, *Neural Computing: an introduction*, 1990.

La red tiene dos entradas porque se tienen dos variables y dos unidades: una oculta y una de salida. El valor del umbral de la capa oculta de 1.5 la mantiene apagada a menos que las dos entradas se activen. Al analizar los valores se observa que si la entrada es (00), la unidad oculta no se activa y, por lo tanto la unidad de salida permanece inactiva al no recibir ningún valor, lo cual coincide con la primera de las soluciones del problema. Si la

entrada es (01) la unidad oculta tampoco se activa porque recibe un valor de 1, pero la de salida sí lo hace al recibir un valor de 1. Lo mismo sucede con el valor (10), teniendo así los resultados correctos del problema.

Cuando el valor de entrada es (11), entonces la unidad oculta se activa puesto que las dos unidades de entrada se activaron y recibe un valor de 2. Esta última unidad pasa un valor de cero a la unidad de salida, al sumar el peso de -2 indicado y, por lo tanto, no se activa, originando con ello el cuarto resultado correcto del problema.

El añadir una unidad a la red neuronal le proporciona la capacidad de resolver problemas que aparentemente no tenían solución con esa herramienta. Este fue uno de los argumentos que en 1986 dieron a conocer David Everett Rumelhart y James Lloyd McClelland y que hizo que resurgiera el interés en este tema.

Otra idea que ayudó a encontrar una mejor solución, fue ocupar en lugar de la función escalón, que sólo nos da 0 o 1 como valor de resultado, una que dé información de cómo se pueden acomodar los pesos para buscar el resultado esperado. Esa función es la sigmoide que, al tener una pendiente, puede orientar si es necesario aumentar o disminuir los pesos relevantes. Si la combinación de pesos dio como resultado un valor mucho mayor o menor, al umbral, la función asegurará que la siguiente combinación de pesos irá en la dirección correcta al estar cercana a los valores esperados. Esto se hace evidente al observar el comportamiento de la función en la Figura 1.12

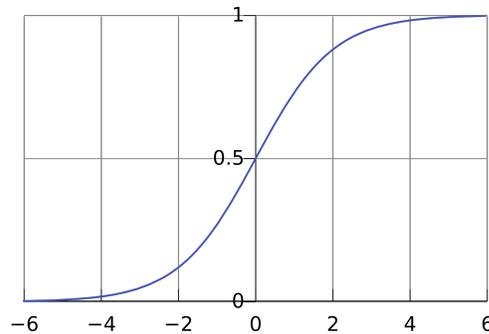


Figura 1.12: La función sigmoide. Al usarse como función de evaluación en una red neuronal, su pendiente da información para conformar la combinación de pesos que se acercan a la solución.

La regla de aprendizaje también cambia en este modelo. En ese sentido, Rumelhart y McClelland propusieron la “regla generalizada delta” o “regla de retropropagación”. Si se tiene un patrón de entrada y se conoce el resultado deseado, se entrena a la red para que la diferencia entre el resultado y el valor esperado se vaya reduciendo conforme se ajustan los pesos; de ahí es que surge el nombre de *aprendizaje supervisado*. La regla de retropropagación calcula el valor de la función de error para una entrada específica y luego regresa o “propaga” ese valor a la capa previa.

1.5. Algoritmo de retropropagación

El algoritmo de retropropagación (AR) busca el mínimo del error en el espacio de pesos usando el método del gradiente descendente; por lo que es necesario que la función de activación sea diferenciable y continua, ya que, en cada paso se calcula el gradiente de la misma. Se considera que la combinación de pesos que minimiza la función de error es una solución al problema de aprendizaje.

En el AR se utiliza el aprendizaje supervisado que es aquel que se utiliza cuando se tiene un conjunto de datos de entrenamiento y otro de resultados reales contra los que se compara el resultado de la aplicación del modelo neuronal. En un modelo típico de retropropagación se tienen tres capas: la de entrada, la oculta y la de salida. Varios autores coinciden que tener más de cuatro capas en una RNA no mejora los resultados [I. Kaastra y M. Boyd, 1996].

En este modelo, primero, los datos son presentados a la red junto con sus pesos; luego se pasan a la capa oculta donde se aplica la función de activación con sus correspondientes pesos y, finalmente, los resultados se pasan a la capa de salida. El resultado de la red se compara con el real y se toma la diferencia. Esta diferencia se multiplica ahora por cada uno de los pesos que participaron en la función de salida y, este nuevo valor, se suma a los pesos de la función de entrada que también se multiplican por el valor de la diferencia.

Como se observa, la diferencia se transmitió a los pesos tomando el camino inverso a la entrada. Ahora, con estos nuevos pesos se vuelven a hacer los cálculos de la RNA, pero la diferencia es que el valor de entrada se multiplica por un factor de ajuste y la derivada parcial de la función de activación; de ahí la necesidad de que sea diferenciable y continua. Con estos nuevos valores ajustados el resultado de la red se aproximará más al valor real y

la diferencia tenderá a disminuir puesto que la estamos haciendo decrecer al tomar el gradiente de la función.

Existen muchas demostraciones que prueban que la función de error, o la combinación de los pesos, tenderá a conseguir el mínimo global de la función, como se explica en el capítulo 7 del libro de Raúl Rojas [R. Rojas, 1996]. También se puede encontrar en internet una explicación gráfica y paso a paso de este algoritmo ¹.

En el próximo capítulo se describirá lo que es un sistema financiero y los integrantes del mismo: bancos, mercados de valores, casas de bolsa, casas de cambio, etcétera. Así como las principales teorías que describen el comportamiento de los mercados financieros. También, se presentará la generación del IPC (Índice de Precios y Cotizaciones) de la Bolsa Mexicana de Valores (BMV), objeto de estudio de este trabajo.

¹Recuperado el 24 de julio de 2012 de: http://galaxy.agh.edu.pl/~vlsi/AI/backp_t_en/backprop.html

Capítulo 2

El sistema financiero

La historia muestra que desde que la humanidad tuvo necesidad de adquirir productos nació el comercio. En un principio, sólo se usaba el trueque o intercambio de bienes en el que se daba una negociación que fuera satisfactoria para ambas partes. Sin embargo este intercambio tenía el inconveniente de que tenía que realizarse con alguien que, específicamente requiriera la mercancía ofrecida y viceversa. Más adelante surge la unidad de intercambio que posteriormente se convertiría en moneda. Estas unidades fueron en un inicio muy sencillas: conchas marinas, cebada, arroz, cacao, etcétera. Aunque en un principio sólo fueron utilizadas como un instrumento para facilitar el trueque, posteriormente se convirtieron en símbolo de poder y riqueza al ser elaboradas, ya como moneda, en diversos metales preciosos.

Las primeras monedas metálicas aparecieron en la antigua Babilonia alrededor del año 3,000 a.C. y sólo eran pequeños trozos de metal de tamaño uniforme. Los romanos y los griegos usaban monedas de forma circular a las cuales les grababan la imagen de sus dioses o de sus monarcas.

Con el nacimiento de la moneda también surgieron los préstamos y los lugares específicos para desarrollar el comercio. El primer banco moderno fue fundado en Génova, Italia en el año de 1406, conocido como el Banco de San Giorgio. El nombre de “banco” deriva de la palabra italiana *banco* que significa “escritorio”, utilizada por los comerciantes judíos florentinos que hacían sus transacciones sobre una mesa con un mantel verde. A través de los siglos y, con el crecimiento de las transacciones en dinero, el banco fue evolucionando hasta ser como lo conocemos hoy en día. Así mismo, evolucionó la forma de manejar el dinero: se desarrollaron nuevos conceptos e instrumentos que fueron satisfaciendo las necesidades de los usuarios de los bancos y del

comercio.

En México el primer banco en establecerse fue el The Bank of London, Mexico and South America en 1864. En 1897 se promulgó la Ley de Instituciones de Crédito que contempló tres modelos bancarios: bancos de emisión, que eran bancos de depósito y descuento con facultad para emitir billetes, bancos hipotecarios y bancos refaccionarios. Con el Artículo 28 de la Constitución de 1917 se estableció que la emisión de billetes sólo podía hacerse por medio de un banco central y con el control del gobierno.

Durante los años cincuenta, la tendencia del crecimiento del sistema financiero en México se orientó a la *banca universal* en la que una sola entidad proveía todos los servicios financieros. El centro de esta entidad era un Banco de Depósito que agrupaba como filiales a una financiera, una hipotecaria, un departamento de ahorro y un departamento de fideicomiso. También, podían adquirir como filiales a entidades de otro tipo, como empresas del sector industrial y de servicios.

En la década de los años setenta se crearon los grupos financieros que, posteriormente, dieron lugar a la creación de la banca múltiple hacia finales de la misma y a principios de los años ochenta. Esta banca múltiple, se creó por la fusión de tres entidades que brindaban tres tipos de servicio diferentes: un banco de depósito, una financiera y una sociedad hipotecaria. Aunque, también surgieron bancas múltiples que no obedecieron a la fusión de otros bancos. Después de la nacionalización de la banca, durante los gobiernos de José López Portillo y Miguel de la Madrid, se inició nuevamente la privatización de la misma y a mediados de 1992 la Secretaría de Hacienda abrió nuevamente la posibilidad para el establecimiento de la banca múltiple.

Actualmente el sistema financiero mexicano está compuesto por: bancos, inversionistas, afianzadoras y mercados financieros, entre otros actores. También se encuentra regulado por leyes y procedimientos que protegen los recursos de los ahorradores y la interacción de éstos distintos actores en sus diversas actividades financieras.

2.1. Los bancos

Los bancos son los principales integrantes de un sistema financiero, ya que son quienes participan activamente en la economía mediante el otorgamiento de créditos, que es su principal función. Estos créditos se otorgan a empresas y a particulares aplicando un interés que obedece a las condiciones del mercado:

a mayor oferta menor precio del insumo y, a mayor demanda, mayor precio del mismo. Siguiendo estas condiciones, si la tasa de interés es alta es porque el insumo principal, el dinero, está escaso; aunque, cabe señalar, que esta no es la única regla a la que obedece el establecimiento de la misma.

Un banco siempre puede prestar más del dinero que conforma su capital aunque con ello conlleva un riesgo ya que sus deudores podrían caer en insolvencia y por lo tanto dejar de cubrir sus adeudos con la institución. Este diferencial en el que el banco se puede exceder al prestar con miras a obtener mayores ganancias está regulado y vigilado por las autoridades financieras. Incluso existen acuerdos internacionales, como el de Basilea¹, en el cual se establecen lineamientos para la regulación, la integración de indicadores de solvencia, solidez y liquidez de los mismos; además de los límites máximos de financiamiento que pueden otorgar. Adicionalmente, cada gobierno garantiza a los inversionistas y usuarios de los bancos que sus recursos están protegidos al crear fondos de reserva y mecanismos de vigilancia de las operaciones bancarias. Aparte de los bancos, también existe el mercado financiero; este está integrado fundamentalmente por los mercados de deuda, el mercado cambiario y los mercados de acciones; en ellos se intercambian activos con el propósito principal de movilizar dinero a través del tiempo.

2.2. Mercados de deuda

Las empresas privadas o los gobiernos pueden valerse de varios instrumentos para contar con recursos económicos. Uno de ellos puede ser el de los préstamos bancarios o los instrumentos de deuda.

Un instrumento de deuda es un título emitido por una entidad gubernamental, o privada, mediante el cual se establece un compromiso de pago por parte del emisor después de haber recibido el recurso solicitado. Este instrumento fija un plazo para su vencimiento considerando un interés; éste, puede ser fijo, variable o indexado al tipo de cambio de una moneda o a la inflación.

Cuando los instrumentos de deuda son emitidos y adquiridos por primera vez por los inversionistas, forman lo que se llama el mercado primario de deuda. Una vez que estos instrumentos se han colocado pasan a formar parte del mercado secundario, en el se demandan y ofrecen libremente a todos los

¹Ver “Basel (II) y (III)” en: <http://www.bis.org/bcbs/index.htm>. Recuperado el 13 de agosto de 2012.

inversionistas. En este mercado los instrumentos se pueden ofrecer de manera pública o privada. Estos instrumentos reciben una calificación por parte de agencias privadas, conocidas como calificadoras. Basándose principalmente en la capacidad de pago por parte del emisor, estos se clasifican en categorías que van desde el menor hasta el más alto riesgo de recuperación de la inversión; siendo los más atractivos aquellos que son calificados con el menor riesgo.

Una de las agencias más conocidas a nivel mundial y en la que se basan muchos inversionistas para realizar sus transacciones es Moody's quién estableció nueve niveles de clasificación: Aaa, Aa, A, Baa, Ba, B, Caa, Ca y C; además de 3 niveles de modificadores que indican el grado de los modificadores dentro de cada nivel. La exposición al riesgo por parte del comprador de estos instrumentos incrementa en forma gradual desde A siendo, entonces, los instrumentos menos seguros para recuperar la inversión los calificados como tipo C.

Estas calificadoras forman un papel importante en las actividades financieras. Como lo establece el documento titulado: "El nuevo acuerdo de capital de Basilea", publicado en enero de 2001 por el Comité de Supervisión Bancaria de Basilea, en el que se establece la importancia del papel de las mismas para calcular el capital regulatorio y, así, se tome más en cuenta el riesgo del que puede ser sujeto.²

En México, el mercado de deuda se inició en 1978 con la emisión por parte del Gobierno Federal de los Certificados de la Tesorería de la Federación (CETES), que, hasta hace poco, eran inaccesibles para personas físicas, ya que solo podían ser adquiridos por empresas o casas de bolsa.

Durante la década de los años 80 el surgimiento de las casas de bolsa amplió el ofrecimiento del mercado de deuda lo cual dio inicio al mercado de deuda privada. En los años 90, la liberalización de las tasas de interés y la decisión del gobierno federal de financiar todo su déficit presupuestal con la emisión de deuda tuvo como consecuencia una mayor participación de los intermediarios en este mercado.

Es hasta el año 2000 que el Gobierno Federal emite los primeros bonos a tasa fija con un plazo de 3 años; actualmente, existen referencias de 3, 5, 10, 20 y 30 años. En materia de deuda privada, en este mismo año, se crearon

²Fuente "Agencias Calificadoras" en <http://www.banxico.org.mx/sistema-financiero/material-educativo/basico/fichas-sistema-financiero.html>. Recuperado el 20 de agosto de 2012

los certificados bursátiles.

En el Cuadro 2.1 se encuentra la clasificación de los certificados de deuda de acuerdo con su emisor.

2.3. Mercado cambiario

El mercado cambiario es la compra-venta de divisas o monedas extranjeras con fines comerciales, con ello se transfiere el poder de compra de una moneda a otra. Dentro de este mercado participan varios actores: bancos comerciales, casas de cambio, exportadores, importadores, turistas nacionales en el extranjero, turistas extranjeros en territorio nacional, trabajadores mexicanos en el extranjero, etcétera.

Éste, se encuentra regido por el tipo de cambio, que no es otra cosa más que la expresión en unidades de la moneda nacional para adquirir una extranjera. Si por ejemplo, para adquirir un dólar estadounidense se requieren trece pesos mexicanos, entonces el tipo de cambio será de trece pesos por dólar. La política cambiaria en México está regida por la Comisión de Cambios; integrada por funcionarios tanto de la Secretaría de Hacienda y Crédito Público como del Banco de México. Dicha comisión puede reunirse en todo momento a solicitud del Secretario de Hacienda y Crédito Público o del Gobernador del Banco de México.

A finales de 1994, la Comisión de Cambios acordó que el régimen cambiario en México fuera flexible. El tipo de cambio flexible se determina libremente y obedece únicamente a ley de la oferta y la demanda; aunque, es sabido, que el Banco de México interviene en el mercado inyectando o comprando dólares para mantener cierta estabilidad. Los establecimientos que participan del mercado cambiario establecen un diferencial entre el tipo de cambio de compra y de venta, siendo este último el más alto; esto ocurre porque no cobran directamente una comisión a sus clientes.

Este mercado está descentralizado y opera las 24 horas del día, los 365 días del año en cualquier parte del mundo.

El Banco de México fija un tipo de cambio conocido como *fix* para las obligaciones que se tengan que cubrir en dólares americanos y que sirve como guía para varias de las operaciones en este mercado; este, se publica todos los días a partir de las 12:00 PM, hora del centro, en su página electrónica y un día hábil después en el Diario Oficial de la Federación. Otro tipo de cambio es el que se usa para las operaciones al mayoreo entre los intermediarios

Emisor	Instrumento
Gobierno Federal	Certificados de la Tesorería (Cetes) Bonos de Desarrollo (Bondes) Bonos M Bonos denominados en UDIs (Udi-bonos)
Instituto para la protección al ahorro bancario	Bonos IPAB (BPA, BPAT y BPA182)
Banco de México	Bonos de Regulación Monetaria (BREM)
Empresas paraestatales e instituciones públicas	Certificados bursátiles y bonos
Banca comercial	Aceptaciones bancarias Certificados de depósito Bonos bancarios Certificados bursátiles Obligaciones bancarias y pagarés
Empresas privadas	Papel comercial Obligaciones privadas Certificados de participación ordinaria (CPO y CPI) Pagarés Certificados bursátiles
Gobiernos estatales y municipales	Certificados bursátiles

Cuadro 2.1: Clasificación de instrumentos de deuda

Los instrumentos de deuda en “El Sistema Financiero.” Recuperado el 20 de agosto de 2012 de: <http://www.banxico.org.mx/divulgacion/sistema-financiero/sistema-financiero.html#Mercadosdedeuda>

financieros, conocido como interbancario. Las denominaciones: “mismo día”, “24 horas” y “48 horas” se refieren a los plazos de liquidación, siendo, el más común el último de ellos, conocido también como *spot*.

2.4. Mercados de acciones o valores

Parte del capital de una empresa, el dinero y recursos disponibles para operar y crecer, está representado por acciones. Éstas, son también consideradas *instrumentos* mediante los cuales una compañía puede obtener recursos al ofrecerlas a los inversionistas ofreciendo una ganancia sobre ellas. Esta ganancia es lo que la compañía reparte de sus utilidades entre sus accionistas después de conservar una parte para el crecimiento e inversión.

Este capital tiene dos clasificaciones: privado, en el que sólo participan los accionistas en la inversión por participación o asociación entre particulares y público, en el que cualquier persona puede comprar o vender acciones de las compañías participantes.

El espacio o lugar en el cual se puede comerciar con estas acciones es el Mercado de Valores, que puede ser una ubicación física o virtual, como es el caso del NASDAQ, definición que se verá más adelante. En estos mercados se opera con reglas transparentes y vigilancia por parte del Gobierno, lo que asegura un trato justo y equivalente para todos los participantes.

Adicionalmente, todas las empresas que ponen a la venta acciones en el mercado público tienen una vigilancia especial por parte de organismos externos que aseguran que la información publicada y la manera de operar de la compañía sea confiable para el público inversionista. Si una compañía mexicana participa en el mercado de otro país, por ejemplo en los Estados Unidos, tiene que apearse a los lineamientos del mismo.

Una ley muy conocida es la Sarbanes-Oxley que fue publicada en el 2002, después del quebranto financiero de la empresa Enron, cuando altos ejecutivos de la empresa ocultaron información financiera importante que revelaba la situación verdadera de la compañía; el cual fue considerado como uno de los peores casos de fraude financiero en la historia de los Estados Unidos ³.

Después de la publicación de la ley referida, las empresas están obligadas a presentar sus estados financieros junto con un reporte interno de auditoría

³Recuperado el 23 de agosto de 2012 de: <http://www.gpo.gov/fdsys/pkg/PLAW-107pub1204/html/PLAW-107pub1204.htm>

que asegura la confiabilidad de los mismos; además, un despacho consultor externo debe avalar dichas cifras.

Dentro de los mercados accionarios más importantes podemos señalar a: New York Stock Exchange (NYSE), National Securities Dealers Automated Quotations (NASDAQ), London Stock Exchange y Tokyo Stock Exchange. En México, el mercado accionario forma parte de la Bolsa Mexicana de Valores (BMV).

Para conocer la historia de la bolsa de valores en México, se cita el cuadro que presenta dicha institución ⁴:

- 1850.- Negociación de primeros títulos accionarios de empresas mineras.
- 1867.- Se promulga la Ley Reglamentaria del Corretaje de Valores.
- 1880-1900.- Las calles de Plateros y Cadena, en el centro de la Ciudad de México, atestiguan reuniones en las que corredores y empresarios buscan realizar compraventas de todo tipo de bienes y valores en la vía pública. Posteriormente se van conformando grupos cerrados de accionistas y emisores, que se reúnen a negociar a puerta cerrada, en diferentes puntos de la ciudad.
- 1886.- Se constituye la Bolsa Mercantil de México.
- 1895.- Se inaugura en la Calle de Plateros, hoy Madero, el centro de operaciones bursátiles Bolsa de México, S.A.
- 1908.- Luego de periodos de inactividad bursátil, provocados por crisis económicas y en los precios internacionales de los metales, se inaugura la Bolsa de Valores de México, SCL, en el Callejón de 5 de Mayo.
- 1920.- La Bolsa de Valores de México, S.C.L. adquiere un predio en Uruguay 68, que operará como sede bursátil hasta 1957.
- 1933.- Comienza la vida bursátil del México moderno. Se promulga la Ley Reglamentaria de Bolsas y se constituye la Bolsa de Valores de México, S.A., supervisada por la Comisión Nacional de Valores, hoy Comisión Nacional Bancaria y de Valores.

⁴“Nuestra Historia.” Recuperado el 20 de agosto de 2012 de: http://www.bmv.com.mx/wb3/wb/BMV/BMV_nuestra_historia

- 1975.- Entra en vigor la Ley del Mercado de Valores, y la Bolsa cambia su denominación a Bolsa Mexicana de Valores, e incorpora en su seno a las bolsas que operaban en Guadalajara y Monterrey.
- 1995.- Introducción del BMV-SENTRA Títulos de Deuda. La totalidad de este mercado es operado por este medio electrónico.
- 1996.- Inicio de operaciones de BMV-SENTRA Capitales.
- 1998.- Constitución de la empresa Servicios de Integración Financiera o SIF, para la operación del sistema de negociación de instrumentos del mercado de títulos de deuda o BMV-SENTRA Títulos de Deuda.
- 1999.- El 11 de enero, la totalidad de la negociación accionaria se incorporó al sistema electrónico. A partir de entonces, el mercado de capitales de la Bolsa opera completamente a través del sistema electrónico de negociación BMV-SENTRA Capitales. En este año también se listaron los contratos de futuros sobre el IPC en MexDer y el principal indicador alcanzó un máximo histórico de 7,129.88 puntos el 30 de diciembre.
- 2008.- El día 13 de junio de 2008, la BMV cambia su razón a Bolsa Mexicana de Valores, S.A.B. de C.V, convirtiéndose en una empresa cuyas acciones son susceptibles de negociarse en el mercado de valores bursátil, llevando a cabo el 13 de junio de 2008 la Oferta Pública Inicial de sus acciones representativas de su capital social; cuya clave de pizarra a partir de esta fecha es “BOLSA”.

La Bolsa Mexicana de Valores concentra el ahorro de los inversionistas hacia la inversión productiva fomentando con esto el crecimiento económico y la generación de empleos. Es también reflejo de la situación económica del país al exponer ante el público y el mundo la capacidad del volumen negociado de las acciones: a mayor volumen, mayor poder económico. Además, es un indicador de las condiciones futuras de los negocios.

Cada mercado de valores tiene un índice que refleja su comportamiento diario y guía muchas de las operaciones del sistema financiero mundial ya que es uno de los índice mas consultados para conocer la situación económica de un país.

Estos son algunos de los indicadores de los mercados de valores en el mundo:

- New York Stock Exchange: NYSE Composite
- National Securities Dealers Automated Quotations (NASDAQ): NASDAQ Composite
- London Exchange: FSTE
- Bolsa de Valores del Estado de Sao Paulo: Bovespa
- Bolsa Mexicana de Valores: IPC

Cada índice se compone de una muestra significativa de las acciones que se comercializan dentro del mercado siguiendo las reglas que se acuerden por los integrantes del mismo en cada país. Los mercados se han tratado de clasificar además por su forma de operar, lo que ha originado una teoría conocida como la de los *mercados eficientes*, que describiremos en la siguiente sección.

2.4.1. Los mercados eficientes y los mercados emergentes

Un mercado eficiente es aquel en el cual toda la información relevante a un título financiero, que pueden ser las acciones de una compañía, se ve reflejada en el precio del mismo, ya sea favorable o desfavorable. Teóricamente, la búsqueda de las oportunidades de ganancia por parte de los inversionistas en las acciones sobrevaluadas o subvaluadas, lleva a estas a alcanzar el precio correcto, generando así un comportamiento eficiente.

La disponibilidad de la información y las regulaciones legales, han propiciado que los inversionistas realicen análisis más profundos y certeros del comportamiento de los precios de las acciones. Como la información disponible puede ser aleatoriamente favorable, o desfavorable, de acuerdo con las expectativas, entonces el cambio en los precios también resulta aleatorio, evitando de esta forma que haya ganancias anormales, que es una característica de un mercado eficiente.

Si el comportamiento de los precios es aleatorio, entonces se puede modelar como una caminata aleatoria, cosa que ya había sido referida por el matemático francés Louis Bachelier en 1900, cuya disertación, sin embargo, fue ignorada hasta los años cincuenta, cuando se retoma el estudio formal del comportamiento de las ganancias de las acciones en el mercado de valores.

Hasta antes de 1953, la apreciación del mercado de valores en ciertos sectores era incluso irrisoria: según John Maynard Keynes, se podía comparar

con un concurso de belleza en el cual cada analista de mercado recomienda no la acción que él piensa que es la mejor, sino la que él considera que los otros analistas ven como la mejor; es decir, los precios estaban basados más en la especulación que en fundamentos económicos. Se consideraba que en el largo plazo los precios guiados por la especulación podrían converger con aquellos fijados por los fundamentos económicos, pero como lo señala el mismo Keynes, “a largo plazo, todos estaremos muertos”.

Maurice Kendall documentó en 1953 la independencia estadística de las ganancias semanales de varios índices británicos y en 1959 Harry Roberts encontró resultados similares en las ganancias del Dow Jones. Adicionalmente, Eugene Fama en 1965 confirmó que varias de las técnicas usadas por los analistas técnicos no tenían un poder de predicción. Todos estos resultados fueron tomados como fundamento para el modelo de la caminata aleatoria en los mercados de acciones, afirmando además que dichos mercados carecían de un significado y contenido económico.

Gracias al trabajo de Paul Samuelson (1965) y Benoit Mandelbrot (1966), se entendió que tal aleatoriedad debe de esperarse de un mercado en buen funcionamiento. Definieron que la competencia implica que el invertir en acciones en un mercado de valores es un “juego justo” en el cual el inversionista no pretende derrotar el mercado sin alguna ventaja informativa. La esencia de este comportamiento es que el precio de la acción del día de hoy refleja las expectativas de los inversionistas, dada la información disponible. El precio de la acción del día de mañana cambiará sólo si las expectativas de los inversionistas cambian también, en forma aleatoria positiva o negativa, y siempre que no intervenga algún sesgo en su comportamiento.

Posteriormente Eugene Fama (1970) clasificó los mercados eficientes en: débil, semi-fuerte y fuerte, surgiendo también otros estudios. Robert Shiller (1981) y Werner DeBondt y Richard Taler (1985), demostraron la volatilidad del comportamiento de los precios a largo plazo, reforzando nuevamente el concepto de que todo está controlado más por la especulación que por el comportamiento aleatorio. La aparente contradicción de los análisis a los mercados eficientes ha originado otra corriente de pensamiento que se ha dado en llamar del “comportamiento financiero”. Este se basa en la evidencia del campo de la psicología que dice que la gente tiende a cometer sistemáticamente errores cognitivos cuando se forma expectativas. Uno de ellos puede ser el identificar erróneamente tendencias en donde no las hay, es decir, suponer que en el futuro cercano todo se comportará como recientemente lo ha hecho.

Las fuertes caídas en las bolsas de valores refuerzan la duda en la teoría

de los mercados eficientes y obligan a buscar otras razones de su comportamiento, basándose en otros fundamentos económicos. Este es un tema que continúa investigándose y que, sin duda, generará otras teorías y polémicas.

Otro tipo de mercado es el emergente, que también podríamos interpretar como el *no maduro*, siendo este el caso de México y otras naciones de economía similar.

En este tipo de mercados, la falta de disponibilidad o la inexactitud de la información financiera puede originar ganancias para los inversionistas, cosa que se puede refutar actualmente por la difusión de las tecnologías de la información y el abaratamiento de los costos para la obtención de datos. Algunos autores coinciden en que los mercados emergentes siguen en cierta forma a los eficientes, lo cual provoca que sean afectados por su cercanía con ellos.

En nuestro caso, tradicionalmente se ha ligado el comportamiento del IPC al del Dow Jones aunque esto no se ha demostrado a cabalidad y no se podría afirmar que tenemos el comportamiento del mercado eficiente porque, incluso, un análisis visual del comportamiento de ambos índices indica que son diferentes.

Las siete economías o mercados emergentes, según una reciente clasificación son las de: China, Brasil, Rusia, India, México, Indonesia y Turquía.

Con esto se puede considerar el aplicar herramientas de análisis para el comportamiento del IPC puesto que no es considerado un mercado eficiente.

2.4.2. El Índice de Precios y Cotizaciones (IPC)

El IPC es una muestra de las acciones negociadas en un determinado día en el mercado accionario mexicano y es un indicador del comportamiento de la economía mexicana. Dicha muestra está determinada por la metodología creada por la Bolsa Mexicana de Valores, [Metodología, 2012] tomando como valor de inicio el del 30 de octubre de 1978. Para entender la estructura de este índice necesitamos unos valores, mismos que se encuentran definidos en la metodología referida.

El primero es el de *acciones flotantes*, que se refiere a aquellas que resultan de restar al total listado de acciones, las que están en posesión de empresas o personas y que lo podríamos interpretar como acciones libres para ser negociadas ya que no se ejerce control en específico sobre ellas que impida su negociación.

Después está el *valor de capitalización flotado de la serie accionaria de la emisora*, que no es más que el resultado de multiplicar el número de acciones flotantes por el precio de mercado de la acción. Se define por:

$$VCF_{it} = (FAF_{it} * Q_{it}) * P_{it} \quad (2.1)$$

Donde:

VCF_{it} : Valor de capitalización flotado

FAF_{it} : Porcentaje de ajuste por acciones flotantes de la emisora i

Q_{it} : Número de acciones inscritas de la serie accionaria de la emisora i

P_{it} : Precio de mercado de la serie accionaria de la emisora i

Se usa también el *porcentaje de acciones flotantes de una serie accionaria de una emisora* que es el porcentaje que representa el valor flotado del valor total de una serie accionaria, también se le conoce como float:

$$\%AF_{it} = \frac{AF_{it}}{AI_{it}} \quad (2.2)$$

Donde:

$\%AF_{it}$: Porcentaje de acciones flotantes

AF_{it} : Acciones flotantes de la serie accionaria de la emisora i

AI_{it} : Acciones inscritas de la serie accionaria de la emisora i

Otro valor a considerar es la *rotación de la serie accionaria de una emisora*, que es el resultante de dividir el número de acciones negociadas en un período entre el número de acciones flotantes de la emisora:

$$R_{it} = \frac{Vol_{it}}{AF_{it}} \quad (2.3)$$

Donde:

R_{it} : Rotación

Vol_{it} : Volumen de acciones negociadas de la serie accionaria i

AF_{it} : Acciones flotantes de la serie accionaria i

Por último, está la suma en un período dado de las acciones negociadas de la serie accionaria de la emisora multiplicado por el precio de la transacción

y que se conoce como el *importe operado de una serie accionaria de una emisora*:

$$Imp_{it} = \sum_j Vol_{it} * P_{it,j}, j = 1 \dots n \quad (2.4)$$

Donde:

Imp_{it} : importe operado de una serie accionaria de una emisora

Vol_{it} : Volumen de acciones negociadas de la serie accionaria i

P_{it} : Precio al que se cerró la transacción

j : número de operaciones realizadas en el período

Para la selección de las emisoras que formarán parte del índice se siguen una serie de pasos o filtros, al final de los cuales se escogen 35 emisoras considerando solamente una serie accionaria por cada una de ellas y que regularmente es, la que califica como más bursátil.

Proceso de selección.

1. **Tiempo mínimo de operación continua.** Serán elegibles aquellas emisoras que tengan al menos tres meses calendario de operación continua, previos al mes de revisión de la muestra.
2. **Porcentaje de acciones flotantes mínimo.** Serán elegibles las emisoras cuyo porcentaje de acciones flotantes sea igual o mayor al 12 % y además, condición no necesaria, que su valor de capitalización flotado sea igual o mayor a 10 mil millones de pesos en la fecha de selección.
3. **Valor de capitalización flotado mínimo.** De las series accionarias de las emisoras que pasaron los dos filtros anteriores, serán seleccionadas aquellas cuyo valor de capitalización sea mayor o igual al 0.1 % del valor de capitalización de la muestra del IPC en la fecha de selección.
4. **Mayor factor de rotación.** Del total de emisoras que pasaron por los filtros anteriores, se elegirán las 45 emisoras cuya serie accionaria mas bursátil tenga el mayor factor de rotación. Para ello se toma la mediana de las 12 medianas mensuales anteriores a la fecha de selección.

5. **Calificación conjunta de indicadores.** Las 45 emisoras anteriores reciben una calificación por cada uno de estos valores: factor de rotación, valor de capitalización y la mediana mensual señalada en el paso anterior. Los valores se ordenan de mayor a menor en tres tablas distintas, otorgando una calificación ascendente para las emisoras que figuren en ellas. Al final se escogen las 35 emisoras con la calificación más baja.

Gráficamente el proceso se puede apreciar en la Figura 2.1:

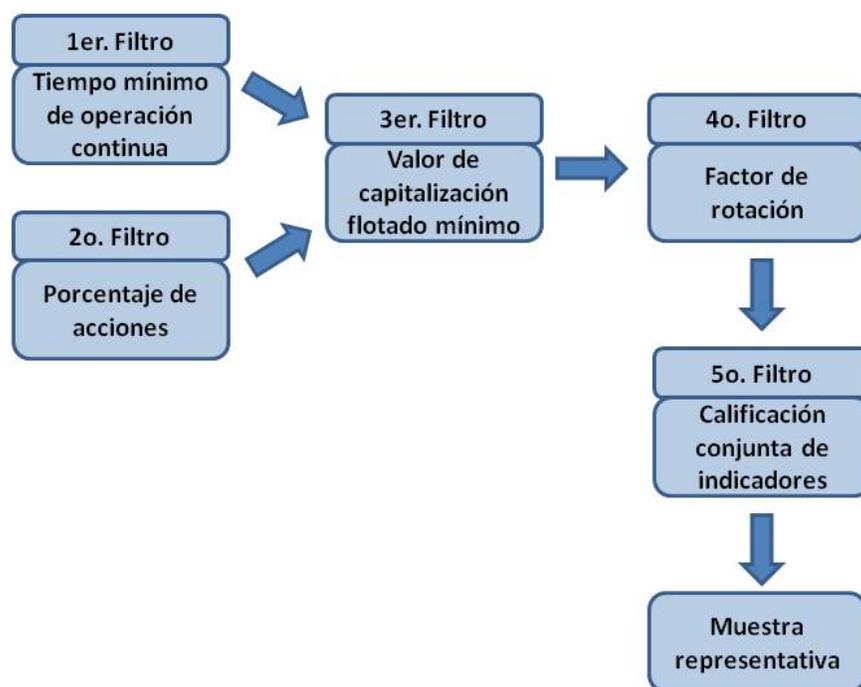


Figura 2.1: Proceso de selección de las emisoras que conforman el IPC

Con las emisoras escogidas anteriormente, diariamente se realiza el cálculo del índice, utilizando la siguiente fórmula:

$$I_t = I_{t-1} \left(\frac{\sum P_{it} * (Q_{it} * FAF_i)}{\sum P_{it-1} * (Q_{it-1} * FAF_i) * f_{it-1}} \right) \quad (2.5)$$

Donde:

I_t : Índice en el día t

P_{it} : Precio de la serie accionaria i el día t

Q_{it} : Acciones inscritas de la serie accionaria i el día t

FAF_i : Factor de ajuste por Acciones Flotantes de la serie accionaria i

f_i : Factor de ajuste por ex-derechos de la serie accionaria i el día t

$i = 1 \dots n$

Base: 0.78 = 30 de octubre de 1978

Este índice se revisa anualmente con los datos de cierre al mes de julio, aplicando los cambios si los hubiera, el primer día del mes de septiembre. Los pesos relativos de las series accionarias dentro de la muestra tienen una revisión trimestral dentro de ese año de vigencia.

Previo al capítulo donde se expondrá el uso de la metodología para la creación de la red neuronal, a continuación, se describirán algunos de los principales usos y aplicaciones de esta teoría.

Capítulo 3

Antecedentes del uso de las redes neuronales artificiales en diversas áreas de negocio y en los índices de los mercados de valores

Las RNA han demostrado su utilidad en diversos campos y aplicaciones: sistemas de identificación y control, reconocimiento de voz, gestos y escritura, finanzas, psicología, minería de datos, procesamiento de imágenes, etcétera. En particular en el área de las finanzas se han usado para predecir bancarrotas, comportamiento de tasas de interés, precios de acciones, volumen de negociación, índices de mercados y bolsas de valores, tipos de cambios de monedas extranjeras, entre otros.

Algunas de las propiedades y ventajas de trabajar con las RNA son señaladas por Simon Haykin [Haykin Simon, 1999]:

- **No linealidad:** si una RNA tiene neuronas que son no lineales, la red es no lineal. Si el fenómeno que origina las señales de entrada para la red es no lineal, la importancia es obvia.
- **Mapeo de entrada-salida:** una técnica muy usada en la aplicación de las RNA es el *aprendizaje supervisado*. A la red se le presentan una serie de entradas o patrones, para los cuales se conoce la respuesta o el resultado. La red se encarga entonces de comparar su propio resultado contra el real, si la diferencia es mayor o menor que un cierto parámetro, entonces se modifican los pesos sinápticos y se vuelve

a aplicar el proceso de aprendizaje de la red hasta que no se observe una mejora significativa en el aprendizaje. Esto se puede lograr incluso cambiando el orden de presentación de las entradas, por lo que la red es capaz de aprender elaborando un mapeo de entrada-salida del problema que se le está presentando. Esto es análogo a la inferencia estadística no paramétrica.

- **Adaptabilidad:** una red neuronal tiene la capacidad de aprender de los cambios en el medio ambiente en que se originó si se entrena nuevamente y si estos cambios además son de menor importancia. Una vez que la red es estable, a través del aprendizaje con distintos cambios, se puede decir que seguirá siendo útil cuando se le presente un sistema no estacionario. Sin embargo hay un dilema conocido como el de *plasticidad-elasticidad* que dice que las constantes de tiempo de un sistema deben de ser lo suficientemente amplias como para ignorar las alteraciones falsas pero también lo suficientemente cortas como para responder a los cambios significativos del ambiente.
- **Respuesta evidencial:** aparte de servir para la clasificación de patrones, una RNA también proporciona información sobre la validez entre un patrón y otro, ayudando a rechazar patrones ambiguos y por lo tanto, mejorando la labor de clasificación de la red.
- **Información contextual:** puesto que la actividad de una red neuronal potencialmente afecta a todas las neuronas de la misma, el manejo de la información contextual es una característica de ella.
- **Tolerancia a fallos:** por el conocimiento que se tiene del estudio anatómico se sabe que un daño parcial a una red de neuronas no afecta su desempeño puesto que la información está distribuida en toda la red. Es necesario que el daño sea mayor para que la red sufra una afectación mayor en su desempeño.
- **Implementación a gran escala:** la integración a gran escala o *Very Large Scale Integration* (VLSI por sus siglas en inglés), es el proceso para crear circuitos integrados combinando miles de circuitos basados en transistores en un solo chip. Por la gran capacidad de una red neuronal para el proceso en paralelo, es adecuado ocuparla en este tipo de desarrollos.

- **Uniformidad de análisis y diseño:** la notación y el lenguaje usado en la teoría de las redes neuronales goza de una amplia aceptación en la comunidad científica por lo que la construcción, desarrollo, aprendizaje y divulgación de la investigación se puede realizar de una forma consistente y sencilla.
- **Analogía neurobiológica:** actualmente se pueden citar una gran cantidad de ejemplos en los que la investigación neurológica inspiró el desarrollo de modelos matemáticos o viceversa. Como lo señala el autor que está siendo citado en esta sección, el cerebro es la prueba viviente de que el proceso tolerante a fallos no solo es físicamente posible sino que además es rápido y poderoso. Esta analogía es una gran fuente de inspiración para la investigación y el desarrollo.

A continuación se expondrán algunos ejemplos de la utilización de las redes neuronales en el campo del mercado de capitales.

Si bien la teoría de las redes neuronales comenzó a desarrollarse desde 1940, el primer trabajo en ocuparlas para el pronóstico de mercados fue el de Halbert White [White H., 1988]. Este investigador se ocupó de encontrar regularidades no lineales en series de tiempo económicas, en particular, con los precios de las acciones de IBM. Sus expectativas fueron demasiado optimistas y concluyó que “la red neuronal no es una máquina de hacer dinero”.

Un estudio que concluye que las redes neuronales con retropropagación son mejores que los modelos de regresión, cuando se tienen datos limitados, es el realizado por Chiang et. al. [Chiang et. al., 1996], al pronosticar los valores netos de los activos al cierre del año de los fondos de inversión.

Problemas	Estudios
Ganancias contables, ganancias sorpresa	Callen et al. (1997), Dar y Chou (2001)
Ciclos de negocio y recesiones	Qi (2001)
Falla de negocios, bancarrota o salud financiera	Yang (1999), Zhang et. al. (1999), Mckee y Greenstein (2000), Anandarajan et. al. (2001), Atiya (2001)
Gastos de los consumidores	Church y Curram (1996)

Continúa en la siguiente página

Cuadro 3.1 – *Continúa de la página anterior*

Problemas	Estudios
Precios de materia prima, elección de precios	Kohzadi et. al. (1996), Yao et. al. (2000)
Preferencia del consumidor, categoría de mercado, segmentos de mercado, segmento de mercado y tendencias	Agrawal y Schorling (1996), West et. al (1997), Aiken y Bsar (1999), Wang (1999), Jiang et. al. (2000), Vellido et. al. (1999)
Demanda de electricidad	Elkateb et. al. (1998), Darbellay y Slama (2000), Hippert et. al. (2001)
Tipos de cambio	Zhang y Hu (1998), Leung et. al. (2000a), Nag y Mitra (2002)
Mercado de Futuros	Kaastra y Boyd (1995), Ntungo y Boyd (1998)
Crecimiento del PIB (Producto Interno Bruto), Inflación, Producción Industrial	Tkacz (2001), Chen et. al. (2001a), Tseng et. al. (2001)
Volumen internacional de pasajeros aéreos, demanda de turistas, demanda de viajes	Nam y Schaefer (1995), de Carvalho et. al. (1998), Law (2000)
Control de inventarios	Bansal y Vadhavkar (1998), Partovi y Anandarajan (2002)
Comportamiento de fusiones	Hu et. al. (1999)
Valores netos de fondos de inversión y su comportamiento	Chiang et. al. (1996), Indro et. al. (1999)
Niveles de concentración de ozono, Predicciones del medio ambiente, calidad del aire	Prybutok et. al. (2000), Ruiz Surez et. al. (1995), Murtagh et. al. (2000), Kolehnmäinen et. al. (2001)
Demanda y venta de productos, Ventas al menudeo	Ansuj et. al (1996), Luxhoj et. al. (1996), Charytoniuk et. al. (2000), Alon et. al. (2001), Kuo (2001), Zhang y Qi (2002)
Crecimiento de Proyectos, Proyectos exitosos	Thieme et. al. (2000), Zhang et. al. (2003)

Continúa en la siguiente página

Cuadro 3.1 – *Continúa de la página anterior*

Problemas	Estudios
Demanda de construcción residencial, valor de vivienda	Hua (1996), Goh (1998), Nguyen y Cripps (2001b)
Ganancias de acciones, Precios de acciones, Tendencias de acciones, Riesgos de acciones	Wang y Leu (1996), Wittkemper y Steiner (1996), Desai y Bharati (1998), Saad et. al (1998), Qi (1999), Leung et. al. (2000), Chen et. al (2003)
Tráfico	Doughetry y Cobbett (1997), Kirby et. al. (1997), Chen et. al. (2001b), Dia (2001), Qiao et. al. (2001)
Bonos del Tesoro de Estados Unidos	Cheng et. al. (1996)

Cuadro 3.1: Aplicaciones de las RNA en áreas de negocio

Una lista de aplicaciones en áreas de negocios es ofrecida por Peter Zhang, [Zhang Peter G., 2004] en el Cuadro 3.1, en donde podemos observar la amplitud de estas.

Por otra parte, según una compilación de artículos en donde específicamente se ocupan las RNA para el pronóstico del comportamiento de los índices y las acciones en los mercados de valores, realizada por David Enke y Suraphan Thawornwong [David Enke y Suraphan Thawornwong, 2004], se comprueba la efectividad de esta herramienta para este tipo de series de tiempo. Un resumen de los trabajos en los cuales se ocupa solamente de un índice se muestra en el Cuadro 3.2.

En esta compilación la mayoría de los artículos basan sus predicciones en los índices de los mercados, que, como ya se había señalado, representan el comportamiento de las acciones en el mismo. La mayoría de los estudios, 33 de 38, reportan que la mejor herramienta de pronóstico son las RNA cuando se comparan contra los indicadores de referencia. Incluso aquellos que realizaron simulaciones de transacción, también reportaron mayores ganancias, 22 de 24, contra los que sólo se basaron en los indicadores de referencia considerados.

Índice	Artículo
Euronext Paris Stock Exchange (CAC)	Refenes et al. (1997)
German Stock Exchange (DAX)	Brown et al. (1998), Gencay (1996), Gencay et al. (1997), Gencay et al. (1998), Gencay (1998)
Dow Jones Industrial Average (DJIA)	Brown et al. (1998), Gencay (1996), Gencay et al. (1997), Gencay et al. (1998), Gencay (1998)
Financial Times Stock Exchange 100 Share (FTSE 100)	Brownstone (1996)
Hang Seng Stock Exchange in Hong Kong (HSI)	Lam et al. (2000)
Korea Stock Exchange (KOSPI)	Kim (1998), Kim et al. (2000), Oh et al. (2002)
Madrid Stock Exchange in Spain	Fernandez-Rodriguez et al. (2000)
Morgan Stanley USA Capital Market (MSCI)	Wood et al. (1996)
New York Stock Exchange (NYSE)	Leigh et al. (2002)
Standard and Poor's 500 Stock Exchange (S& P 500)	Austin et al. (1997), Chenoweth et al. (1996a, 1996b), Desai et al. (1998a), Dorsey et al. (1998), Qi (1999), Qi et al. (1999), Tsaih et al. (1998)
Singapore Stock Price	Kim et al. (1998)
Tokyo Stock Exchange Price in Japan (TOPIX)	Kohara et al. (1996, 1997)
Taiwan Stock Exchange Weighted Index (TSEW)	Kuo (1998)
Wilshi 5000 Stock Exchange (Wilshi)	Chandra et al. (1999)
Warsaw Stock Exchange (WSE)	Zemke (1999)

Cuadro 3.2: Artículos que tratan de un sólo índice

Algunos autores resaltan la importancia de pronosticar la dirección de cambio del índice más que del valor específico, (Aggarwal & Demaskey, 1997; Wu & Zhang 1997). Steven Walczack [Walczak, 2004] encontró que la predicción del signo de cambio de los índices de los mercados emergentes como México y Brasil, mejora significativamente al ligarlos con los de economías más poderosas. Cuando considera los mercados emergentes por sí solos, siguen un comportamiento muy parecido al de la caminata aleatoria, comprobando con esto la teoría de los mercados eficientes.

Capítulo 4

Metodología

Diversos esfuerzos se han realizado para utilizar las redes neuronales en la predicción o estimación de los valores futuros en finanzas. Estos, van desde estudios de una sola variable hasta combinaciones y cruces con varias de ellas. Estas variables pueden ser: el precio de una o varias acciones en un mercado específico, el tipo de cambio de una moneda en varios mercados o el índice de una bolsa de valores.

Además de la decisión o atracción por trabajar con ciertas variables, está el problema de decidir la metodología a utilizar, el modelo y el tratamiento de la o las variables en sí. Así que, una vez tomada la decisión del caso a analizar, el siguiente paso es definir el tamaño de los datos que se quiere considerar. Surge la pregunta: ¿es adecuado tomar una gran cantidad de datos? o ¿sólo una parte de ellos?. Lo que dice una regla heurística en series de tiempo es que entre más datos se tengan, mejor será la estimación. Sin embargo hay estudios que demuestran que tomando solo los datos más cercanos, se puede generar una mejor predicción. Esto es lo que se llama el “efecto de cercanía en las series de tiempo”, señalado por [Walczak, 2001].

4.1. Metodología

Como ya se ha señalado, la presente investigación se realizará con base en los datos del cierre diario del Índice de Precios y Cotizaciones (IPC) de la Bolsa Mexicana de Valores (BMV), del 3 de enero de 2000 hasta el 6 de julio de 2012, con un total de 3,151 registros.

La metodología usada para crear este índice, busca que este número sea

un indicador del comportamiento diario del mercado y como tal se tomará, sin ningún análisis adicional de la validez de la representatividad del mismo.

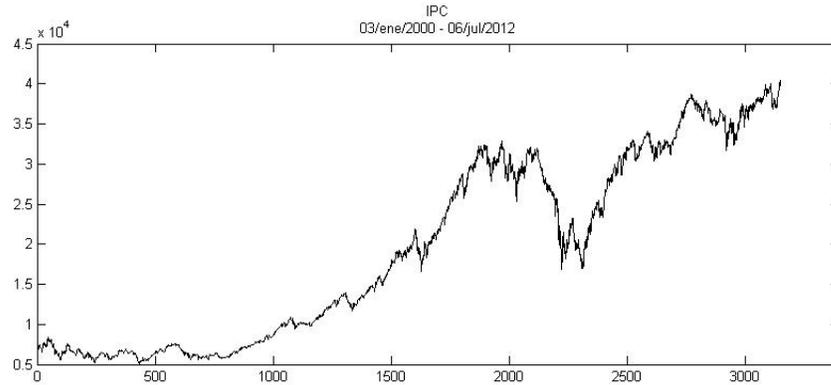


Figura 4.1: Índice de Precios y Cotizaciones del 3 de enero del 2000 al 6 de julio de 2012.

Fuente: elaboración propia con los datos del IPC de la BMV

Si se observa la gráfica de la Figura 4.1, esta sugiere definir alguna estacionalidad o comportamiento cíclico y quizá tratarlo de ligar con alguna otra variable económica, como el índice de otra bolsa de valores por ejemplo. Sin embargo, como ya se señaló, este trabajo se centra solamente en el IPC, haciendo un análisis homogéneo.

Una medida interesante para este índice, es la variación que tiene contra el cierre del día inmediato anterior ya que esto refleja si el mercado operó con ganancia o pérdida en el día de referencia. Siendo este un mercado donde se opera con recursos económicos, cuando dicha diferencia es positiva significa que al menos para un cierto porcentaje de los participantes hubo alguna ganancia. De manera que si podemos predecir el signo que tendrá para un día determinado de operación, tomando como referencia los valores anteriores, se podría contar con la información que serviría como un indicador adicional para ayudar en la toma de decisiones de los inversionistas.

Sea V_t el número que representa la variación del índice en el tiempo t al tomar la diferencia de este contra el del día hábil inmediato anterior:

$$V_t = I_t - I_{t-1} \tag{4.1}$$

Obviamente esta variación puede ser negativa o positiva y se acostumbra expresarla en unidades por ejemplo, si en determinado día ésta es de -28.14, entonces se dice “el índice tuvo una variación de menos 28.14 unidades”. Una gráfica de estas variaciones es la que se muestra en la Figura 4.2:

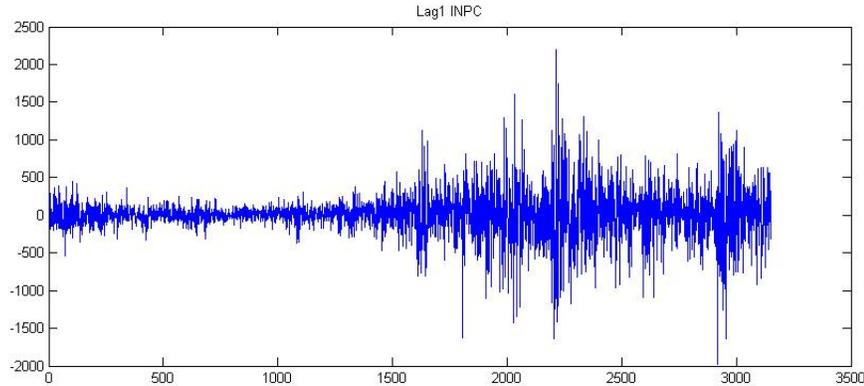


Figura 4.2: Variación del IPC contra el cierre del día hábil inmediato anterior.

Fuente: elaboración propia con los datos del IPC de la BMV

Este retraso se puede ampliar y tomar la variación a 2, 3 o más días, sin embargo, cuando se requiere tener una variación confiable para la predicción del signo de cambio del día siguiente, se hace necesario considerar los retrasos a 1, 2 y 5 días, tomando como referencia los trabajos de [Walczak et al, 1998], [Walczak, 2001], [Walczak, 2004], en donde demuestra que dichos retrasos son los más adecuados para este tipo de series de tiempo, incluido el IPC, que en particular se analiza en [Walczak, 2004].

Iebeling Kaastra y Milton Boyd [I. Kaastra y M. Boyd, 1996], proponen un procedimiento para desarrollar un modelo de red neuronal para pronosticar series de tiempo financieras y económicas. Para la construcción del presente modelo se usara este procedimiento como referencia, complementándolo con investigaciones más recientes. Este procedimiento se basa en 8 pasos o etapas:

1. Selección de las variables.
2. Recolección de los datos.
3. Preparación de los datos.

4. Conjuntos de entrenamiento, prueba y validación.
5. Aplicación de los paradigmas de las redes neuronales.
6. Criterios de evaluación.
7. Entrenamiento de la red neuronal.
8. Implementación.

4.1.1. Selección de las variables

La variable del presente trabajo será una variable técnica, que es la que se calcula de los retrasos tomados de la misma, de un indicador financiero. En este caso, como ya se señaló, es el Índice de Precios y Cotizaciones (IPC), de la Bolsa Mexicana de Valores.

Otro tipo de variable es la fundamental: un indicador económico que se supone tiene influencia sobre la variable dependiente. Por ejemplo: el tipo de cambio peso-dólar, el índice de la bolsa de Nueva York, el valor futuro del precio del petróleo, etcétera.

Aún cuando se ha referido el efecto de cercanía en las series de tiempo [Walczak, 2001], que establece que se dan mejores resultados cuando se consideran conjuntos de datos pequeños, por ejemplo, no más de dos años hacia atrás del punto de pronóstico, se tomarán los índices del 3 de enero de 2000 al 6 de julio de 2012, con un total de 3,151 registros.

4.1.2. Recolección de los datos

Todos los días en los que tiene operación el mercado de valores en México, este índice se hace del conocimiento público al cierre de las operaciones a través del sitio oficial de la BMV y con el envío de los reportes oficiales. Se puede tener disponible mediante suscripción gratuita via SRS (Really Simple Syndication), por sus siglas en inglés. Los datos que aquí se usarán se obtuvieron de esta fuente y de la página citada al pie ¹. Estos datos están disponibles en forma gratuita, sin embargo si se quiere obtener los datos históricos directamente de la BMV, éstos tienen un costo que en algunos casos no es tan accesible.

¹Histórico de cotizaciones México IPC <http://www.labolsa.com/indices/MX/historico/>

El costo de la obtención de los datos es algo que hay que considerar al momento de crear el modelo de red neuronal, porque también se suma el tiempo invertido en ello, el software que se usará para el análisis y el tiempo de proceso que se consumirá en la computadora. Si bien, con el poder de cómputo que se tiene actualmente disponible esto podría pasar a segundo plano, al diseñar redes complejas sí puede ser un elemento importante del costo.

Una parte esencial de los datos que se analizarán es asegurarse que se seguirán publicando con la misma periodicidad con la que se obtuvieron al momento del estudio. Como se señaló anteriormente, esto se garantiza vía la suscripción SRS al sitio de la BMV.

4.1.3. Preparación de los datos

La lista de los datos obtenidos se recopiló en una hoja de cálculo electrónica. En ella se organizó por columnas la información de los registros diarios del IPC en su valor mínimo, máximo y de cierre, siendo este último el de interés para este estudio. Los datos se ordenaron por fecha en forma ascendente y en una celda contigua y, por fórmula, se obtuvo el día de la semana para revisar la consistencia en los días hábiles de lectura, con la finalidad de que no existiera un registro de un día no hábil.

Se calcularon los retrasos de la variable de 1, 2 y 5 días, por considerarse lo más adecuado para esta serie de tiempo, como se estableció al principio de este capítulo. Dichos retrasos quedaron registrados en columnas contiguas a la serie original, dando lugar a las matrices de 3 columnas por el número de renglones a ocupar para cada análisis. En este caso, como la intención es el registro de retrasos hasta de 5 días de la muestra, eso ocasiona tener renglones completos a partir del número 5, por lo que la muestra se reduce en el mismo número, quedando un total de 3,146 registros.

El tratamiento posterior de los datos: cálculo de valores absolutos, cambio de escala, graficación y la aplicación del modelo de red neuronal, se realizó en el software Matlab versión 7.9.0.529.

4.1.4. Conjuntos de entrenamiento, prueba y validación

El conjunto de datos para armar el modelo neuronal quedó formado por 2,546 registros abarcando desde el 10 de enero de 2000 hasta el 17 de febrero de 2010, dejando los 600 registros restantes para el conjunto de prueba con

datos reales. La definición de los porcentajes para el entrenamiento, prueba y validación del conjunto de datos, se realizó dentro de la herramienta, observando los recomendados por el método señalado: 70 % para el entrenamiento, 20 % de prueba y el restante 10 % para validación.

Con esto se realizó la recomendación de conservar los datos de validación lo más cercano a la fecha en la que se utilizó el modelo para el pronóstico. Para el uso posterior de este modelo las fechas se pueden ir corriendo conforme se tengan más datos, descartando los más antiguos para conservar la proporción propuesta. Gracias al registro en la hoja de cálculo y el software utilizado para la fabricación de la red neuronal, se logró un procedimiento sencillo que permite evaluar la efectividad del modelo construido.

4.1.5. Aplicación de los paradigmas de las redes neuronales

Los modelos de redes neuronales han probado ser buenos aproximadores de funciones y para descubrir la relación de variables de comportamiento complejo. En particular, el modelo de retropropagación ha demostrado ser útil para el pronóstico de variables de tipo financiero, como se puede observar en los trabajos de [Fu, 1994], [Walczak, 1999] y [Widrow et al, 1994].

El software señalado en la sección de la preparación de los datos, facilitó la construcción del modelo de red neuronal. La mayoría de los autores coinciden en que encontrar la red más adecuada para un problema específico es un proceso de prueba y error hasta dar con la combinación óptima de neuronas de entrada, de salida, número de capas y funciones de transferencia. Para este estudio se eligió un modelo de dos capas, entrada y oculta, con la sigmoide como función de transferencia a la capa oculta y la identidad como función de transferencia para la salida. La gráfica que proporciona el software para el modelo se aprecia en la Figura 4.3.

En el artículo que se usó como base para la construcción del modelo, se cita a los autores que han encontrado que los mercados financieros son no lineales y tienen una memoria, lo cual da soporte a la idea de que el uso de las funciones de transferencia no lineales son las más apropiadas, [Richard Levich y Lee Thomas, 1993] y [Kao Wenchi y Ma Christopher, 1992].

En cuanto al número de neuronas de entrada, se realizaron varias pruebas para encontrar el número que mejor proporcionó la estimación para la dirección de cambio del signo del IPC. El software utilizado tiene la facilidad

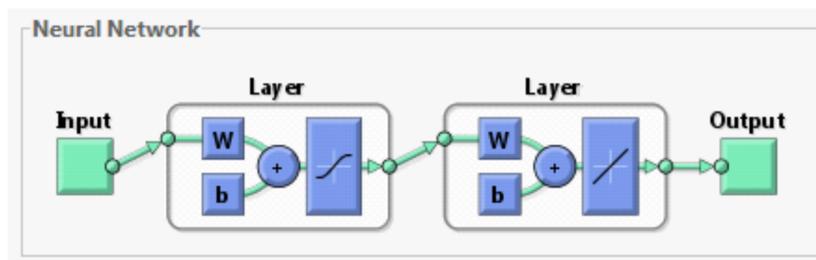


Figura 4.3: Modelo de la red neuronal generada por Matlab. Se ocuparon dos capas, entrada y oculta, con la sigmoide como función de transferencia a la capa oculta y la identidad como función de transferencia para la salida.

Fuente: software Matlab versión 7.9.0.529

de salvar las instrucciones tecleadas para crear el modelo y posteriormente usarlas sólo cambiando los parámetros que se deseen modificar. Este código de Matlab ayudó a evaluar los cambios en el modelo escogido y con ello se llegó a que el número de neuronas apropiado fue de 5. El código comentado se encuentra en el apéndice B.

4.1.6. Criterios de evaluación

En esta sección se definen los criterios de evaluación para los resultados obtenidos al aplicar el modelo neuronal elegido. Si el objetivo de éste es pronosticar la dirección de cambio en el cierre diario del IPC, el vector resultante después de presentar los nuevos datos al modelo es una serie de ceros y unos, de los cuales interesa el último de ellos. Se define entonces que un cero indica que el cambio para el día siguiente será negativo y un uno, lo contrario.

Así, se tiene que si después de presentar los datos hasta el día t el último valor del vector es 1, indica que el índice al día $t + 1$ será mayor que el registrado al día t . El modelo se centra sólo en predecir la dirección de cambio no el índice en sí, por lo que si esta dirección es positiva, implicará un incremento en unidades del mismo. Esta predicción de la dirección aunada a otras herramientas o predicciones y al conocimiento de la situación del mercado financiero, podría sentar las bases para formular una estimación confiable del comportamiento del IPC.

4.1.7. Entrenamiento de la red neuronal e implementación

A la última sección, la de entrenamiento, se decidió unir la de implementación, ya que como se señaló el software utilizado facilitó estas tareas.

En el modelo desarrollado, el número de iteraciones para la búsqueda de los pesos óptimos de la red se fijó en 1000, modificándose solamente en caso de que la red no lograra encontrar esta combinación de los mismos. En particular no fue necesario incrementar este número ya que el mayor número de iteraciones estuvo por debajo de las 100.

El software también puede entrenar a la red una vez que se encuentran los parámetros óptimos para los datos presentados ya que cuenta con la capacidad de guardar el escenario logrado para ocuparlo posteriormente con los datos que se presentarán para el pronóstico.

Para la creación del modelo neuronal, se presentó a la herramienta una matriz de 3 X 2,546, que corresponde a la de las columnas que muestran los retrasos de 1, 2 y 5 días con renglones que representan a los días registrados hasta el 17 de febrero de 2010. El vector de los resultados de salida o destino, es el compuesto por ceros y unos, presentados en una matriz de una columna por 2,546 renglones. Para la estimación de la variación del IPC del día siguiente se consideró el resultado esperado para el día actual que es el que se obtiene de la variación del siguiente día. Es decir, que para obtener ésta se hace un cálculo basado en el resultado conocido para entrenar a la red a predecir el resultado que en determinado momento se desconoce.

Ambas matrices se presentaron a la herramienta para generar el modelo de red neuronal comentado anteriormente. En la Figura 4.4 se presenta el reporte de este proceso.

Posteriormente para los casos de prueba se consideraron los 600 registros restantes de la muestra, del 18 de febrero de 2010 al 6 de julio de 2012, que luego fueron complementados con 1,946 registros de fechas anteriores a la referida para cumplir con el tamaño en renglones del modelo generado.

Se crearon 600 vectores de prueba, combinando aleatoriamente los datos de los registros recién descritos para presentarlos al modelo y realizar la simulación con cada uno de ellos. Los componentes a en los vectores generados son valores que están en el rango $[0,1]$ y para efectos del modelo se aplicó la siguiente regla:

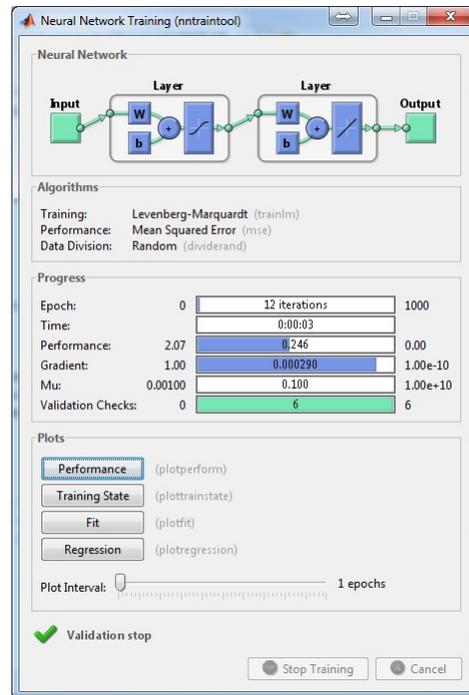


Figura 4.4: Reporte de resultados de la generación de la red neuronal según Matlab.

Fuente: software Matlab versión 7.9.0.529

$$\begin{aligned}
 a &= 0 & a &\leq 0,5 \\
 a &= 1 & a &> 0,5
 \end{aligned}$$

Una vez transformados los valores de los vectores generados se compararon contra los vectores destino otorgando un cero para los valores diferentes y un uno para los aciertos de manera que, al final, se obtuvo un porcentaje de aciertos o efectividad del modelo generado.

Para comprobar la efectividad del modelo, se generaron 10 vectores con datos de cierre recientes, del 24 de enero 2013 al 7 de febrero de 2013, tomando los 2,546 renglones que requiere el modelo en forma contigua y sin ninguna generación aleatoria, de manera que la fecha más antigua considerada para este caso fue el 18 de diciembre de 2002. Este resultado y el señalado en el párrafo anterior, será comentado en el capítulo siguiente.

C	Esperado	Generado	Calificación	Aciertos
1	1	0.5484	1	1
2	0	0.5837	1	0
3	0	0.6138	1	0
4	0	0.5618	1	0
5	1	0.5109	1	1
6	0	0.5010	1	0
7	0	0.6279	1	0
8	0	0.5563	1	0
9	0	0.5184	1	0
10	1	0.4745	0	0

Cuadro 4.1: Cuadro comparativo datos recientes

Fuente: datos propios

4.2. Resultados

La aplicación del modelo a los 600 casos de prueba, generó los *vectores resultado* que se compararon contra los vectores destino, aplicando la regla de clasificación antes descrita. Por cada vector generado se registró un porcentaje de efectividad, o de aciertos, no mayor al 54%, el cual se conservó en toda la muestra ya que la moda de todos los resultados fue de 54.54 y el promedio de 54.34. Una descripción detallada del archivo generado para la comparación de estos vectores se muestra en el apéndice A.

En este punto, con la finalidad de conocer la efectividad del modelo de experimentación y saber si era posible lograr al menos una tasa de aciertos igual a la registrada en los casos anteriores, se realizó una prueba más del modelo ingresando diez casos de valores que no habían sido presentados anteriormente. Considerando la capacidad de la red para aprender de los resultados que se le presentan y la historia previa a estos, como lo apunta la teoría las redes neuronales, se esperaba lograr, de ser posible, una tasa de aciertos igual a la registrada en los casos de prueba.

En el Cuadro 4.1 se presentan los resultados para los 10 casos con los datos recientes utilizados en este punto, del 6 de julio de 2012 al 19 de julio de 2012 y se observa que solo se obtuvo una eficiencia del 20%. Este porcentaje es menor al obtenido en los casos de prueba y no ofrece una confiabilidad en la red generada por la herramienta.

Conclusiones

Los resultados generados por la red neuronal usada en este trabajo sugiere que el comportamiento del mercado accionario mexicano, identificado por el IPC, tiene un comportamiento aleatorio, en cuanto a la generación de su índice de medición. La tasa de acierto de predicción del signo del índice señalado no excedió al 54 %, confirmando en cierta forma que no es posible hacer una predicción tomando únicamente esta variable técnica aislada.

Recordando la teoría de los mercados eficientes, que establece que todas las acciones adquieren su valor correcto a través de la búsqueda de las oportunidades de ganancia por parte de los inversionistas, se puede decir que es el comportamiento que sigue el mercado mexicano, sin embargo, en el ámbito financiero mundial es considerado como un mercado emergente, vulnerable y factible de ser influenciado por los mercados externos, como el de los Estados Unidos de Norteamérica, por su cercanía geográfica y poderío económico.

Como se citó anteriormente, [Walczak, 2004] encontró que al combinar los índices de ciertos mercados emergentes con aquellos más estables, como el de Estados Unidos o Inglaterra, mejora la tasa de predicción de los modelos ocupados hasta más de un 60 %.

Aunque los resultados encontrados obtenidos en la presente investigación no ofrecieron una herramienta de predicción confiable que usando valores de sí misma permitieran determinar el comportamiento del índice bursátil, se podría considerar de utilidad ya que da pie a nuevas investigaciones que con otras combinaciones de variables podrían arrojar nueva información que permita acercarse a la construcción de una herramienta mucho más precisa, capaz de adaptarse a las cambiantes necesidades del mundo financiero global y, al mismo tiempo, ser capaz de generar resultados significativos para el mercado local.

Por otra parte, el proceso de elaboración de este modelo, así como el conjunto de instrucciones para la herramienta utilizada, podrían ser consid-

erados de gran utilidad para todos los interesados en el estudio de las redes neuronales y su aplicación a datos de toda índole, no solo de índole financiera, en tanto se presenten como series de tiempo.

Esta es, sin duda, y más allá de los resultados obtenidos, una de las ventajas de las redes neuronales: con ellas se puede conceptualizar un problema en específico y lograr una predicción, si se presenta de la forma en la que una red neuronal puede interpretar.

La investigación y la exploración de las posibilidades de las redes neuronales, aunada a la capacidad que ofrecen las nuevas tecnologías de la información para el procesamiento de datos, invita a continuar las investigaciones en este campo y a mantener la búsqueda de nuevas aplicaciones para muchas más áreas de la vida práctica del ser humano. Este tema resulta exponencialmente inagotable y, seguramente, continuará siendo del interés de muchos especialistas que desde que comenzó el estudio del funcionamiento del cerebro humano han vislumbrado como replicable, y aplicable, en los diferentes campos de la ciencias.

Para finalizar, además de la descripción del modelo y el conjunto de instrucciones de la herramienta, presentados en los apéndices correspondientes, el autor ofrece a la comunidad interesada el total de los datos utilizados, ordenados en la forma en la que se presentó a la red neuronal.

Apéndice A

Descripción detallada de los datos

A.1. Datos fuente IPC

Orden	Día	Fecha	Máximo	Mínimo	Cierre	Calif. Dic.12	Lag1	Lag2	Lag5
3063	miércoles	29-feb-12	38282.18	37800.99	37816.69	1.00	-204.73	30.97	-95.86
3064	jueves	01-mar-12	38063.72	37706.48	38033.45	1.00	216.76	12.03	5.61
3065	viernes	02-mar-12	38332.36	37914.88	38327.43	0.00	293.98	510.74	382.21
3066	lunes	05-mar-12	38332.74	38019.06	38155.27	0.00	-172.16	121.82	369.55
3067	martes	06-mar-12	38155.22	37380.63	37487.63	1.00	-667.64	-839.80	-533.79
3068	miércoles	07-mar-12	37674.71	37472.22	37548.54	1.00	60.91	-606.73	-268.15
3069	jueves	08-mar-12	37823.67	37561.62	37815.43	0.00	266.89	327.80	-218.02
3070	viernes	09-mar-12	37938.10	37682.06	37691.01	0.00	-124.42	142.47	-636.42
3071	lunes	12-mar-12	37742.30	37501.18	37590.17	1.00	-100.84	-225.26	-565.10
3072	martes	13-mar-12	38048.62	37549.47	38019.76	0.00	429.59	328.75	532.13
3073	miércoles	14-mar-12	38082.75	37898.67	37971.44	1.00	-48.32	381.27	422.90
3074	jueves	15-mar-12	38142.88	37919.38	38134.57	1.00	163.13	114.81	319.14
3075	viernes	16-mar-12	38310.56	38057.31	38258.45	0.00	123.88	287.01	567.44
3076	martes	20-mar-12	38275.65	37979.87	38055.27	1.00	-203.18	-79.30	465.10
3077	miércoles	21-mar-12	38443.54	38050.30	38434.51	0.00	379.24	176.06	414.75
3078	jueves	22-mar-12	38446.02	38256.61	38323.51	1.00	-111.00	268.24	352.07
3079	viernes	23-mar-12	38422.74	38262.49	38334.90	1.00	11.39	-99.61	200.33
3080	lunes	26-mar-12	38942.41	38325.40	38863.15	1.00	528.25	539.64	604.70
3081	martes	27-mar-12	39134.88	38818.68	38956.32	0.00	93.17	621.42	901.05
3082	miércoles	28-mar-12	39033.33	38685.99	38910.68	1.00	-45.64	47.53	476.17
3083	jueves	29-mar-12	39199.68	38655.57	39125.35	1.00	214.67	169.03	801.84
3084	viernes	30-mar-12	39567.99	39100.66	39521.24	1.00	395.89	610.56	1186.34

Figura A.1: Datos fuente IPC

Los datos del cierre diario del IPC de la BMV, se recopilaron en una hoja de cálculo Excel, ordenados en forma ascendente por fecha, como se observa en la Figura A.1.

De izquierda a derecha, esta es la descripción de los campos:

1. **Orden.** Campo numérico progresivo que se ocupa para revisiones y confirmaciones de orden en los distintos procesos del modelo.
2. **Día.** Día de la semana en el que se está registrando el dato, sirve para revisar la consistencia de los días laborables.
3. **Fecha.** Fecha de registro de los datos.
4. **Máximo.** Valor máximo alcanzado por el IPC en el día indicado.
5. **Mínimo.** Valor mínimo alcanzado por el IPC en el día indicado.
6. **Cierre.** Valor del cierre del IPC, dato fuente del presente estudio.
7. **Calif.Dic.12.** Valor de retraso tomando el registro del día siguiente y ocupado para generar los vectores destino del modelo.
8. **Lag1.** Retraso de 1 día para el valor de cierre.
9. **Lag2.** Retraso de 2 días para el valor de cierre.
10. **Lag5.** Retraso de 5 días para el valor de cierre.

A.2. Hoja para generación aleatoria

En la Figura A.2 se muestra la estructura de la hoja utilizada para la generación aleatoria de los vectores de prueba. Los datos fuente están colocados de la columna 2 a la 6 y la primera columna contiene números generados aleatoriamente. En cada iteración para generar un vector de prueba con su vector destino, los datos se ordenan en forma ascendente y el resultado se copia hacia la derecha, datos de prueba y destino, volviendo a generar los números aleatorios y ordenando nuevamente para la siguiente iteración.

Si se desea comprobar la correcta ejecución del ordenamiento, se usa el campo Orden para localizar los datos fuente y comparar el resultado, tanto los datos de prueba como los de destino deben coincidir.

Esta es la descripción de los campos:

1. **Aleatorio.** Números generados aleatoriamente en cada iteración de la creación de vectores de prueba.
2. **Orden.** Campo que guarda el consecutivo del dato en la tabla de origen.

Aleatorio	Orden	Lag1	Lag2	Lag5	VD	Vector Generado	n	Tn
0.985925	1197	103.37	224.26	494.03	0	103.37,224.26,494.03,	-111.89,-38.06,7.95,	0
0.549025	1960	462.61	362.10	1244.75	1	462.61,362.1,1244.75,	59.19,149.23,234.65,	0
0.173680	153	-150.85	-116.00	-24.93	0	-150.85,-116,-24.93,	120.07,137.23,434.95,	1
0.337287	1076	0.50	-76.17	-103.86	1	0.5,-76.17,-103.86,	19.06,102.58,286.95,	1
0.395137	2210	-864.39	-2104.76	-4004.16	0	-864.39,-2104.76,-4004.16,	-79.36,-144.28,-385.82,	1
0.220546	359	-56.91	-7.26	164.06	0	-56.91,-7.26,164.06,	284.31,176.65,8.67,	1
0.555935	438	-18.76	1.70	224.16	1	-18.76,1.7,224.16,	11.61,34.68,31.18,	1
0.619958	2371	78.20	852.12	543.75	0	78.2,852.12,543.75,	-93.48,-428.32,130.49,	0
0.282261	1423	58.18	117.83	4.94	1	58.18,117.83,4.94,	-33.05,-25.92,24.43,	1
0.398546	918	64.36	66.75	41.93	1	64.36,66.75,41.93,	-250.01,-331.41,-361.56,	1
0.445551	1733	345.50	197.57	883.69	1	345.5,197.57,883.69,	68.99,274.08,90.71,	0
0.170576	2095	-737.21	-838.96	-786.66	0	-737.21,-838.96,-786.66,	-117.78,13.58,118.49,	0
0.140509	2042	-157.68	-1500.64	-334.66	1	-157.68,-1500.64,-334.66,	-73.28,97.66,439.74,	0
0.199760	382	-148.90	-144.99	-304.75	0	-148.9,-144.99,-304.75,	78.17,78.47,672.7,	0
0.667390	150	73.50	38.74	238.64	1	73.5,38.74,238.64,	823.09,925.94,1278.83,	1
0.884185	1638	-469.93	-524.26	240.19	1	-469.93,-524.26,240.19,	-1638.1,-1689.83,-1176.2	1
0.154722	2394	-6.07	-478.62	-412.84	0	-6.07,-478.62,-412.84,	91.03,162.41,342.83,	0
0.806123	2515	98.30	308.21	600.12	1	98.3,308.21,600.12,	304.83,848.81,1064.44,	1
0.300380	1465	27.11	-170.78	8.79	0	27.11,-170.78,8.79,	145.81,-209.6,173.61,	0

Figura A.2: Generación aleatoria de vectores de prueba

3. **Lag1.** Retraso de 1 día para el valor de cierre.
4. **Lag2.** Retraso de 2 días para el valor de cierre.
5. **Lag5.** Retraso de 5 días para el valor de cierre.
6. **VD.** Vector destino donde se ubican los valores para cada caso de prueba.
7. **Vector Generado.** Valores para el vector de prueba a presentar a la red neuronal.
8. **n.** Vector generado en la iteración n .
9. **Tn.** Vector destino en la iteración n .

Apéndice B

Código Matlab

```
%Aquí se indica el nombre del archivo que contiene los datos de entrada  
o de entrenamiento. Debe ser un archivo de texto CSV.  
Ent=importdata ('Training1.txt')
```

```
%Aquí se indica el nombre del archivo que contiene los datos de prueba  
o destino. Debe ser un archivo de texto CSV.  
Pru=importdata ('Pru1.txt')
```

```
%Estás son las órdenes para formar la transpuesta de la matriz formada  
por la lectura de los datos.  
%Fue necesario hacer esta transformación ya que los datos estaban  
ordenados por fecha en forma ascendente  
%y el modelo a ocupar toma los renglones de una matriz como los casos a  
presentar.  
TR=Ent'  
PR=Pru'
```

```
net = newfit(TR,PR,5) %Esta es la orden para crear la red neuronal:  
%           TR = Datos de entrada o de entrenamiento  
%           PR = Datos de prueba o destino  
%           nnn = Número de neuronas en la capa de entrada,  
en este caso es 5
```

```
net.divideParam.trainRatio = 70/100; % Porcentaje de los datos que se  
ocuparán para el entrenamiento
```

```
net.divideParam.valRatio = 20/100; % Porcentaje de los datos que se
    ocuparán para la validación

net.divideParam.testRatio = 10/100; % Porcentaje de los datos que se
    ocuparán para la prueba

[net,tr] = train(net,TR,PR); % Orden para el entrenamiento de la red
    con los parámetros obtenidos

Y=SIM(net,RE) % Aplicación del modelo a un vector RE presentado a la
    red
```

Bibliografía

- [Anderson, 2007] Anderson, James A. *Redes Neurales*, Alfaomega Grupo Editor, S.A. de C.V., 2007.
- [Chiang et. al., 1996] W.-C. Chiang, T.L. Urban, G.W. Baldrige *A neural network approach to mutual fund net asset value forecasting* Omega, Volume 24, Issue 2, April 1996.
- [David Enke y Suraphan Thawornwong, 2004] Enke David y Thawornwong Suraphan *Forecasting Stock Returns with Artificial Neural Networks*, en *Neural Networks in Business Forecasting*, Idea Group Publishing Inc. 2004.
- [Fu, 1994] Fu, Limin *Neural Networks in Computer Intelligence*, New York: McGraw-Hill. 1994.
- [Haykin Simon, 1999] Haykin Simon *Neural Networks, A Comprehensive Foundation*, Prentice Hall. 1999.
- [Kao Wenchi y Ma Christopher, 1992] Wenchi Kao y Christopher Ma, *Memories, heteroscedasticity and prices limit in currency futures markets*, The Journal of Futures Markets, Volumen 12, número 6, 1992.
- [I. Kaastra y M. Boyd, 1996] I. Kaastra y M. Boyd, *Designing a neural network for forecasting financial and economic time series*, en *Neurocomputing*, Elsevier Science B.V. 1996.
- [Metodología, 2012] *Nota Metodológica del Índice de Precios y Cotizaciones*, Bolsa Mexicana de Valores, abril 2012.
- [M. Minsky y S. Papert, 1969] M. Minsky y S. Papert *Perceptrons*, 1st. Ed. Cambridge: MIT Press, Tercera Edición, 1988.

- [Nilsson Nils J., 1965] Nilsson N. *Learning Machines*, New York: McGraw-Hill (reimpreso por Morgan-Kaufman Publishers, Los Altos, CA)
- [Peretto, 1992] Peretto, Pierre. *An Introduction to the Modeling of Neural Networks*, Press Syndicate of the University of Cambridge, 1992.
- [R. Beale y T. Jackson, 1990] Russell Beale and Tom Jackson, *Neural Computing: an introduction*, Institute of Physics Publishing, Bristol and Philadelphia, 1990.
- [Richard Levich y Lee Thomas, 1993] Levich Richard & Thomas Lee, *The significance of technical trading rule profits in the foreign exchange market: A bootstrap approach*, en *Strategic Currency Investing - Trading and Hedging in the Foreign Exchange Market*, Probus, Chicago, 1993.
- [R. Rojas, 1996] Rojas, *Neural Networks, A Systematic Introduction* Springer-Verlag, Berlín, 1996.
- [Shapiro A.] Arnold F. Shapiro *Capital Market Applications of Neural Networks, Fuzzy Logic and Genetic Algorithms* Penn State University, PA, 2003.
- [Steven L. Jones and Jeffry M. Netter, 2008] Steven L. Jones and Jeffry M. Netter, *Efficient Capital Markets. The Concise Encyclopedia of Economics* Library of Economics and Liberty. Recuperado de: <http://www.econlib.org/library/Enc/EfficientCapitalMarkets.html>
- [Walczak et al, 1998] Steven Walczak, Ali Tahai, and Khondkar Karim *Improved Cash Flows Using Neural Network Models for Forecasting Foreign Exchange Rates*, en *P. Siegel, K. Omer, A. deKorvin, & A. Zebda (Eds.) Applications of Fuzzy Sets and The Theory of Evidence to Accounting II*, Stamford, Connecticut: JAI Press, pp. 293-310.
- [Walczak, 2001] Walczak, Steven. *An Empirical Analysis of Data Requirements for Financial Forecasting with Neural Networks*, Journal of Management Information Systems, Vol 17, No. 4, 2001, Pags. 203-222.
- [Walczak, 2004] Walczak, Steven. *Forecasting Emerging Market Indexes with Neural Networks*, en *Neural Networks in Business Forecasting*, Idea Group Publishing Inc. 2004.

- [Walczak, 1999] Walczak, S. & Cerpa, N. *Heuristic principles for the design of artificial neural networks*, Information And Software Technology, Volumen 41, número 2. (1999).
- [White H., 1988] White Halbert. *Economic prediction using neural networks: the case of IBM daily stock returns* Neural Networks, 1988., IEEE International Conference on , vol.2, 24-27 July 1988
- [Widrow et al, 1994] Widrow, B.,Rumelhart, D. E., & Lehr, M. A. *Neural networks: Applications in industry, business and science*, en *Communications of the ACM*, Volumen 37, número 3. 1994.
- [W. S. McCulloch y W. H. Pitts, 1943] W.S. McCulloch y W.H. Pitts. *A Logical Calculus of the Ideas Immanent in Nervous Activity*, Bulletin of Mathematical Biophysics, 5, 115-133.
- [Zhang Peter G., 2004] Zhang Peter G., *Business Forecasting with Artificial Neural Networks: An Overview*, en *Neural Networks in Business Forecasting*, Idea Group Publishing Inc. 2004.

Índice de figuras

1.1.	Esquema sencillo de una neurona	8
1.2.	Dibujo del neocórtex de una rata	9
1.3.	Modelo básico de la neurona de McCulloch-Pitts	12
1.4.	Esquema del perceptrón desarrollado por Frank Rosenblatt	13
1.5.	Perceptrón temprano	14
1.6.	Perceptrón simplificado	15
1.7.	Separabilidad lineal	16
1.8.	Múltiples planos de solución	17
1.9.	Región de solución	18
1.10.	El problema XOR	20
1.11.	Una solución al problema XOR	21
1.12.	La función sigmoide	22
2.1.	Proceso de selección de las emisoras que conforman el IPC	39
4.1.	Índice de Precios y Cotizaciones de la BMV	50
4.2.	Variación a 1 día del IPC	51
4.3.	Modelo de red neuronal según Matlab	55
4.4.	Resultados de la generación de la red neuronal	57
A.1.	Datos fuente IPC	61
A.2.	Generación aleatoria de vectores de prueba	63

Índice de cuadros

1.1. Disyunción inclusiva	10
1.2. Conjunción	11
1.3. Disyunción exclusiva	20
2.1. Clasificación de instrumentos de deuda	30
3.1. Aplicaciones de las RNA en áreas de negocio	45
3.2. Artículos que tratan de un sólo índice	46
4.1. Cuadro comparativo datos recientes	58