



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
PROGRAMA DE MAESTRÍA Y DOCTORADO EN INGENIERÍA
INGENIERÍA DE SISTEMAS – OPTIMIZACIÓN FINANCIERA

IDENTIFICACIÓN DE CICLOS FINANCIEROS EN TIEMPO REAL PARA EL IPC MEDIANTE
LA RED NEURONAL ARTIFICIAL LVQ

TESIS

QUE PARA OPTAR POR EL GRADO DE:

MAESTRO EN INGENIERÍA

PRESENTA:

ING. JUAN CARLOS DE LA TORRE HINOJOSA

TUTOR:

DR. FEDERICO HERNÁNDEZ ÁLVAREZ
FACULTAD DE INGENIERÍA

CIUDAD UNIVERSITARIA, CD. MX. SEPTIEMBRE 2016

JURADO ASIGNADO:

Presidente: Dr. Sánchez Lara Benito
Secretario: Dr. Martínez Miranda Elio Agustín
Vocal: Dr. Hernández Álvarez Federico
1^{er.} Suplente: M.I. Figueroa Palacios José Domingo
2^{do.} Suplente: M.I. Rodríguez Rubio Jorge

Lugar donde se realizó la tesis:

Universidad Nacional Autónoma de México
Posgrado de Ingeniería
Ciudad de México, México

TUTOR DE TESIS:

Dr. Federico Hernández Álvarez

FIRMA

Índice General

Índice General	i
Índice de tablas y figuras	ii
Resumen	iii
Glosario de términos	iv
1. Introducción	1
1.1. Antecedente y propósito de la investigación	1
1.2. Resumen capitulario	3
2. Ciclos económicos y financieros	5
2.1. ¿Qué son los ciclos económicos y financieros?	5
2.2. Historia moderna de los ciclos económicos y su definición en México	6
2.3. Ciclos clásicos y ciclos de crecimiento	7
2.4. Fundamentos para trabajar con ciclos clásicos	9
3. Cuantificador de vectores de aprendizaje	11
3.1. Redes neuronales artificiales	11
3.2. Tipos de redes neuronales artificiales	11
3.3. Redes neuronales competitivas	12
3.4. Regla de aprendizaje de Kohonen	13
3.5. El por qué de los sesgos (<i>bias</i>) en las redes competitivas	14
3.6. Introducción al algoritmo LVQ	14
3.7. Regla de aprendizaje LVQ	16
4. Implementación del algoritmo LVQ	18
4.1. Determinación de parámetros en el algoritmo	18
4.2. Datos empleados como variables de entrada	20
4.3. Base para clasificación de series: el modelo de Bry y Boschan	22
4.4. Medidas para evaluar los resultados del algoritmo LVQ	27
5. Resultados de la aplicación del algoritmo LVQ	29
5.1. Selección de mejor combinación de variables de entrada	29
5.2. Seguimiento esquemático de pasos realizados	30
6. Ejercicio práctico: adecuación de un portafolio con base en el estado del IPC	36
6.1. Planteamiento del ejercicio	36
6.2. Resultado de los portafolios comparados	38
7. Conclusiones y recomendaciones	41
Referencias	45

Índice de tablas y figuras

Capítulo 2

Figura 2.1. Etapas de un ciclo económico con contracción	5
Figura 2.2. PIB de E.U.A. de 1972 a 2000, expresado como ciclo clásico o de negocios	8
Figura 2.3. PIB de E.U.A. de 1972 a 2000, expresado como ciclo creciente	9

Capítulo 3

Figura 3.1. Arquitectura de una red competitiva	13
Figura 3.2. Arquitectura de un algoritmo LVQ	15

Capítulo 4

Tabla 4.1. Opciones para cada uno de los parámetros de entrada	19
Tabla 4.2. Índices bursátiles incluidos en el análisis	21
Figura 4.3. Gráfica del comportamiento del IPC	25
Tabla 4.4. Cronología de los puntos de giro para el IPC	25
Tabla 4.5. Distribución de periodos recesivos y expansivos	26
Tabla 4.6. Matriz de contingencia 2x2 adaptada a este trabajo	28

Capítulo 5

Tabla 5.1. Mejores variables de entrada. Ordenadas por AC evaluación	30
Tabla 5.2. Mejores parámetros de entrada. Ordenadas por AC evaluación	30
Tabla 5.3. Resumen de entradas para algoritmo LVQ	31
Figura 5.4. Representación gráfica del LVQ en etapa inicial	32
Figura 5.5. Representación del proceso de entrenamiento del LVQ	32
Figura 5.6. Posicionamiento de las neuronas luego del periodo de entrenamiento	33
Figura 5.7. Matriz de confusión 2x2	34
Figura 5.8. Gráfica comparativa entre predicción del algoritmo LVQ y el modelo base	35

Capítulo 6

Figura 6.1. Comparativa Esquema activo-LVQ y Esquema Pasivo-IPC	39
Tabla 6.2. Medidas de retorno y riesgo para esquema pasivo y activo	40

Capítulo 7

Figura 7.1. Diagrama de flujo propuesto para futuras investigaciones concerniente a la selección de parámetros y variables de entrada	43
---	----

Resumen

El propósito de esta tesis es identificar oportunamente las etapas recesivas y expansivas, además de los puntos de giro en el Índice de Precios y Cotizaciones del mercado accionario mexicano (IPC). Como se prueba lo largo del documento, la identificación oportuna de los ciclos financieros puede ser de gran utilidad para inversores y, en general, para cualquier operador al momento de integrar un portafolio. Para tal efecto, se elaboró un indicador a partir de un tipo de red neuronal conocida como cuantificador de vectores de aprendizaje (LVQ), introducida a la economía por Giusto y Piger en 2014. La serie del IPC se presenta en forma de ciclo clásico, respetando la metodología propuesta por Burns y Mitchell, y el consecuente desarrollo estadístico para el fechado de series promulgado por Bry y Boschan en su artículo de 1971.

Para evaluar la asertividad del sistema planteado, se emplean criterios estadísticos comunes para esquemas binarios. Finalmente, se agrega un caso práctico que consiste en la integración de un portafolio en que las decisiones de compra-venta están basadas en el resultado obtenido de la red neuronal LVQ. Para este ejercicio, se aplican medidas de rendimiento y de riesgo de amplio uso en el ámbito financiero, especialmente en el que respecta al mercado bursátil. Las conclusiones resultan ser claras: por una parte, la red neuronal es capaz de acertar en el 79.2% de las ocasiones cuál es el estado del IPC. Por otra parte, a pesar del buen resultado obtenido por la red neuronal, éste no se traduce en un incremento de rendimientos en el portafolio que contempla esta información, aunque sí en una reducción del riesgo.

Glosario de términos (en orden alfabético)

Adaptación: modificación automática, usualmente para maximizar alguna medida de rendimiento. En evolución: ajuste de las especies a su ambiente a través de la selección natural o de cambios en su comportamiento.

Algoritmo: conjunto ordenado de operaciones sistemáticas que permite hacer un cálculo y hallar la solución de un tipo de problemas.

Aprendizaje: Nombre genérico para cualquier cambio de comportamiento que dependa de experiencias y que mejore el rendimiento de un sistema. En este sentido, el aprendizaje es igual a la adaptación: modificaciones selectivas en los parámetros de un sistema.

Aprendizaje sin supervisión: aprendizaje sin conocimiento previo acerca de la clasificación de las muestras.

Aprendizaje supervisado: esquema de aprendizaje en donde las diferencias entre los resultados obtenidos y los reales se disminuye; todo ello durante el periodo de entrenamiento.

Arquitectura: organización general de una red neuronal sobre su transferencia, almacenamiento y manejo de información.

Atributo: componente de un vector de datos que tiene cierta propiedad descriptiva.

Calibración: aumento de la correspondencia entre los resultados obtenidos con un instrumento y los valores reales que se miden con él.

Capa de entrada: capa de neuronas que está conectada directamente a los datos entrantes.

Capa de salida: capa de neuronas encargada de generar los señales de salida.

Capa oculta: una capa intermedia de neuronas en una red neuronal multicapa que no tiene conexión directa con las entradas o las salidas.

Característica distintiva: característica para la que se puede hacer una separación en clases sin riesgo a equivocarse.

Categoría: conjunto de elementos que comparten propiedades similares para una característica particular.

Clasificador no paramétrico: método de clasificación que no se basa en alguna función matemática para determinar las regiones o clases, sino que se remonta directamente a los datos disponibles.

Clasificación respecto al vecino más cercano (*K nearest-neighbor classification*): clasificación de nuevos eventos en lo que el sistema recurre a experiencias previas y determina qué clase es la más cercana. También puede ser visto como un método en el que el sistema identifica a qué punto previo, el nuevo dato es más semejante, para incluirlo en la misma categoría. De ahí que también sea llamado algoritmo perezoso (*lazy algorithm*).

Clasificador paramétrico: método de clasificación en que las regiones o clases están delimitadas por funciones matemáticas.

Contexto: información que describe la situación o instancia en la que un evento ocurre.

Cuantificador de vectores de aprendizaje (*learning vector quantization*): método de aprendizaje supervisado en el que una red neuronal es capaz de agrupar vectores entrantes dentro de clases previamente definidas por el desarrollador.

Datos de entrenamiento: conjunto de datos usados como entradas para el proceso de aprendizaje de una red neuronal artificial.

Determinístico: experimento o fenómeno que da lugar a un resultado cierto o seguro. Partiendo de las mismas condiciones iniciales se tiene la certeza de lo que va a suceder.

Diagrama de árbol: Estructura de ramificaciones en la que los nodos inferiores están subordinados a los nodos superiores.

Dimensionalidad: número de características o atributos con los que cuenta una matriz de datos.

Distancia: medida específica de separación entre un par de vectores. En este trabajo de investigación se hace referencia a la distancia euclidiana que toma propiedades geométricas específicas.

Emulación: realización de una tarea imitando otra que ya fue hecha.

Época: conjunto finito de datos de entrada presentados secuencialmente. Cada vez que se presenta el total de vectores de entrada, durante el periodo de entrenamiento, representa una época.

Espacio euclidiano: conjunto matemático de elementos vectoriales y su relación de acuerdo a distancias euclidianas.

Estado de transición: transformación de un estado a otro en un sistema dinámico durante un intervalo de tiempo.

Estímulo: algo que temporalmente influencia la actividad de un organismo.

Estocástico: aleatorio, habiendo cierto probabilidad de ocurrencia.

Estratificación: confinamiento en capas.

Excitación: acción sináptica que incrementa la activación de una neurona.

Frontera de decisión: límite definido matemáticamente en el espacio, para hacer una separación de clases en reconocimiento de patrones o decisiones estadísticas.

Función del sombrero mexicano: función usualmente simétrica definida en el espacio (o en un plano) que tiene valores positivos en la región central de su dominio y valores negativos alrededor de él.

Función de la neurona ganadora que se lleva todo (*winner take-all function*): principio computacional en el que las neuronas de una capa compiten entre si para ser activadas. En su forma clásica, sólo la neurona que recibe el mayor estímulo se activa, mientras el resto permanecen apagadas.

Grupo (*cluster*): conjunto de puntos muestrales cercanos entre si dentro de las entradas de una red neuronal.

Heurística: relacionado con la toma de decisiones en los seres humanos; basado en la intuición.

Indicador de rendimiento: medida o función objetivo que describe, global y analíticamente, el rendimiento de un sistema. Algunos ejemplos son la exactitud de la clasificación o el error cuadrático medio.

Inferencia deductiva: conclusiones que se obtienen a partir de premisas o principios generales; yendo de lo general a lo particular.

Inteligencia: habilidad para realizar nuevas tareas que resulta directa o indirectamente vital para la supervivencia; habilidad para resolver problemas de los que no tenemos precedente alguno.

Inteligencia artificial: habilidad de un sistema artificial para realizar tareas que, bajo la definición popular, requieren de inteligencia.

Iteración: repetición de un procedimiento como, por ejemplo, el entrenamiento. Cada vez que se presenta un vector de entrada, durante el periodo de entrenamiento, representa una iteración.

Red neuronal competitiva: tipo de red neuronal perteneciente al conjunto de clasificadores, diseñada especialmente para extraer características comunes de una base de datos. También se puede presentar como capa, sólo formando parte de una red neuronal con características más amplias.

Memoria asociativa: nombre genérico para cualquier función en que la memoria retoma información almacenada y la asocia con eventos que se suscitan en el presente.

Medianas de k (*k-medoids*): al igual que las medias de k , es un método de agrupamiento que tiene como objetivo la partición de un conjunto de n observaciones en k grupos. Pretende minimizar la distancia entre las observaciones y los grupos a los que fueron asignadas. El centroide queda ubicado en el punto central para cada grupo de observaciones.

Medias de k (*k-means*): es un método de agrupamiento que tiene como objetivo la partición de un conjunto de n observaciones en k grupos. Cada observación pertenece al grupo en el que su media es más cercana. Primero se generan los grupos; sus centroides se ubican cerca del centro del conjunto de datos o de forma aleatoria a lo largo de las observaciones (1). Luego las observaciones son asociadas con la media (o centroide) más cercana (2). De acuerdo a las observaciones que se tengan, el centroide se mueve hacia el promedio de ellas (3), y estos dos últimos pasos se repiten hasta que se logre la convergencia.

Neurona: cualquiera de los numerosos tipos de células especializadas en el cerebro o sistema nervioso encargadas de transmitir y procesar señales neurales. En el ámbito artificial, una neurona es una función matemática que tiene como modelo una neurona biológica. Son la base de las redes neuronales.

Neurona ganadora: en una capa o red neuronal competitiva es aquella neurona que reporta el máximo grado de activación o la menor diferencia en su distancia euclidiana respecto a un vector de entrada.

Neurona muerta: término empleado en el argot de las redes neuronales artificiales para llamar a aquellas neuronas que se ubican en lugares distantes a los que registran los vectores entrantes y, por ende, no se activan en ninguna ocasión durante el periodo de entrenamiento. No tienen función útil por lo que se busca la implementación de sistemas que reduzcan, si no eliminen, su existencia.

Nodo: ubicación del elemento de procesamiento, o neurona, en una red neuronal artificial.

Norma euclidiana: largo geométrico de un vector en el espacio euclidiano.

Ortonormalidad: propiedad conjunta de ortogonalidad y normalización.

Ortogonalidad: propiedad de un par de vectores cuando su producto interno es cero. Dicho de otra forma, es una generalización de la noción geométrica de perpendicularidad.

Pre-procesamiento: conjunto de operaciones realizadas a los datos entrantes antes de que sean presentados a la red neuronal.

Problema de agente viajero (*traveling-salesman problem*): problema de optimización en el que una red neuronal debe encontrar el recorrido más corto entre una serie de puntos.

Procesamiento en grupo (*Batch processing*): conjunto de tareas realizadas en un solo proceso computacional. Para el caso del entrenamiento neuronal se presentan todos los vectores de entrada (de una época) y no se realiza ninguna modificación en la posición de las neuronas. Una vez que termina esa época específica, el algoritmo determina cuál fue la neurona ganadora para cada uno de los vectores entrantes. La neurona, entonces, actualiza su ubicación hacia la posición promedio de los vectores entrantes para los que fue ganadora. El propósito fundamental de este tipo de algoritmo es el de reducir el tiempo de entrenamiento, y por tanto mejorar el rendimiento del sistema computacional. Mientras que en circunstancias normales, la red neuronal actualizaría la ubicación neuronal en cada iteración, en el batch algorithm lo hace sólo una vez por época.

Proceso de Markov: proceso estocástico en el que la nueva situación del sistema depende exclusivamente de sus estados previos.

Programación dinámica: método de optimización algorítmica que encuentra el patrón óptimo como una expresión recursiva.

Propagación hacia atrás (*Back-propagation*): método de optimización usado en redes neuronales multicapa, en el que hay un proceso de corrección en la posición de las neuronas siguiendo una secuencia inversa (de la capa de salida hacia la capa de entrada).

Prototipo: muestra típica de una clase usada para entrenar un sistema.

Rango de aprendizaje: factor o velocidad a la cual se modifica un sistema con el fin de adaptarse.

Razonamiento probabilístico: conjunto de decisiones basadas en el cálculo probabilístico.

Reconocimiento de patrones: ciencia que tiene como propósito extraer información de objetos físicos o abstractos para identificar y establecer propiedades comunes. Integra procesos de ingeniería, computación y matemáticas.

Red neuronal artificial: las redes neuronales son conjuntos de elementos de cálculo simples, generalmente adaptativos, interconectados masivamente en paralelo y con una organización jerárquica que le permiten interactuar con algún sistema del mismo modo que lo hace el sistema nervioso biológico.

Regla de aprendizaje de Hebb's: La más citada regla de aprendizaje en la teoría de redes neuronales. Plantea que si la neurona i participa repetidamente en la activación de la neurona j , su conexión sináptica se ve fortalecida y, conforme avanza el tiempo, la neurona j se vuelve más sensible al estímulo de la neurona i .

Retroalimentación: conjunto de señales conectadas entre las capas de salida y las de entrada.

Segmentación: conjunto de operaciones realizadas a un grupo de datos para dividirlos en grupos más pequeños que tengan características semejantes.

Sesgo (*bias*): valor agregado con la finalidad de interferir, ya sea favoreciendo o perjudicando, el resultado de un proceso.

Vecindario: conjunto de neuronas ubicadas dentro de un radio delimitado alrededor de la neurona ganadora.

Capítulo 1: Introducción

1.1. Antecedentes y propósito de la investigación

El mercado accionario ha sido objeto de estudio por parte de investigadores, académicos y profesionales de distintas áreas. En torno a él, se han desarrollado numerosas teorías y modelos con el propósito de explicar y, quizás, de predecir su comportamiento.

Uno de los tópicos fundamentales en la comprensión del mercado accionario es la existencia de ciclos, los cuales se entienden como periodos en los que se registra una concentración de resultados positivos, conocidos como expansiones o, por el contrario, una acumulación de rachas negativas o recesiones. Este principio es aplicado tanto para la economía como para las finanzas.

Gran parte de los estudios respecto a la existencia de ciclos se ha centrado, específicamente, en ciclos económicos debido a que su identificación, en un tiempo oportuno, resulta ser de gran utilidad para la toma de decisiones en política monetaria, como lo señalan Anas y Ferrara [1]. Los países enfrentan la dificultad de mantener a sus economías en sendas de crecimiento estable y sostenible a largo plazo. De ahí que, tanto bancos centrales como secretarías de estado, se hayan dado a la tarea de monitorear minuciosamente muchas actividades económicas con el fin de identificar la existencia de tendencias comunes que alerten de posibles fluctuaciones en el estado de la economía.

Sin embargo, parece ser que los ciclos financieros no han llamado en gran medida la atención de los investigadores. Una razón de peso para este suceso es que existe un consentimiento general en que los índices financieros anteceden a los económicos. Esto significa que, si ya de por sí resulta bastante complejo identificar la situación actual y los puntos de giro para la economía, como lo han demostrado Berge [2] y Katayama [3] en sus trabajos, cuanto más difícil es realizar esta labor para series financieras.

Esto es justamente el objetivo general de esta investigación: identificar oportunamente las etapas recesivas y expansivas, además de los puntos de giro en el Índice de Precios y Cotizaciones del mercado accionario mexicano (en adelante nos referiremos a él como IPC, por sus siglas).

La hipótesis central plantea que, mediante un tipo de red neuronal conocida como cuantificador de vectores de aprendizaje (LVQ), apenas introducido en economía por Giusto y Piger en 2014 [4], es posible la identificación oportuna de etapas de expansión y recesión, así como los puntos de giro en el IPC.

En cuanto a los objetivos particulares se tienen los siguientes; en primer lugar, se pretende revisar si el algoritmo LVQ es capaz de identificar el mercado bajista que se registró de 2007 a 2009 y, de ser así, con qué prontitud lo hace. En segundo lugar, se plantea corroborar los trabajos que estudian las correlaciones existentes entre mercados accionarios de varios países, los cuales se detallan en el apartado 4.2., y el IPC. Todo esto ya que, para esta investigación, se utilizaron índices financieros de los principales mercados accionarios y, con ellos, se identificó el estado del IPC.

Como tercer y último punto, con lo que se cierran los objetivos particulares, se prueba si el conocimiento del estado en que se encuentra el IPC genera beneficios, entendidos como un mayor rendimiento y/o un menor riesgo o, si por el contrario, este conocimiento no genera diferencia alguna. Para ello, se plantea un caso práctico que consiste en la integración de un portafolio en que las decisiones de compra-venta están basadas únicamente en este criterio, y se compara con un portafolio que hace caso omiso de esta información.

Las diferencias de este trabajo, respecto a otros que abordan la identificación de ciclos, son las siguientes: en primer lugar, en esta tesis se trabajó con series financieras, mientras que el resto de los trabajos se han enfocado únicamente en series económicas. En segundo lugar, las series están expresadas en forma de ciclos clásicos, mientras el resto de los trabajos se vislumbran desde la perspectiva de ciclos de crecimiento. Como tercer lugar se tiene la metodología empleada para la detección de ciclos; dadas las características del objeto de estudio, una gran variedad de modelos paramétricos no lineales se han propuesto. Por mencionar algunos, tenemos los modelos regresivos de transición suavizada (*smooth-transition autoregressive models*), propuestos por Ferrara y Guegan [5]; los modelos no lineales inversos de distribución (*non-linear probit models*), de Liu y Moench [6]; y los modelos de cadenas de Markov (*Markov switching models*), planteados por Guidolin en 2011 [7]. Para este trabajo se ha empleado, como ya se ha dicho con anterioridad, el LVQ: un tipo de red neuronal capaz de clasificar y reconocer patrones. Este conjunto de cambios, además de señalar el hecho de que no se ha realizado trabajo similar para el caso de México, repercute en una serie de variantes, beneficios en su empleabilidad y ventajas que se detallan a lo largo del documento.

Para quienes invierten en el mercado bursátil, el conocer el estado del IPC es trascendente para balancear sus inversiones entre títulos de deuda, que forman parte de los instrumentos de bajo riesgo, y títulos de propiedad o acciones que son más riesgosos. Si a ello se le agrega que es necesaria una forma más rápida de identificar periodos en un ciclo financiero debido a que, actualmente, el método existente -detallado en el apartado 3.3.- contempla al menos una tardanza de 6 meses, se tiene entonces una buena justificación para este trabajo de tesis.

Tomando como base lo dicho en el párrafo anterior, parte de esta investigación está enfocada a generar un indicador, usando el algoritmo LVQ, para identificar en el tiempo en que suceden (*nowcasting*) los puntos de giro en un esquema de ciclos clásicos para el IPC.

Su propósito es obtener señales que puedan ser usadas de forma fácil e inmediata por inversionistas y operadores bursátiles.

Los resultados de este trabajo muestran que, en efecto, el esquema propuesto es capaz de identificar el estado actual del IPC y los puntos de giro, entre recesiones y expansiones, en el momento en que se presentan, ya que se cuenta con un indicador global de aciertos de 79.2% respecto a los valores reales conocidos. Sin embargo, no se puede aseverar que el conocimiento de esta información sea de utilidad al momento de tomar decisiones de compra – venta en el mercado bursátil mexicano ya que se generan señales inversas: por una parte, al final del periodo de evaluación que involucró 120 meses, se tuvo un decremento en el rendimiento de 3.28% respecto al esquema que se usó para hacer la comparativa; por otra parte, se logró una reducción en el riesgo de inversión de 1.41% de forma anualizada. La conclusión entonces: no es posible obtener mejores rendimientos y , a la vez, un menor riesgo para inversiones bursátiles en el mercado accionario mexicano sólo con saber la situación actual del índice IPC.

1.2. Resumen Capitular

Esta tesis se desarrolla en siete capítulos de los cuales el capítulo 1 está enfocado en dar la introducción a esta investigación. En el capítulo 2, se aborda qué es un ciclo, ya sea económico o financiero; también se ahonda un poco en la historia contemporánea de este tema; asimismo se enuncian los tipos de ciclos que existen, cuál se ha decidido emplear en este trabajo y por qué, así como cuál es la preferencia de las instituciones en nuestro país respecto a su uso.

El capítulo 3 inicia abordando la definición de red neuronal artificial. Igualmente se muestra cómo pueden ser clasificadas de acuerdo a su utilidad; también se profundiza en un tipo de red neuronal, que es la competitiva, debido a que comparte muchos elementos con el sistema que se estará empleando en este trabajo. Finalmente, se llega al algoritmo que nos atañe: el LVQ. Tanto el capítulo 2 como el capítulo 3 servirán de marco teórico para el resto del trabajo de investigación.

En el capítulo 4 se detalla la metodología empleada. Puntos como qué parámetros se eligieron para entrenar al algoritmo, cuáles fueron los datos empleados y qué criterios se toman en cuenta para realizar la clasificación son algunos de los puntos que se encuentran en este apartado. De igual modo, en este capítulo se expone el modelo base con el cual se hacen las comparaciones y se define una serie de métricos que servirán para establecer la calidad de los resultados generados por nuestro sistema.

Para el capítulo 5 se reservan las soluciones obtenidas y, a fin de evaluar la confiabilidad del sistema que se propone, se revisan los resultados de medidas estadísticas empleadas especialmente para esquemas binarios. También se detalla cuáles han sido los pasos seguidos en la parte de programación que incluye este trabajo.

En el capítulo 6, para mostrar de una manera más tangible el beneficio del esquema que se sugiere, se ha desarrollado un caso práctico donde se monetizan los beneficios de conocer en tiempo real el estado del IPC. La pretensión es mostrar a los investigadores y estudiosos del tema cuál es el efecto de conocer la situación del IPC y si esta información puede repercutir en ganancias o pérdidas adicionales.

Finalmente, en el capítulo 7 se exponen las conclusiones obtenidas de este trabajo, así como las áreas de oportunidad que se detectan y que abren las puertas para realizar futuras investigaciones.

Adicionalmente, se ha incorporado un glosario de términos al inicio de esta tesis. Su propósito es explicar brevemente muchos de los conceptos que se tocan a lo largo del trabajo de investigación sin tener que detenerse para ello. Cabe señalar que este glosario no es exhaustivo y sirve sólo para la lectura de este trabajo; se ha pretendido que las definiciones sean, ante todo, explicativas. En caso de que una palabra tenga varios significados, sólo se han incluido aquellos que tengan relación con las redes neuronales, ya sea artificiales o biológicas. En algunos casos se indica, entre paréntesis, el término usado en inglés. El compendio de términos y definiciones se ha obtenido de libros especializados [8] y [9], y se ha complementado por el autor de este trabajo.

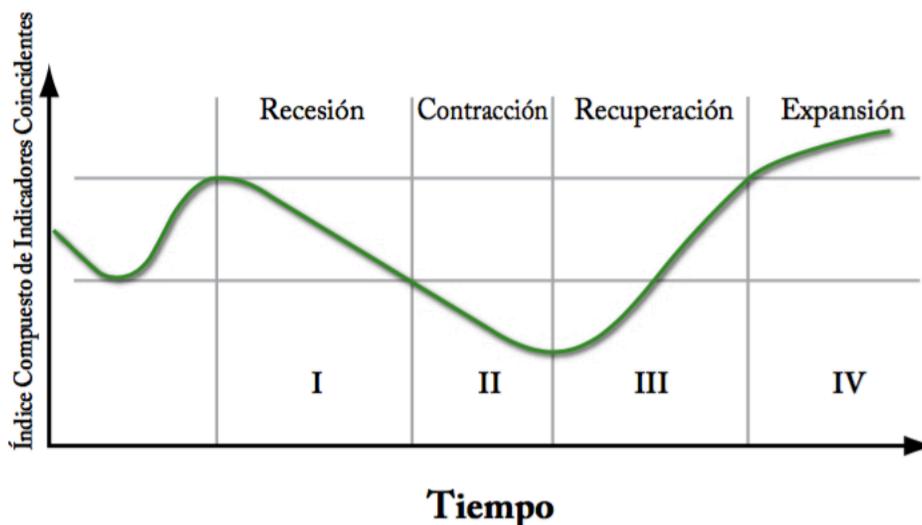
Capítulo 2. Ciclos económicos y financieros

2.1. ¿Qué son los ciclos económicos y financieros?

En la economía se experimentan periodos de oscilación, mismos que se ven representados en periodos de expansión, recesión o estabilidad económica. Citando a Burns y Mitchell [10] tenemos que “...los ciclos económicos son un tipo de fluctuación que se encuentra en la actividad económica agregada de las economías (países) que organizan su trabajo principalmente mediante empresas de negocios. Un ciclo económico consiste en expansiones que ocurren aproximadamente al mismo tiempo en muchas actividades económicas, seguidas de recesiones, contracciones y recuperaciones igualmente generales que se consolidan en la fase de expansión del ciclo siguiente. La secuencia de las fases es recurrente pero no periódica. En cuanto a su duración, los ciclos varían de más de un año a periodos de diez o doce años y no son divisibles en ciclos más cortos de características similares.” Así, se puede decir que se denomina ciclo económico a los periodos de expansión y contracción que experimenta el nivel general de actividad económica de un país en horizontes de mediano plazo.

Una manera más clara de entender este fenómeno es a través de una representación gráfica como la que se muestra a continuación. En ella se observa que existen cuatro periodos clave en los ciclos: recesión, contracción, recuperación y expansión.

Figura 2.1. Etapas de un ciclo económico con contracción.



Fuente: Adaptado de Jonathan Heath [11].

De forma general, la **recesión** se entiende como una caída sin que esta llegue a estar por debajo del punto mínimo anterior. Cuando el indicador ha descendido lo suficiente para estar por debajo del mínimo alcanzado con anterioridad, se habla entonces de una

contracción. Por otro lado, cuando el índice se encuentra en una etapa alcista, pero está no ha superado el máximo anterior se denomina **recuperación**, para que, finalmente, se hable de una **expansión** cuando, en una racha alcista, se haya podido superar el máximo registrado con anterioridad.

Las fases anteriormente mencionadas también se presentan en series financieras. Burns y Mitchell [10] consideraron el análisis de estas fases de la siguiente forma: tanto los periodos de recesión como los de contracción quedan englobados en **recesiones**, mientras que tanto los periodos de recuperación como los de expansión se agrupan simplemente en **expansiones**. Dicho de otra forma, cualquier serie, a mediano plazo, que tuviera una racha bajista queda comprendida como recesión, mientras que cualquiera con una racha alcista es denominada expansión. Así, la expansión es la fase que se presenta entre un valle y un pico y, por el contrario, la transición entre un pico y un valle es llamada recesión.

Es importante señalar lo anterior ya que este estudio tomará esta última nomenclatura y sólo se hará distinción entre ciclos expansivos y recesivos, por lo cual se estará trabajando con un modelo binario. Con el fin de manejar una terminología financiera, en algunas ocasiones, nos referiremos a la etapa de expansión como *bull market* y a la etapa de recesión como *bear market*.

Respecto a la duración de cualquiera de los eventos antes mencionados, no existe un consenso general respecto al tiempo mínimo para que una etapa sea declarada como de expansión o de recesión. Existe una regla de aproximación, más no una definición exacta, de que el indicador debe estar al menos dos trimestres consecutivos en un estado, es decir que su duración se debe prolongar por al menos seis meses antes de que pueda existir un nuevo giro. En caso de que haya movimientos de menor duración, estos no se consideran cambios de estado, sino simplemente movimientos respecto a alguna tendencia o variaciones estacionales. Sin embargo, como lo señalan Achutan y Banerji [12], esta regla es simplista y puede llevar a conclusiones erróneas. Para este estudio, a pesar de que no se establece esta regla para la clasificación de ciclos, el periodo de recesión más corto que se registró fue, precisamente, de seis meses. En el Capítulo III se verán los detalles.

2.2. Historia moderna de los ciclos económicos y la definición en México

Se puede decir que la historia moderna para la medición de ciclos económicos comienza con la fundación, en 1920, de la Oficina Nacional de Investigación Económica de los E.U.A. (*National Bureau of Economic Research* o NBER por sus siglas en inglés), la cual publicó por primera vez, en 1929, las fechas del ciclo económico. La NBER fue en realidad la primera institución en encargarse, de manera oficial, de investigar y analizar ciclos económicos. Se ocupa, entre otras actividades, de anunciar la ocurrencia de las distintas fases, de mostrar las características de la economía a lo largo de las mismas y de dar las justificaciones de los anuncios. Cabe señalar que, a la fecha, mantiene un record de todos los ciclos económicos de E.U.A. a partir de Diciembre de 1854.

Para 1978, la NBER estableció el *Comité de Fechas de los Ciclos Económicos* y, desde entonces, ha existido un proceso formal para anunciar los picos y valles en esta actividad. De manera universal, avalado por académicos, economistas, funcionarios públicos y empresas, se reconoce a la NBER como el árbitro oficial de los ciclos económicos en los E.U.A.

En el caso de México, como lo señala Jonathan Heath [13], no debería existir una definición distinta para los términos recesión o expansión, y mucho menos, más de una. Por ende, lo lógico sería aplicar la terminología clásica y oficial de la NBER a los ciclos económicos de México. Sin embargo, surgen una serie de limitantes que se deben de tomar en cuenta. En primer lugar, los datos oficiales de 1980 a 1992 no son comparables a los que parten de 1993, ya que se cambió la base y se realizaron adecuaciones metodológicas. Por otra parte, fue hasta finales de los años 90 cuando el INEGI comenzó a publicar datos ajustados por la estacionalidad, lo que permitió realizar comparaciones entre trimestres subsecuentes o posteriores. De la misma forma, hasta mediados de los años 80, el INEGI se dio a la tarea de elaborar la mayoría de los indicadores económicos relevantes, conocidos como indicadores de coyuntura. De ahí que, sólo a partir de inicios de este siglo podamos definir, con relativa facilidad, las fechas de los ciclos económicos para México.

2.3. Ciclos clásicos y ciclos de crecimiento

Otro tema de relevancia para este capítulo es la existencia de diferentes perspectivas para visualizar los datos y para plasmar los resultados. Por una parte tenemos el ciclo clásico; Víctor M. Guerrero [14] lo describe como el espacio de tiempo donde la actividad de negocios disminuye y después se recupera en relación al nivel alcanzado. La otra forma de presentar la información es usando el ciclo de crecimiento, que resulta de tomar una serie de tiempo para luego retirarle la tendencia a largo plazo. Este tipo de ciclo busca así, representar las fluctuaciones alrededor de ella.

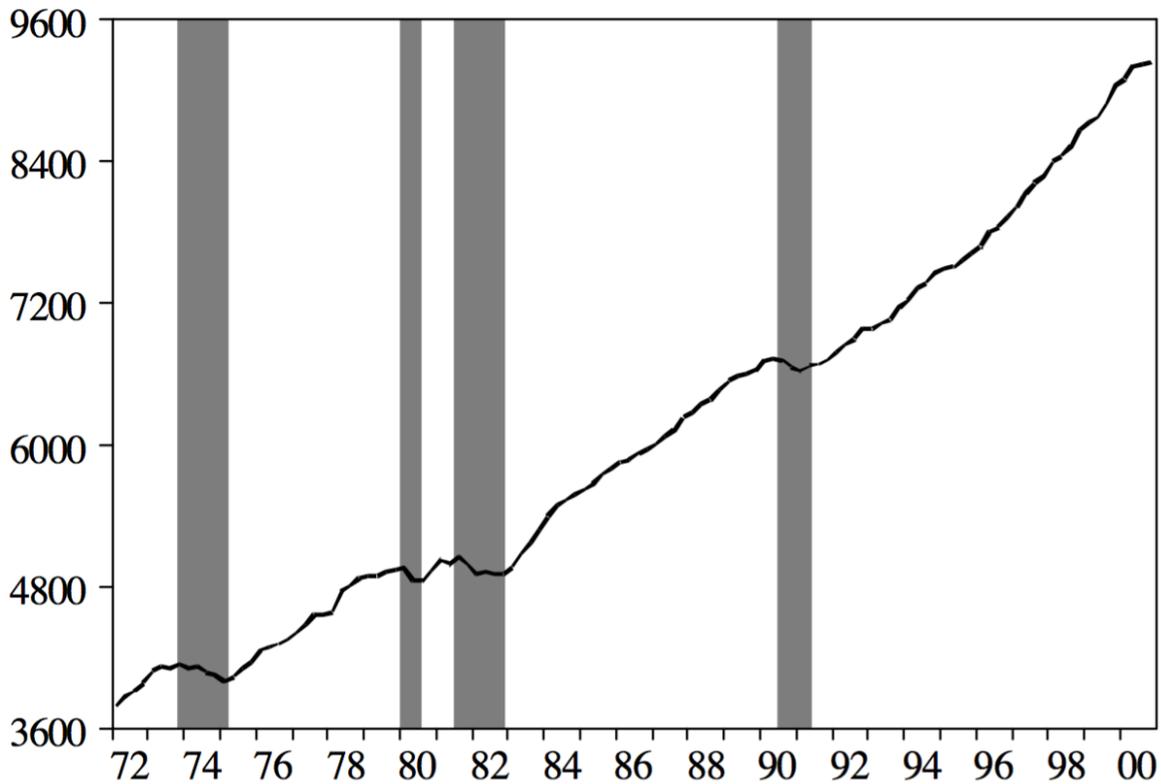
Para poder explicar esto de una forma más clara, se toman las figuras 2.2. y 2.3 como apoyo. La primera de ellas muestra una serie de tiempo expresada en forma de ciclo clásico o de negocios. La característica más significativa para este tipo de ciclo es que los datos se presentan eliminando solamente el componente estacional. Cabe señalar que todos los bancos de información, como el Banco Mundial, y las instituciones encargadas de generar la misma en cada país, como el INEGI, ya han retirado este factor de estacionalidad, por lo que no es necesario realizar algún paso adicional. Para este caso, se cuenta con el ejemplo de la figura 2.2. En su eje vertical se puede visualizar el Producto Interno Bruto a valores de Paridad de Poder Adquisitivo (PIB PPA) expresado en millones de dólares, mientras que en el eje horizontal aparecen los últimos dos dígitos del año estudiado (desde 1972 y hasta el año 2000).

Por otro lado, en la figura 2.3. se muestra esta misma serie histórica sólo que presentada en forma de ciclo de crecimiento. Para ello, ha tenido que ser removida la tendencia a largo plazo y su factor estacional. Por ende, el eje vertical indica el porcentaje de desviación del

PIB PPA, respecto a su tendencia a largo plazo, que queda representada por el cero a lo largo de la serie.

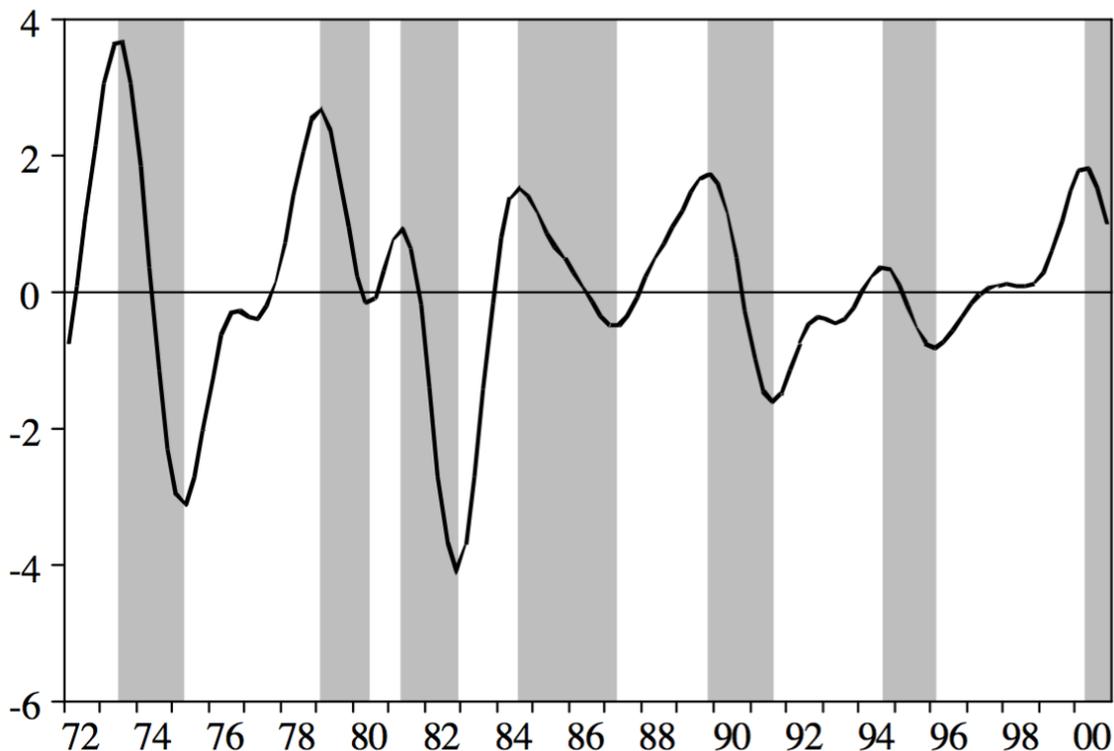
Es importante señalar que, para el ciclo clásico, los puntos de giro separan los periodos de **recesión** de los de **expansión** mientras que, para el ciclo de crecimiento, lo que se distingue es una separación entre periodos de **desaceleración** y **aceleración**. Para explicar la diferencia entre estos conceptos se han colocado intencionalmente las figuras 2.2. y 2.3. alineadas verticalmente. Como el lector puede observar, existen tres periodos de desaceleración que se indican en la figura 2.3. (el primero del 84-87, el segundo del 94-96 y el tercero a partir del 2000) que no se traducen en etapas recesivas para la figura 2.2. El significado de esto es bastante sencillo: todas las recesiones involucran desaceleraciones, pero no todas las desaceleraciones son, necesariamente, recesiones. De ahí que los ciclos de crecimiento sean siempre más numerosos que los ciclos de negocios.

Figura 2.2. PIB de E.U.A. de 1972 a 2000, expresado como ciclo clásico o de negocios.



Fuente: Adaptado de Anas y Ferrara [1].

Figura 2.3. PIB de E.U.A. de 1972 a 2000, expresado como ciclo creciente.



Fuente: Adaptado de Anas y Ferrara [1].

Tanto el ciclo clásico como el ciclo de crecimiento se usan en diversos países para referirse al estado de la economía. Tanto el *Conference Board* (por sus siglas en inglés, CB) como el NBER usan la definición de ciclo clásico o de negocios, mientras que la Organización para la Cooperación y el Desarrollo Económico (OCDE) utiliza el ciclo de crecimiento. Como ya hemos dicho con anterioridad, en este trabajo se estará haciendo referencia al ciclo clásico o de negocios. El siguiente apartado detalla las razones de ello.

2.4. Fundamentos para trabajar con ciclos clásicos

Existen varias razones que soportan la decisión de haber elegido trabajar con ciclos clásicos. La primera de ellas es que tiene la aceptación de la gran mayoría de los economistas de E.U.A., en especial de los académicos. La segunda razón es que, por el otro lado, los ciclos de crecimiento han sido criticados por varios investigadores, entre los que destacan Achuthan y Banerji [12], ya que sostienen que esta metodología ha creado confusión empleando, por ejemplo, la misma terminología que se usa en los ciclos clásicos, y generando así malentendidos. Un caso de este tipo es el uso de la palabra *recesión* cuando hay un crecimiento por debajo de la tendencia a largo plazo, contradictorio al uso tradicional.

Asimismo, un punto que hace descartar el uso del ciclo de crecimiento es que no se puede conocer la tendencia de largo plazo, misma que se retira para poder observar únicamente

el componente cíclico, ya que la tendencia retirada es un concepto dinámico y cambiante a lo largo del tiempo. Un ejemplo de ello son los reportes del ciclo de crecimiento. En ellos, se dice que la desviación cíclica respecto a la tendencia a largo plazo es de un determinado porcentaje y, sin embargo, no podemos conocer a ciencia cierta cuál es la tendencia sobre la que están fluctuando los valores publicados.

Capítulo 3. Cuantificador de vectores de aprendizaje

3.1. Redes Neuronales artificiales

Existen varias definiciones acerca del concepto de red neuronal. En esta ocasión se enuncia una de las más aceptadas, propuesta por Teuvo Kohonen, quién ha hecho importantes contribuciones en este campo: “Las redes neuronales son conjuntos de elementos de cálculo simples, generalmente adaptativos, interconectados masivamente en paralelo y con una organización jerárquica que les permite interactuar con algún sistema del mismo modo que lo hace el sistema nervioso biológico...” [8].

En otras palabras, una red neuronal artificial, por sus siglas en inglés ANN, es un conjunto de algoritmos matemáticos capaces de encontrar relaciones, incluso no lineales, para una gran cantidad de datos. También puede ser vista como un sistema de cómputo, con componentes interconectados, que procesan la información y producen respuestas dinámicas frente a variables externas.

Se denominan redes neuronales porque intentan replicar el funcionamiento de una neurona biológica cuando esta procesa información. Desde finales de la Segunda Guerra Mundial se han venido desarrollando modelos computacionales que persiguen este propósito, por ejemplo el expuesto por McCulloch y Pitts en 1943 [15]. Todos estos modelos tienen un elemento en común: el uso de una estructura en forma de red, en la cual sus nodos, o neuronas, son en realidad procesos matemáticos.

A través de los años, el uso de las redes neuronales artificiales, como herramienta de simulación, modelación y descripción de sistemas dinámicos, se ha ido consolidando como una técnica eficaz y relativamente rápida, gracias al gran desarrollo experimentado en los sistemas de cómputo. Debido a sus características, se ha logrado que esta técnica de uso común en ingeniería aplicada se haya extendido a otras áreas como la biología, la industrial, la militar y recientemente a la financiera. En el apartado 3.6. se enuncian investigaciones científicas que abordan exclusivamente el tipo de red neuronal artificial usada y el área de estudio que atañe a esta investigación, es decir: el algoritmo LVQ en las finanzas bursátiles.

3.2. Tipos de redes neuronales artificiales

En términos generales existen cuatro tipos básicos de redes neuronales artificiales, cada una con diferente propósito y forma de trabajar. A continuación describiremos brevemente cada una de ellas.

(1) De aproximación de funciones y regresiones no lineales: su propósito es crear una red neuronal capaz de generar relaciones no lineales, entre las entradas y las salidas. Están calificadas para ajustar los datos a cierta distribución o comportamiento. Un ejemplo de su

uso sería el pronosticar el precio de una casa, de acuerdo a un conjunto de características geográficas y de bienes raíces. Para ello, entrenaríamos a la red con una base de datos que contenga, por un lado, las características de la casa, y por el otro, el precio de la vivienda. Luego, tan solo proporcionándole las particularidades de la nueva propiedad, la red neuronal nos puede dar un estimado de su precio.

(2) De agrupamiento: la pretensión aquí es congregar los datos de acuerdo a sus similitudes. La información introducida puede tener cuantas características o dimensiones se desee y el sistema se encargará de separarlos en tantas clases como número de neuronas tenga la red. Un ejemplo de su uso sería la segmentación de mercado de acuerdo a los patrones de compra de un grupo de personas. En ellas, es necesario que la red pase por un proceso previo de entrenamiento sin supervisión. Los principales exponentes de este tipo de red neuronal son: la red neuronal competitiva (a la que se hace referencia en el apartado siguiente) y el llamado Mapa Auto-Organizado (*Self-Organizing Feature Map*, que por sus siglas llamaremos SOFM).

(3) De series de tiempo dinámicas: el propósito de este tipo de red es hacer pronósticos basados en una o más series de tiempo. Funcionan como filtros dinámicos. Existen muchas aplicaciones como pronosticar el precio futuro de una acción con base a su precio histórico y a, por ejemplo, el número de puntos de venta que la compañía tiene. Otro ejemplo pudiera ser el pronosticar la tasas de mortalidad en una población de acuerdo a su serie histórica y a los niveles de contaminantes (monóxido de carbono, dióxido de azufre, ozono, etc.) que se tienen en el ambiente.

(4) De reconocimiento de patrones y clasificación: aquí se persigue que la red neuronal realice una clasificación de las entradas de acuerdo a un grupo de categorías que el propio programador impone. Antes de que la red pueda hacer una clasificación correcta de las entradas, es necesario que pase por un proceso de entrenamiento supervisado. Un ejemplo de su uso es la clasificación de tumores, en malignos o benignos, basado en la uniformidad de las células, su dureza y espesor. El principal exponente de este tipo de red neuronal es el llamado algoritmo LVQ, mismo que estaremos usando a lo largo de este proyecto.

Antes de adentrarnos en el LVQ -del tipo “de reconocimiento de patrones y clasificación”- se pasan a revisar, primeramente, las redes neuronales competitivas –del tipo “de agrupamiento”-. La razón para ello es que el LVQ posee dos capas y la primera de ellas es de carácter competitivo.

3.3. Redes neuronales competitivas

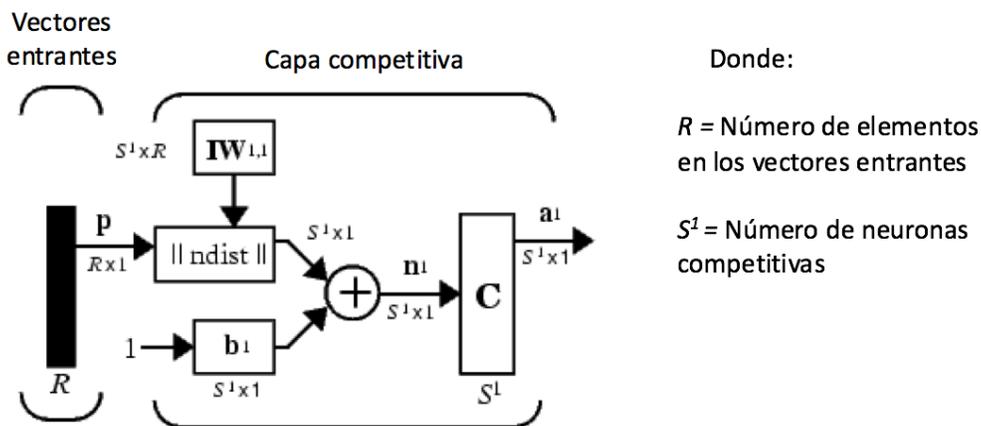
Los redes neuronales competitivas pueden aprender a detectar regularidades, correlaciones y grupos similares de entradas para, en el proceso, adaptar su respuesta a ellos. Esto es logrado mediante el movimiento de las neuronas, durante el entrenamiento, hacia los lugares donde se registre una mayor concentración de vectores entrantes.

La figura 3.1. nos ofrece una visión más clara de la estructura de esta red neuronal. La caja denominada $\|ndist\|$ recibe un vector p perteneciente a los datos de entrada, así como las posiciones de las neuronas ($IW_{1,1}$) y produce un vector S^1 con las distancias euclidianas negativas entre el vector y cada una de las neuronas. La distancia euclidiana entre dos puntos p y q es igual al largo del segmento de línea que los conecta (\overline{pq}), de tal forma que:

$$d(p, q) = d(q, p) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \dots + (q_n - p_n)^2} = \sqrt{\sum_{i=1}^n (q_i - p_i)^2}$$

Luego, continuando con la explicación, se agrega el sesgo (b^1), en caso de existir, y se define cuál es la neurona que está más cerca del vector de entrada p . Aquella neurona que resulte más próxima es la ganadora y podrá acercarse hacia la posición donde se registró el vector entrante p .

Figura 3.1. Arquitectura de una red competitiva.



Fuente: Adaptado de Hudson, Hagan y Demuth [9].

Un punto relevante es que solo una neurona resulta ganadora (la más cercana) para cada uno de los vectores entrantes y solo ella es la que actualiza su posición para acercarse. Esto se conoce como la función de la neurona ganadora que se lleva todo (*winner take-all function*).

La actualización en la posición de las neuronas se realiza a través de reglas de aprendizaje. Para este tipo de red, se emplea la regla de aprendizaje de Kohonen, misma que detallaremos en el siguiente apartado.

3.4. Regla de aprendizaje de Kohonen

Suponga que la neurona i es la ganadora, entonces los elementos de la misma son actualizados de la siguiente forma:

$${}_iIW^{1,1}(q) = {}_iIW^{1,1}(q - 1) + \alpha(p(q) - {}_iIW^{1,1}(q - 1))$$

Con ello, la neurona i será más propensa a ganar la competencia la próxima vez que un vector similar a p se presente, así como menos proclive a ganar cuando un vector muy distinto aparezca. De esta forma, mientras más vectores entrantes sean presentados, cada una de las neuronas se acercará más a un grupo en particular. Eventualmente, si hay suficientes neuronas, cada agrupación de vectores entrantes similares tendrán una neurona asignada. Así, la red competitiva ha aprendido a categorizar los vectores entrantes que recibe.

3.5. El porqué de los sesgos (*bias*) en las redes competitivas

Una de las limitaciones de las redes competitivas es que algunas de sus neuronas pueden no ser reubicadas. En otras palabras, la posición inicial de algunas neuronas puede estar lejos de donde se registran los vectores entrantes y nunca ganar competencia, sin importar cuan largo sea el entrenamiento. Estas son las llamadas **neuronas muertas**, y no desempeñan una función útil. Para evitar este suceso, se les da una ventaja a las neuronas que no han ganado o que lo han hecho en pocas ocasiones, para que tengan mayores posibilidades de ganar. Lo anterior se realiza mediante el seguimiento de un contador que registra la cantidad de veces que cada neurona ha ganado y, con base a éste, se asigna un sesgo que varía para cada iteración. Este mecanismo tiene las ventajas siguientes: en promedio, todas las neuronas ganan en la misma cantidad de ocasiones; se eliminan la neuronas muertas y mientras más vectores entrantes haya en una región del espacio euclidiano más neuronas tendrá el mismo. Esto último tiene por consecuencia que el espacio sea clasificado en secciones más pequeñas y especializadas.

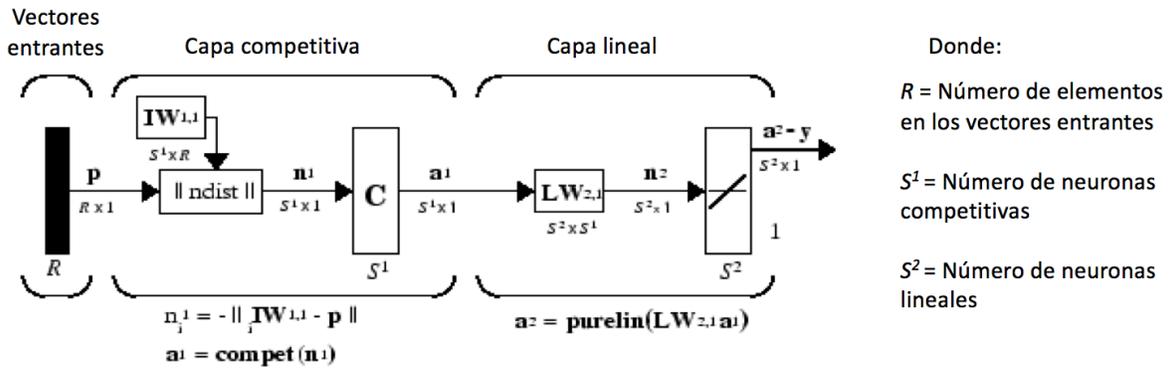
3.6. Introducción al algoritmo LVQ

El LVQ, introducido por Kohonen [8], es un algoritmo de clasificación no paramétrico que funciona con una primera capa competitiva y segunda capa lineal. La primer capa realiza la clasificación de forma muy similar a la de las redes neuronales competitivas, que se describe en el apartado 3.3. de este trabajo. Por otra parte, la segunda capa, se encarga de transformar los resultados de la primera a clases que son definidas por el usuario.

Este algoritmo ha sido ampliamente usado en tareas de clasificación en tiempo real. Destacan como ejemplos de su uso: el monitoreo de la calidad en líneas de producción, el reconocimiento de voz en dispositivos electrónicos, la clasificación de olores y la detección de correos no deseados, entre otros. Respecto a los usos en el área financiera, se encuentran: la asignación de calificaciones crediticias para bonos emitido en el mercado, y que tienen una gran concordancia con las emitidas por calificadoras especializadas [16], la predicción de una bancarrota corporativa, de gran interés bancario debido a que ofrece una medida del riesgo crediticio [17], entre otros.

Para poder tener una mejor perspectiva, a continuación se incluye la figura 3.2., la cual muestra la arquitectura de un algoritmo LVQ. En ella se puede ver que está compuesto por dos capas: la primer capa clasifica los vectores entrantes en subclases, mientras que la segunda transforma estas subclases en clases objetivo definidas por el usuario. En adelante, las clases de la primer capa son nombradas *subclases*, mientras que las clases de la segunda capa son nombradas *clases objetivo*.

Figura 3.2. Arquitectura de un algoritmo LVQ.



Fuente: Adaptado de Hudson, Hagan y Demuth [9].

Para que el LVQ pueda funcionar requiere de un entrenamiento supervisado, en el que se le diga cuál es el resultado que debe de obtener para cada uno de los vectores entrantes que se le presentan. Así, esta red neuronal adecúa sus resultados para descubrir patrones durante la fase de entrenamiento con el objeto de que, finalmente, sea capaz de clasificar por si misma, entradas que no se le habían mostrado anteriormente.

Giusto y Piger [4] describen el algoritmo del LVQ de la siguiente forma: Sea X una colección de N observaciones $x_n \in \mathbb{R}^m, n = 1, \dots, N$ para los que la clasificación en la clase objetivo $\{C_k\}_{k=1}^K$ es conocida. Sean entonces $\bar{N} \in [K, N]$ neuronas $m_i \in \mathbb{R}^m, i = 1, 2, \dots, \bar{N}$ con determinadas posiciones iniciales. Finalmente, sea $g = 1, 2, \dots, G$ el conjunto de iteraciones del algoritmo y sea α^g una secuencia decreciente de números reales entre 0 y 1. Dadas las posiciones iniciales de las \bar{N} neuronas, el algoritmo del LVQ hace ajustes a sus posiciones a través de los siguientes pasos.

Paso 1: Permitir que $g = 1$ y $n = 1$

Paso 2: Identificar la neurona m_g^c más cercana al valor ingresado x_n :

$$c = \operatorname{argmin}_{i \in \{1, \dots, \bar{N}\}} \|x_n - m_i^g\|$$

Paso 3: Ajustar la ubicación la neurona de acuerdo a la siguiente regla:

$$\begin{cases} m_c^{g+1} = m_c^g + \alpha^g [x_n - m_c^g] & \text{si } x_n \text{ y } m_c^g \text{ pertenecen a la misma clase objetivo} \\ m_c^{g+1} = m_c^g - \alpha^g [x_n - m_c^g] & \text{para cualquier otro caso} \end{cases}$$

Paso 4: Si $n + 1 \leq N$ sea entonces $n = n + 1$ y pase a repetir el procedimiento a partir del paso 2. De otra forma, sea $n = 1$ y $g = g + 1$. Si $g \leq G$, repita a partir del paso 2 y detenga para cualquier otro caso.

Como se aprecia en el paso 3 de la secuencia para realizar ajustes en la posiciones de las neuronas, durante el entrenamiento del algoritmo LVQ es posible que la neurona ganadora se acerque o se aleje del vector entrante. Esto es diferente a lo que observado en las redes neuronales competitivas, donde la neurona ganadora sólo podía acercarse. Para ahondar en esta pequeña diferencia, se describe a continuación la regla de aprendizaje que usan los algoritmos LVQ.

3.7. Regla de aprendizaje LVQ

El aprendizaje del LVQ en la capa competitiva está basado en pares que se integran por vectores de entrada (1) y por sus correspondientes clases objetivo (2):

$$\{p_1, t_1\}, \{p_2, t_2\}, \dots, \{p_q, t_q\}$$

Para entrenar la red neuronal, un vector entrante p es presentado y la distancia entre éste y cada una de las neuronas de la primer capa $IW_{1,1}$ es computada con distancias euclidianas. Se comparan las distancias y se asigna la neurona que está más cerca del vector de entrada. La neurona ganadora en la primer capa, se encuentra, a su vez, dentro del espacio de una de las clases objetivo de la segunda capa, mismo que puede coincidir con el que el usuario había indicado al momento de cargar la base de datos para el entrenamiento o no. Si coinciden:

$$i_k^{2,1} = t_k$$

La ubicación de la neurona ganadora i se actualiza acercándose al vector entrante p :

$${}_iIW^{1,1}(q) = {}_iIW^{1,1}(q - 1) + \alpha(p(q) - {}_iIW^{1,1}(q - 1))$$

Si, de lo contrario, discrepan:

$$i_k^{2,1} \neq t_k$$

La ubicación de la neurona ganadora i se actualiza alejándose del vector entrante p :

$${}_iIW^{1,1}(q) = {}_iIW^{1,1}(q - 1) - \alpha(p(q) - {}_iIW^{1,1}(q - 1))$$

Estas correcciones son realizadas por el sistema automáticamente a través de un mecanismo llamado propagación hacia atrás (*back-propagation*) en donde la señal de salida se compara con la salida deseada.

Esto mismo ocurre para cada uno de los vectores entrantes mientras dure el entrenamiento. Si, en efecto, el vector entrante fue clasificado correctamente, entonces la ubicación de la neurona se modifica para acercarse a él. Por el contrario, si el vector entrante fue erróneamente clasificado, entonces la neurona se aleja del mismo. Cabe aclarar que α (el ritmo de aprendizaje) tiene un valor entre 0 y 1, y que éste se mantiene constante a lo largo de todo el periodo. Una vez que el entrenamiento ha culminado, el clasificador puede ser usado para predecir la clase objetivo de las nuevas observaciones.

La predicción de la clase objetivo es expresada de la siguiente forma. Sea x_{N+1} un nuevo vector para el que su clasificación es desconocida. Su subclase es predicha encontrando cuál es la neurona m_c más cercana a x_{N+1} de acuerdo a la métrica euclidiana:

$$c = \operatorname{argmin}_{i \in \{1, \dots, \bar{N}\}} \|x_{N+1} - m_i\|$$

Para que finalmente sea asignado x_{N+1} a la misma clase objetivo a la que esa subclase pertenece.

Como lo menciona Kohonen [8], los métodos LVQ fueron intencionalmente desarrollados para el reconocimiento de patrones estadísticos, especialmente cuando se trabaja con información estocástica multidimensional susceptible de gran cantidad de ruido blanco. La idea general sobre la que reside el algoritmo LVQ es el aprendizaje supervisado, bajo el esquema de premio y castigo.

Otro punto que también es importante mencionar es que existen varias versiones de este prototipo tales como OLVQ, LVQ2 y LVQ3; a pesar de que, como el propio creador apunta, las tres opciones tienen casi la misma exactitud en la mayoría de las tareas de reconocimiento de patrones, éstas se soportan en diferentes formas de abordar los eventos de análisis. Para esta investigación se optó por el uso de LVQ dado que LVQ2 sólo debe ser empleado una vez que se haya ejecutado LVQ primeramente [9]. Se recomienda para futuros trabajos la aplicación de LVQ3 y de OLVQ, y su consiguiente comparativa con los resultados obtenidos en este trabajo.

Capítulo 4. Implementación del algoritmo LVQ

Luego de haber estudiado cómo funciona el algoritmo LVQ, es momento de entrar de lleno al modelo empleado en este trabajo. En esta sección se observan qué parámetros fueron definidos para moldear la red neuronal al problema específico de identificación de ciclos. También se abordan las bases de datos que sirvieron de entrada, el modelo base que funge como guía para entrenar al algoritmo LVQ, así como ciertas mediciones para evaluar el desempeño del sistema.

4.1. Determinación de parámetros en el algoritmo LVQ

Para implementar el algoritmo LVQ, primeramente es necesario definir ciertos parámetros. A saber: las clases objetivo sobre las que el sistema hará la clasificación t_c (1); el número de neuronas para la primer capa, \bar{N} (2); el ritmo de aprendizaje de las neuronas, α (3); y el número de épocas u ocasiones que se presentan los vectores de entrada durante el entrenamiento, G (4). Para algunos de estos criterios, se tomaron las recomendaciones hechas por Kohonen [8], que están soportadas por un gran número de aplicaciones empíricas.

Los valores que emplearemos para cada uno de los criterios sirven solamente para definir los parámetros iniciales. Es decir, serán aquellos con los que iniciaremos el experimento; más adelante, estos serán probados en una serie de iteraciones para seleccionar aquellos que resulten los mejores para este trabajo en específico.

t_c = Clases Objetivo. Las clases objetivo son solamente dos: recesión (0), que representa la transición entre un pico y un valle y, por el otro lado, expansión (1), que es la fase que se presenta entre un valle y un pico.

\bar{N} = Número de neuronas en la primera capa. De acuerdo a la información del citado anteriormente, un buen número de neuronas en la primer capa es igual a la cantidad de vectores entrantes que, en el periodo de entrenamiento, se ubicaron en la clase que registró menos frecuencia. Dado que, en nuestro experimento, los vectores entrantes que pertenecen a la clase objetivo recesión son menos, entonces el número total de neuronas que asignaremos en la primer capa es igual al doble de los vectores entrantes que están clasificados como recesión.

Concretamente, nuestro periodo de entrenamiento solamente registro 26 meses con recesión de un total de 153 meses evaluados. Así podemos concluir que el número de neuronas iniciales que tendrá nuestro algoritmo LVQ en la primer capa es de 52 ($26*2$).

α = **Ritmo de aprendizaje de las neuronas.** Hemos decidido iniciar el experimento con el propuesto por Giusto y Piger [4], equivalente a 0.3. Conforme avancemos en el trabajo éste será modificado hasta llegar al ritmo de aprendizaje óptimo.

G = **Número de épocas ingresadas durante el entrenamiento.** Usualmente ocurre que una red neuronal, como la que tenemos en cuestión, “sobre-aprende”. Este hecho sucede cuando un gran número de épocas se presentaron durante el periodo de aprendizaje. Una posible explicación a este suceso es que las neuronas se vuelven demasiado especializadas respecto a la información contenida en el entrenamiento, causando así que la habilidad del algoritmo para generalizar eventos se pierda. De ahí que sea importante establecer una cantidad de épocas adecuada. Kohonen [8] resalta que es necesario detener el proceso de aprendizaje una vez que se hayan completado un máximo de iteraciones de entre 50 y 200 veces el número de neuronas en la primer capa, dependiendo por supuesto del ritmo de aprendizaje.

Dado que nuestro algoritmo cuenta con 52 neuronas, esto nos hace pensar que un número apropiado de iteraciones es de al menos 2,600 (52*50). Como el número de observaciones con las que contamos para nuestro entrenamiento es de 153, esto nos lleva a concluir que realizaremos al menos 17 ciclos o épocas (que redondearemos a 20) para comenzar con el ejercicio.

Exceptuando las clases objetivo (t_c), se realizaron un conjunto de pruebas para seleccionar los mejores parámetros para el algoritmo LVQ. De esta forma: ensayaremos con cada uno de los puntos muestrales que se generen partiendo de la tabla 4.1. Como podemos ver existen 6 opciones para cada uno de los 3 parámetro de entrada. Así, aplicando una regla de multiplicación, podemos concluir que habrá 216 puntos muestrales posibles ($6*6*6 = 6^3 = 216$).

Tabla 4.1. Opciones para cada uno de los parámetros de entrada.

Arbol de decisiones			
Número de Neuronas	Ritmo Aprendizaje	Periodos Entrenamiento	Punto Muestral (NN, RA, PE)
50	0.01	30	50, 0.01, 30
60	0.05	40	.
70	0.11	50	.
80	0.16	60	
90	0.21	70	
100	0.26	80	100, 0.26, 80

Fuente: Elaboración propia.

El último punto a tomar en cuenta es el acomodo inicial de las neuronas en la primera capa, mismo que puede tener efectos en los resultados. El trabajo de Giusto y Piger [4] resalta una serie de pruebas para su acomodo inicial como medias de k (*k-means*), medianas de k (*k-medoids*) o incluso empleando un mapa auto-organizado (SOFM), que a su vez usa alguna topología. Por su parte, Raffinot [18] prefiere valerse de un acomodo aleatorio. Para este trabajo, luego de varias pruebas, se optó por un acomodo inicial -para todas las neuronas- en un punto central en función de la posición de los vectores entrantes.

4.2. Datos empleados como variables de entrada

En lo que corresponde a la selección de variables que sirven como datos de entrada, inicialmente se consideraron tanto series económicas como financieras; sin embargo, una investigación más a fondo hizo que se descartaran las de índole económica. Las razones para tomar esta decisión fueron, en primer lugar, que los indicadores económicos están atrasados con respecto a los indicadores financieros y, en segundo lugar que, una vez que son publicados pueden presentar varias correcciones incluso de varios puntos porcentuales.

Estos dos puntos repercuten en serias limitantes para usar variables económicas como datos de entrada. Por ello, para el presente análisis se decidió usar sólo variables financieras y omitir por completo las variables económicas.

Hay multitud de trabajos que estudian las correlaciones existentes entre mercados accionarios de varios países. Benelli y Ganguly [19], por ejemplo, realizaron trabajos para determinar la relación entre el mercado financiero de E.U.A. y el de siete países latinoamericanos entre los que se incluye México. De la misma manera, Valdés [20] encuentra una relación de cointegración en los mercados accionarios de Argentina, Brasil, Chile y México, con el de E.U.A. para el periodo de 1991 a 2005. Y, a pesar de que se pudiera pensar que la relación existe solo entre mercados de Latinoamérica, Naranjo [21] encuentra una relación de largo plazo entre el IPC y el índice de E.U.A. con el de Reino Unido.

Con fundamento en los trabajos anteriores, podemos ver relaciones profundas entre los mercados financieros del mundo. De ahí que, para el presente trabajo, se haya optado por seleccionar, como variables de entrada para identificar el estado del IPC, los principales índices financieros globales. La forma en la que estos fueron elegidos contempla, primeramente, que formen parte de las áreas económicas más importantes a nivel global (Norteamérica, Europa y el Este Asiático) y, en segundo término, que son los que reúnen a las empresas con más capitalización y volúmen en sus mercados. Como lo denota Joan Restá [22], “estos índices suelen ser los termómetros de las bolsas de cada país...” “Son una muestra representativa de todo el universo analizado.” Las compañías que forman parte de los índices son, entonces, las más negociadas y, por tanto, las más líquidas o, lo que es lo mismo, las que despiertan mayor interés de compra y/o de venta entre los accionistas. La tabla 4.2. señala qué índices fueron considerados para esta investigación.

Tabla 4.2. Índices bursátiles incluidos en el análisis.

Europa	Norteamérica	Asia	América Latina
CAC 40	NASDAQ 100	Nikkei 225	BOVESPA
Paris, Francia	NY, EEUU	Tokio, Japón	Sao Paulo, Brasil
DAX 30	NYSE	Hang Seng	
Frankfurt, Alemania	NY, EEUU	Hong Kong, China	
FTSE 100	S&P 500		
Londres, Reino Unido	NY, EEUU		

Fuente: Elaboración propia.

Ahora, se pasa a comentar brevemente las características principales de cada uno de ellos:

CAC 40: toma su nombre del primer sistema de automatización de la Bolsa de París (*Cotation Assistée en Continu*). El índice es una medida ponderada según la capitalización de los 40 valores más significativos de entre las 100 mayores empresas negociadas en la bolsa anteriormente señalada.

DAX 30: el índice *Deutscher Aktienindex* es el índice bursátil de referencia de la Bolsa de Fráncfort y el más conocido de la Bolsa alemana. Su cálculo consiste en una media aritmética ponderada por capitalización. Está compuesto por las 30 empresas con más capitalización de Alemania.

FTSE 100: el *Financial Times Stock Exchange* está compuesto de los 100 principales valores de la Bolsa de Londres (*London Stock Exchange*). La capitalización de las empresas que componen el índice supone el 70% del valor total del mercado de valores de Londres.

NASDAQ (Composite) 100: Incluye todos los valores (tanto nacionales como extranjeros) que cotizan en el mercado *Nasdaq*, y que reúnen a más de 5,000 empresas. La alta ponderación de valores tecnológicos dentro del mercado *Nasdaq* lo ha hecho muy popular y representativo para esta área.

NYSE: también conocido como el “*Big Board*”, es una bolsa de valores americana situada en Manhattan, NY. Es, por mucho, la más grande bolsa de valores por capitalización de mercado, ya que sus compañías listadas, valían alrededor de \$19.69 trillones de dólares en Mayo de 2015 de acuerdo al reporte mensual de la Comisión Internacional para las transacciones bursátiles (*World Federation of Exchanges*) [23]. Tenía un promedio de transacciones diarias de \$169 billones de dólares en 2013.

S&P 500: este índice es uno de los más importantes de Estados Unidos. Se le considera el más representativo de la situación real del mercado.

NIKKEI 225: es el índice bursátil más popular del mercado japonés, lo componen los 225 valores más líquidos que cotizan en la Bolsa de Tokio. Desde 1971, lo calcula el periódico *Nihon Keizai Shinbun* (Diario Japonés de los Negocios), de cuyas iniciales proviene su

nombre.

HANG SENG: es el principal índice bursátil chino en el *Hong Kong Stock Exchange*. Es usado para grabar y monitorizar diariamente los cambios de las más grandes compañías de Hong Kong en el mercado de acciones. Consiste en 33 compañías representando el 65% del *Hong Kong Stock Exchange*. Sus valores se ponderan por el criterio de capitalización. Para formar parte del índice, el valor debe estar dentro del 90% de empresas con mayor capitalización y volumen, y haber cotizado en la Bolsa de Hong Kong durante al menos 24 meses.

BOVESPA: el índice BOVESPA está compuesto por acciones que representan el 80% del volumen de las transacciones realizadas en la bolsa de valores de Sao Paulo en los últimos 12 meses. Representa aproximadamente el 70% del valor de todas las acciones que se comercian en esta bolsa.

Para todos los indicadores se extrajeron las siguientes series de tiempo: primeramente la que corresponde a los indicadores netos, en segundo lugar, la que resulta de las diferencias porcentuales respecto al mes anterior y, finalmente, una tercera serie que tiene como base el indicador neto al que se le retira la tendencia a largo plazo para después ser sometido a un proceso de normalización $[N(\mu, \sigma)]$. Es importante señalar que, luego de varias pruebas, la serie de tiempo que se terminó por emplear fue el indicador neto, debido a que con él se generaban los resultados que mayor coincidencia tenían con los reales durante el periodo de evaluación.

4.3. Base para clasificación de series: el modelo de Bry y Boschan

Para entrenar al algoritmo LVQ es necesario contar con una clasificación previa de las etapas de expansión y recesión. Para tal efecto, empleamos el método para fechado de series propuesto por Bry y Boschan [24], el cual logra acercarse con relativo éxito a la cronología publicada periódicamente por el NBER y que resulta de aplicar la metodología y los principios de Burns y Mitchell [10].

El trabajo de Bry y Boschan [24] es, precisamente, la codificación más conocida de los métodos de fechado de Burns y Mitchell. En él, se reflejan nociones de tipo estadístico, puramente instrumental y sin concepciones emanadas de la teoría económica.

A pesar de que el método de Bry y Boschan no incluye el análisis complementario y la intuición que agrega el grupo de expertos en los pronunciamientos del NBER, captura de manera relativamente eficiente los puntos definidos por esta institución. En un ejercicio realizado por los mismos autores, sobre 50 series desde 1947 y hasta 1966 para E.U.A., encontraron que su técnica capturaba el 95% de los puntos de quiebre identificados por el NBER, 90% de los cuales se ajustaba a la cronología ajustada.

El método de Bry y Boschan consiste en la aplicación sucesiva de un algoritmo de identificación de puntos de giro a una secuencia de series filtradas, evolucionando el grado

de suavidad de mayor a menor. Así, se utiliza en primer lugar una media móvil de 12 términos (MM(12)), a continuación un filtro Spencer, después una media móvil de 3 términos (MM(3)) y, finalmente, la serie original sin suavizar. Los puntos de giro finales son determinados en esta última serie, manteniendo la coherencia con los que se han determinado previamente en sus versiones suavizadas y asegurando el cumplimiento de una serie de restricciones. Debe resaltarse que la serie de partida a la que se le aplica el procedimiento debe de carecer de oscilaciones estacionales.

Siguiendo a King y Plosser [25], los principales pasos de dicha metodología son los siguientes:

1) Determinación y suavización de los valores extremos: tenemos diferentes formas de hacerlo. Artis et al. [26] reemplazan el valor que se identifica como extremo, x_t , por el promedio de las dos observaciones adyacentes, es decir: $(x_{t-1} + x_{t+1})/2$. Everts [27], por su parte, considera que dichas observaciones deben ser eliminadas y reemplazadas por su valor en la curva de Spencer. Finalmente, Inklaar [28] considera que no es necesario eliminar dichos valores atípicos (*outliers*), ya que estos forman parte del comportamiento de la serie.

2) Determinación de los ciclos en la serie MM(12), previa eliminación de extremos.

a) Identificación de máximos y mínimos tomando como rango cinco meses de cada lado del punto analizado.

b) Asegurar la correcta alternancia de los puntos de giro, seleccionando solo los picos más altos y los valles más bajos.

3) Determinación de los correspondientes puntos de giro en la curva de Spencer, previa eliminación de extremos. Una curva de Spencer es un promedio móvil ponderado de 15 meses donde a las observaciones se les asignan las siguientes ponderaciones: [-3, -6, -5, 3, 21, 46, 67, 74, 67, 46, 21, 3, -5, -6, -3] /320.

a) Identificación del mayor (o menor) valor en un entorno de +/- 5 meses

b) Imposición de que la duración mínima del ciclo sea de 15 meses, eliminando los picos bajos y los valles altos que definen ciclos menores.

4) Determinación de los correspondientes puntos de giro en la serie MM(3) o MM(6), dependiendo de los meses de dominancia cíclica (*Months for cyclical dominance* o MCD por sus siglas en inglés).

a) Identificación del mayor (o menor) valor en un entorno de +/- 5 meses de los cambios de dirección seleccionados en la curva de Spencer.

5) Determinación de los puntos de giro en la serie no suavizada.

a) Identificación de los puntos más altos (o bajos) en un entorno de +/- 4 meses (o de MCD meses) de los puntos de giro seleccionados previamente en la serie MM(3) o MM(6).

- b) Eliminación de los puntos de giro que se encuentran a menos de 6 meses del comienzo y del final de la serie.
- c) Eliminación de los picos o valles en ambos extremos de la serie que sean menores (o mayores) que los valores próximos a los extremos.
- d) Eliminación de los ciclos cuya duración es inferior a 15 meses.
- e) Eliminación de las fases cuya duración es inferior a 5 meses.

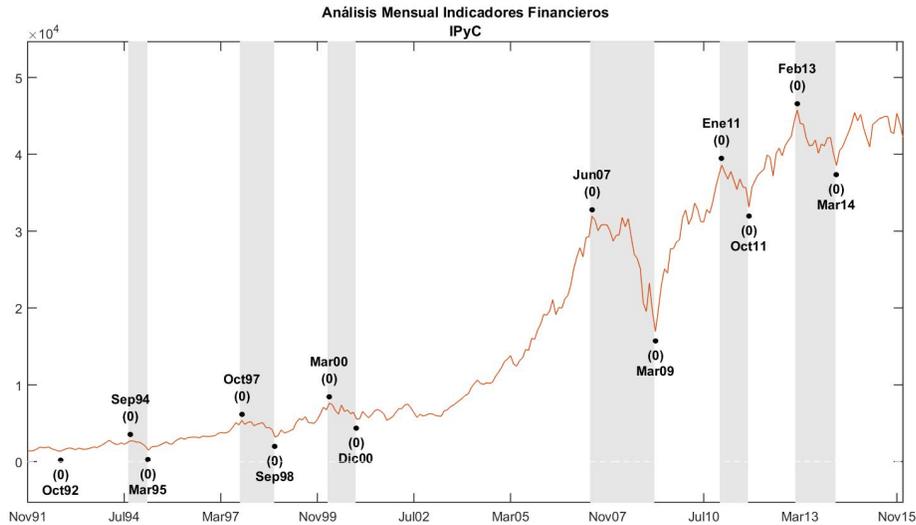
6) Presentación final de los puntos de giro identificados.

En el procedimiento de Bry y Boschan existen tres elementos fundamentales: (a) una identificación preliminar de los puntos de giro en la serie MM(12), (b) una proyección temporal de dichos puntos sobre secuencias de series transformadas mediante filtros de paso bajo hasta llegar a la serie no transformada, y (c) la imposición final de todas las restricciones de duración y ubicación que aseguren que los puntos de giro identificados son compatibles con la definición habitual de ciclo.

Se aclara que este trabajo de investigación no pretende ahondar ni en la metodología y principios propuestos por Burns y Mitchell, ni en el método para fechado de series de Bry y Boschan. Para tal efecto, se ha recurrido a los resultados que arrojan los programas computacionales de fechado de series, que toma como base los trabajos de los autores citados previamente. Tales resultados han servido como base para el entrenamiento de nuestra red neuronal.

Los resultados obtenidos para la serie del IPC se muestran gráficamente en la figura 4.3. En esta figura se separan los periodos de expansión (*bull markets*), que se mantienen con un fondo blanco, de los de recesión (*bear markets*) que, en cambio, llevan un fondo gris. Las fechas señaladas indican los picos y los valles contenidos a lo largo de toda la serie. Estos mismos resultados también se muestran en la tabla 4.4. Se indican secuencialmente los periodos de expansión y de recesión, indicando en las columnas el mes y año de inicio y de culminación para cada uno de ellos.

Figura 4.3. Gráfica del comportamiento del IPC.



Fuente: Elaboración propia con base en metodología para fechado de series de Bry y Boschan [24].

Tabla 4.4. Cronología de los puntos de giro para el IPC.

Índice de precios y cotizaciones (IPC)		
Periodo de entrenamiento (Abril 1993 - Diciembre 2005)		
	Inicio	Fin
Expansión	-	Septiembre 1994
Recesión	Septiembre 1994	Marzo 1995
Expansión	Abril 1995	Octubre 1997
Recesión	Noviembre 1997	Septiembre 1998
Expansión	Octubre 1998	Marzo 2000
Recesión	Abril 2000	Diciembre 2000
Expansión	Enero 2000	-
Periodo de evaluación (Enero 2006 – Diciembre 2015)		
	Inicio	Fin
Expansión	-	Junio 2007
Recesión	Julio 2007	Marzo 2009
Expansión	Abril 2009	Enero 2011
Recesión	Febrero 2011	Octubre 2011
Expansión	Noviembre 2011	Febrero 2013
Recesión	Marzo 2013	Marzo 2014
Expansión	Abril 2014	-

Fuente: Elaboración propia con base en figura 4.3.

En cuanto a nuestro estudio, las clases objetivo (t_c) quedan expresadas siguiendo el esquema siguiente:

$$E_t = \begin{cases} 1, & \text{si la etapa es de expansión} \\ 0, & \text{para cualquier otro caso (recesión)} \end{cases}$$

De esta forma, por ejemplo, los meses que se encuentran entre Septiembre de 1994 y Marzo de 1995 llevan consigo la clase objetivo 0, mientras que los contenidos entre Abril de 1995 y Octubre de 1997 serán considerados clase objetivo 1.

Cada mes, el algoritmo LVQ ensaya con la información que le brindamos. Para ello, fue necesario dividir la serie de tiempo en dos periodos: el periodo de entrenamiento, que va de Abril de 1993 a Diciembre de 2005 y que comprende un total de 153 meses (doce años y nueve meses); y el periodo de evaluación, que es donde se muestra si la red neuronal artificial aprendió lo suficiente para poder catalogar datos que no se le han presentado anteriormente. Este último periodo inicia en Enero de 2006 y culmina en Diciembre de 2015, por un total de 10 años o 120 meses.

Durante el periodo observado en esta investigación, de acuerdo al modelo base, el 75% de los datos encajaron como etapa de expansión, mientras que el restante 25% fueron de recesión. Por otra parte, se contabilizaron un total de 12 puntos de giro, de los cuales 6 se encuentran en la etapa de entrenamiento y el resto en la etapa de evaluación. Parte del propósito de seleccionar estos periodos fue revisar si el algoritmo LVQ era capaz de identificar el mercado bajista que se registró de 2007 a 2009 y, de ser así, con qué prontitud lo hacía. Los detalles respecto a la distribución de los periodos en las dos clases objetivo se muestran en la tabla 4.5.

Tabla 4.5. Distribución de periodos recesivos y expansivos.

	Número de periodos (meses)		
	Entrenamiento	Evaluación	Total
Expansión	127	79	206
Recesión	26	41	67
Total	153	120	273

Fuente: Elaboración propia con base en figura 4.3.

El siguiente paso en nuestro ejercicio fue reducir la cantidad de variables de entrada que estamos empleando y quedarnos solamente con aquellas relevantes. Esto con fundamento en trabajos similares como el de Serena Ng [29], donde identifica qué variables funcionan como señales de alarma antes de que detonen periodos de recesión. Su estudio comienza con un total de 132 variables de entrada para luego concluir que menos de 10 son significativas. Aunado a esto, se tiene como fundamento el trabajo empírico realizado donde, al intentar incluir la totalidad de variables, se reducía la capacidad del LVQ para clasificar adecuadamente los periodos.

El propósito, entonces, fue mantener solamente con un conjunto de cuatro variables de entrada, después de probar con todas las posibles combinaciones ($9C4 = 126$) para determinar el grupo de variables que mejores resultados dieran. Cada combinación fue ingresada en el LVQ, para luego medir su desempeño en el periodo de evaluación. Los métricos que se mencionan en el apartado 4.4. sirvieron como base para determinar las mejores combinaciones.

Se ha de hacer énfasis en que es totalmente arbitraria la decisión de trabajar con conjuntos de 4 variables; sin embargo, es consistente con el número de series que emplearon Stock y Watson [30] para estimar los parámetros en su modelo de clasificación de ciclos de negocios para la economía de E.U.A., por parte del *National Bureau of Economic Research* (NBER).

4.4. Medidas para evaluar los resultados del algoritmo LVQ

Con el fin de evaluar la calidad de los resultados del algoritmo LVQ, tres métricos fueron empleados. Dos de ellos son medidas estadísticas comúnmente usadas en modelos binarios, mientras que el tercero puede ser aplicado en cualquier sistema de predicción. El primero de ellos es el *Brier's Quadratic Probability Score* (QPS) cuyo objetivo es medir la exactitud de las predicciones probabilísticas. Fue propuesto por Glenn W. Brier en 1950 [31] y es aplicable a tareas en que los resultados son mutuamente excluyentes.

Se define de la siguiente forma:

$$QPS = \frac{1}{T} \sum_{t=1}^T (\hat{E}_t - E_t)^2$$

Donde $t = 1, \dots, T$ es el número de pronósticos. \hat{E}_t es el pronóstico que arroja nuestro modelo y E_t es el resultado verdadero. Es, en sí, la sumatoria de los errores cuadráticos medios. En cada iteración, un resultado de 0 es el mejor posible mientras que 1 es el peor a obtener. Un modelo más efectivo apunta a minimizar el QPS.

La segunda medida usada es la correlación de coeficientes de Matthews (*Matthews correlation coefficient*, que en adelante llamaremos por sus siglas en inglés, MCC). Fue introducido por el bioquímico Brian W. Matthews en 1975 [32] y recibe también el nombre de coeficiente *phi* (ϕ ó r_ϕ). Es similar al coeficiente de correlación de Pearson en su interpretación. De hecho, como lo demuestran Everitt y Skronnal [33], un coeficiente de correlación de Pearson estimado para dos variables binarias nos dará como resultado el coeficiente *phi*.

Citando el trabajo de Matthews [32]: "El MCC es, en esencia, una correlación de coeficientes entre las clasificaciones pronosticadas y las observadas en un sistema binario. Regresa valores de +1, para una predicción perfecta, de -1 para un total desacuerdo entre el

pronóstico y el observado, y de 0 para una predicción no mejor que una hecha aleatoriamente.” Uno de los puntos interesantes por los que se emplea esta medida es que toma en cuenta falsos y verdaderos positivos, así como falsos y verdaderos negativos. Se define de la siguiente forma:

$$MCC = \frac{T_p T_n - F_p F_n}{\sqrt{(T_p + F_n)(T_n + F_p)(T_p + F_p)(T_n + F_n)}}$$

Donde T_p son verdaderos positivos, es decir verdaderas expansiones; T_n representan los verdaderos negativos o verdaderas recesiones; F_p son los falsos positivos o falsas expansiones y finalmente F_n son los falsos negativos o falsas recesiones. Para un mejor entendimiento de los posibles resultados se muestra la tabla 4.6.

Tabla 4.6. Matriz de contingencia 2x2 adaptada a este trabajo.

		Realidad	
		Recesión (0)	Expansión (1)
Pronóstico	Recesión (0)	Verdadero Negativo T_n	Falso Negativo F_n
	Expansión (1)	Falso Positivo F_p	Verdadero Positivo T_p

Fuente: Elaboración propia con base en MCC [32].

La tercer y última medida que emplearemos es el porcentaje de aciertos obtenidos. Aquí, simplemente dividimos el número de aciertos, correspondientes al número de verdaderos positivos (V_p) más el número de verdaderos negativos (V_n), entre el total de periodos evaluados, para luego ser multiplicado por 100 y obtener así una porcentual. Si bien, esta medida es bastante sencilla da una idea global que resulta más fácil de visualizar.

Capítulo 5. Resultados de la aplicación del algoritmo LVQ

Para este capítulo, se inicia indicando cuál fue el conjunto de cuatro variables de entrada – dentro de las nueve candidatas- que ofreció la mejor respuesta, y cuál fue el procedimiento que realizado para llegar a dicha conclusión. También se muestran los resultados para cada uno de los métricos datallados en el apartado 4.4. y se observan gráficamente las similitudes y diferencias entre el modelo base y el algoritmo LVQ respecto a la identificación de ciclos recesivos y expansivos.

5.1. Selección de mejor combinación de variables de entrada

La selección de las variables de entrada es uno de los puntos que mayor importancia tienen para que el algoritmo LVQ arroje buenos resultados. Para poder llevar esto a cabo, primeramente es necesario ingresar los parámetros de entrada que permiten al algoritmo LVQ comenzar con sus funciones. Los parámetros son los siguientes: número de neuronas (1), épocas de entrenamiento (2) y ritmo de aprendizaje (3). Una vez introducidos estos datos, se pasa a crear un ciclo que permita probar con cada una de las 132 combinaciones de variables de entrada posibles ($9C4 = 132$).

Los resultados (métricos del apartado 3.4.) para cada opción son guardados y, una vez que el sistema ha terminado de probar con todas las combinaciones, se pasa a elegir la que haya generado los mejores resultados. Ya con el grupo de 4 variables de entrada seleccionado, creamos un segundo ciclo que elija la mejor combinación de parámetros de entrada. Ellos fueron decididos probando cada uno de los 216 puntos muestrales resultantes de las 6 opciones para cada uno de los 3 parámetros ($6*6*6 = 6^3 = 216$).

De esta forma se abordan dos puntos importantes del trabajo: la mejor combinación de variables de entrada (que resultaron ser NIKKEI 225, HANG SENG, DAX 30 y CAC 40) (1), y el mejor grupo de parámetros de entrada (número de neuronas: 80, ritmo de aprendizaje: 0.21 y épocas de entrenamiento: 10) (2). Esto último es muy relevante ya que la mayoría de los trabajos se basan en otras investigaciones para elegirlos y no tienen la certeza de que, en su experimento, sean los idóneos. En este caso, a pesar de que nuestro proceso de selección de variables y parámetros de entrada no sigue del todo los mecanismos para ser considerado “de optimación” y por tanto no es exhaustivo, la serie de pruebas hace que se tenga una selección más apropiada de los parámetros de entrada. Así sabemos, por ejemplo, que 80 neuronas con un ritmo de aprendizaje de 0.21 y 10 entrenamientos son los que mejores resultados arrojan tratándose del estudio de ciclos financieros para el IPC. Más detalles al respecto los podremos encontrar en el capítulo VI, en el apartado de recomendaciones para futuras investigaciones.

La tabla 5.1. muestra las cinco combinaciones, para las variables de entrada, que arrojaron los mejores resultados (mayor porcentaje de aciertos, menor QPS y mayor MCC), así como

los resultados que arroja para cada métrico mencionado anteriormente. La columna intitulada *Línea* es en realidad la fila, dentro de las 132 estudiadas, donde se registró tal resultado.

Tabla 5.1. Variables de entrada que arrojan mejores resultados en los criterios evaluados. Ordenadas por AC (%) evaluación.

Línea	Combinación				Entrenamiento			Evaluación		
					QPS	MCC	AC (%)	QPS	MCC	AC (%)
109	NIKKEI 225	DAX 30	CAC 40	HANGSENG	0.059	0.331	94.118	0.258	0.257	74.167
59	S&P 500	NIKKEI 225	NASDAQ	CAC 40	0.059	0.553	94.118	0.275	0.776	72.500
69	S&P 500	NIKKEI 225	CAC 40	FTSE 100	0.065	0.514	93.464	0.292	0.891	70.833
26	NYSE	NIKKEI 225	NASDAQ 100	HANG SENG	0.052	0.441	94.771	0.308	-0.101	69.167
8	NYSE	S&P 500	NASDAQ 100	DAX 30	0.137	0.730	86.275	0.325	0.888	67.500

Fuente: Elaboración propia.

La tabla 5.2. muestra los cinco conjuntos de parámetros de entrada que produjeron los resultados más altos en los criterios evaluados. Estos siguen el mismo formato de la tabla anterior. La columna llamada *Línea* es la fila, dentro de las 216 estudiadas.

Tabla 5.2. Parámetros de entrada que arrojan mejores resultados en criterios evaluados. Ordenadas por AC (%) evaluación.

Línea	Núm. Neuronas	Factor Aprendizaje	Periodos entrenamiento	Entrenamiento			Evaluación		
				QPS	MCC	AC (%)	QPS	MCC	AC (%)
169	80	0.21	10	0.085	0.549	91.503	0.208	0.063	79.166
17	40	0.11	50	0.065	0.444	93.464	0.258	-0.024	74.167
63	50	0.21	30	0.072	0.413	92.810	0.258	0.498	74.167
157	80	0.11	10	0.092	0.339	90.850	0.258	-0.064	74.167
20	40	0.16	20	0.059	0.405	94.118	0.275	0.304	72.500

Fuente: Elaboración propia.

En el siguiente apartado, se muestran los pasos realizados para la parte de programación que incluyó este trabajo. Ellos son presentados en formato gráfico para facilitar su comprensión. Es importante mencionar que la información mostrada forma parte de los resultados obtenidos.

5.2. Seguimiento esquemático de pasos realizados

El primer paso corresponde a la elaboración de la red neuronal artificial. Para ello fue necesario insertar los vectores de entrada (1), a qué clase objetivo pertenece cada uno de ellos (2), la totalidad de las neuronas que estarán trabajando en la primer capa (3), además del ritmo de aprendizaje (4) y, finalmente, el número de épocas durante las que será

entrenada la red neuronal (5). Los datos ingresados corresponden a aquellos que arrojaron los mejores resultados en las pruebas. A saber:

Tabla 5.3. Resumen de entradas para algoritmo LVQ.

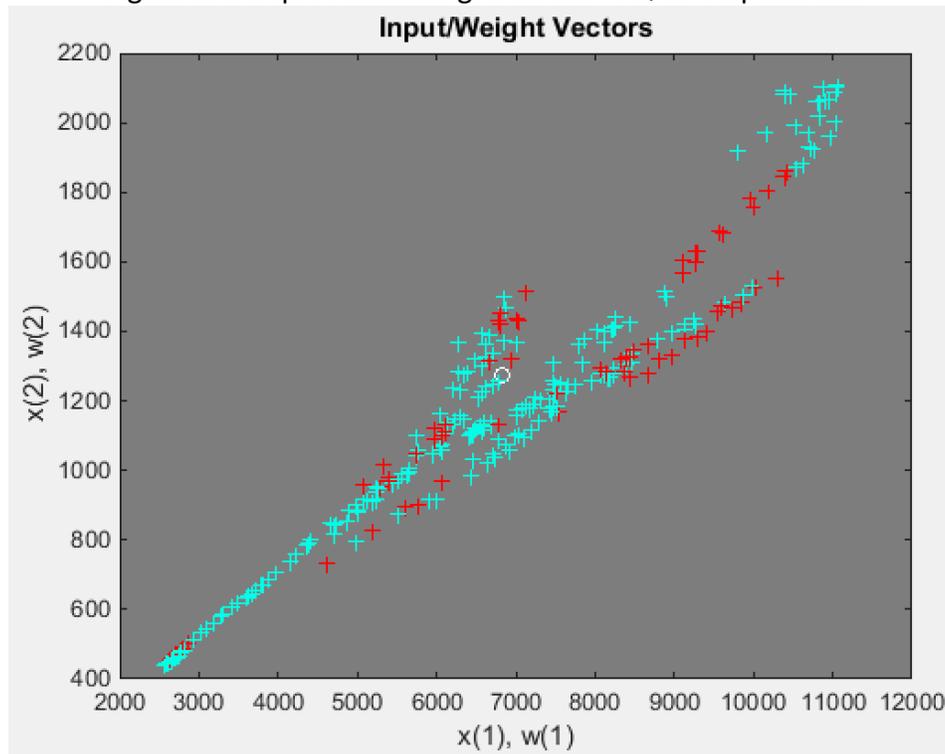
Entradas	Valor	Fuente
Vectores de entrada	NIKKEI 225, HANG SENG , DAX 30, CAC 40	Pruebas realizadas (mejor conjunto tabla 5.1.)
Clases Objetivo	0 Recesión - 1 Expansión	Resultados figura 4.3. – tabla 4.4.
Número de neuronas	80	Pruebas realizadas (mejor conjunto tabla 5.2.)
Ritmo de aprendizaje	0.21	Pruebas realizadas (mejor conjunto tabla 5.2.)
Épocas de entrenamiento	10	Pruebas realizadas (mejor conjunto tabla 5.2.)

Fuente: Elaboración propia.

En la figura 5.4., las cruces (+) representan los vectores de entrada. Se puede observar que las hay en dos colores; las azul claro son expansiones, mientras que las rojas son recesiones. Al centro aparece un círculo (que en realidad son 80 círculos ubicados en la misma posición) de color blanco, que representa la ubicación de las neuronas en la primer capa (competitiva). Esto demuestra que la red neuronal artificial está configurada y lista para comenzar a ser entrenada.

Es importante mencionar que esto es sólo una representación gráfica en dos dimensiones. Dado que están siendo ingresadas cuatro variables de entrada, el sistema LVQ está tratando, en realidad, con cuatro dimensiones.

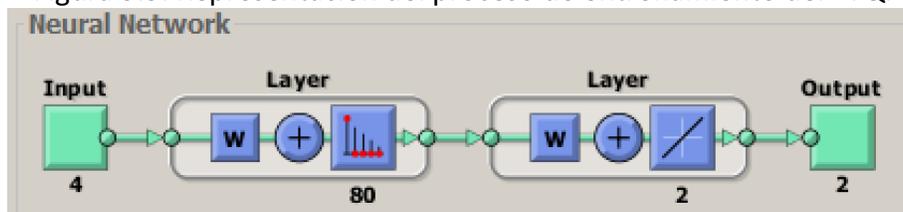
Figura 5.4. Representación gráfica del LVQ en etapa inicial.



Fuente: Elaboración propia empleando paquetería de Matlab.

El siguiente paso fue entrenar la red neuronal y la figura 5.5. es la encargada de representarlo. Como se aprecia, el extremo del lado izquierdo indica el ingreso de cuatro variables; cada una de ellas es, en sí, una serie de tiempo. Todos estos datos pasan a una primera capa, la competitiva, donde se indica que hay un total de 80 neuronas. Los resultados de esta capa van, a su vez, hacia una segunda capa, de tipo lineal, que cuenta solamente con dos neuronas puesto que solo son dos clases las que estamos diferenciando. Con ello, el resultado es precisamente la elección de una de esas dos clases.

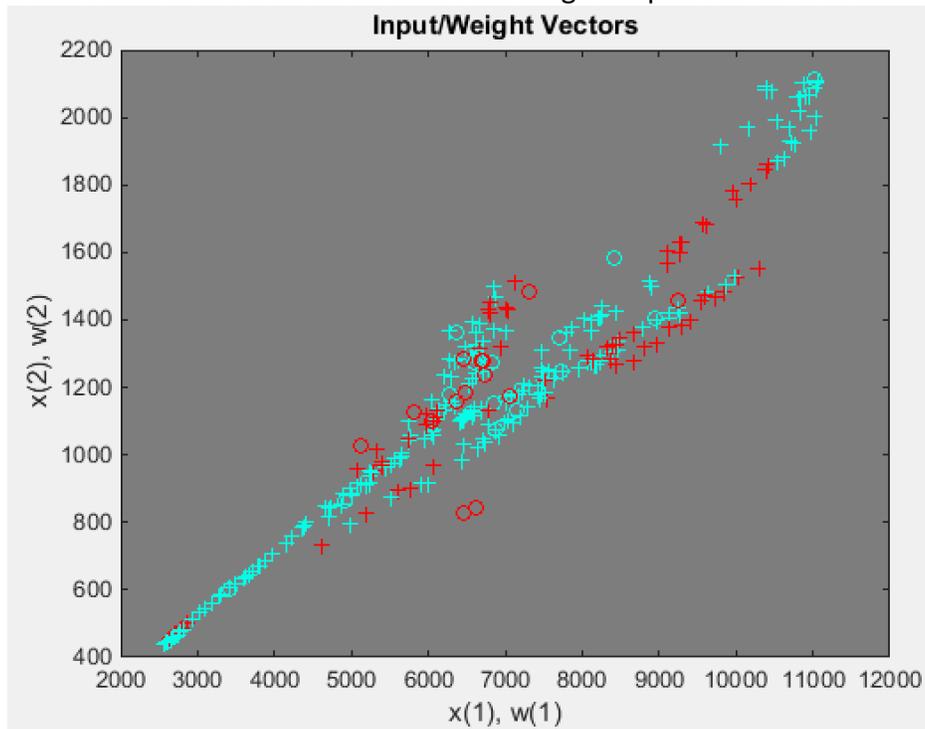
Figura 5.5. Representación del proceso de entrenamiento del LVQ.



Fuente: Elaboración propia empleando paquetería de Matlab.

Durante el proceso de entrenamiento, la posición de las neuronas cambia y comienzan a especializarse en algún conjunto específico de datos. Ello se puede observar en la figura 5.6., donde es claro que la mayoría de las neuronas han salido del punto central y han intentado esparsirse a lo largo y ancho del espacio muestral.

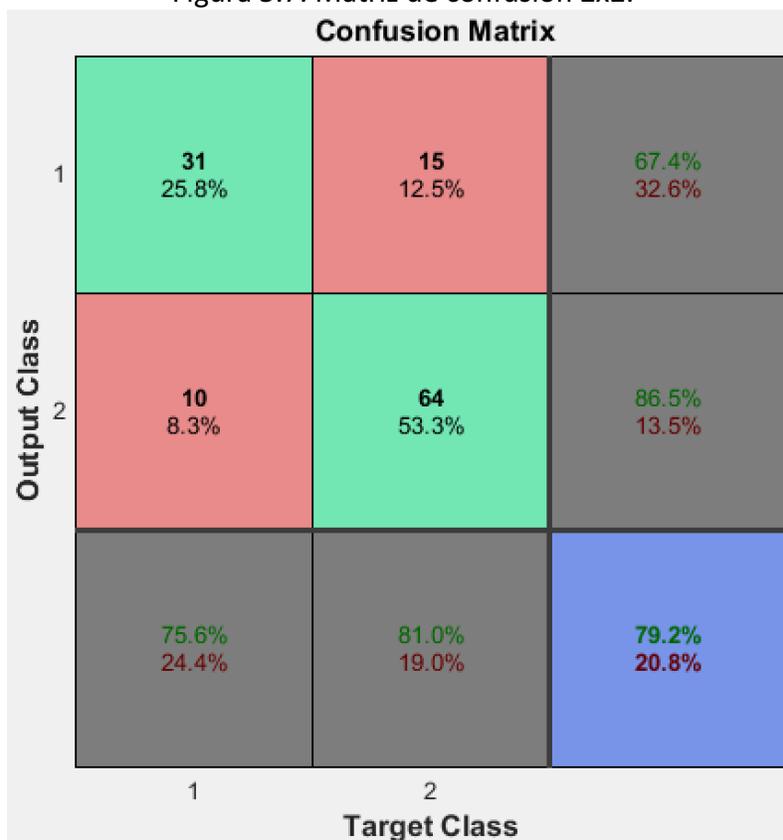
Figura 5.6. Posicionamiento de las neuronas luego del periodo de entrenamiento.



Fuente: Elaboración propia empleando paquetería de Matlab.

Una vez terminado el entrenamiento, se puede entonces recurrir a los métricos establecidos para averiguar qué tan confiables han sido los resultados. Además de ellos, una de las tablas más representativas al respecto es la matriz de confusión 2x2 (figura 5.7.), en la cual se observan el número de aciertos y errores, y el porcentaje que representan, para las dos clases evaluadas.

Figura 5.7. Matriz de confusión 2x2.



Fuente: Elaboración propia empleando paquetería de Matlab.

Como último paso, queda ingresar la información correspondiente al periodo de evaluación. En él, aún se desconoce a qué tipo de clase pertenecen los vectores entrantes pero la pretensión es que, con lo que ha aprendido el sistema LVQ, él mismo puede hacer la clasificación.

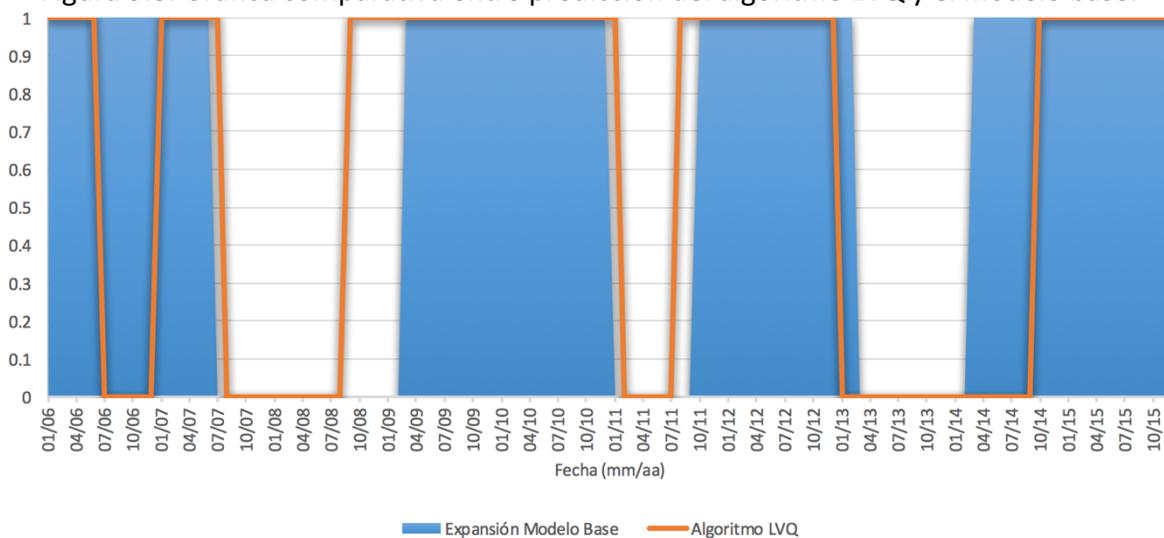
Finalmente, La figura 5.8. ilustra qué tan parecidos fueron los pronósticos que arrojó el algoritmo LVQ respecto al modelo base. En esta figura, los periodos de expansión se encuentran con fondo azul, mientras que los periodos recesivos tienen un fondo blanco. Asimismo, la línea color naranja marca cuál fue el pronóstico del algoritmo LVQ para tal periodo. Como se puede observar, el algoritmo resulta ser muy confiable en términos de rapidez e identificación de los periodos recesivos y expansivos, e incluso suele adelantarse al modelo base por algunos meses. Así, por ejemplo, se muestra que la recesión que el modelo base registra de Junio de 2007 a Marzo de 2009, es identificada por el algoritmo LVQ justo un mes después, en Julio de 2007, a pesar de que en lo que respecta a la recuperación, el algoritmo LVQ manda una señal con 6 meses de anticipación (en Septiembre de 2008) cuando, en realidad, la recuperación comienza en Marzo de 2009.

La segunda recesión, que el modelo base registra de Enero de 2011 a Septiembre del mismo año, es detectada por el algoritmo LVQ de Febrero a Julio de 2011. De nueva cuenta la

detección de la etapa recesiva es bastante acertada, con solo un mes de diferencia, aunque la recuperación se anticipa dos meses respecto al modelo base. Finalmente, la última etapa recesiva de la que se tiene registro durante el periodo de evaluación, es marcada por el modelo base de Marzo de 2013 a Febrero de 2014. El algoritmo LVQ, por su parte, la indica de Enero de 2013 a Septiembre de 2014 lo cual muestra un anticipo de dos meses en la detección de la recesión y un retraso de siete meses en la detección de la recuperación. Finalmente, por parte del algoritmo LVQ, existe un breve periodo de recesión que no concuerda con el modelo base. Este corre de Julio a Diciembre de 2006.

Todo lo anterior es mostrado en la figura 5.8., la cual es una forma clara de exponer qué tan acertado resulta ser el algoritmo LVQ en la detección de los puntos de giro del IPC.

Figura 5.8. Gráfica comparativa entre predicción del algoritmo LVQ y el modelo base.



Fuente: Elaboración propia empleando paquetería de Matlab.

Capítulo 6. Caso práctico: adecuación de un portafolio con base en el estado del IPC

6.1. Planteamiento del ejercicio

Con la finalidad de mostrar, de forma práctica, la utilidad de conocer en tiempo real el estado del IPC, se decidió implementar una estrategia de inversión, que llamamos *activa*, para evaluar las utilidades de un portafolio. Ésta consiste en lo siguiente: a partir de la fecha en que comenzamos a evaluar el desempeño del algoritmo LVQ (principio del periodo de evaluación: Enero 2006), integramos un portafolio por un total de \$100 MXP. Luego, simulando que el inversor corriera mensualmente el algoritmo LVQ, éste realiza una de las siguientes acciones:

- 1) Si el sistema indica una etapa de expansión, entonces invierte el 80% del portafolio en el emulador del índice del IPC y el 20% en CETES a 28 días.
- 2) Si el sistema indica una etapa de recesión, entonces invierte 80% del portafolio en CETES a 28 días y 20% en el emulador del índice del IPC.

La estrategia contra la que haremos la comparación será llamada *pasiva*, es decir: comprar y mantener el 100% del portafolio hasta el fin de nuestro ejercicio (fin del periodo de evaluación: Diciembre 2015).

Adicionalmente, y con la finalidad de proporcionar al lector información acerca del entorno financiero que ha tenido lugar durante este periodo, se ha llevado un registro del rendimiento de los CETES a 28 días. Con ella, el lector puede visualizar que hubiera sucedido de haber elegido una inversión libre de riesgo.

Los datos empleados del índice IPC se obtuvieron de la página de Yahoo Finanzas México [34]. Para evaluar el rendimiento mensual, simplemente dividimos el valor publicado en el mes t (x_t), entre el del valor del mes inmediato anterior (x_{t-1}), para luego restarle un entero. Respecto a los CETES a 28 días, estos son los resultados de la subasta que realiza el Banco de México y que son mostrados en su página oficial [35]. Dado que los datos publicados corresponden a los rendimientos anualizados, fue necesario convertirlos a rentabilidades mensuales de la forma siguiente:

$$C_m = (1 + C_a)^{1/12} - 1$$

Donde C_m representa el rendimiento mensual de los CETES a 28 días, y C_a el rendimiento anual publicado por el Banco de México.

La forma de evaluar los resultados será realizando la comparación de ambos portafolios utilizando las siguientes medidas:

a) Retorno anualizado promedio: resultante de aplicar la fórmula de interés compuesto:

$$S = C(1 + i)^n$$

De la cual queremos obtener la equivalencia de i , que representa la tasa de interés anual:

$$i = \sqrt[n]{\frac{S}{C}} - 1 \quad \text{ó} \quad i = \left(\frac{S}{C}\right)^{1/n} - 1$$

Donde S es el monto al final del periodo, C el monto al inicio del mismo y n el tiempo transcurrido expresado en años.

b) Desviación estándar de los rendimientos mensuales: es la medida de dispersión de los valores respecto a la media. Se usa como cuantificador de la frecuencia e intensidad en los cambios de una serie de valores en un horizonte temporal específico. En este caso la usaremos para medir la dispersión de los rendimientos mensuales empleando la fórmula poblacional:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}$$

Donde x_i es la observación en el tiempo i , \bar{x} es el promedio aritmético del total de observaciones y n es el total de datos que contiene la población.

c) Razón o ratio de Sharpe: originalmente propuesta por William Forsyth Sharpe [36], es una medida del exceso de rendimiento por unidad de riesgo en una inversión. Se define como:

$$S = \frac{E[R_p - R_f]}{\sigma_p}$$

Donde R_p es el rendimiento de la inversión en cuestión, en este caso del portafolio; R_f es el rendimiento de una inversión de referencia que tenga como característica la ausencia de riesgo, que en este ejercicio serán los CETES a 28 días y σ_p es el riesgo del portafolio.

La ratio de Sharpe se utiliza para mostrar hasta qué punto el rendimiento de una inversión compensa el riesgo asumido; una razón por encima de 1 es considerada buena.

d) Pérdida máxima esperada (MDD por sus siglas en inglés): Es la caída más grande de un pico a un valle, antes de que un nuevo pico sea alcanzado. Es expresado en términos porcentuales negativos y se calcula de la siguiente manera:

$$MDD = \frac{(V_v - V_p)}{V_p}$$

Donde V_v es el valor alcanzado en el valle y V_p es el valor en el pico. Es de utilidad para un inversionista porque con él puede conocer cuánto ha sido la máxima pérdida, respecto a su valor mayor previo [37].

Con la finalidad de simplificar el ejercicio, no se consideraron costos de transacción. Así, se da término a lo concerniente a métricos para realizar la comparativa entre los portafolios activo y pasivo. En el siguiente apartado veremos los resultados obtenidos.

6.2. Resultado de los portafolios comparados

Para mostrar los resultados, se ha recurrido a la tabla 6.2. La primera columna muestra cada una de las medidas con las que se evalúan los portafolios. La segunda columna muestra los resultados del portafolio bajo la estrategia activa que toma como base los resultados del algoritmo LVQ. La tercera columna es un ejemplo de la aplicación de la estrategia activa (80%-20%), si tomásemos como base los resultados generados empleando la metodología para fechado de series Bry y Boschan. Es importante señalar que este ejemplo se agregó sólo con la finalidad de ver el potencial que tiene saber, en el momento justo, si el estado del IPC es expansivo o recesivo. Representa entonces un esquema utópico puesto que saber sus resultados requiere de un tiempo considerable y ahí habremos perdido la oportunidad de realizar esa inversión; se presenta, más bien, para mostrar hasta dónde se puede llegar de realizar un buen pronóstico con el *esquema activo del algoritmo LVQ*.

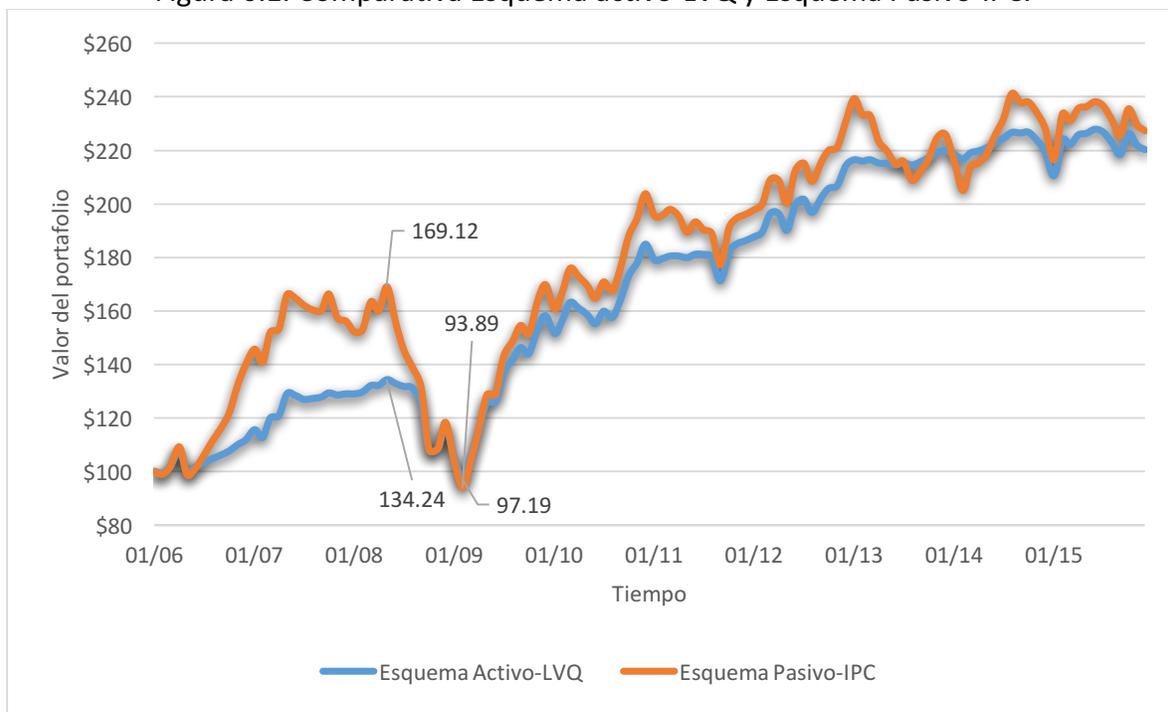
Retomando la explicación de las columnas que aparecen en la tabla 6.2., la cuarta columna se enfoca en el esquema pasivo (qué hubiera pasado si se dejara la inversión durante todo el lapso en el emulador del IPC) y, finalmente la quinta columna, representa un último ejercicio con los resultados obtenidos en caso de dejar la inversión de inicio a fin en CETES a 28 días. Como se observa, este último ejercicio no pretende competir con el resto de los esquemas propuestos. Se mantiene sólo para plasmar cuál hubiese sido el resultado de incurrir en una opción que no generase riesgo alguno.

Dicho esto, los resultados que realmente se deben de comparar son los obtenidos del *esquema activo para el algoritmo LVQ* y el *esquema pasivo para el IPC*. Como se puede ver en la figura 6.1. y en la tabla 6.2., el algoritmo LVQ genera un valor final menor que simplemente dejar la inversión durante todo el periodo en el emulador del IPC (*Pasivo-IPC*). Existe una diferencia negativa en los valores finales (de \$7.24) y en los retornos anualizados promedio (0.36% de margen), lo cual indica que hay desventaja si vemos la estrategia de

inversión propuesta bajo la óptica del rendimiento. Por el otro lado, si se observan los resultados bajo la óptica del riesgo, la desviación estándar de los rendimientos y la pérdida máxima esperada se ven reducidas indicando que, en este rubro, sí existe una ventaja empleando el portafolio propuesto.

La figura 6.1. tiene etiquetados los picos y valles con mayor diferencia antes de que un nuevo pico haya sido alcanzado. El esquema que involucra al LVQ tiene un MDD de -27.59%, resultante de restar el valor del valle menos el valor del pico, para luego ser dividido entre el valor del pico ($[\$97.19-\$134.24]/\$134.24$), mientras que el *esquema pasivo del IPC*, por su parte, tiene un MDD de -44.48% ($[\$93.89-\$169.12]/\$169.12$). Existe entonces una diferencia significativa, de 16.89%, entre la pérdida máxima esperada que pudiera suponer un inversionista entre ambas opciones.

Figura 6.1. Comparativa Esquema activo-LVQ y Esquema Pasivo-IPC.



Fuente: Elaboración propia.

Finalmente, siguiendo la directriz de la razón de Sharpe, estaríamos frente a una mejor inversión si se tratase de emplear el algoritmo LVQ ya que tiene un ratio de 0.95; por su parte, el esquema pasivo del IPC solo logró un ratio de 0.75. Aun así no hay que olvidar que esta razón considera como buenas inversiones a aquellas que sobrepasan un ratio de 1, por lo que ninguno de los portafolios comparados son aceptables.

En resumen, solo bajo la óptica del riesgo, el sistema propone mejoras con tan solo hacer una revisión mensual y tomando en cuenta sólo una variable. En lo que respecta al rendimiento, la propuesta ofrecida no resulta favorable.

Tabla 6.2. Medidas de retorno y riesgo para esquema pasivo y activo.

Comparativo esquema activo vs pasivo (Enero 2006 - Diciembre 2015)				
Medidas portafolio	Esquema Activo		Esquema Pasivo	
	Algoritmo LVQ	Modelo base	IPC	CETES
Valor final (\$)	220.16	400.43	227.40	162.01
Retorno anualizado promedio	8.28%	15.02%	8.64%	5.02%
Desviación estándar de los rendimientos	3.42%	2.90%	4.83%	0.14%
Razón de Sharpe	0.95	3.44	0.75	0
MDD	-27.59%	-8.05%	-44.48%	0

Fuente: Elaboración propia.

Capítulo 7. Conclusiones y recomendaciones

En conclusión, esta tesis ha probado su hipótesis central la cual es la identificación del estado, expansivo o recesivo, del IPC así como sus consecuentes puntos de giro mediante el uso del algoritmo LVQ. Esto queda plasmado en la figura 5.8, de forma gráfica, y en la primera fila de la tabla 5.2., con una serie de medidas estadísticas que hacen una comparativa entre los resultados obtenidos por el algoritmo y los valores obtenidos del modelo base.

Igualmente, se han logrado cumplir los objetivos particulares planteados. En primer lugar, el algoritmo fue capaz de identificar, con solo un mes de retraso, el mercado bajista ocurrido de 2007 a 2009 a pesar de que, en lo que respecta a la recuperación, manda una señal con 6 meses de anticipación (en Septiembre de 2008) cuando, en realidad, la recuperación comienza en Marzo de 2009.

En segundo lugar, siguiendo con los objetivos particulares, se ha corroborado la estrecha correlación existente entre los índices de los mercados accionarios de varios países y el principal índice bursátil mexicano. Esto una vez que, a través de ellos, ha sido posible identificar el estado del IPC con un grado de efectividad del 79.2%.

Finalmente, en tercer lugar y como último objetivo particular, no ha sido posible aseverar que conocer el estado en que se encuentra el IPC genere beneficios adicionales en comparación con un portafolio que hace caso omiso de esta información toda vez que, mediante el caso práctico, se pudo constatar el descenso en las medidas de rendimiento – Valor final del portafolio: \$7.24 menor, Retorno anualizado promedio: 0.36% menor- pero, por otro lado, también hubo una reducción en las medidas correspondientes al riesgo – Desviación estándar de los rendimientos: 1.42% menor y MDD: 16.89% menor-. Esto nos indica que el portafolio propuesto tiene, por una parte, un rendimiento menor pero por otra genera una menor exposición al riesgo. La medida final, la Razón de Sharpe, resulta 0.20 mayor en el esquema propuesto que en comparado, lo cual también es favorable.

Inicialmente se tenían un total de 9 variables de entrada, que son los índices financieros globales más importantes de acuerdo a su capitalización y volumen de mercado (ver tabla 4.2.), y se logró obtener un conjunto de solo 4 variables que, juntas, dieran los mejores resultados para el pronóstico. Para este fin, se midió la eficacia y asertividad de cada uno de los 132 conjuntos posibles ($9C4$) a través de medidas estadísticas como QPS (*Quadratic probability score*), MCC (*Matthews correlation coefficient*) y el porcentaje de aciertos que registraba cada prueba. Igualmente, se realizó un proceso para la elección de los tres parámetros de entrada -el número de neuronas en la primer capa, el ritmo de aprendizaje y el número de periodos que duró el entrenamiento de la red neuronal-. En el proceso de elección se tomaron 6 opciones posibles para cada una de los 3 parámetros (ver tabla 3.1.), y se probó con cada uno de los 216 conjuntos, resultando en la prevalencia de los que ofrecieron mejores resultados.

Ya que se contaba con el mejor modelo posible –que involucra 4 variables de entrada (NIKKEI, HANG SENG, DAX 30 y CAC 40) y 3 parámetros de entrada (Número de neurona: 80, Ritmo de aprendizaje: 0.21, y Periodos de entrenamiento: 10)-, se realizó una prueba la cual fue cotejada contra un modelo base. La comparativa queda plasmada en la figura 5.8 y los resultados de los métricos –QPS = 0.208, MCC = 0.063 y Porcentaje de aciertos = 79.2%- respecto al modelo base son mostrados en la primer fila de la tabla 4.2.

Una vez realizado lo anterior, se procedió a usar el algoritmo LVQ en un caso práctico (capítulo 6) en donde, primeramente, se definieron métricos –Valor final del portafolio, Retorno anualizado promedio, Desviación Estándar de los rendimientos, Razón de Sharpe y MDD- para evaluar el rendimiento y la aversión al riesgo de un portafolio que tomara como base las salidas del algoritmo LVQ en comparación con otro que no las considerara.

Los resultados (ver tabla 6.2.) muestran que aquel portafolio que tomó en cuenta los resultados del algoritmo LVQ (Esquema Activo – Algoritmo LVQ) tiene menores rendimientos –Valor final del portafolio: \$7.24 menor, Retorno anualizado promedio: 0.36% menor-, pero también menores riesgos -Desviación estándar de los rendimientos: 1.42% menor, MDD: 16.89% menor- que el esquema de comparación propuesto (Esquema Pasivo – IPC). De lo anterior deducimos, primeramente, que el algoritmo LVQ ha sido capaz de identificar el estado del índice IPC, y sus consecuentes puntos de giro, con un nivel de asertividad bastante bueno. Sin embargo, en segundo lugar, se deduce del capítulo 6 que el conocer el estado del IPC no es suficiente para obtener beneficios, en forma de rendimientos, superiores a los de un portafolio de inversión que no toma en cuenta esta información. Lo anterior deja un espacio para que, en futuras investigaciones, se agreguen otros criterios antes de la toma de decisiones en un portafolio.

Por otra parte, en referencia a las recomendaciones para futuras investigaciones, me gustaría agregar como primer punto, la realización de este mismo ejercicio con otros tipos de redes neuronales especializadas en el reconocimiento de patrones como lo son *LVQ2*, el *LVQ3* y el *OLVQ*, que en cierta parte son muy similares al usado.

Otra sugerencia está relacionada con variaciones en cuanto a los parámetros de la red neuronal como, por ejemplo, el acomodo inicial de las neuronas. Como se señaló anteriormente, este ejercicio se realizó colocando a las neuronas en un punto de partida central en referencia a los vectores entrantes, pero esto puede ser modificado. La ubicación inicial puede hacerse siguiendo una topología hexagonal, en forma de cuadrícula o incluso de forma aleatoria. Sería interesante observar si los resultados se mejoran con alguno de estos acomodos.

Otro parámetro que puede ser modificado es la función de aprendizaje. Aquí se puede configurar al sistema de tal forma que, por ejemplo, se modifiquen los movimientos de las neuronas en grupo, hasta una vez que cada época de entrenamiento haya terminado. Más

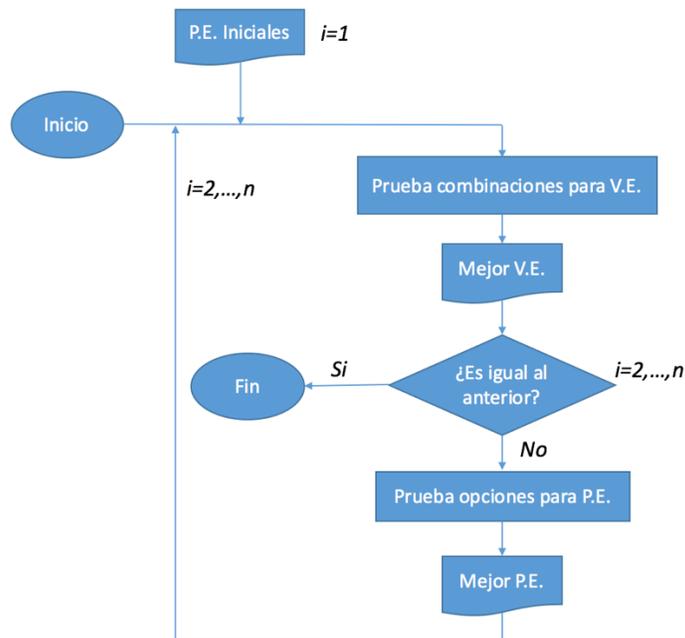
información sobre esta función de aprendizaje se puede encontrar en el glosario de términos bajo la palabra “procesamiento en grupo (*batch processing*)”.

En lo que respecta a los parámetros que dan forma al sistema, en resumen, son muchas las modificaciones que se pueden realizar. Será tarea de futuras investigaciones el probar con alternativas diferentes a las usadas en esta tesis.

Existe una última recomendación para futuras investigaciones: durante este trabajo se probó con una gran cantidad de opciones con la finalidad de integrar el mejor conjunto de variables y parámetros antes de poder comenzar con el experimento. En concreto, fueron probadas 132 combinaciones de variables de entrada ($9C4$) y 216 conjuntos de parámetros de entrada (6^3). Sin embargo, este procedimiento no puede ser considerado exhaustivo y deja la puerta abierta a un proceso iterativo de optimización lineal que tenga como función objetivo la maximización del MCC, el porcentaje de aciertos o, desde una perspectiva contraria, la minimización del QPS.

El proceso iterativo pudiera seguir el flujo propuesto en la figura 7.1.: primeramente un ciclo que evalúe cada una de las variables de entrada (V.E. en la tabla 4.2.), luego un segundo ciclo que haga lo propio con los parámetros de entrada (P.E. en la tabla 4.1.) y de nueva cuenta regresará a la realización de estos dos procesos hasta que el sistema converja a ciertas variables y parámetros de entrada idóneos para el trabajo en cuestión. En esta investigación, sobre todo por falta de recursos computacionales, no fue posible seguir este diagrama de flujo de forma iterativa.

Figura 7.1. Diagrama de flujo propuesto para futuras investigaciones concerniente a la selección de parámetros y variables de entrada.



Fuente: Elaboración propia.

Se dan por terminadas las conclusiones y recomendaciones y, de la misma forma, por culminado este trabajo.

Referencias (en orden de aparición)

- [1] Anas, J. y Ferrara, L. (2004). Detecting Cyclical Turning Points: The ABCD Approach and Two Probabilistic Indicators. *Journal of Business Cycle Measurement and Analysis*, 2, 193-225.
- [2] Berge, T. (2013). Predicting recessions with leading indicators: model averaging and selection over the business cycle. Research Working Paper RWP 13-05, Federal Reserve Bank of Kansas City.
- [3] Katayama, M. (2010). Improving recession probability forecast in the U.S. economy. Technical report, Louisiana State University.
- [4] Giusto, A. y Piger, J. (2014). Identifying Business Cycle Turning Points in Real Time with Vector Quantization. Technical report.
- [5] Ferrara, L. y Guegan, D. (2005). Detection of the industrial business cycle using sector models. *Journal of Business Cycle Measurement and Analysis*, 3, 353-371.
- [6] Liu, W. y Moench, E. (2014). What predicts U.S. recessions? Staff Reports 691, Federal Reserve Bank of New York.
- [7] Guidolin, M. (2011). Markov Switching Models in Empirical Finance. Technical report.
- [8] Kohonen, T. (2001). Self-organizing maps. Espoo, Finland: Springer Series in Information Sciences.
- [9] Hudson, M., Hagan, M. y Demuth, H. (2015). Self-Organizing and Learning Vector Quantization Networks en Neural Network Toolbox. *Matlab User's Guide (189-232)*. Natick, MA: The MathWorks, Inc.
- [10] Burns, A. y Mitchell, W. (1946). *Measuring Business Cycles*, NBER, ISBN: 0-870-14085-X.
- [11] Heath, J. (2011). Identificación de ciclos económicos en México: 30 años de evidencia. Realidad, datos y espacio. *Revista Internacional de Estadística y Geografía*. INEGI, 2, núm. 2, 18-31.
- [12] Achuthan, L. y Banerji, A. (2004). *Beating the Business cycle*, a Currency Book, published by Double-day, a Division of Random House, 52-72.
- [13] Heath, J. (2012). Lo que indican los indicadores. Cómo utilizar la información estadística para entender la realidad económica de México. INEGI, 31-57.

- [14] Guerrero, V. (2013). Capacidad predictiva de los índices cíclicos compuestos para los puntos de giro de la economía mexicana. *Revista economía mexicana nueva época*, vol. XXII, núm. 1, 1er semestre, 47-99.
- [15] McCulloch, W. y Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biology*, 5, núm. 4, 115-133.
- [16] Kyung-Sup, K. y Ingo, H. (2001). The cluster-indexing method for case-based reasoning using self-organizing maps and learning vector quantization for bond rating cases. *International Journal of Expert systems with applications*, Editorial Elsevier.
- [17] Neves, J. y Vieira, A. (2006). Improving bankruptcy prediction with hidden layer learning vector quantization. *European Accounting Review*, 15, núm. 2, 253-271.
- [18] Raffinot, T. (2014). Can macroeconomists get rich nowcasting economic turning points? Technical report. BNPP Cardiff Investment & Asset Management.
- [19] Benelli, R. y Ganguly, S. (2007). Financial linkages between the U.S. and Latin America: evidence from daily data. *International Monetary Fund. Working Paper*, núm. 262.
- [20] Valdés, A. (2006). Modelos de corrección de error no lineal entre mercados accionarios latinoamericanos y el mercado accionario de Estados Unidos. *Revista de análisis económico*, 21, núm. 1, 117-129.
- [21] Naranjo, A. (1997). Relación de largo plazo del indicador bursátil mexicano con los de Estados Unidos, Japón, Reino Unido y Singapur: un análisis bivariado de cointegración y mecanismos de corrección del error. *Ejecutivos de Finanzas*, 69-80.
- [22] Joan Restá. (28/05/2015). ¿Cuáles son los principales índices bursátiles del mundo?. 30/08/2016, de Instituto Superior de Empresa y Finanzas, SL Sitio web: <http://isefi.es/cuales-son-los-principales-indices-bursatiles-del-mundo/>
- [23] World Federation of Exchanges (2015). Monthly reports.
- [24] Bry, G. y Boschan, C. (1971). Cyclical analysis of Time Series: Selected Procedures and Computer Programs. *NBER Technical Report*, 7-63.
- [25] King, R. y Plosser, C. (1994). Real Business Cycles and the Test of the Adelmans. *Journal of Monetary Economics*, núm. 33, 405-438.
- [26] Artis, M., Kontolemis, Z. y Ozbörn, D. (1997). Business Cycles for G7 and European Countries. *The Journal of Business*, 70, núm. 2, 249-279.

[27] Everts, M. (2006). Duration of Business Cycles. Department of Economics, University of Bern. Technical report (preliminar e incompleto).

[28] Inklaar, R., Jacobs, J. y Romp, W. (2003). Business cycle indexes: Does a heap of data help? Discussion Paper 2003-12. University of Groningen, CCSO Centre of Economic Research.

[29] Ng, S. (2014). Viewpoint: Boosting recessions. Canadian Journal of Economics, 47, núm. 1, 1-34.

[30] Stock, J. y Watson, M. (1989). New Indexes of Coincident and Leading Economic Indicators. NBER Macroeconomics Annual 1989, 4, 351-409.

[31] Brier, G. (1950). Verification of forecasts expressed in terms of probability. Monthly weather review, 78, núm. 1. Weather Bureau, Department of Commerce.

[32] Matthews, B. (1975). Comparison of the predicted and observed secondary structure of T4 phage lysozyme. Biochimica et Biophysica Acta (BBA) – Protein Structure, 405, núm. 2, 442-451.

[33] Everitt, B. y Skrondal, A. (2010). The Cambridge Dictionary of Statistics, 4th Edition. ISBN 9780521766999.

[34] Yahoo! Finanzas México. [En línea] <http://mx.finance.yahoo.com/>.

[35] Banco de México. [En línea] <http://www.banxico.org.mx/>.

[36] Sharpe, W. (1994). The Sharpe Ratio. The Journal of Portfolio Management. Stanford University.

[37] Magdon-Ismail, M. y Atiya, A. (2004). Maximum Drawdown. Risk Magazine.