UNIVERSIDAD TECNICA FEDERICO SANTA MARIA

Repositorio Digital USM **https://repositorio.usm.cl**

Tesis USM TESIS de Pregrado de acceso ABIERTO

2018

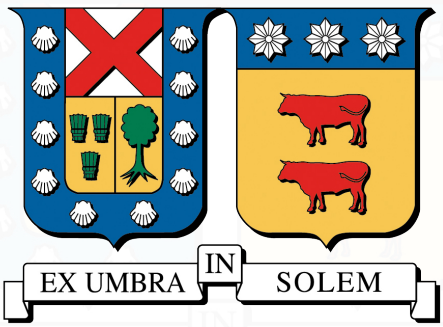
PROPUESTA DE METODOLOGÍA PARA MEJORAR ESTRATEGIA DE INVERSIÓN DE MOMENTUM EN EL MERCADO BURSÁTIL CHILENO POR MEDIO DE UN MODELO BASADO EN REDES NEURONALES ARTIFICIALES

ROJAS MERANI, CARLOS RONY

https://hdl.handle.net/11673/46600

Repositorio Digital USM, UNIVERSIDAD TECNICA FEDERICO SANTA MARIA

UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE INDUSTRIAS
VALPARAÍSO - CHILE



PROPUESTA DE METODOLOGIA PARA MEJORAR ESTRATEGIA DE INVERSIÓN DE MOMENTUM EN EL MERCADO BURSÁTIL CHILENO POR MEDIO DE UN MODELO BASADO EN REDES NEURONALES ARTIFICIALES

CARLOS RONY ROJAS MERANI

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

PROFESOR GUÍA	:	DR. WERNER KRISTJANPOLLER R.
PROFESOR CORREFERENTE	:	KEVIN MICHELL V.

Noviembre 2018

Resumen Ejecutivo

Las estrategias de inversión no están ajenas a los nuevos tiempos y deben adaptarse al entorno dinámico del último tiempo y es tarea y desafío de las nuevas generaciones propuestas de metodologías que permitan una correcta toma de decisiones al invertir.
Esta memoria propone la implementación de un modelo de Redes Neuronales Artificiales en la estrategia de momentum, capaz de mejorar el desempeño, logrando mayores retornos a un menor riesgo.

Los resultados obtenidos muestran sinergia en dos modelos propuestos: un modelo que se alimenta de las rentabilidades que siguen unas determinas acciones y proyecta las rentabilidades acumuladas futuras y otro modelo que recibe en forma de inputs estas rentabilidades acumuladas y tiene como target entregar directamente las acciones en las que se debería invertir. Este segundo modelo logra una estrategia que domina al resto, comparándola además con otras estrategias y un modelo econométrico.
Ambos modelos tienen como mejor configuración de red 1 capa oculta y 5 neuronas y superan estadísticamente a la estrategia de momentum.

La metodología propuesta cuenta con la ventaja de ser de simple programación y entrega resultados claros.

Índice de Contenidos

Resumen Ejecutivo I

1. Problema de Investigación 1

2. Objetivos 3

2.1. Objetivo General 3

2.2. Objetivos Específicos 3

3. Marco Teórico 4

3.1. Mercados 4

3.1.1. Inversión e introducción a la teoría de portfolio 4

3.2. Series de tiempo 5

3.2.1. Estacionariedad 6

3.3. Modelo econométrico 7

3.3.1. ARIMA 8

3.4. Valor en Riesgo 9

3.5. Prueba Estadística 9

3.6. Antecedentes Generales Momentum 10

3.6.1. Causas del momentum 11

3.6.2. Validación y evolución 13

3.7. Fundamentos estrategia 52 Week High 14

3.8. Inteligencia Artificial 15

3.9. Redes Neuronales Artificiales 16

3.9.1. Neurona Artificial 17

3.9.2. Comportamiento de información de entrada y salida 23
3.9.3. Tipos de Redes Neuronales 24
3.9.3.1. Red neuronal monocapa 24
3.9.3.2. Red neuronal multicapa 24
3.9.3.3. Red neuronal recurrente 25
3.9.4. Clasificación de conexión 25
3.9.5. Aprendizaje de una red 26
3.9.5.1. Aprendizaje supervisado 28
3.9.5.2. Aprendizaje no supervisado 28
3.9.5.3. Aprendizaje híbrido 28
3.9.5.4. Aprendizaje reforzado 28
3.9.6. Backpropagation 29

4. Metodología **31**
4.1. Datos 31
4.2. Procedimiento 31
4.2.1. Posiciones financieras 32
4.2.2. Aplicación estrategia momentum 32
4.2.3. Aplicación estrategia 52 Week High 34
4.2.4. Estructura de la Red Neuronal Artificial 35
4.2.5. Implementación RNA 35
4.2.5.1. Modelo 1: Rentabilidad de cada periodo 36
4.2.5.2. Modelo 2: Rentabilidad acumulada 37
4.2.5.3. Modelo 3: Clasificación ganadoras 38

5. Resultados **41**

6. Conclusión y Discusión **54**

7. Bibliografia **58**

Índice de Tablas

4.1. Ejemplo para modelo 1 de entrada la rentabilidad de una empresa y salida la rentabilidad del período siguiente de la misma empresa. Fuente: Elaboración propia 37

4.2. Ejemplo para modelo 2 de entradas la rentabilidad logarítmica de una empresa y salida la rentabilidad logarítmica del periodo siguiente para la misma empresa. Fuente: Elaboración propia 38

4.3. Ejemplo para modelo 3 de entradas considerando la rentabilidad acumulada de la acción y target la decisión en forma binaria si es que correspondió a una acción en la se debía invertir. Fuente: Elaboración propia 39

5.1. Rentabilidad promedio mensual para cada modelo y estrategia calculado con las series proyectadas 45

5.2. Estadística descriptiva de los resultados de los 3 modelos propuestos de redes neuronales y de las 3 estrategias alternativas convencionales de inversión 46

5.3. Ranking modelos según retorno promedio mensual 47

5.4. VaR para cada modelo y estrategia calculado con la serie de rentabilidades proyectadas utilizando un nivel de confianza del 95 % 48

5.5. Ranking modelos según VaR Histórico 49

5.6. Indice de Sharpe para cada modelo y estrategia calculado con la serie de rentabilidades proyectadas 52

5.7. Ranking modelos según Indice de Sharpe 53

5.8. Indice de Sharpe global para cada modelo considerando la serie de rentabilidad promedio de cada mes al usar todas las estrategias en forma simultanea 53

Índice de Figuras

3.1. Estructura Red Neuronal Artificial Fuente: https://victoryepes.blogs.upv.es 19
3.2. Función lineal Fuente: https://http://recursostic.educacion.es 20
3.3. Función escalón unitario Fuente: Matlab 21
3.4. Función Sigmoide Fuente: https://es.kisspng.com/ 22
3.5. Función Tangente Hiperbólica Fuente: https://www.researchgate.net/ 22
3.6. Estructura Neurona Artificial Fuente: http://www.http://grupo.us.es/ 23
3.7. Red Neuronal Monocapa Fuente: http://grupo.us.es/ 24
3.8. Red Neuronal Multicapa Fuente:http://grupo.us.es/ 25
3.9. Red Neuronal Recurrente Fuente: http://grupo.us.es/ 25

4.1. Ejemplo Estrategia Momentum 3x3 Fuente: Elaboración Propia 33

5.1. Rentabilidad promedio mensual para cada estrategia del modelo 1 42
5.2. Rentabilidad promedio mensual para cada estrategia del modelo 2 42
5.3. Rentabilidad promedio mensual para cada estrategia del modelo 3 43
5.4. Rentabilidad promedio mensual para cada estrategia de los modelos propuestos 44
5.5. Gráfico comparativo Indice de Sharpe por estrategia entre los modelos de redes neuronales propuestos 50
5.6. Gráfico comparativo Indice de Sharpe por estrategia, agrupando estrategias con período de mantención común entre los modelos de redes neuronales propuestos. Con *n* tomando los periodos de 3, 6, 9 y 12 meses de formación 51

1 | Problema de Investigación

¿Qué tan buenos somos para predecir el desempeño que tendrá una acción a futuro?. Amplia es la literatura que respalda que la rentabilidad de las acciones son predecibles en función de su información pasada y que mediante una *estrategia momentum*, estudiada por primera vez por Jegadeesh & Titman (1993), es posible obtener retornos anormales en el corto - mediano plazo. Vasta es también la literatura para los mercados desarrollados, sin embargo para el caso de los mercados emergentes, como el chileno y el resto de los mercados accionarios latinoamericanos la información es más escasa y no tan profunda. No obstante, una manera buscar nuevas estrategias y corroborar la existencia del principio de impulso de estas estrategias de inversión y aprovechando las nuevas tecnologías y conocimientos actuales, es mediante la aplicación de unas de las ramas de la Inteligencia Artificial (IA), como lo es las Redes Neuronales Artificiales (RNA).

Una definición de economía financiera es el estudio del comportamiento de los individuos en la asignación intertemporal de sus recursos en un entorno incierto (Marín y Rubio, 2001). De esta definición podemos identificar la presencia del concepto de intertemporalidad y de incertidumbre. De la primera idea podemos destacar que la economía financiera tiene dependencia de sucesos que no son solo del momento presente que estamos viviendo. Del segundo elemento, la incertidumbre, podemos esperar que se refiere a lo que sucederá a futuro, ya que es algo que hasta ahora, el hombre no conoce a cabalidad lo que sucederá, mas ¿qué nos dice el pasado?.

Los inversionistas de todo el mundo tienen como objetivo generar ganancias en base a distintas estrategias de inversión, así como profesores, académicos e investigadores intentan

entender y explicar las causas del comportamiento financiero. Algunos inversionistas más ligados a la teoría financiera tradicional, piensan que el precio de la acción depende únicamente de la información presente que está presente en el mercado sumado a las especulaciones que tienen los inversores a futuro. Por otro lado, están los accionistas que creen en el desempeño pasado de las acciones tiene su continuidad a futuro, basándose en la evidencia que las acciones que han logrado un buen desempeño pasado en el pasado continúa obteniendo buenos resultados en el futuro y análogamente, las acciones con deficiente desempeño pasado, mantienen la misma línea a futuro.

Un tipo de estrategia para este último tipo de inversores es la de de *52 Week High* propuesta por George y Hwang (2004), la cual está basada en la estrategia de momentum de Jegadeesh & Titman (1993) que consiste en mantener posición larga en una cartera compuesta por las acciones dentro del primer decil (ganadoras) ordenadas por rendimiento y tener posición corta con las acciones de peor desempeño (perdedoras) en un horizonte de corto - mediano plazo. Esta estrategia planteada por George y Hwang, propone, a diferencia del momentum tradicional que considera la rentabilidad como indice predictor, que el poder de predicción de los rendimientos de las acciones es más elevado a medida que se aproxima al precio más alto que ha alcanzado el activo financiero en las últimas 52 semanas.

Está estrategia a superado las rentabilidades logradas con el momentum tradicional para algunos mercados desarrollados y emergentes (Ding Du,2007), sin embargo, siempre ha sido un objetivo del humano el superarse y obtener cada vez mejores resultados.

El enfoque de este análisis está centrando en las estrategias de inversión de Momentum y su variación de 52 Week High para el mercado latinoamericano y en mejorar su rentabilidad obtenida mediante la aplicación de un modelo con Redes Neuronales Artificiales, utilizando las nuevas herramientas tecnológicas y conocimientos que van a la vanguardia, con la intención de aportar a futuras investigaciones y a inversores, buscando recomendar o descartar objetivamente su uso para mercados emergentes.

2 | Objetivos

2.1. Objetivo General

Realizar una propuesta de metodología para potenciar el desempeño de una estrategia de inversión basada en Momentum mediante Redes Neuronales capaz de ser una alternativa real para un inversionista en el mercado accionario chileno.

2.2. Objetivos Específicos

- Aplicar la estrategia tradicional de momentum para las 40 principales acciones del mercado bursátil chileno, correspondientes al IPSA, para un periodo comprendido entre los años 2010 y 2018.
- Aplicar la estrategia de 52 Week High utilizando la misma data usada para el momentum.
- Potenciar la aplicación de la estrategia de inversión de Momentum al combinarla con pronósticos por medio de un modelo que utiliza Redes Neuronales Artificiales.
- Potenciar la estrategia de inversión de momentum al utilizar un modelo econométrico de referencia y realizar una comparación objetiva de los resultados obtenidos para cada estrategia.
- Pronosticar rentabilidades para 24 meses con horizonte de tiempo común y realizar comparación objetiva entre modelos.

3 | Marco Teórico

3.1. Mercados

Desde el punto de vista económico, un mercado es el conjunto de transacciones de procesos o intercambio de bienes o servicios entre demandantes (compradores) y oferentes (vendedores). En forma particular, se identifica al mercado financiero, el cual se define como el espacio, pudiendo ser físico o virtual, en el cual los compradores y vendedores están interesados en realizar comercio por instrumentos financieros.
Dentro de los mercados financieros existen distintos tipos de submercados, donde se encuentran el mercado de capitales, correspondiente a la negociación financieros de mediano y largo plazo de títulos, en su mayoría acciones de empresas que cotizan en Bolsa. Por su parte, se encuentra el mercado de valores, siendo este un tipo de mercado de capitales, el cual posee un rol más operativo, ya que corresponde a donde se realiza la negociación de los valores negociables, los cuales son los instrumentos del mercado monetario líquidos, que pueden convertirse fácilmente en efectivo. En el mercado de valores se registra, monitorea y liquida todas las operaciones que se llevan a cabo entre los compradores y vendedores.

3.1.1. Inversión e introducción a la teoría de portfolio

Una cartera o portafolio (más comúnmente porftolio en inglés) en finanzas, corresponde a una combinación de inversiones en activos financieros realizadas por alguna persona o

institución. Su conformación se lleva a cabo en marco de una estrategia de diversificación en las inversiones realizadas y controlar el riesgo.

En 1952, Harry Markowitz, revolucionó el campo financiero al proponer un enfoque que planteaba del como se debía conformar un portfolio de inversión de tal manera que sea eficiente. Según, Markowitz, una cartera eficiente, es aquel conjunto de acciones que minimiza el riesgo, para un retorno dado, o de forma equivalente, el conjunto de acciones que maximiza el retorno para un nivel de riesgo dado.
Este enfoque y principios entregados para un portfolio eficiente, están presentes en gran cantidad de modelos de conformación de carteras de inversión, tanto lineales como no lineales, los cuales entregan el porcentaje de inversión que se le debe asignar a cada acción. Estos modelos siguen la corriente de la teoría financiera tradicional.

A la hora de conformar una cartera de inversión, uno de los índices de performance más famoso y utilizado a nivel mundial es el Indice de Sharpe, el cual se utiliza como ratio que sirve para comparar portfolios en relación con el retorno esperado de la cartera (r_p), el riesgo asociando en forma de la desviación estándar (σ) y la tasa libre de riesgo (r_f):

$$\text{Indice de Sharpe}\,(I_{sharpe}) = \frac{r_p - r_f}{\sigma_p} \tag{3.1}$$

3.2. Series de tiempo

Una serie de tiempo corresponde a un secuencia de datos u observaciones medidos en instantes de tiempo determinados y que se encuentran ordenados cronológicamente. La importancia de las series de tiempo en el ámbito financiero está estrechamente ligado a las decisiones de inversión dado el interés en tratar de predecir valores o rentabilidades futuras de tal manera de obtener beneficios, esto debido a que el precio de las acciones en Bolsa

de los distintos mercados se puede modelar en base a una serie de tiempo al igual que los retornos generados por dichas acciones a lo largo del tiempo, esto debido a que no es posible predecir de manera precisa los valores futuros, lo cual entregaría un retorno conocido. Para lo anterior, la motivación radica fundamentalmente en hallar modelos que proporcionen proyecciones que se ajuste a la realidad, los cuales permitan realizar inferencias de los valores futuros que se debiesen obtener.

3.2.1. Estacionariedad

La gran mayoría de los modelos econométricos, para realizar pronósticos sobre las series de tiempo, tienen como condición el que la serie sea estacionaria, es decir, que a lo largo del tiempo existirá una reversión a la media y las variaciones entorno a la media mantendrá una amplitud similar.
En finanzas, bajo el supuesto de mercado eficiente, los precios de las acciones reflejan toda la información disponible y salvo variaciones de los precios que ocurren en el corto plazo, no existe otra relación entre si para estos cambios. Las series no estacionarias se dirá que presentan Camina Aleatoria y que no es posible predecir sus precios futuros.

Para decir que una serie es estacionaria, el criterio de evaluación de existencia de raíces unitarias. Un test usado es el de Dickey-Fuller, para el cual se tiene como punto inicial el proceso estocástico:

$$\Delta Y_t = \delta Y_{t-1} + u_t \tag{3.2}$$

Donde Y corresponde a la serie de tiempo, Δ al operador de primera diferencia, t al periodo y u_t es error ruido blanco.
Si $\delta = 0$implica que existe raíz unitaria, por lo cual la serie es no estacionar. Si δ toma valor negativo, la serie sí es estacionaria.

Es posible que una serie sea no estacionaria, pero su primera diferencia sí lo sea. A estás series las llamaremos de orden 1 o *I(1)*.

El error ruido blanco, mencionado anteriormente, se define bajo el cumplimiento de 3 supuestos: media del error igual a cero, varianza constante y covarianza igual a 0 entre errores correspondientes a observaciones de distintos periodos.

3.3. Modelo econométrico

Una de las maneras más comunes para realizar predicciones en economía es a través del análisis e interpretaciones de modelos econométricos, los cuales presentan variables dependientes que se buscan explicar (variable endógena) por medio de variables explicativas (variable exógena) las cuales son de carácter independiente. Estas variables son acompañadas de parámetros mediante una formulación matemática y análisis estadístico que dan lugar a una estimación numérica de dichos parámetros, los cuales permiten realizar una proyección de los valores futuros de la serie de tiempo.

3.3.1. ARIMA

ARIMA significa a Modelo Autorregresivo Integrado de Media Móvil (Autoregressive Integrated Moving Average por sus siglas en inglés). Este modelo es una combinación de un proceso autorregresivo, en el cual la variable endógena en un periodo t es explicada por las observaciones de periodos anteriores de la misma, donde además, se le agrega un término que hace referencia al error. Estos modelos autorregresivos se representan como *AR(p)*, donde p corresponder al *orden* del modelo (1,2,....etc.), donde dicho *orden* expresa la cantidad de observaciones con rezago que están presentes en la ecuación matemática explicando a la variable endógena.

Este modelo autorregresivo es el que se combina con el de Medias Móviles, en el cual la variable endógena es explicada por medio de un término independiente en un periodo t con una sucesión de errores de acuerdo a periodos anteriores que son ponderados de manera conveniente. La abreviatura de este Modelo de Media Móvil es *MA(q)*, donde q, al igual que en el modelo autorregresivo, corresponde al *orden*.

La mezcla de estos dos modelos forma un único modelo econométrico llamado ARMA, sin embargo, con el correr de los años se le agregó el concepto de *Integrado(d)*, con d que expresa el número de diferencias que se le deben aplica a serie para que quede estacionaria. De manera genérica un modelo ARMA(pq) es:

$$Y_t = \theta + \sum_{1}^{p} \alpha_i Y_{t-i} + \sum_{1}^{q} \beta_i \epsilon_{t-i} + \epsilon_t \tag{3.3}$$

Donde θ es una constante y ϵ_t es error ruido blanco.

Además podemos generalizar aún más con ARIMA(p,d,q) teniendo p términos autorregresivos, q términos de media móvil y su diferencia d es estacionaria.

3.4. Valor en Riesgo

El Valor en riesgo, más conocido por *Value at Risk* (o VaR por sus siglas en inglés), es una técnica o método estadístico para cuantificar el riesgo financiero de una inversión en el mercado bursátil. El VaR indica la pérdida que se podría sufrir en condiciones normales de mercado en un horizonte de tiempo y a un cierto nivel de probabilidad o de confianza. Una forma de calcular el VaR, es de manera histórica con la serie de tiempo de los retornos, ordenando de manera ascendente las rentabilidades de un intervalo de tiempo y fijar un nivel de confianza, por ejemplo de un 90 %, el último dato ubicado en el primer decil de la serie ordenada corresponderá al VaR y se puede decir que al nivel de confianza dado (90 %) hay un 10 % de probabilidad que las pérdidas excedan al VaR, en otras palabras, en un mes la probabilidad de que la pérdida sea igual o menor al VaR es de un 90 %.

3.5. Prueba Estadística

Para realizar comparaciones entre distintas datas, pudiendo ser series de tiempo, es necesario tener criterios de discriminación o herramientas que permitan rankear, con tal de evaluar de manera objetiva la diferencia en resultados que se pueden obtener de distintos modelos y confirmar con un cierto nivel de confianza, que existe tal diferencia. Es por esto que es posible utilizar indicadores de performance y pruebas estadísticas que permitan la validación estadística que a su vez confirmen significativamente la ventaja en rentabilidades de un modelo sobre otro. Una prueba comúnmente utilizada es la *t* de Student, la cual en el campo de las finanzas e inversiones, utiliza como hipótesis nula que las medias de los retornos son iguales entre dos modelos y como hipótesis alternativa que un modelo tiene un retorno mayor en promedio al otro modelo.

3.6. Antecedentes Generales Momentum

Jegadeesh & Titman (1993) fueron los primeros en proponer el concepto de momentum en el área financiera, observando que no pocos inversionistas aplicaban una estrategia de invertir en activos financieros en función de su fortaleza con respecto al mercado, por lo que los autores en primer lugar tuvieron como objetivo el evaluar la rentabilidad de esta estrategia y en segundo lugar las razones que expliquen este fenómeno. Su primer objetivo se logró y evidenciaron la presencia de este efecto y que al seguir esta estrategia se generan retornos anormales[1] en el corto y mediano plazo. Esta consiste en una estrategia de inversión de mantener posiciones largas en una cartera de acciones que ha tenido mayor desempeño en el último periodo de tiempo, llamada en adelante cartera ganadora y posiciones cortas para aquellas acciones con menor rentabilidad, desde ahora cartera perdedora, ya que existe evidencia que estas acciones mantienen la tendencia que llevan. Su segundo objetivo de identificar las razones que explican el efecto no se cumplió, no pudiendo detectar las causas de su descubrimiento.

En el estudio de Jegadeesh y Titman, donde aplicaron la estrategia para un periodo desde 1927 hasta 1965 en el mercado de Estados Unidos se evidenció que las carteras ganadoras obtenían retornos altamente significativos, sobretodo en el primer año posterior a su formación. Sin embargo, evidenciaron que el retorno pierde su intensidad a partir del segundo año de mantención en aproximadamente la mitad de los retornos extraordinarios obtenidos en el primer año durante los dos años subsiguientes. Otro hallazgo de estos profesores norteamericanos,fue que este retorno no se veía reflejada compensación por un mayor riesgo sistémico ni una reacción retardada a factores que afectaban a todo el mercado sino una reacción tardía a la información particular a cada acción.

[1] Se entenderá por retornos anormales a aquellos retornos que no son capaces de explicarse por medio de valorización sistemática ya sea por ICAPM (Merton, 1973) o Modelos Multifactores de Fama & French (1998) entre otros.

3.6.1. Causas del momentum

No existe consenso sobres las fuentes del efecto de momentum, pero existen 2 corrientes teóricas que intentan identificar las razones del fenómeno. Una de estas teorías dentro de la literatura está ligada a las finanzas conductuales, la cual asocia el sesgo presente en los inversores como a la irracionalidad que provoca una ineficiencia en el mercado, explicando los retornos anormales. La segunda corriente hace relación que se produce el efecto del momentum por una compensación por riesgo no diversificable, sustentada en gran medida por la teoría tradicional financiera.

Lo primero que se viene a la mente para explicar el efecto momentum desde la perspectiva de la teoría tradicional financiera, es el modelo CAPM[2] (Sharpe et al,1964), un modelo de valoración del precio de activos financieros, siendo el más famoso del mundo por su sencillez y facilidad empírica permite estimar su rentabilidad esperada en función del riesgo sistemático. La formulación de este modelo viene dada por un mercado que tiene una competencia perfecta, existiendo homogeneidad de información para todos los inversores y que estos no pueden afectar el precio, además las personas deberían comportarse como seres racionales, es decir, en cada decisión de elección o consumo deberían maximizar sus propias funciones de utilidad y por ende su bienestar. Esto se lograría gracias a que se cuenta con toda la información requerida disponible, sin incurrir en costos para obtener esta última. Sin embargo, estas se sugieren como limitantes del modelo CAPM y de sus evoluciones como con el ICAPM (Merton,1973) que no permiten que se apeguen a la realidad y no han podido explicar las causas de los retornos sobre normales que conlleva seguir la estrategia de momentum.

Otros interesados fueron Fama & French(1996), destacados profesores e investigadores, que dentro de su modelo de tres factores analizaron el efecto momentum para EE.UU.

[2]Capital Asset Pricing Model

El modelo de los autores para la valorización de activos se basa en evaluar los retornos de las acciones tomando en consideración una prima por riesgo sistémico, una prima por tamaño de la empresa y una prima por valor (value). Este modelo ha sido capaz de explicar más del 90 % del retorno de las carteras diversificadas, sin embargo, los mismo autores afirman que su modelo no es capaz de explicar los retornos sobre normales causados por el efecto momentum, mas si corroboran su existencia, por lo cual sugieren que el momentum pudo deberse por 3 razones:

- Casualidad de una muestra en especifico, basado en el no cumplimiento de la estrategia aplicada en la década de los '30 en Estados Unidos de los resultados obtenidos por Jegadeesh & Titman (1993). Sugieren Fama & French ampliar la investigación a mercados internacionales y periodos de tiempo con la intención de robustecer el modelo.
- Plantean que una posible irracionalidad de los inversores en base a las reacciones que tengan estos antes las noticias sería la causante del momentum.
- Finalmente, plantean que el mercado sí es racional y que el no poder explicar los retornos anormales generados por el momentum es una limitación de su modelo de 3 factores, por lo cual potencialmente se podría corregir agregando factores al modelo[3] y que estos capturen la fuente de la anomalía.

Dentro de la otra corriente que busca explicar esta anomalía del mercado, correspondiente a las finanzas conductuales, siendo éstas un campo de las finanzas que nace de la unión de la psicología, la economía tradicional y la neuroeconomía. Esta rama analiza las finanzas desde un punto de vista psicológico. Describe cómo se comportan las personas y como éstas toman las decisiones, admitiendo que al momento de tomar decisiones hay sesgos emocionales y cognitivos. Pudiendo el juicio final del inversor verse alterado según

[3]Fama & French realizan evolución a este modelo y postulan el modelo de 5 factores y de multifactores

la forma en que se presenta la información , siendo esta una de las alternativas planteadas por Fama y French en 1996. Otras investigaciones importantes en este campo estuvo a cargo de Kahneman y Tversky en 1979 con avances sobre técnicas de psicología cognitiva, para explicar un cierto número de anomalías documentadas en la toma de decisiones económicas racionales. George & Hwang(2004) siguiendo esta corriente argumentan que cuando una buena/mala noticia ha empujado al precio de una acción cerca/lejos de el punto de referencia,como lo puede ser del precio más alto de una acción, los inversores son reacios a hacer una oferta por el precio más alto/bajo, aunque la información lo justificase. Pero con el tiempo los inversores corrigen este sesgo inicial sin tener una reacción exagerada para esto. Específicamente para el momentun proponen que se produce por este efecto de anclaje y ajuste.

3.6.2. Validación y evolución

A pesar de no existir consenso de la causa del momentum, éste desde la década de los ’90 ha sido ampliamente reconocido por inversores y académicos especialmente en los mercados desarrollados como lo es el de Estados Unidos (Jegadeesh & Titman, 1993) y los mismos autores han mostrado que no existe la evidencia para corroborar que el efecto no ha desaparecido. También se ha detectado este efecto en otros mercados desarrollados de Europa: Rouwenhorst (1998) propone evidencia que valida la existencia del momentum en 12 mercados desarrollados de Europa durante los períodos comprendidos entre 1980 y 1995, y está es ratificada en periodos posteriores: Hon y Tonks (2003) lo hace para el Reino Unido, Glaser y Weber para Alemania el 2001 y España el 2006. Además, Hameed y Kusnadi (2002) encuentra evidencia del efecto para 6 países asiáticos. Para mercados emergentes se ha evidenciado también la presencia del momentum con estudios de Rouwenhorst (1999), Van de Hart et al. (2003) y Griffin et al. (2003), los que concluyeron en general que a pesar de estar presente el efecto en estos mercados, lo está en menor intensidad que los mercados desarrollados. Sin embargo, estudios de Muga y Santamaría (2007) sugieren que con su evaluación del momentum, con una data más recientes que los tres estudios anteriores,

está presente en igual intensidad para mercados emergentes[4] como para los desarrollados. González (2006) evidenció la presencia de momentum en el mercado chileno, evidenciando que no es tan marcado como en países con economías consolidadas.

Una evolución de la estrategia de momentum es la estrategia de 52 Week High propuesta por George & Hwang(2004), en la cual los autores plantean que la proximidad al máximo de precio de 52 semanas domina los rendimientos pasados en términos de poder predictivo y explica en gran medida las ganancias de momentum.

3.7. Fundamentos estrategia 52 Week High

George & Hwang(2004) fueron los que propusieron la estrategia de 52 Week High, quienes en sus estudios plantean que los inversores poseen un sesgo a la hora de tomar decisiones financieras, ya sea reaccionando en forma tardía o bien en forma exagerada. Esto expone una irracionalidad que desafía a la teoría tradicional financiera que se fundamenta en la eficiencia de los mercados. Esto presenta una explicación a los retornos extraordinarios que se obtienen al seguir una estrategia de momentum basada en sesgo de anclaje y ajuste planteada por Kahneman, Slovic y Tversky (1982), los cuales consisten en un perjuicio cognitivo provocado por la tendencia humana a darle un mayor peso a la primera noticia, argumento o evidencia que se tiene sobre algo a la hora de tomar una decisión[5].

George y Hwang sugieren que una acción cuyo precio está en o cerca de su precio más alto en 52 semanas es una acción para la que han llegado recientemente buenas noticias y además debido a la cercanía del precio de una acción a su máximo de 52 semanas se encuentra entre la información que es más fácilmente disponible para los inversores, por lo cual, estos utilizan el precio máximo de 52 semanas como un punto de referencia con el cual evalúan el potencial impacto de las noticias e informaciones que reciben.

[4] Analizaron específicamente los mercados de Chile, Argentina, Brasil y México

[5] Ejemplo famoso de este sesgo es el propuesto por los autores ante la pregunta de cuántos países africanos hay en la ONU justamente después de lanzar una bola en un juego de ruleta

Cuando llegan buenas noticias que elevan el precio de una acción cerca o a un nuevo máximo en sus últimas 52 semanas, los inversores son reacios a comprar más acciones o subir más aún el precio de la acción, aún cuando la información justifique un alza, no obstante, está noticia eventualmente prevalece y el valor de la acción sigue subiendo. De igual forma, cuando se recibe una noticia negativa que empuja el precio de una acción lejos del precio máximo alcanzado en las últimas 52 semanas, los traders o inversionitas se muestran reacios a vender sus acciones a un precio considerado bajo aún cuando la noticia más bien lo valga. Con lo anterior se puede visualizar una tardía respuesta de los inversores ante las noticias mostrándose reacios a los ajustes dependiendo del nivel de precios cuando estos están en su punto más cerca o más lejos del máximo precio de la acción alcanzado en las últimas 52 semanas. Para los precios que no están cerca ni lejos del máximo de 52 semanas, los ajustes previos se realizan más rápidamente y no hay una predicción tan clara de los valores futuros al momento de llegar la información. Los resultados propuestos por George & Hwang son consistentes con la evidencia encontrada por Grinblatt y Keloharju (2001), los que encuentran efectos de nivel de precios que afectan a los inversores, evidenciando patrones en el mercado. Utilizando datos del mercado accionario finlandés, encuentran que es mucho más probable que los inversionistas vendan (en vez de mantener o comprar) una acción cuyo precio esté cerca de un máximo histórico, y que es más probable que compren (en vez de vender) una acción que está cerca de una bajo histórico.

En comparación con la estrategia tradicional de momentum, George & Hwang encontraron que la cercanía a la máxima de 52 semanas es un mejor predictor del rendimiento futuro que los retornos pasados.

3.8. Inteligencia Artificial

No existe una definición exacta para lo que corresponde a la Inteligencia Artificial (IA), sin embargo, una definición tradicional de IA hace referencia a la capacidad de una

máquina de poseer inteligencia comparable con la del ser humano. Discrepancias sobre *qué es la inteligencia* han dado lugar a muchas interrogantes a que si una máquina puede pensar y ser inteligente. Allan Turing, hace más de 3 décadas fue uno de los pioneros y padre de la computación que conocemos hoy en día y planteó el debate si es posible un pensamiento humano en un cerebro inerte. Hoy en día, con un enfoque moderno, vastos son los intentos de probar está hipótesis y han surgido distintas ramas para alcanzar la IA.

El porqué aplicar Inteligencia Artificial en inversiones, hace referencia a que la mayor parte de los modelos de predicción de series de tiempos (como la de las rentabilidades de las acciones), se basan en modelos lineales o casi lineales, sin embargo pueden existir patrones no lineales ocultos en las series de tiempo que no permiten un ajuste óptimo del modelo. Por lo cual, el uso de IA, por medio de auto aprendizaje con Redes Neuronales Artificiales, entre otras, pretenden capturar estos patrones no lineales y mediante un algoritmo inteligente, entregar predicciones de las rentabilidades que posean un mejor ajuste que con los modelos tradicionales, sin la necesidad de conocer previamente la estructura de datos con los que se alimenta.

3.9. Redes Neuronales Artificiales

Se comienza a desarrollar la inteligencia artificial desde la década de los '40 con la investigación propuesta por McCulloch y Walter Pitts, los cuales en 1943 mostraron que cualquier función lógica (AND, NOT, OR) podía ser calculada a partir de una red de neuronas interconectadas de manera sencilla, además propusieron que una red bien conectada tendría la capacidad de aprender.

Inspiradas en las redes neuronales de la biología humana, las Redes Neuronales Artificiales (RNA o ANN[6]) están tomando un creciente interés por parte de investigadores

[6] Abreviadas así por las siglas en inglés para Artificial Neutral Networks

que han desarrollado cada vez propuestas con RNA más complejas, completas y sofisticadas redes. Las RNA más detalladamente, corresponden a una técnica de Inteligencia Artificial, de tipo parámetrica. Estas redes intentan recrear el sistema de aprendizaje sobre estructuras de datos y patrones propio de humanos en computadores, tomando en consideración percepciones, estímulos externos y experiencias, con lo cual logran un aprendizaje adaptativo y un autoajuste, es decir, las RNA pueden modificar su comportamiento como respuesta a su entorno y poder producir respuestas consistentes a entradas y salidas que ha aprendido luego de un periodo de entrenamiento, lo cual las convierte en una herramienta totalmente flexible a la información con la cual se opera.

La implementación de la IA y RNA en el área financiera no está exenta, con lo que se ha visto transformaciones desde las finanzas tradicionales, con su fuerte base descriptiva, a las finanzas modernas, la cual posee un enfoque más aplicado y llevado a la práctica, tratando de explicar los fenómenos o anomalías financieras, como lo es la de 52 Week High.

3.9.1. Neurona Artificial

Las neuronas artificiales se basan en las biológicas, siendo estás últimas en la unidad nerviosa básica funcional y estructural del sistema nervioso, siendo una célula de especializada en captar los estímulos provenientes del ambiente y de transportar y transmitir impulsos nerviosos, conectándose con otras células en el proceso llamado sinapsis, para luego entregar una respuesta.

De forma análoga, la neurona artificial es la unidad básica del sistema que compone la RNA, en la cual se pueden diferenciar 3 tipos de roles de estás unidades:

- Neurona de entrada: encargada de recibir señales (de manera similir al trabajo de las dendrinas en la neurona biológica) y datos del entorno que son proporcionados a la red y que emiten la información a las unidades intermedias. El conjunto de neuronas o nodos de entrada conforman la capa de entrada (Inputs). El número de unidades de entrada dependerá del caso en estudio, ya que no existe algún método matemático o una regla general que permita encontrar la cantidad óptima a colocar.

- Neurona oculta: unidad que no tiene conexión directa con el entorno. Su labor es el procesar la información (soma de la neurona biológica), por medio de la función de activación y envía a las neuronas terminales o de salida los datos procesados. El conjunto de neuronas intermedias conforma una capa oculta, pudiendo existir más de una de estas capas en un sistema. A diferencia de la capa de entrada, el número de unidades en las capas ocultas y la cantidad de capas tienen relación con la complejidad y el grado de linealidad que presentan los datos, con lo cual al aumentar el número de capas ocultas, es posible capturar de mejor manera la componentes no lineales que presentan dichos datos. Sin embargo, no es posible tampoco conocer dicho limite de capas a colocar para encontrar una solución óptima y el agregar capas extras tiene un costo computacional elevado.

- Neurona terminal o de salida: tienen como misión entregar una solución y enviarla al entorno fuera de la red (análoga a los axones de la neurona biológica). El conjunto de unidades terminales se denomina capa de salida (Outputs).

En forma general, se conoce como capa nivel a un conjunto de neuronas o unidades cuyas entradas provienen de la misma fuente y cuyas salidas se dirigen al mismo destino. Las neuronas en una capa suelen ser del mismo tipo.

Se denomina arquitectura a la estructura a la constitución de distintas neuronas que forman un conjunto de capas, lo cual se denomina *red neuronal*, en la cual los nodos se conectan entre si por medio de sinapsis. Las conexiones sinápticas son direccionales, es decir, la información fluye en un solo sentido, siendo este desde una neurona presináptica a una

postsináptica.

Se puede visualizar un esquema de las neuronas y capas en la Figura 3.1

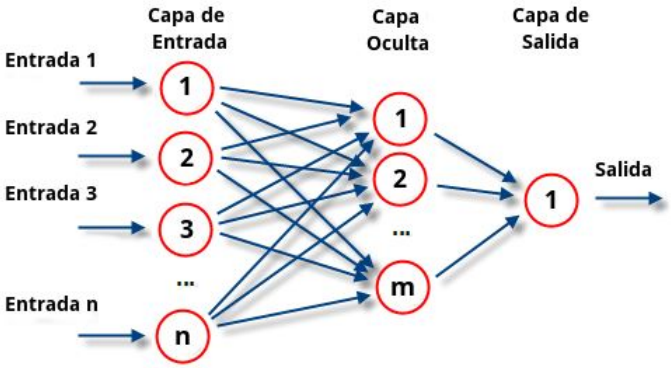


Figura 3.1: Estructura Red Neuronal Artificial
Fuente: https://victoryepes.blogs.upv.es

Dentro de otras características que poseen las RNA, se encuentra:

- Peso sináptico: corresponde a una magnitud que hace referencia a la fuerza de conexión entre 2 neuronas. La importancia de esta fuerza se debe a que algunas entradas pueden ser más relevantes que otras y el peso asignado funciona como una magnitud determinante para definir la excitación de la neurona. El peso, se representará en términos de vectores de la forma:

$$w_{ki} = (w_{k1}, w_{k2}, ..., w_{kn}) \tag{3.4}$$

Donde k corresponde a la *k-ésima* neurona y existen n señales de entrada como se ejemplifica en la Figura 3.1, por lo cual los pesos respectivos de las m entradas a la *k-ésima* neurona son de los asociados a los componentes del vector de la Ecuación 3.4.

La entrada total a la neurona *k-ésima* corresponde a la sumaproducto de todas las entradas que se dirigen a la neurona con sus pesos respectivos.

Si llamamos e_k a la entrada neta a la neurona k, el inputs neto a dicha unidad vendrá

dada por la expresión de la :

$$e_k = \sum_{i=1}^{n} w_{ki} \cdot x_i \tag{3.5}$$

Donde x_i corresponde a la entrada *i-ésima*, con $i = 1, ..., m$

- Estado de activación: corresponde a un valor numérico propio de cada neurona, el cual indica el comportamiento o estado de la neurona i en un tiempo determinado t. Se presenta con el vector $A_i(t)$ y puede tomar valores continuos o discretos. Lo más simple es considerar solo valores binarios 1 y 0, como *on / off* ó *activo / pasivo*.

- Función de activación o transferencia: a cada unidad se le asocia una función de transferencia que se encarga de transformar la entrada e_k en una salida. Existen varios tipos de funciones de transferencia, peor las más comunes son las siguientes:

  - Función lineal o identidad: Usada cuando se pretende una salida con valores iguales a los de entrada. Gráficamente es de la forma:

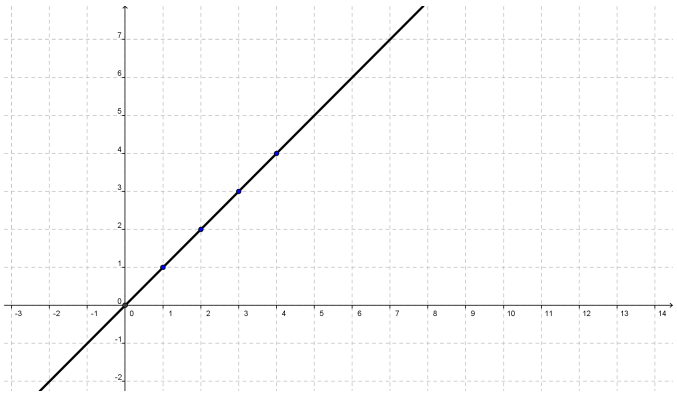


Figura 3.2: Función lineal
Fuente: https://http://recursostic.educacion.es

Y la expresión matemática que la representa es:

$$f(x) = x \tag{3.6}$$

- Función Escalón Unitario: Función no lineal de transferencia usada para el modelo original de una neurona de McCulloch and Pitts (1943). Se aplica para este tipo de modelos debido a que presenta una discontinuidad. Su gráfica la podemos observar en la Figura 3.3

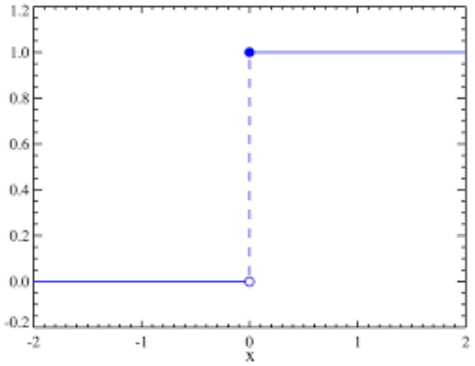


Figura 3.3: Función escalón unitario
Fuente: Matlab

Matemáticamente queda expresado como:

$$f(x) = \begin{cases} 1 & ; \quad x > 0 \\ 0 & ; \quad x \leq 0 \end{cases} \tag{3.7}$$

- Función Sigmoide: También llamada función logística, tiene como salida valores comprendidos entre 0 y 1. Es una de las más usadas en el campo de las redes neuronales para la predicción. Su representación gráfica es de la forma:

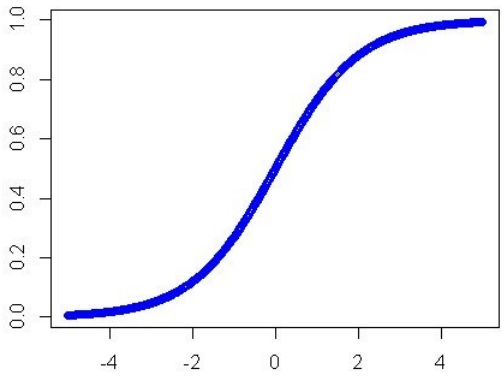


Figura 3.4: Función Sigmoide
Fuente: https://es.kisspng.com/

Con expresión matemática de la forma:

$$f(x) = \frac{1}{1 + e^{-x}} \tag{3.8}$$

- Función Tangente Hiperbólica: La forma de esta función es bastante similar a la sigmoide, sin embargo presenta como diferencia que las salidas toman valores entre -1 y 1.

  La gráfica de esta función corresponde a:

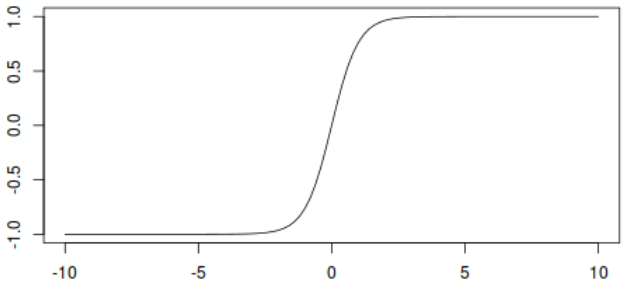


Figura 3.5: Función Tangente Hiperbólica
Fuente: https://www.researchgate.net/

Con expresión matemática de la forma:

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{3.9}$$

Con lo definido hasta ahora, podemos esquematizar a una neurona artificial de la siguiente manera:

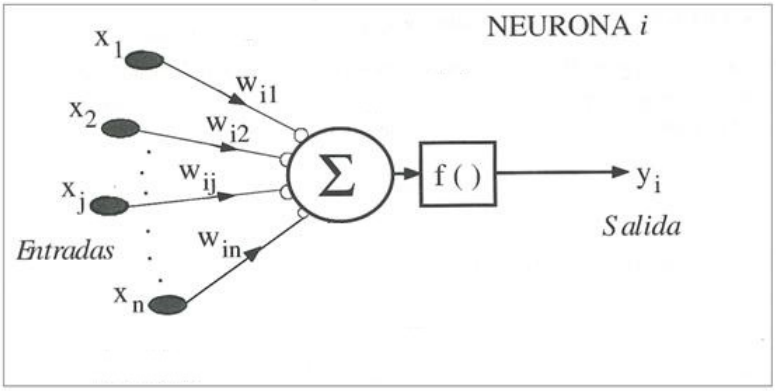


Figura 3.6: Estructura Neurona Artificial
Fuente: http://www.http://grupo.us.es/

3.9.2. Comportamiento de información de entrada y salida

De acuerdo al tipo de valores que se le entregan a la red neuronal desde al exterior a la capa de entrada y a la salida calculada por la RNA, podemos identificar 3 tipos de comportamientos:

- Continua: Es el más común de las redes, en la cual los datos de entrada y los de salida corresponden a valores reales continuos y generalmente se encuentran normalizados, es decir, su valor absoluto es menor a la unidad. Para las redes que aceptan este comportamiento, la función de activación de las neuronas serán también continuas, ya sea del tipo lineal, sigmoidal ó de tangente hiperbólica.

- Discreta: Estas redes solo aceptan como inputs solo valores discretos y binarios, lo cual genera que la salida computada sea del tipo [0,1]. La función de activación que

acepta estos valores discretos son del tipo escalón unitario.

- Híbrida: Mezcla del continuo y el discreto. La entrada a la red es del tipo continua y la respuesta es del tipo binaria.

3.9.3. Tipos de Redes Neuronales

3.9.3.1. Red neuronal monocapa

Esta red corresponde a la más sencilla, ya que solo contiene 2 tipos de capas de neuronas: de entrada, recibiendo los inputs y la capa de salida, donde se realizarán los distintos cálculos y entregará una solución.
La representación de este tipo de red es la que se muestra en la Figura 3.7 :

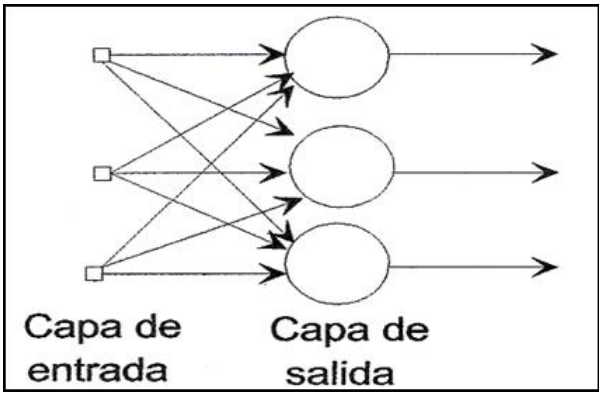


Figura 3.7: Red Neuronal Monocapa
Fuente: http://grupo.us.es/

3.9.3.2. Red neuronal multicapa

Su diferencia con la red neuronal monocapa, es que en esta existe un conjunto de capas intermedias (ocultas) entre la capa de entrada y salida, y es en estas capas ocultas donde se procesa la información y se realizan los cálculos.
Gráficamente es de la forma:

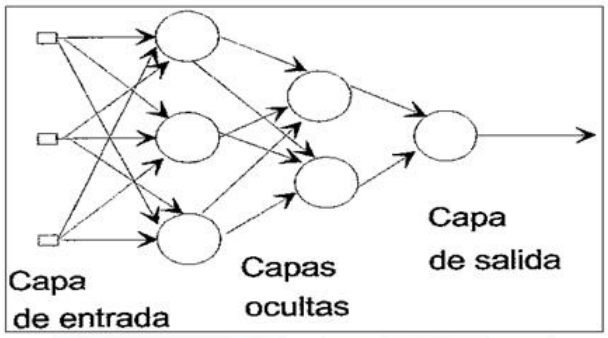


Figura 3.8: Red Neuronal Multicapa
Fuente:http://grupo.us.es/

3.9.3.3. Red neuronal recurrente

Tipo de red en la que su estructura permite la existencia de lazos de retroalimentación. Los lazos puede de ser de una neurona consigo misma, con otras neuronas de la misma capa o, incluso entre neuronas de distintas capas. La mayor utilidad de este tipo de red, es para estudiar la dinámica de los sistemas no lineales, ya que capta mejor esta componente.

Su gráfica asociada es:

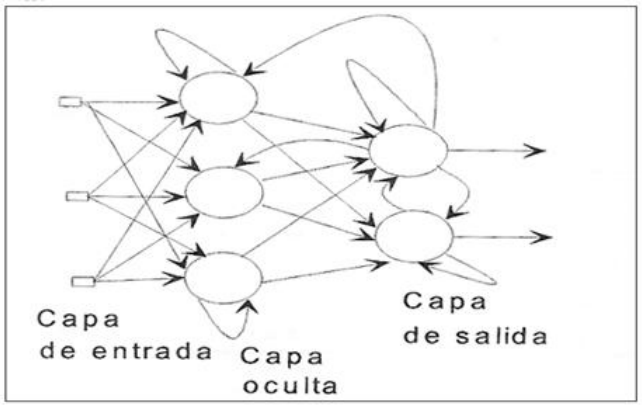


Figura 3.9: Red Neuronal Recurrente
Fuente: http://grupo.us.es/

3.9.4. Clasificación de conexión

Es posible identificar dos tipos de conexiones generales existentes entre las distintas neuronas y capas que conforman la red neuronal artificial:

- Feedforward: La información corre en una sola dirección y es hacia adelante, es decir, desde la capa de entrada a la capa de salida, pasando por la o las capas ocultas (si es que las hay). Es el tipo de conexión más simple, no existiendo ciclos de información entre neuronas ni capas. Es el tipo de conexión que utiliza la red neuronal monocapa y multicapa.
- Feedback: La información corre hacia atrás. Se forman bucles o ciclos, donde la información puede retroceder a una capa anterior. Este tipo de conexión es el utilizado por las redes neuronales recurrentes.

3.9.5. Aprendizaje de una red

El aprendizaje, en el contexto de redes neuronales, es posible definirlo como el proceso en el cual se genera una adecuación de los parámetros no rígidos de la red a partir de la estimulación que genera el entorno sobre la red. La intención del aprendizaje es que la red aprenda cuales son los parámetros que permitirán realizar correctamente el tipo de procesamiento deseado. En la gran mayoría de los casos de requerimiento de aprendizaje consiste en poder determinar cual es el conjunto de pesos sinápticos óptimos, ya que inicialmente, al poseer un sistema neuronal, que contenga la red neuronal con sus inputs, este sistema contiene una determinada estructura de red y tipo de neurona, sin embargo los pesos sinápticos en este entonces son desconocidos, por lo cual se les asigna un valor nulo o aleatorio.

Es posible identificar dos manera posibles para que el aprendizaje o entrenamiento se lleve a cabo:

1) Modelación de las sinapsis: Corresponde generalmente a una optimización de una función de error. Una regla común de aprendizaje, es aquella en la que se minimiza el error, con lo cual se va aumentando la eficacia de operación que posee la red neuronal, por lo que se va modificando los pesos sinápticos generalmente en un proceso iterativo.,

hasta lograr un error arbitrariamente aceptable.

Teniendo un algoritmo para el aprendizaje, donde sea $w_{ij}(t)$ el peso de conexión entre una neurona presináptica j con otra postsináptica i en la iteración en t, de acuerdo a la información nueva que ingresa en ese instante a la red procedente del entorno, se formará un diferencial $\delta w_{ij}(t)$, que sirve como parámetro corrector al peso sináptico inicial, con los cual se actualiza para una nueva iteración de la forma:

$$\delta w_{ij}(t+1) = w_{ij}(t) + \delta w_{ij}(t) \tag{3.10}$$

La actualización de los pesos sinápticos se realiza hasta lograr un error arbitrariamente aceptable, logrando de esta manera el rendimiento deseado.

Generalmente los algoritmos de aprendizaje se basan en métodos numéricos y al ser estos iterativos, pueden caer en algunos casos en problemas de convergencia, donde esta última corresponde a una comprobación al sistema neuronal para ver su capacidad de resolver el problema.

2) Creación o destrucción de neuronas: Consta de una modificación a la misma estructura de la red neuronal

Dentro de los algoritmos de aprendizaje, existen cuatro tipos que se pueden identificar:

- Supervisado
- No Supervisado
- Híbrido
- Reforzado

Los 2 primeros aprendizajes, supervisado y no supervisado, corresponden a los aprendizajes básicos y las otras dos reglas (híbrido y reforzado) son variaciones de los primeros.

3.9.5.1. Aprendizaje supervisado

Suele ser el más exacto en los resultados que obtiene, pero también el más complejo. Para este tipo de aprendizaje, se le entrega a la red un conjunto de patrones y señales como inputs, además de la salida deseada u objetivo denominada *target*. Se realiza una iteración de los pesos comparando la salida computada con el target y se ajustan los pesos hasta alcanzar un error aceptable arbitrario entre ambas salidas, de acuerdo a un parámetro de error entregado previamente.

3.9.5.2. Aprendizaje no supervisado

Se diferencia del aprendizaje supervisado, en que este no posee guía, es decir, no se entregan las salidas deseadas, por lo cual la red, por medio de esta rega de aprendizaje, estima los pesos sinápticos a partir de regularidades y patrones del conjunto de datos de entrada, con lo cual la salida computada busca que sea consistente y más similar según rasgos o agrupaciones de los datos de entrada.

3.9.5.3. Aprendizaje híbrido

Es una regla de aprendizaje presente en la red, en la cual se mezcla el aprendizaje supervisado y no supervisado, generalmente en distintas capas de neuronas de la red.

3.9.5.4. Aprendizaje reforzado

También corresponde a una mezcla del aprendizaje supervisado y no supervisado. En este tipo de aprendizaje, se le entrega información respecto a la salida, pero solamente mediante un indice global si se ha cometido error o no, indicando solamente si está actuando de buena manera o no, sin suministrar explícitamente la salida deseada.

En forma general,el error entre la salida deseada y la computada se calcula con la siguiente función, equivalente al error cuadrático medio:

$$E(\vec{w}) = \frac{1}{2}\sum_{d\in D}(\tau_d - O_d)^2 \tag{3.11}$$

Donde:

- $\vec{w}$: vector de pesos sinápticos
- τ_d : salida deseada o target en el ejemplo de entrenamiento d
- O_d: salida computada entregada por la red al aplicarle el entrenamiento d
- D : conjunto de ejemplos de entrenamiento
- d : ejemplo de entrenamiento puntual

Para que la red resulte operativa es necesario que aprenda mediante el entrenamiento luego de definir la arquitectura y tipo de aprendizaje. Esta etapa permite que la red neuronal mediante repetición, gane experiencia y pueda mejorar más. No existe modelo matemático que permita determinar el tiempo requerido de aprendizaje o entrenamiento óptimo y queda a criterio del usuario darle el tiempo o iteraciones que se crea mejor.

3.9.6. Backpropagation

Corresponde al algoritmo de entrenamiento de redes neuronales más utilizado, con el cual busca ajustar los pesos definitivos de las neuronas que ofrezcan mejores resultados. Es un algoritmo de aprendizaje supervisado el cual no solo considera que la salida de cada neurona dependa unicamente de las entradas que se le suministren a la red, sino que también depende del resto de las salidas que entreguen las demás neuronas. Por lo cual, el desajuste que cometa una neurona no solo hace referencia a que los pesos asignados a esa dicha neurona sean correctos, sino que también afectará el error acumulado que traigan el resto de las neuronas que precendan en la red a la de salida.

Para controlar el error que se va acumulando, se ajusta la función de error de la ecuación 3.11, para dejarlo como:

$$E(\vec{w}) = \frac{1}{2} \sum_{d \in D} \sum_{k \in salidas} (\tau_k d - O_k d)^2 \tag{3.12}$$

Con lo cual se agregan unos nuevos parámetros:

- Salidas: conjunto de neuronas de salida (están en la cada de salida)
- k: corresponde a la k-ésima neurona de salida

Y se redefinen otros parámetros:

- τ_d : salida deseada o target en el ejemplo de entrenamiento d que debería dar la neurona k
- O_d: salida computada entregada por la red al aplicarle el entrenamiento d calculada por la neurona k

4 | Metodología

4.1. Datos

La data utilizada en el estudio corresponde a los precios mensuales de cierre de las 40 principales acciones de la bolsa del mercado accionario chileno, correspondiente al IPSA. La data abarca los periodos comprendidos desde enero del 2010 hasta marzo del 2018, abarcando una cantidad de 3960 datos.

4.2. Procedimiento

A continuación se presentan los modelos y procedimiento para aplicar la estrategia de inversión de *Momentum* y *52 Week High* para igual horizonte de evaluación, de tal manera de ser consistente los resultados para la comparación. Los intervalos de mantención a comparar son de 3, 6, 9, 12 y 24 meses para ambas estrategias.

De manera adicional a las estrategias de inversión mencionadas, la red neuronal artificial será utilizada como una herramienta de proyección de rentabilidades para cada acción, con lo cual el criterio de ranking quedará dado por aquellas acciones que tengan mayor retorno.

4.2.1. Posiciones financieras

Se pueden identificar tres posiciones financieras generales a las cuales los inversionistas deben tomar postura al momento de buscar alternativas de inversión:

- No invertir: Corresponde a la alternativa más simple, la cual descarta la opción de buscar rentar con los activos financieros en el momento actual. No quiere decir que "no se haga nada", ya que se le puede hacer un seguimiento para una inversión en un periodo futuro.
- Posición larga: Es aquella en donde el inversor espera que un valor o activo financiero crezca en el tiempo, por lo cual toma una posición de compra para adquirir el valor y obtener beneficios al momento de que este valor o activo aumente su precio.
- Posición corta: Contraria a la posición larga, corresponde a aquella estrategia en la cual el inversor tiene la creencia de que un valor o activo financiera caerá en su valor. Para llevarla a cabo, se realiza una venta en corta, es decir, se venderán las acciones que se crean que bajarán, por haber rentado menos en comparación a las otras, sin tener las acciones en el momento necesariamente. Al momento de bajar el precio, lo cual no necesariamente ocurre, se compra la acción a este menor precio y se obtiene beneficios de la diferencia entre la compra y la venta.

4.2.2. Aplicación estrategia momentum

Se conforman carteras de inversión según lo planteado por el momentun tradicional de Jegadeesh & Titman (1993) y de 52 Week High de George & Hwang (2004). Para la primera estrategia se realiza un ranking de las rentabilidades acumuladas durante X meses correspondientes a un período de evaluación, y para el tercio de acciones de mayor rentabilidad se procederá a tomar una posición larga para este portfolio de acciones consideradas ganadoras, en donde para cada acción, se considera que se debe invertir con igual peso en cada una. Para esta cartera conformada con las ganadoras, se procede a mantenerlas por Y meses, correspondientes al periodo de mantención. Las acciones perdedoras por su parte conformarán la cartera perdedora, con la cual se tomará una posición en corto.

Aspecto importante a considerar es que se utiliza un mes de desfase entre el período X y el Y con tal de evitar presiones en el precio y reacciones retardadas encontradas por Jegadeesh. Además, se utilizarán carteras superpuestas mes a mes como se visualiza a continuación en un ejemplo de estrategia momentum de 3 meses de formación y 3 de mantención (3x3):

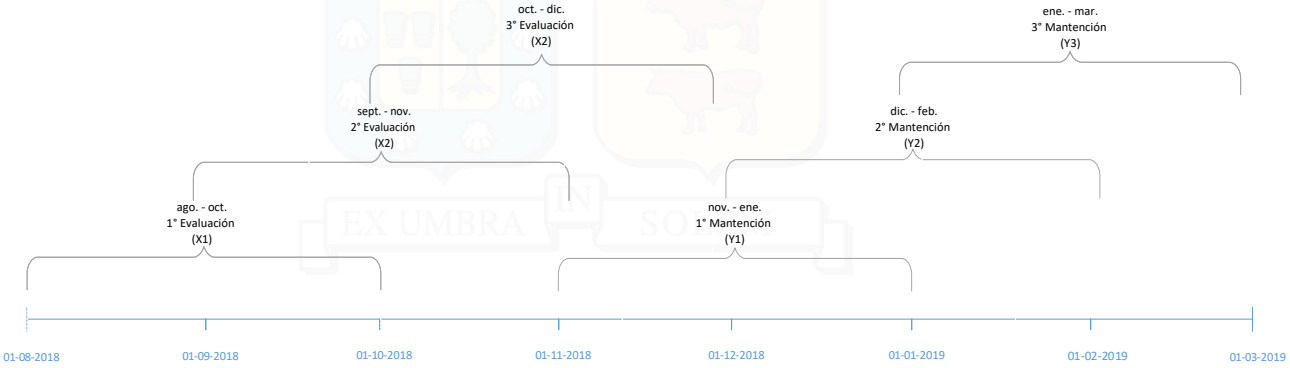


Figura 4.1: Ejemplo Estrategia Momentum 3x3
Fuente: Elaboración Propia

La rentabilidad es calculada de manera logarítmica, es decir se considera continua. La razón de esto es debido al mejor ajuste que tiene a una distribución normal de los datos, a diferencia de un cálculo de rentabilidad discreto, como el porcentual. Por lo cual, el cálculo de la rentabilidad queda dado como:

$$\text{Rentabilidad} = R_t = \ln(\text{Precio acción}_t) - \ln(\text{Precio acción}_{t-1}) \tag{4.1}$$

La rentabilidad de cada acción i en el mes t para el ejemplo de estrategia 3x3, se calcula como:

$$R_i, t = \sum_{t=1}^{X} R_i, t \quad \forall i \tag{4.2}$$

En donde X $\in 1, 2, 3$ correspondientes a los primeros 3 meses dados por el periodo de formación. Luego, los R_i, t se proceden a ser rankeados según mayor rentabilidad y luego se arma una cartera con el tercio de acciones que más rentan en este intervalo de tiempo,

conformado por los X meses de formación. Para estas acciones del portfolio ganador se calcula la rentabilidad de cada acción *i* que la conforma considerando los Y meses de mantención, de la forma:

$$R_i, t = \sum_{t=1}^{Y} R_i, t \quad \forall i \tag{4.3}$$

En donde Y $\in 4, 5, 6$ lo cual entrega la rentabilidad de cada acción *i* al tomar una posición larga de la cartera conformada con las acciones ganadoras y una posición corta con las perdedoras.

4.2.3. Aplicación estrategia 52 Week High

La estrategia de 52 Week High, se basa en el mismo principio de impulso que la estrategia de momentum, evaluando las acciones en un periodo de formación, correspondiente a 52 semanas e invirtiendo durante un periodo de mantención las que tengan mejor performance y tomando posición corta con las de peor rendimiento. Sin embargo, el criterio para rankear las acciones, en cada periodo, no será la rentabilidad media, sino que se considerará un ratio que toma en consideración el precio de la acción en un periodo t y el precio más alto de la misma acción en un periodo de 52 semanas (12 meses) hasta el periodo t.

En forma matemática, la expresión que representa a este índice de performance es:

$$\frac{P_{i,t}}{Max_{i,t}} \tag{4.4}$$

Donde $P_{i,t}$ corresponde al precio de la acción *i* en el periodo *t* y $Max_{i,t}$ es el precio de la acción *i* durante el periodo de 12 meses que terminan en el instante *t*.

4.2.4. Estructura de la Red Neuronal Artificial

Cada red neuronal artificial posee una arquitectura la cual conlleva a como se modelará la red y el rendimiento que esta obtendrá. En la arquitectura de la red se deben definir ciertos parámetros dentro de los que se encuentran generalmente el número de neuronas por capas, cantidad de capas intermedias u ocultas, pesos sinápticos iniciales, entre otros. Para esta investigación, los parámetros de interés que se tomarán en cuenta serán la cantidad de capas ocultas, el número de capas que contenga cada capa, los inputs que se le entregan a la red neuronal y las salidas deseadas (target).

El modelo propuesto corresponde a una red neuronal multicapa feedforward, el cual poseerá un aprendizaje del tipo supervisado, es decir, se le guiará a la red y tendrá como parámetro la salida deseada o target. Además, se alimenta la red en base al principio de ventana móvil o corrediza, la cual en forma dinámica cada vez que predice un periodo, 'avanza' un mes en su data para ajustarse temporalmente a los periodos pronosticados.

El entrenamiento utilizado se lleva a cabo mediante un algoritmo backporpagation, el cual no debe ser excesivo, ya que al entrenar una red neuronal con unos cuantos ejemplos e intentar minimizar el error a unas cuotas ínfimas, se corre en el riesgo de especializar demasiado la red neuronal, la cual se comportará de una manera óptima con los ejemplos con los cuales ha sido entrenada, pero para ejemplos con los que no se ha entrenado, se pueden producir errores considerables, ya que la especialización conlleva una pérdida de generalización, por lo cual a caídas de precios anómalas, como lo puede ser una crisis financiera, la red puede no saber como responder ante eso.

4.2.5. Implementación RNA

Se proponen 3 modelos de redes neuronales, probando distinta información suministrada a la red en los inputs y target. Los inputs de cada modelo, corresponden a rentabilidades rezagadas y se evalúan con 3, 6, 9 y 12 meses de información pasada para poder generar una proyección, lo cual es un simil al periodo formación o evaluación de la estrategia de momentum. Además, se consideran para todos los modelos propuestos, periodos

de 3, 6, 9 y 12 meses para intervalos de mantención, con los cuales se calcula la rentabilidad acumulada, estos intervalos de formación y mantención darán lugar a 16 estrategias posibles: $3x3, 3x6, 3x9, 3x12, 6x3, 6x6, 6x9, 6x12, 9x3, 9x6, 9x9, 9x12, 12x3, 12x6, 12x9$ y $12x12$. El algoritmo de la red neuronal se aplica para las 40 acciones que componen el IPSA de manera independiente, utilizando la información de las rentabilidades logarítmicas de cada una para predecir la salida deseada.

La salida computada por la red, se utilizará como criterio de selección para escoger las acciones con mejor proyección de performance y que se espera que obtengan mayor retorno de acuerdo al intervalo de mantención. Por lo cual, la respuesta que entregue la red, es ordenada en un ranking y se procede a formar portfolio con las acciones ganadoras del percentil del tercio superior, con la cual se tomará una posición larga y otra cartera perdedora, con el tercio de acciones de peor desempeño, para realizar venta en corto, al igual que la estrategia de momentum.

Además, cada uno de los modelos, se prueba con 1 y 4 capas ocultas y con 5, 10 y 15 neuronas cada una.

El criterio para escoger el mejor modelo y configuración de los 3 propuestos, es la que obtenga en promedio la mayor rentabilidad final para un horizonte de evaluación común, comparando mediante Indice de Sharpe y VaR. Además, se comparan los 3 modelos de redes neuronales con la estrategia de momentum tradicional, la estrategia de 52 Week High y con la rentabilidad pronosticada mediante el modelo ARIMA. Son pronosticados valores para 24 meses comunes por cada red y modelo ARIMA, además de utilizar esos mismos 24 meses para aplicar la estrategia convencional de momentum y 52 Week High. A continuación se detallan los modelos propuestos de las RNA:

4.2.5.1. Modelo 1: Rentabilidad de cada periodo

Para cada mercado y acción se procede a ingresar a la red los inputs correspondientes a las series de rentabilidades logarítmica rezagadas y se utiliza un 70 % de la data para entrenar la red. El target suministrado a la RNA consiste en la rentabilidad logarítmica del

periodo $t + 1$ de la acción respectiva.
El resto de la data es utilizada para testear un aprendizaje correcto y luego se procede a utilizar las rentabilidades proyectadas entregadas por la red para conformar un ranking de desempeño, ordenando las acciones de manera descendente, utilizándolas como indicador para la estrategia de momentum.

A modo de ejemplo, suponiendo una cantidad de 3 periodos o meses suministrados (rezago) de información a la red en forma de inputs, se visualiza en la Tabla 4.1 la salida deseada de la red y los inputs:

Empresa A	Target	Inputs Empresa A		
Periodo	t+1	t	t-1	t-2
1	-0,0945	0,1301	0,1521	0,1276
2	0,1121	-0,0945	0,1301	0,1521
3	0,1521	0,1121	-0,0945	0,1301
4	0,1347	0,1521	0,1121	-0,0945
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
n	0,1433	0,1964	-0,1234	0,1783

Tabla 4.1: Ejemplo para modelo 1 de entrada la rentabilidad de una empresa y salida la rentabilidad del período siguiente de la misma empresa.
Fuente: Elaboración propia

4.2.5.2. Modelo 2: Rentabilidad acumulada

El procedimiento es similar al del modelo 1, sin embargo, para esta segunda propuesta de modelo, se ingresa como target la rentabilidad acumulada de la acción de una empresa para un intervalo de mantención dado. Las salida entregada por la red, correspondiente a la serie de rentabilidades proyectadas acumuladas y al igual que en el modelo anterior, se ordenan en forma descendente para utilizarlas como criterio de selección para clasificar las acciones ganadoras y perdedoras en cada periodo. Luego, se procede a conformar las carteras de acuerdo a la estrategia de momentum.

Utilizando los datos del ejemplo del modelo 1 y una estrategia 3x3, obtendríamos una Tabla como la 4.2.

Empresa A	Target	Inputs Empresa A		
Periodo	t+3	t	t-1	t-2
1	0,1697	0,1301	0,1521	0,1276
2	0,2332*	-0,0945	0,1301	0,1521
3	0,0071*	0,1121	-0,0945	0,1301
4	0,2002*	0,1521	0,1121	-0,0945
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
n	0,5328*	0,1964	-0,1234	0,1783

Tabla 4.2: Ejemplo para modelo 2 de entradas la rentabilidad logarítmica de una empresa y salida la rentabilidad logarítmica del periodo siguiente para la misma empresa.
Fuente: Elaboración propia

El Target del periodo 1, al considerar como instante *t* al periodo actual , se calcula con la suma de las rentabilidades en *(t+1) + (t+2) + (t+3)*, correspondiente a -0,0945; 0,1121; 0,1521 respectivamente, siendo estas las rentabilidades logarítmicas futuras, de acuerdo a la ecuación 4.3.

(*) Los valores mostrados del target no pueden ser calculados solo con la información proporcionada en la tabla, ya que son necesario más datos de las rentabilidades logarítmicas para calcular la rentabilidad acumulada en el intervalo de mantención (3 meses en el ejemplo dado).

4.2.5.3. Modelo 3: Clasificación ganadoras

Este modelo propone proporcionar como target de la RNA, la matriz correspondiente a las acciones que han estado dentro del tercio de mejor desempeño, según su rentabilidad acumulada para el intervalo de mantención. La matriz target se conforma de forma binaria, donde los 1 representan que la empresa está, en el periodo analizado, dentro del tercio ganador con mejor rendimiento y 0 en caso contrario.

Por el tipo de target entregado, el problema se convierte en uno de clasificación, en el cual se espera que la red proporcione como respuesta si es conveniente o no invertir en una acción. No obstante, la salida computada no necesariamente corresponde a una respuesta binaria, más aún con una data "pequeña"que no permite que la red aprenda completamente que el target sea también binario. Por lo tanto, el output entregado por la red, se utiliza al igual que en los modelos propuestos anteriores, como indicador de desempeño, en donde las acciones que tengan en cada periodo mayor valor en este output, serán consideradas como las ganadoras.

Además, se propone para este modelo, suministrar a la red en forma de inputs, la rentabilidad acumulada del periodo de mantención, de manera de tener mayor relación con el target, ya que la decisión de invertir viene dada por el portafolio que logró mayor retorno acumulado en cada una de sus acciones para un intervalo de tiempo.

Siguiendo contexto y ejemplos anteriores, la representación de este modelo se visualiza en las Tabla 4.3

Empresa A	Target	Inputs Empresa A		
Periodo	t+1	t	t-1	t-2
1	0	0,1301	0,1521	0,1276
2	1	-0,0945	0,1301	0,1521
3	1	0,1121	-0,0945	0,1301
4	0	0,1521	0,1121	-0,0945
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
n	1	0,1964	-0,1234	0,1783

Tabla 4.3: Ejemplo para modelo 3 de entradas considerando la rentabilidad acumulada de la acción y target la decisión en forma binaria si es que correspondió a una acción en la se debía invertir.
Fuente: Elaboración propia

Para que cada modelo funcione, se genera un algoritmo computacional que se divide en 3 partes:

- Preprocesamiento: sección inicial que utiliza la data de los precios de las acciones para crear matrices correspondientes a los inputs y target que alimentan a la red neuronal.
- Predicción RNA: sección en la cual se ejecuta la red neuronal artificial sensibilizando cantidad de capas y neuronas ocultas. La salida computada corresponde a la predicción para periodos siguientes y es la utilizada como el criterio de ranking del principio de impulso o momentum.
- Procesamiento: Parte final del algoritmo que se alimenta de los resultados de la red neuronal y elabora las carteras de inversión, entregando además la rentabilidad promedio mensual para cada estrategia (3x3, 3x6, 6x9, etc).

Además de escoger cual de los tres modelos de RNA propuestos es el que logra mejores resultados, se hace necesario realizar una comparación con otras metodologías y estrategias de inversión, ya que es esta comparación la que finalmente le interesa a un inversor si le conviene o no utilizar la metodología con redes neuronales por sobre otros modelos más convencionales. Es por esto, que la comparación de los modelos de redes se realiza con la estrategia tradicional de momentum de Jegadeesh & Titman, 52 Week High de George & Hwang y con el retorno que se logra con el modelo econométrico ARIMA, con el cual se proyectan las rentabilidades para el mismo horizonte de tiempo que las redes y se utilizan dichas rentabilidades como criterio de selección para conformar las carteras ganadoras y perdedoras.

Para hallar el mejor modelo ARIMA, se trabaja suponiendo la serie de rentabilidades como de orden 0 y bajo el criterio del akaike se buscan los mejores valores para p y q correspondientes al orden de los modelos autorregresivos y media móvil respectivamente. El modelo ARIMA, al no necesitar periodo de formación, es comparado considerando solo el periodo de mantención.

5 | Resultados

Siguiendo la metodología señalada, se ejecuta el algoritmo para los 3 modelos propuestos de redes neuronales, en los que se espera encontrar la configuración de capas y neuronas que entreguen mejores resultados y la estrategia recomendada para el inversor.

Se aplica un test de medias para la serie de rentabilidades mensuales proyectadas para cada modelo y estrategia para cada una de sus 6 configuraciones posibles de capas ocultas y neuronas: 1 capa/5 neuronas; 1 capa/10 neuronas; 1 capa/15 neuronas; 4 capas/5 neuronas; 4 capas/10 neuronas; 4 capas/15 neuronas.

El criterio para escoger la mejor configuración fue la de que obtuviera mayor cantidad de estrategias de las 16 posibles con un retorno medio superior con una significancia del 5 % sobre las otras.
Los resultados obtenidos dan como mejor configuración 1 capa oculta (más la capa de entrada y salida) y 5 neuronas por capa para los tres modelos propuestos.

Escogida la configuración óptima, se extraen los retornos promedios mensuales para cada estrategia. Se visualizan las rentabilidades logradas para cada modelo propuesto en las gráficas de las figuras 5.1, 5.2 y 5.3:

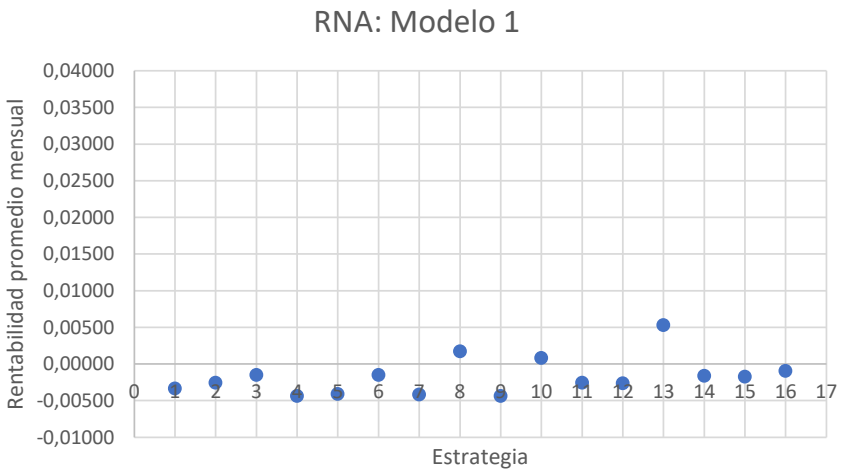


Figura 5.1: Rentabilidad promedio mensual para cada estrategia del modelo 1

Las coordenadas del eje de las abscisas corresponden a las distintas combinaciones estrategias formadas con los 3, 6, 9 y 12 periodos de formación y mantención, siendo 1 = 3x3, 2 = 3x6, 3 = 3x9, 4 = 3x12, 5 = 6x3, 6 = 6x6, 7 = 6x9, 8 = 6x12, 9 = 9x3, 10 = 9x6, 11 = 9x9, 12 = 9x12, 13 = 12x3, 14 = 12x6, 15 = 12x9 y 16 = 12x12

Descata del gráfico del Modelo 1 que existe poca variabilidad en los retornos que consigue cada estrategia, sin embargo, las rentabilidades tampoco son altas, siendo las más bajas de los 3 modelos.

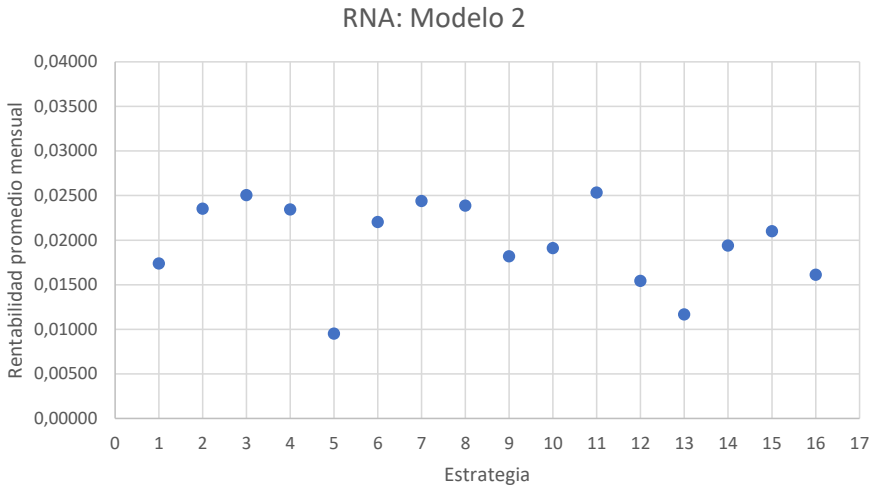


Figura 5.2: Rentabilidad promedio mensual para cada estrategia del modelo 2

Es posible observar que existe una mayor dispersión en las rentabilidades por es-

trategia en este modelo, sin seguir una tendencia clara. Por otro lado, se ve que alcanza rentabilidades más altas que el modelo 1, pero levemente menores que el modelo 3.

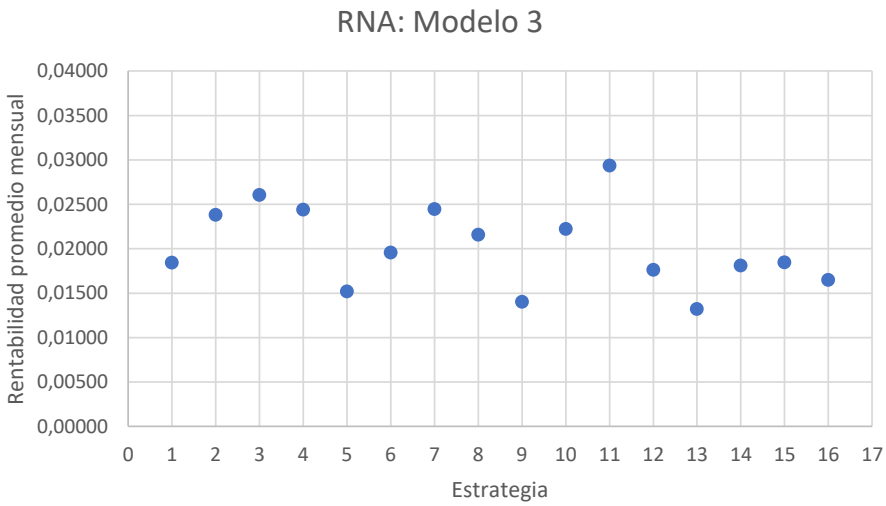


Figura 5.3: Rentabilidad promedio mensual para cada estrategia del modelo 3

El Modelo 3, tiene cierta similitud al del modelo 2, presentando mayor variabilidad que el modelo 1, sin embargo, a diferencia del modelo 1, logra retornos mayores que este último. También es posible notar que los retornos son levemente superiores que los modelos anteriores.

Para una comparación general entre los 3 modelos, podemos observar la gráfica 5.4 y la Tabla 5.1 en la cual se agregan además las rentabilidades que son capaces de lograrse con el momentum tradicional, el modelo econométrico ARIMA y la metodología propuesta de George & Hwang de 52 Week High:

Figura 5.4: Rentabilidad promedio mensual para cada estrategia de los modelos propuestos

Se desprende de la gráfica de los retornos, que la rentabilidad del Modelo 1 es menor en todas las estrategias en comparación a los otros dos modelos. Además, se visualiza que para las estrategias con período de mantención de 3 y 12 meses existe un menor nivel de retorno en forma general por el modelo 2 y 3, y con intervalo de mantención de 6 y 9 meses se registran las mayores rentabilidades. Además se visualiza que el máximo en rentabilidad se logra con la estrategia 9x9 del modelo 3 y que para el resto de las estrategia existe cierta paridad entre el modelo 2 y 3.

Ahora, al incluir al Momentum tradicional, ARIMA y 52 Week High a la comparación, dado que los valores de la tabla 5.1 corresponden todos a retornos mensuales, son comparables entre distintas estrategias, a pesar de poseer distinto intervalo de mantención, y según lo anterior, es posible observar que el mayor retorno logrado es de 0,02937 correspondiente a la estrategia 9x9 del Modelo 3.

Retorno	Modelo 1	Modelo 2	Modelo 3	Momentum	ARIMA	52WH**
3x3	-0,00333	0,01736	0,01846	0,00997	0,00204	-
3x6	-0,00253	0,02352	0,02385	0,01134	0,00046	-
3x9	-0,00147	0,02504	0,02606	0,01329	-0,00005	-
3x12	-0,00434	0,02344	0,02443	0,01186	0,00026	-
6x3	-0,00408	0,00952	0,01521	0,01335	0,00204	-
6x6	-0,00150	0,02203	0,01959	0,01827	0,00046	-
6x9	-0,00414	0,02438	0,02448	0,01855	-0,00005	-
6x12	0,00172	0,02385	0,02159	0,01681	0,00026	-
9x3	-0,00437	0,01819	0,01406	0,02402	0,00204	-
9x6	0,00083	0,01908	0,02226	0,02412	0,00046	-
9x9	-0,00257	0,02531	**0,02937**	0,02397	-0,00005	-
9x12	-0,00261	0,01542	0,01766	0,02275	0,00026	-
12x3	0,00531	0,01167	0,01323	0,02387	0,00204	0,02581
12x6	-0,00161	0,01939	0,01815	0,02458	0,00046	0,02752
12x9	-0,00174	0,02098	0,01848	0,02366	-0,00005	0,02450
12x12	-0,00094	0,01612	0,01651	0,02286	0,00026	0,02212

Tabla 5.1: Rentabilidad promedio mensual para cada modelo y estrategia calculado con las series proyectadas

(**) La estrategia de 52 Week High (52WH) se compara solo con periodo de formación de 12 meses, para ser consistente con lo planteado con George & Hwang

Es posible además notar de la tabla que los retornos promedio mensuales de los modelos 2 y 3 son claramente superiores que los del modelo 1, que a su vez se asemeja bastante a los conseguidos con el modelo econométrico ARIMA. En cuanto a la superioridad en retorno del modelo 2 y 3 con los modelos mencionados, no ocurre tan notoriamente al comparar estos modelos con los retornos del momentum y 52 Week High.

No obstante, dado que es necesaria una prueba estadística que valide un retorno promedio mensual sea significativamente mayor, se utiliza una prueba t de student con un nivel de confianza de un 95 %. De lo anterior resultó que los modelos 2 y 3 logran en promedio un mayor retorno que el Modelo 1 y ARIMA con valor t de $5,92104$ y $5,98090$ respectivamente para el modelo 2 y valor t de $6,0127$ y $6,0183$ para el modelo 3 con un 5 % de significancia, con lo cual se rechaza la hipótesis nula que plantea que las medias son

iguales y se acepta la hipótesis alternativa de que la media es significativamente mayor. Al comparar el Modelo 2 con el 3, se obtiene un valor t de $1,42402$ y un valor crítico de $1,67247$, con lo cual se acepta la hipótesis nula y en promedio, el retorno promedio del modelo 3 no es mayor que el que se puede logra en promedio con el modelo 2. En otros resultados, se acepta la hipótesis nula al comparar el modelo 2 con el momentum, no pudiendo ganarle en rentabilidad, cosa que si logra el modelo 3 con un valor t de 1,76532 y valo crítico de 1,68322. Finalmente, ninguno de los modelos de redes propuestos, pudo ganarle en rentabilidad en promedio a la estrategia de 52 Week High.

De un análisis estadístico descriptivo de la rentabilidad promedio mensual global inter-estrategia aplicado a los modelos de redes y a los de comparación presentados en la Tabla 5.2, es posible extraer que el Modelo 3 es el que logra en promedio mayor retorno en sus estrategias al igual que mediana, sin embargo, es el modelo 2 el que presenta menor variabilidad, obteniendo la menor desviación estándar. Otro punto a destacar, es que el Momentum y 52 Week High son las que presentan mayor curtosis, es decir, poseen el mayor grado de concentración de los retornos en la región central de su distribución, pudiendo considerarse la más estable entorno a la media, en lo cual superan a los modelos propuestos.

	Media	**Mediana**	**Desv estándar**	**Curtosis**
Modelo 1	0,00531	0,00833	0,01642	0,19729
Modeo 2	0,02531	0,02449	0,00924	-0,71234
Modelo 3	0,02937	0,03015	0,01047	0,38037
Momentum	0,02458	0,02671	0,01243	2,96209
ARIMA	0,00204	-0,00225	0,01553	-0,56677
52 WH**	0,02752	0,02877	0,01035	1,61241

Tabla 5.2: Estadística descriptiva de los resultados de los 3 modelos propuestos de redes neuronales y de las 3 estrategias alternativas convencionales de inversión

Los resultados mostrados en la Tabla 5.2 son en base a la rentabilidad promedio mensual obtenida para cada estrategia en cada modelo, con lo cual se puede considerar como llevar a cabo, en forma paralela, una inversión en las estrategias 3x3, 3x6, 3x9, 3x12,

6x3, 6x6, 6x9, 6x12, 9x3, 9x6, 9x9, 9x12, 12x3, 12x6, 12x9 y 12x12.

Volviendo a los resultados individuales, se realiza un ranking por modelo con la estrategia que logró mayor retorno, quedando de la siguiente manera:

Ranking	**Modelo**	**Estrategia**	Indicador: **Retorno**
1	Modelo 3	9x9	0,02937
2	52 Week High	12x3	0,02752
3	Modelo 2	9x9	0,02531
4	Momentum	12x6	0,02458
5	Modelo 1	9x6	0,00531
6	ARIMA	9x3	0,00204

Tabla 5.3: Ranking modelos según retorno promedio mensual

Además aplicando test de medias, esta vez a las estrategias ganadoras de cada modelo entre si, se obtiene que la rentabilidad promedio mensual con un 5 % de significancia del modelo 3 es mayor que la mejor estrategia de todos los otros modelos, a excepción del 52 Week High, con un valor *t* de 6,7035 con el modelo econométrico ARIMA y valor *t* de 1,70854 con el modelo 2, siendo estos con el que tiene mayor y menor diferencia estadística respectivamente, rechazando la hipótesis nula de rentabilidades iguales y aceptando la hipótesis alternativa. La estrategia del Modelo 2 por su parte, rechaza la hipótesis nula para su prueba con el Modelo 1 y Arima y acepta para el resto. La estrategia del Modelo 1 en tanto solo rechaza la hipótesis nula en su comparación de medias con el modelo ARIMA con un valor *t* de 1,83076 con un 95 % de nivel de confianza y valor crítico de 1,67236. Finalmente, los retornos de todas las estrategias, a excepción del modelo ARIMA, superan al retorno promedio mensual del mercado que es de un 0,00313.

El riesgo financiero por su parte, medido con el Valor en Riego (VaR) histórico, arroja que con un 95 % de confianza que el mejor escenario es con el Modelo 3 utilizando la estrategia 3x9 con un VaR de 0,01636, con la cual con un 5 % de probabilidad la rentabilidad será a lo más de un 0,01636 y de forma similar, que con un 95 % de confianza, el retorno esperado sea mayor al VaR.

El segundo mejor valor encontrado para el VaR es para la estrategia 12x9 correspondiente a la estrategia de 52 Week High, la cual había obtenido el segundo mayor retorno promedio como se observó anteriormente, con un Valor de Riesgo de un 0,01436 al mismo nivel de confianza y horizonte de 24 meses.

El mejor VaR para el Modelo 2 se encuentra en la estrategia 3x9 y es de 0,01312, el cual está por sobre el mejor VaR de momentum, Modelo 1 y ARIMA respectivamente, lo cual indica que los modelos con redes neuronales 2 y tienen menor riesgo financiero menor que el momentum.

VaR	Modelo 1	Modelo 2	Modelo 3	Momentum	ARIMA	52WH**
3x3	-0,02908	-0,00996	-0,00973	-0,03047	-0,01699	-
3x6	-0,02409	0,01110	0,00901	-0,00810	-0,02026	-
3x9	-0,01677	0,01312	**0,01636**	-0,00033	-0,01459	-
3x12	-0,02013	0,00934	0,01210	-0,00101	-0,01582	-
6x3	-0,03457	-0,02126	-0,01496	-0,02359	-0,01699	-
6x6	-0,01917	-0,00298	0,00477	-0,00007	-0,02026	-
6x9	-0,02012	0,00911	0,01384	0,00552	-0,01459	-
6x12	-0,01034	0,01035	0,01046	0,00279	-0,01582	-
9x3	-0,03665	-0,00793	-0,00459	-0,01287	-0,01699	-
9x6	-0,01496	0,00292	0,00185	0,00686	-0,02026	-
9x9	-0,02026	0,00869	0,01253	0,00919	-0,01459	-
9x12	-0,01741	-0,01174	-0,00273	0,00682	-0,01582	-
12x3	-0,02733	-0,01318	-0,01356	-0,01828	-0,01699	0,00668
12x6	-0,01912	0,00400	-0,00172	0,00241	-0,02026	0,01436
12x9	-0,01411	0,00818	-0,00017	0,00613	-0,01459	0,01650
12x12	-0,02020	-0,00053	-0,00330	0,00539	-0,01582	0,01342

Tabla 5.4: VaR para cada modelo y estrategia calculado con la serie de rentabilidades proyectadas utilizando un nivel de confianza del 95 %

Con lo cual, el ranking de estrategias por modelo según VaR es:

Ranking	**Modelo**	**Estrategia**	Indicador: **VaR**
1	Modelo 3	3x9	0,01636
2	52 Week High	12x9	0,01436
3	Modelo 2	3x9	0,01312
4	Momentum	9x9	0,00919
5	Modelo 1	6x12	-0,01034
6	ARIMA	9x3	-0,1459

Tabla 5.5: Ranking modelos según VaR Histórico

Utilizando uno de los indicadores más famosos del mundo, el Indice de Sharpe, ocupando una tasa de libre de riesgo constante igual a 0,165 % mensual (2 % anual) para todas las estrategias y modelos, se realiza en primer lugar la comparación entre los resultados de los algoritmos de las redes neuronales, en donde es posible apreciar en el gráfico de la figura 5.5 que para un intervalo de periodo de mantención de 3 meses y 12 meses, es el modelo 3 el cual posee un mayor del Indice de Sharpe, el cual mantiene casi en la misma proporción para los tres modelos en estos intervalos de tiempo. Desde un periodo de mantención de 6 y 9 meses el Modelo 3 presenta un notorio aumento en el ratio llegando a alcanzar un máximo absoluto.
Por su parte el Modelo 1 sigue la misma tendencia de tener el peor desempeño que los otros propuestos, teniendo un Indice menor al Modelo 2 y 3 en todos los periodos de mantención.

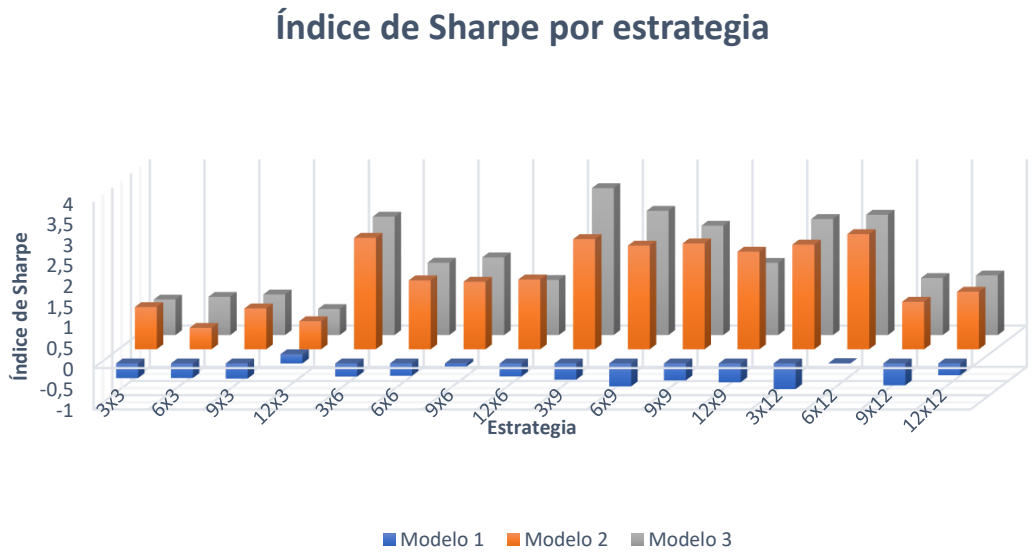


Figura 5.5: Gráfico comparativo Indice de Sharpe por estrategia entre los modelos de redes neuronales propuestos

Dado que se visualiza tramos marcados para los distintos periodos de mantención, es que se realiza un agrupamiento de las estrategias que posean el mismo intervalo de tiempo en el cual se conservan las acciones.

Se puede distinguir en la gráfica 5.6 que no existe un dominio claro en promedio de los Indices de Sharpe en las estrategias del Modelo 2 con el 3 para todos los intervalos

de mantención, al igual de lo visualizado a simple vista para el análisis individual de estrategias. Para el periodo incial de análisis, de 3 meses de mantención, el indicador resulta claramente menor para todos los modelos y por el contrario, para el intervalo de mantención de 9 meses, se observa el mayor Indice de Sharpe global considerando la agrupación de las estrategias por tramos y además se visualiza la mayor diferencia entre modelos 2 y 3 que en cualquier otro tramo, sin embargo está sigue sin ser amplia. El modelo 1 en ningún tramo se acerca a los otros modelos.

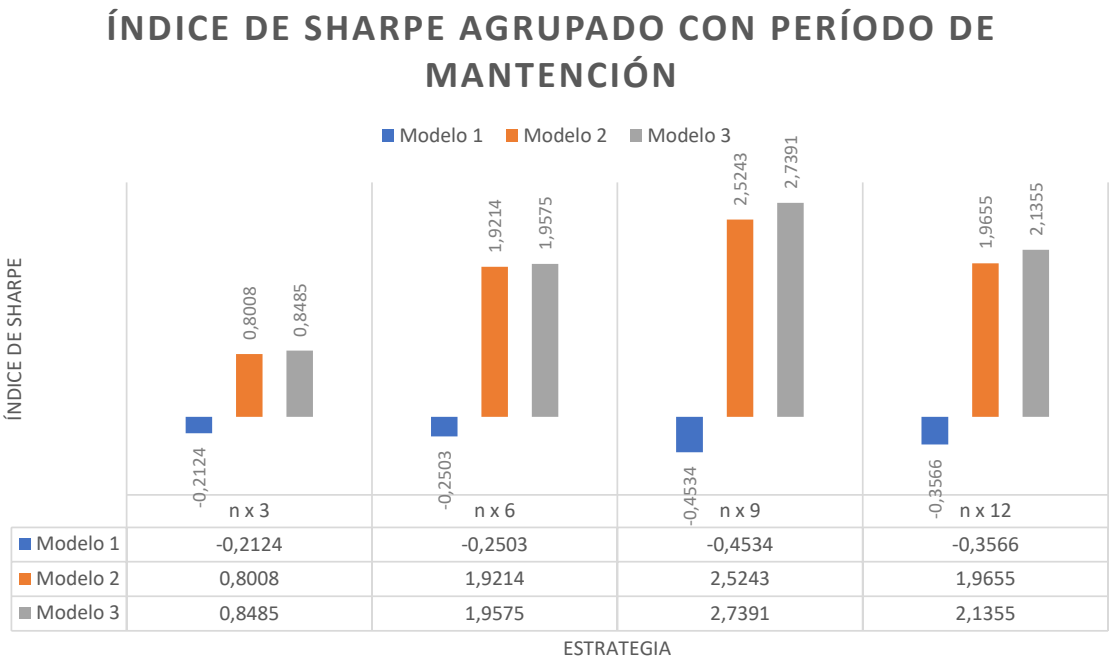


Figura 5.6: Gráfico comparativo Indice de Sharpe por estrategia, agrupando estrategias con período de mantención común entre los modelos de redes neuronales propuestos. Con *n* tomando los periodos de 3, 6, 9 y 12 meses de formación

Además, de manera análoga al retorno, realizando un análisis individual por estrategia y por modelo, se compara también con los Indice de Sharpe que obtienen las estrategias de comparación utilizadas anteriormente. Los valores obtenidos, con los cuales se construyeron las gráficas anteriores, se observan en la Tabla 5.6:

I Sharpe	Modelo 1	Modelo 2	Modelo 3	Momentum	ARIMA	52WH**
3x3	-0,35437	1,02044	0,85867	0,39729	0,02522	-
3x6	-0,31745	2,69669	2,86515	0,80398	-0,09891	-
3x9	-0,39246	2,66532	**3,55435**	1,35562	-0,19865	-
3x12	-0,62079	2,53277	2,80885	1,24721	-0,17001	-
6x3	-0,35282	0,51674	0,92469	0,59633	0,01996	-
6x6	-0,29794	1,66720	1,75176	1,61500	-0,11582	-
6x9	-0,55516	2,50849	3,00704	2,44083	-0,24646	-
6x12	0,00864	2,78640	2,91138	2,27571	-0,20883	-
9x3	-0,36521	0,98788	0,98118	1,16518	0,02041	-
9x6	-0,07066	1,63353	1,87952	2,37401	-0,12589	-
9x9	-0,40957	2,56106	2,64733	2,77624	-0,21223	-
9x12	-0,52975	1,14945	1,37959	2,67662	-0,17648	-
12x3	0,22262	0,67825	0,62962	1,12476	0,01983	2,39185
12x6	-0,31518	1,68805	1,33356	1,84427	-0,09584	2,49982
12x9	-0,45655	2,36249	1,74787	2,19784	-0,17035	2,33130
12x12	-0,28434	1,39327	1,44221	2,45952	-0,16135	2,71858

Tabla 5.6: Indice de Sharpe para cada modelo y estrategia calculado con la serie de rentabilidades proyectadas

En primer lugar, lo más destacado es que el mayor valor de Indice de Sharpe que se alcanzó, fue el de un 3,55435 el cual correspondió a la estrategia 3x9 del Modelo 3, obteniendo el mejor resultado, al igual que con el VaR. Esto es posible a que pese no tener la mayor rentabilidad (la cual se encuentra en otra estrategia del Modelo 3), posee menor volatilidad en su serie de retornos mensuales (intra-estrategia) que aquella estrategia 9x9 del mismo modelo. Los valores obtenidos del Indice de Sharpe, son bastante elevados en comparación con el -0,15712 correspondiente al del mercado, invirtiendo en igual proporción en todas las acciones que componen el IPSA para mismo periodo y tasa libre de riesgo, lo cual indica buen desempeño de los modelos.

A modo más general, se visualiza una gran paridad entre los índices de los modelos 2, 3, Momentum y 52 Week high para casi todas las estrategias, quedando relegadas las estrategias del modelo 1 y ARIMA, las cuales entre si tienen resultados similares. Otro aspecto notorio, es que los mayores valores de Sharpe para todos los modelos de redes neuronales se dan para mayores periodos de mantención (9 y 12 meses, en especial este último).

Con lo cual, el ranking de estrategias por modelo con este ratio es:

Ranking	**Modelo**	**Estrategia**	Indicador: **I Sharpe**
1	Modelo 3	3x9	3,55435
2	Modelo 2	6x12	2,78640
3	Momentum	9x9	2,77624
4	52 Week High	12x12	2,71858
5	Modelo 1	9x9	0,22262
6	ARIMA	9x3	0,0255

Tabla 5.7: Ranking modelos según Indice de Sharpe

Otra manera que se puede analizar según Indice de Sharpe, es el considerar un solo indicador global por modelo, para aquellos inversores que prefieran llevar a cabo más de un tipo de estrategias de manera paralela, donde los resultados se observan a continuación:

I Sharpe	**Modelo 1**	**Modelo 2**	**Modelo 3**	**Momentum**	**ARIMA**	**52WH****
Global	-0,99192	3,80468	4,68417	4,55736	-0,08906	3,16837

Tabla 5.8: Indice de Sharpe global para cada modelo considerando la serie de rentabilidad promedio de cada mes al usar todas las estrategias en forma simultanea

De la tabla se desprende que a nivel global, es decir, considerando todas las estrategias de un modelo, sacando el promedio de rentabilidad para cada mes de las rentabilidades promedio mensual por estrategia, el Modelo 2 es el que mayor índice alcanza.

6 | Conclusión y Discusión

Los inversionistas se enfrentan a la disyuntiva de en qué invertir y en qué momento deben hacerlo y para ello se enfrentan a la necesidad de guiarse por alguna estrategia de inversión que satisfaga sus requerimientos de nivel de retorno y tolerancia al riesgo que están dispuestos a mantener.
Es en este contexto en el que herramientas o modelos de pronóstico enfocados al área financiera son de gran ayuda y juegan un rol que cada vez es más primordial en las decisiones de los inversores.

Desde ya hace un par de décadas que existe la estrategia de inversión basa en el impulso o momentum propuesta por Jegadeesh y Titman, sin embargo, al igual que otros aspectos, es necesario adaptar dichas estrategias utilizando conocimientos y tecnologías que van a la vanguardia que se acoplen mejor a la actualidad y permitan lograr mejores resultados.
Es por lo anterior, que se propone utilizar Redes Neuronales Artificiales (RNA) para generar un modelo que siga el principio del momentum y logre potenciar sus resultados de modo de hacerlo más llamativo para un inversionista en un corto-mediano plazo.

Utilizando una ventana móvil que se ajuste de mejor manera a los cambios en los datos y generar un modelo dinámico, para predecir 24 periodos, en meses, se plantean 3 modelos que consideran RNA: Un primer modelo que pronostica la rentabilidad mensual de un periodo siguiente, un segundo modelo que tiene como salida la proyección de rentabilidad acumulada para un periodo de mantención dado y un tercer modelo que tiene

como objetivo entregar directamente las acciones en las cuales se debería invertir. Las redes de los dos primeros modelos propuestos se alimentan con la rentabilidad logarítmica de cada acción en el análisis y el tercer modelo tiene como inputs la rentabilidad acumulada del periodo de mantención, calculada de igual forma que la planteada por Jegadeesh y Titman. Las salidas computadas por cada red, son utilizadas como criterio de selección para conformar carteras de inversión con acciones del IPSA en las cuales se debe tomar posición larga llamadas ganadoras y en cuales se deberia tomar posición corta, llamada cartera perdedora, según tengan mayor o menor desempeño según el criterio utilizado.
Cada modelo por su parte se aplica para 16 estrategias correspondientes a utilizar 3, 6, 9 o 12 meses en forma de rezagos con los que se alimentan a la red como inputs y que cumplen análogo rol que el periodo de formación que el momentum. Por su parte se mantienen las carteras ganadoras y perdedoras por periodos de mantención de 3, 6, 9 y 12 meses, generando así 16 posibles combinaciones junto a los periodos de rezago, dando lugar a las 16 estrategias mencionadas.
Los resultados de los modelos propuestos son comparados con la estrategia original de momentum, con la estrategia de inversión de 52 Week High planteada por George & Hwang basada de igual manera en el principio de impulso y con el modelo econométrico ARIMA.

Los resultados obtenidos muestran que con 1 capa oculta y 5 neuronas es la configuración que mayor cantidad de estrategias con rentabilidad promedio se logran versus otras configuraciones analizadas.
En cuanto a los pronósticos de los modelos propuestos, el modelo 3 proyecta la rentabilidad acumulada para el periodo de mantención logrando el mayor retorno promedio mensual en su estrategia que considera 9 meses de rezagos de inputs en la red y 9 meses de mantención (9x9), alcanzando una rentabilidad promedio mensual de 0,02937 para el horizonte de tiempo de 24 meses analizados. Las rentabilidades mensuales conseguidas por esta estrategia del modelo 3 son estadísticamente mayor al resto de los modelos propuestos, salvo al 52 Week High y superan en gran medida al momentum, y al modelo econométrico. Sin embargo, esta mejor estrategia 9x9, presenta un mayor variabilidad que otras estrategias del mismo modelo y que otros, lo cual se ve reflejado en el Indice de Sharpe de 2,64733 para la estrategia 9x3, tomando en consideración el riesgo en forma de la desviación estándar

de las rentabilidades mensuales. Este ratio está por debajo de lo que se consigue con el mismo tercer modelo propuesto, que busca predecir las acciones en las cuales se debería invertir, el cual en su estrategia 3x9, logra un Indice de Sharpe de un 3,55435, superando enormemente al Indide de Sharpe del mercado igual a 0,04814 y pese a que tiene una rentabilidad promedio igual a un 0,02606, siendo menor que la de la estrategia 9x9 sigue estando elevada en comparación a la tasa libre de riesgo del 2 % anual utilizada, y por otro lado, supera estadísticamente al momentum y a todas las otras mejores estrategias de cada modelo (con excepción de la estrategia 12x6 del 52 Week High). Analizando un tercer indicador, se utiliza el Value at Risk (VaR) de un 5 % de significancia, para medir el riesgo financiero de las estrategias de los distintos modelos. Los resultados que se obtuvieron en este caso fue que la estrategia 3x9 del modelo 3 vuelve a tener el mejor desempeño con un VaR de 0,01636 y en segundo lugar se encuentra la estrategia 12x9 del 52 Week High con un 0,01436 teniendo esta última, mayor riesgo financiero. El modelo 1 por su parte, no obtiene resultados concluyentes para una estrategia que destaque sobre otra, ya que no repite ninguna estrategia con mejores resultados y que repita una buena performance en los indicadores utilizados, obteniendo para todos los análisis, resultados cercanos al modelo econométrico ARIMA

Para un análisis más global de las estrategias, el cual se puede aplicar invirtiendo de manera paralela en todas las acciones de un modelo, el retorno promedio mensual más alto lo consigue el modelo 3 seguido del momentum, modelo 2, 52 Week High y finalmente ARIMA , misma situación al considerar un Indice de Sharpe global por modelo.

Con los resultados obtenidos, y recalcando que cada inversionistas tiene sus preferencias en cuanto al riesgo considerado tolerable para explicar el retorno, a nivel individual, escogiendo una estrategia vencedora absoluta, la ganadora resultaría el modelo 3 con la estrategia 3x9, por ser la que alcanza un alto nivel de rentabilidad y un riesgo considerablemente bajo en comparación con el resto.

A nivel más global considerando el total de estrategias, los modelos propuestos no superan considerablemente al 52 Week High, pero el modelo 3 sigue logrando el mejor desempeño, sin embargo, cabe destacar que en la realidad se incurre en costos transaccionales y complejidad de manejar el portfolio mucho mayores, lo cual puede tener alteraciones.

Finalmente, 2 de los 3 modelos propuestos en base a algoritmos de redes neuronales artificiales han probado tener un buen desempeño en el pronóstico y complementarse efectivamente mostrando sinergia junto al momentum, logrando capturar de buena manera el principio de impulso que lleva esta estrategia.

Desafío para investigaciones a futuro es utilizar estos resultados y agregar costos transaccionales y operacionales al manejar distintas carteras de inversión cada mes. Además, el optimizar los tiempos de ejecución de los algoritmos de redes neuronales de manera de disminuir el costo computacional y poseer resultados más oportunos que no tarden demasiadas horas en obtenerse.

7 | Bibliografia

Van de Hart, J., Slagter, E., & Van Dijk, D. (2003). Stock selection strategies in emerging. *Journal of Empirical Finance*, 105-132.

Fama, E., & French, K. (1996). Multifactor Explanations of Asset Pricing Anomalies. *The Journal of Finance*, 55-84.

George, T., & Hwang, C.-Y. (2004). The 52-Week High and Momentum Investing. *The Journal of Finance*, 2145-2176.

Griffin, J., Ji, X., & Martin, J. (2003). Momentum investing and business cycle risk: evidence.*The Journal of Finance* , 2515-2547.

Grinblatt, M., & Keloharju, M. (2001). What makes investors trade?.*Journal of Finance*, 589–616.

Hameed, A., & Kusnadi, Y. (2002). Momentum Strategies: Evidence from Pacific Basin Stock. *Journal of Financial Research*, 383-397.

Hon, M., & Tonks, I. (2003). Momentum in the United Kingdom Stock Market. *Journal of Multinational Financial Management*, 43-70.

Jegadeesh, N., & Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *The Journal of Finance*, 65-91.

Kahneman, D., Tvesky, A., & Slovic, P. (1982). Judgment under Uncertainty: Heuristics. (Cambridge University Press, New York).

Marin, G., & Marin, J. (2001). Economía financiera . Barcelona: Antoni Bosch.

Merton, R. (1973). An Intertemporal Capital Asset Pricing Model.*Econometrica* ,

867-887.

Muga, L., & Santamaria, R. (2007). The momentum effect in Latin American emerging. *Emerging Markets Finance and Trade*, 24-45.

González (2006). Estrategia contrarian y momentum: Evidencia en Chile. *Revista Estudios de Administración Vol. 13 Nº 1, pp 67-*.

Rouwenhorst, K. G. (1998). International momentum strategies. *The Journal of Finance*, 267-284.

Rouwenhorst, K. G. (1999). Local return factors and turnover in emerging stocks markets. *The Journal of Finance*, 1439-1464.

Sharpe, W. (1964). Capital Asset Prices: A Theory of market Equilibirum under Conditions of Risk.*The Journal of Finance* , 424-442.