

Universidad de Buenos Aires
Facultad de Ciencias Económicas
Escuela de Negocios y Administración Pública

**MAESTRÍA EN GESTIÓN ECONÓMICA Y
FINANCIERA DE RIESGOS**

TRABAJO FINAL DE MAESTRÍA

**Análisis comparativo de modelos paramétricos y no paramétricos para la
predicción de valores futuros de acciones financieras de distintos sectores de la
economía Estadounidense durante el período 2019-2020**

AUTOR: ECONOMISTA IGOR BELISARIO

DIRECTOR: PROF. MAG. EMANUEL J. AGÚ

FEBRERO 2023

A mis hermanos Isaac e Iván, mis padres Magaly y Argenis, a mi esposa, Moirenys Vásquez.

Resumen

En el presente trabajo se realiza un análisis comparativo del poder predictivo del modelo Movimiento Browniano Generalizado en contraposición a dos modelos de aprendizaje automático, llamados “potenciación del gradiente extremo” (XGBoosting) y “redes neuronales” (NN) aplicados a los precios de acciones financieras. Para ello, en una primera instancia se desarrolla una vista general del comportamiento de los mercados financieros, de los modelos aplicados en finanzas, de su importancia en el análisis de riesgo de mercado y se da una muestra de la disciplina del aprendizaje automático. Luego, cada uno de los modelos se emplean para estudiar 15 acciones de la economía estadounidense en el periodo 2019-2020 y evalúa el beneficio de incorporar nuevos modelos al análisis de riesgo de mercado.

Abstract

This study deployed a comparative analysis of the financial stock price prediction power between Generalized Brownian Motion and two machine learning models called Extreme Gradient Boosting and Neural Networks. In order to show different performances firstly developed financial markets general insight, the models applied in finance, and is given a brief of the machine learning techniques. Finally, each model is implemented using as input 15 different stocks from the American market between 2019 and 2020 to evaluate the individual performance and compare each other.

Índice

Resumen.....	2
Introducción	6
Capítulo 1: Mercados financieros: el paradigma del aprendizaje automático.....	9
1.1. Comportamientos de los mercados financiero.....	9
Capítulo 2: Modelos: aspectos teóricos	14
2.1. Riesgo de mercado.....	14
2.2. Modelo de Valoración de Activos Financieros (CAPM).....	14
2.3. Proceso estocástico.....	15
2.3.1. Movimiento Browniano Generalizado	16
2.3.2. Lema Ito de dimensión 1.....	16
2.4. Modelos de aprendizaje automático	17
2.4.1. Error cuadrático medio residual	17
2.4.2. Tasa de aprendizaje	17
2.4.3. Regularización	17
2.4.4. Función de pérdida	17
2.4.5. Hiperparámetros.....	18
2.4.6. Sobreajuste (overfitting).....	18
2.4.7. Grupo de entrenamiento y de prueba.....	18
2.5. Extreme Gradient Boosting (XGBoosting).....	18
2.6. Red neuronal multicapas (<i>Multilayer Neural Network</i>).....	20
2.6.1. Proceso de aprendizaje.....	21
2.6.2. Propagación hacia adelante y hacia atrás.....	21
2.6.3. Función de activación	22
2.6.4. Capas de una red neuronal	22
Capítulo 3: Análisis de resultados.....	24
3.1. Datos	24
3.2. Resumen estadístico	26
3.3. Movimiento Browniano Generalizado (MBG)	26
3.3.1. Propiedad lognormal	27

3.3.2. Simulación de Montecarlo	28
3.4. Modelos de Aprendizaje Automático	30
3.4.1. Características extraídas de los datos	30
3.4.2. Extreme Gradient Boosting (XGBoosting/XGB)	31
3.4.3. Redes Neuronales (RN)	33
3.5. Comparación entre modelos	35
Capítulo IV: Conclusiones y futuras líneas de investigación	39
Referencias.....	43
Anexos	48
1. Código Python – Movimiento Browniano Generalizado	48
2. Código Python – Extreme Gradient Boosting	48
3. Código Python – Redes Neuronales.....	48

Introducción

Se tiene conocimiento que la primera bolsa de valores se creó en el siglo XVII en Ámsterdam en el año 1602 (Stringham y Curott 2015) y que un siglo más tarde, fue fundado el primer mercado de futuros en Japón en 1730 (Dojima Rice Exchange - World's First Futures Exchange. s.f.). En ambos casos su constitución fue generada por la necesidad de crear un intercambio de instrumentos diferentes a los bienes de consumo. En los mercados financieros, la dinámica está basada en la interacción entre quienes requieren capital y quienes poseen sobrante. Esta interacción empieza en la negociación y culmina cuando se llega a un acuerdo e intercambian temporalmente capacidad de dinero por rendimiento. Estos instrumentos utilizados para el intercambio en su mayoría son ofrecidos en mercados primarios y secundarios, tales como la bolsa de valores o en mercados *over-the-counter*. Los instrumentos que son comercializados en mercados secundarios cobran un interés particular sobre los otros, debido a la búsqueda de beneficios a través de estrategias de arbitraje o de la cobertura.

De ahí surge la idea de encontrar algún modelo que permita reconocer patrones en los mercados financieros con el fin de generar beneficios monetarios, lo cual implica que la participación de carteras de inversión incurra en riesgo de mercado. En este ámbito existen posiciones diferentes acerca de la posibilidad de encontrar dicho modelo que pueda anticiparse al mercado. Siguiendo a (Eugene Fama 1970), Premio Nobel en Economía en 2013, publicó la teoría de los “mercados eficientes” postulando que los mercados incorporan rápidamente nueva información sobre el desempeño de una empresa para asignar precios a las acciones y esto deja sin efecto cualquier intento de análisis de precios basado en series históricas. Por lo tanto, las fluctuaciones de precios

de acciones no pueden modelarse como un proceso determinístico, sino como uno cuyos movimientos sean aleatorios, lo cual es representado como un paseo aleatorio (*random walk*) (Fama, 1965).

Contrario a la posición de Fama, se encuentran perspectivas del análisis de precio como la aplicada por Jim Simmons y James Ax, dos matemáticos creadores de técnicas de análisis de patrones de movimiento de los precios de las acciones, que fundaron en 1988 una firma de inversiones llamada *Renaissance Technologies*, que ha aplicado algoritmos de análisis de precios utilizando datos de la Reserva Federal y del Banco Mundial desde el año 1700 (Mondragón 2021). Así, lograron a través de una técnica llamada *Trading de Alta Frecuencia* (HFT) aplicar este tipo de algoritmos y convertir a *Renaissance Technologies* en una de las empresas de inversión más exitosas de los últimos años. Por lo tanto, es necesario preguntarse cuál debe ser la correcta estrategia para un analista de riesgo para analizar los movimientos de su cartera y así disminuir el riesgo de mercado. Si bien los resultados y las teorías más aceptadas e incluso enseñadas en maestrías como la de este tipo, se encuentran del lado del tipo de modelo propuesto por Fama, la actualidad muestra cada vez con mayor preponderancia la facilidad de la aplicación de modelos más del estilo algorítmico. Si bien estos modelos no son nuevos, por ejemplo, la invención de las redes neuronales se remonta a 50 años atrás, el avance de la computación ha generado que sean más fácil su aplicación por la capacidad de cómputo de los ordenadores (Patterson y Gibson 2017). Esta revolución del mundo computacional ligado al financiero ha llegado incluso a la computación cuántica, por ejemplo (Rebentrost, Gupt y Bromley 2018) presentaron un estudio sobre el cálculo del precio de un derivado financiero usando la computación cuántica

empleando algoritmos cuánticos. Lo que indica que la tendencia no se detiene y se debe indagar más al respecto.

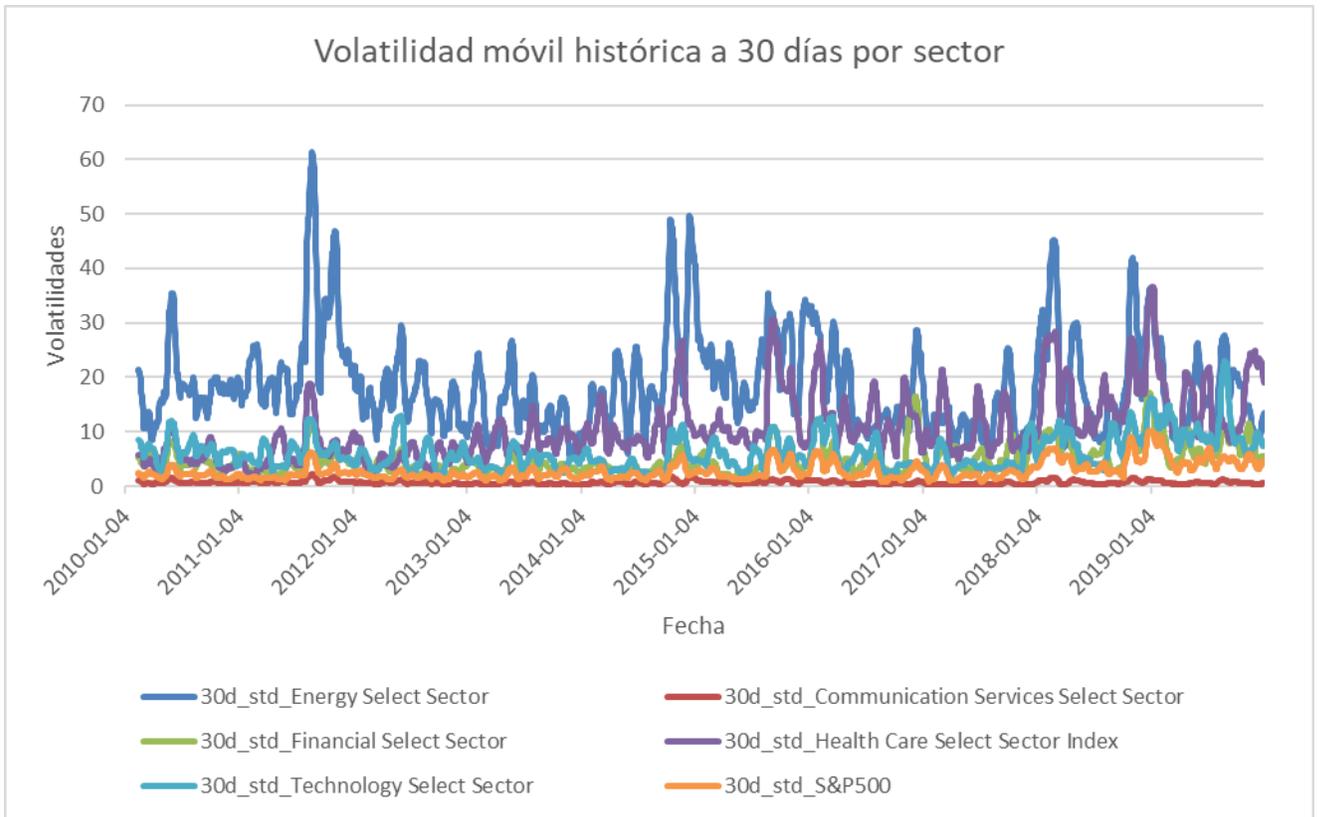
A partir de esta problemática expuesta, se presenta en el primer capítulo el comportamiento del mercado financiero y sus variaciones entre diferentes rubros de la economía estadounidense, mostrando distintos niveles de volatilidad para cada uno de ellos. Luego se introduce el concepto de aprendizaje automático y se realiza un pasaje por diversas aplicaciones de estos modelos en el análisis de precios del mercado financiero y se presentan visiones al respecto de diferentes autores. En el segundo capítulo se da un aspecto más preciso sobre el modelo que sigue la teoría de los mercados eficientes y los modelos de aprendizaje automático, las NN y el XGBoosting propuestos como herramienta de análisis. En el tercer capítulo se aplican todos los conceptos desarrollados con anterioridad a 15 acciones de diferentes rubros de la economía estadounidense con el objetivo de comparar el desempeño de los distintos modelos entre distintos rubros con diferentes niveles de volatilidad y entre distintas ventanas temporales, utilizando una técnica llamada error cuadrático medio residual. Finalmente, en la cuarta sección se presentan las conclusiones del trabajo en función a los rendimientos de los modelos y potenciales usos para distintos análisis del tiempo, así como también se indican futuras investigaciones y usos de otros posibles algoritmos a este tipo de análisis con el fin de incrementar recomendar una expansión del conjunto de herramientas que son enseñadas y empleadas en la actualidad en el análisis de riesgo de mercado.

Capítulo 1: Mercados financieros: el paradigma del aprendizaje automático

1.1. Comportamientos de los mercados financiero

Los individuos que incursionan en este tipo de mercados constituyen portafolios de inversión que les permiten obtener rendimientos, que expuestos a la pérdida de valor en el paso del tiempo constituyendo un riesgo. Por lo tanto, los inversores se ven en la necesidad de utilizar herramientas que les permitan analizar el comportamiento de precios y en lo posible predecirlos y de esta manera tener un mejor panorama sobre el comportamiento de sus inversiones.

En el año 1990 el Profesor Harry Max Markowitz es laureado con El Premio Nobel de Economía por sus investigaciones en el campo de las finanzas y específicamente en el área de la elección eficiente de un portafolio o cartera de inversión. En (Markowitz, 1952) se demuestra que en la relación retorno-riesgo, un portafolio eficiente es aquel cuyo nivel de diversificación es amplio, ya que permite una reducción de la volatilidad (riesgo) a un nivel aceptable de rendimiento. Si bien la diversificación de una cartera de inversión reduce el riesgo, al mismo tiempo algunos activos de la cartera son más volátiles que otros. La volatilidad también depende de los sectores de la economía en donde se desempeñan los mismos, los cuales tienden a comportarse de forma diferente. Marc (2021) expone a través de un estudio sobre la desviación estándar móvil de 30 días de los sectores de la economía estadounidense entre los años 2010 y 2020, que las volatilidades de los sectores (por ejemplo: sector energético, sector financiero, sector tecnológico, sector de comunicaciones y el sector salud) se comportan de manera distinta siendo para algunos mayores a la volatilidad del mercado y otros incluso menores.



Fuente: Elaboración propia en base a Yahoo Finance

Dada la evidencia empírica, se destaca la necesidad de tomar en cuenta la procedencia sectorial de los activos en lo momento de la composición de la cartera ya que la diversificación no es solo una cuestión de seleccionar distintos activos de diferentes al azar con el objeto de generar diversidad de procedencia de los mismo, sino también debe ser estudiada la volatilidad según el sector de donde provienen ya que esta puede incorporar información necesaria para el análisis de riesgo de mercado que genera la cartera.

A la par de los análisis de sectores, es fundamental para el análisis y construcción de los portafolios de inversión el tipo de modelo de predicción utilizado como mecanismo de toma de decisión. El tipo de riesgo financiero asociado a la inversión en carteras constituidas por instrumentos del mercado bursátil es el riesgo de mercado, el cual está

fuertemente relacionado con los problemas de predicción (Khaidem, Saha y Dey 2016). Por lo tanto, encontrar la mejor herramienta para reducir el riesgo de mercado, es menester y para ello deben ser utilizados diferentes tipos de modelos y evidenciar cuál posee el mayor nivel de precisión en la predicción y de esta forma contribuir a la mitigación del riesgo.

Modelos financieros: el nuevo paradigma de “machine learning” En el análisis y predicción de precios de acciones bursátiles existen estrategias de diversas procedencias, destaca el uso de modelos econométricos tal como los análisis de series de tiempo, modelización de la varianza y modelos logit. De igual manera destacan los modelos de análisis técnico y fundamental, así como los modelos estocásticos (modelo de Heston y modelización de proceso browniano). En cuanto a los modelos utilizados para el análisis de precios de opciones financieras, hay una amplia variedad de modelos, entre los más conocidos están Black and Scholes (1973), modelo de Merton (1973) y Corrado-Su (1996). Estos tipos de modelos estocásticos sirvieron de base para la modelización de opciones en el siglo XX, siendo sus inventores los Profesores Black, Merton y Scholes, galardonados con el Nobel de Economía 1997.

Sin embargo, a pesar de su popularidad y buen desempeño, los supuestos en los que se basan estos modelos de tipo paramétrico son fuertes y poco reales, tales como: volatilidad constante, deriva constante, principio de no arbitraje y tasa de retornos log normales (Ivaşcu 2020, Hull 2018, Park 2014, Amilon 2003). Ivaşcu (2020) también puntualiza que los modelos paramétricos son computacionalmente muy costosos y requieren una alta calibración de parámetros implícitos, así mismo no presentan una forma de solución de la evaluación de ecuaciones, requiriendo numerosos procesos de

optimización. Otro aspecto de este tipo de modelo es que, debido a la estimación de parámetros pierden versatilidad, ya que los mismos están en constante revisión, lo que los hace menos eficientes y se exponen a una mayor cantidad de errores de estimación.

La revolución informática ha dado mucho de qué hablar en las últimas décadas ya que ha permeado a diversas áreas de investigación científica. La velocidad de sus avances ha sido relevante de modo tal que existen tópicos subyacentes que toman posición cada día con más frecuencia, tales como los relacionados con la inteligencia artificial, el aprendizaje automático y el *Big Data*. Esto se debe a que el progreso tecnológico de la ciencia computacional se ha convertido en una aliada de otras como la matemática, biología, física, medicina e incluso en la ciencia económica y especialmente en el espectro de las finanzas por lo que el uso de la computación para la resolución problemas en el análisis de riesgo financiero se ha convertido punto de interés, ya que puede dar nuevas perspectivas y mejoras en el campo.

Por ello, el conocimiento de este tipo de herramientas dirigidas a la generación de conocimiento y resolución de problemas a través del uso de modelos que vayan de la mano con lo propuesto por la nueva era digital no debe sorprender dentro del área de las finanzas y es un hecho que es un tópico de suma importancia para la actualidad y los años venideros.

En contraposición a los modelos paramétricos propuestos en la década de 1990 están los de corte no paramétricos, que poseen la ventaja de ser más adaptativos y reducir errores de especificación (Hutchinson 1994). Una clase de modelos no paramétricos utiliza algoritmos de aprendizaje automático (machine learning, ML), los cuales, al no estar sujetos a la estimación de parámetros son más flexibles ante cambios en los datos

y al análisis de tendencia. Estos modelos están basados en la búsqueda de patrones en los datos y no en la inferencia a partir de un valor estimado. Así mismo, los supuestos que deben ser respetados en los modelos estocásticos, dejan de ser pilares, ya que los no paramétricos encuentran la información necesaria en las variables que les son dadas como insumo para su entrenamiento.

De este modo y más allá de la bondad de los modelos de ML, los modelos paramétricos continúan predominando en el análisis de precios de opciones financieras. Debido a esta realidad, surge la necesidad de comparar el desempeño entre las distintas metodologías y considerar cuál es el más adecuado en el análisis de precios en diferentes situaciones.

Capítulo 2: Modelos: aspectos teóricos

2.1. Riesgo de mercado

En el campo del análisis de riesgos una de las herramientas más importantes para el seguimiento de riesgo es el capital económico. Siguiendo la definición del Banco Central de la República Argentina (BCRA), el capital económico es aquel capital requerido por la entidad para afrontar pérdidas inesperadas por riesgo de mercado y otros riesgos asociados al sistema financiero (BCRA, comunicación "A" 5398). Así mismo, el BCRA define el riesgo de mercado como la posibilidad de incurrir en pérdidas en posiciones de balance o fuera de él, debido a fluctuaciones adversas en los precios de mercado de diversos activos. Estas pérdidas están asociadas principalmente a las carteras que las entidades financieras mantengan con diversos títulos que estén expuestos a dichas fluctuaciones.

Una vez definidos los tipos de riesgo, el BCRA exige a las entidades financieras la gestión de estos, en el caso del riesgo de mercado se entiende como gestión la evaluación y seguimiento y el control y mitigación de este tipo de riesgo asociado a la cartera de negociación. Por lo tanto, se puede afirmar que el eje central del analista y gestor de riesgo de mercado se encuentra en el seguimiento y correcta confección de la cartera compuesta por títulos de mercado.

2.2. Modelo de Valoración de Activos Financieros (CAPM)

Para el análisis de los precios de las acciones y de las opciones de acciones financieras, es esperado que los precios de estas varíen. Para los modelos paramétricos, es necesario modelizar ese cambio esperado, el cual puede calcularse a través del

modelo de Valoración de Activos Financieros (CAPM por sus siglas en inglés). De esta forma se obtiene la deriva (μ) del precio de la acción.

El modelo CAPM supone que el retorno esperado de un activo financiero i será igual a la tasa libre de riesgo, más un valor beta multiplicado por la prima de riesgo de mercado.

$$E[R_i] = R_f + \beta_i(E[R_M] - R_f)$$

Donde

$E[R_i]$: es el retorno esperado del activo financiero i

R_f : es la tasa libre de riesgo

β_i : es un parámetro que muestra cómo varía el precio del activo i cuando varía el mercado

$E[R_M] - R_f$: es la prima de riesgo de mercado que está compuesta por el valor esperado del mercado y la tasa libre de riesgo.

Siguiendo a Loffler y Posch (2007), μ no es representada por el valor esperado del retorno, sino por el retorno logarítmico:

$$\mu = \ln(1 + E[R_i])$$

La tasa libre de riesgo utilizada en este modelo será la publicada por la Reserva Federal de los Estados Unidos con capitalización diaria.

2.3. Proceso estocástico

Un proceso estocástico es un conjunto de variables aleatorias tiempos dependientes filtradas en un espacio muestral, los cuales pueden ser de tiempo discreto o continuo.

2.3.1. Movimiento Browniano Generalizado

Un proceso de Wiener o Movimiento Browniano Generalizado es un proceso estocástico continuo que se obtiene como resultado de tener paseos aleatorios escalados $W^{(n)}(t)$ cuando $n \rightarrow \infty$ con deriva μ y una varianza infinitesimal σ^2 (o volatilidad σ) si el mismo satisface las siguientes dos condiciones:

Tiene incrementos independientes. Es decir, para todos los valores futuros del proceso se cumple la propiedad de Márkov. Para todo par ordenado de valores $s > t$, la variable aleatoria $W_{\mu,\sigma}(s) - W_{\mu,\sigma}(t)$ tiene:

Condición 1:

$$E\{W_{\mu,\sigma}(s) - W_{\mu,\sigma}(t)\} = \mu(s - t)$$

Condición 2:

$$Var\{W_{\mu,\sigma}(s) - W_{\mu,\sigma}(t)\} = \sigma^2(s - t)$$

2.3.2. Lema Ito de dimensión 1

Sea $\{X(t)\}$ la solución a la ecuación diferencial estocástica

$$dX = \alpha(t, X)dt + \sigma(t, X)dW,$$

y $g(t, x)$ a una función determinística la cual es continuamente diferenciable por el tiempo t y es doblemente diferenciable por x . Entonces, el proceso estocástico $\{g(t, X(t))\}$ es una solución de la siguiente ecuación diferencial estocástica:

$$dg(t, X) = \left[\frac{\partial g(t, X)}{\partial t} + \alpha(t, X) \frac{\partial g(t, X)}{\partial x} + \frac{1}{2} \sigma^2(t, X) \frac{\partial^2 g(t, X)}{\partial x^2} \right] dt + \sigma(t, X) \frac{\partial g(t, X)}{\partial x} dW$$

2.4. Modelos de aprendizaje automático

2.4.1. Error cuadrático medio residual

El error cuadrático medio residual se define como la raíz cuadrada de la desviación media de los valores estimados y los valores observados.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=0}^n (y_i - \hat{y}_i)^2}$$

2.4.2. Tasa de aprendizaje

Es un coeficiente que escala el tamaño de las actualizaciones que toma el vector de parámetros de una red neuronal de esta forma afecta el espacio de la función de pérdida. La tasa de aprendizaje afecta la cantidad por la cual los parámetros se ajustan durante la optimización para minimizar los errores de las estimaciones hechas por la red neuronal.

2.4.3. Regularización

Regularización es el nombre que se le da en el ámbito de aprendizaje automático, a los procesos de control del sobreajuste, que se basa en minimizar el tamaño parámetros en el tiempo. Esta visión lleva la idea de penalizar los modelos en la medida en que se tornen más complejos. Así, se ayuda a evitar el sobreajuste penalizando las ponderaciones con grandes magnitudes.

2.4.4. Función de pérdida

Las funciones de pérdida miden qué tan cerca se encuentran las estimaciones de la red neuronal al valor exacto para el cual ha sido entrenada la red. Las funciones de pérdida pueden variar dependiendo del tipo de resultado que se quiera obtener de la red

neuronal ya sea, si se desea obtener valores reales (regresión) o si se desean probabilidades (clasificación).

2.4.5. Hiperparámetros

En el ámbito del Machine Learning se tienen ambos tipos, los parámetros del modelo y los parámetros que se calibran para hacer que los algoritmos entrenen y generen el aprendizaje del modelo. Esos parámetros calibrados son llamados “hiperparámetros” y ellos trabajan en el control de las funciones de optimización y la selección durante el proceso de entrenamiento.

2.4.6. Sobreajuste (overfitting)

El sobreajuste es un problema que afecta a los modelos de entrenamiento, el cual radica en que el modelo no aprende la señal contenida en los datos, sino que tampoco es capaz de aprender a evidenciar los errores.

2.4.7. Grupo de entrenamiento y de prueba

En el ámbito del aprendizaje automático los modelos son entrenados para luego obtener las predicciones. Los modelos se entrenan utilizando los datos de insumo, pero no en su totalidad, ya que para evitar problemas de sobre ajuste se hace una división de los datos. Grupo de entrenamiento es el nombre que se le da a la porción de todos los datos que son usados solo para entrenar el algoritmo del modelo y estimar los hiperparámetros. Y el grupo de prueba son los datos con los que se realizan las estimaciones, una vez se ha calibrado el modelo.

2.5. Extreme Gradient Boosting (XGBoosting)

Es un método de ensamble de algoritmos desarrollado Chen y Guestrin (2016) que propone un sistema de aprendizaje automático escalable para el ensamble de árboles.

Este modelo propone una mejora a diferencia de otros modelos de Gradient Boosting y es la incorporación de la función objetivo ya regularizada.

Siguiendo a Chen y Guestrin (2016), la derivación del modelo sigue la misma idea del enfoque inicial dado por Friedman et al. (2000) con la particularidad de la modificación de la función objetivo. Por lo tanto, se puede expresar que dado un conjunto de datos con n datos y m características, sea

$$D = \{(\mathbf{x}_i, y_i)\} (|D| = n, \mathbf{x}_i \in R^m, y_i \in R)$$

Un modelo de árbol ensamblado, que usa K funciones aditivas para predecir el resultado:

$$\hat{y}_i = \phi(\mathbf{x}_i) = \sum_{k=1}^K f_k(\mathbf{x}_i), f_k \in F$$

Donde:

$$F = \{f(\mathbf{x}) = w_{q(\mathbf{x})}\}(q: R^m \rightarrow T, w \in R^T)$$

Es el espacio de regresión del árbol (también conocido como CART). Aquí q representa la estructura de cada uno de los árboles correspondientes a cada hoja índice. T es el número de hojas en el árbol. Cada f_k corresponde a una estructura de árbol independiente q , y hojas ponderadas y las ponderaciones de las hojas w .

Finalmente, para el aprendizaje y conjunto de funciones usadas en el modelo, se minimiza la siguiente función objetivo regularizada.

$$L(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k)$$

Donde

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \|w\|^2$$

Aquí, l es una función de pérdida convexa que mide la diferencia entre el valor predicho \hat{y}_i y el objetivo y_i .

El segundo término omega penaliza la complejidad del modelo. El término de regularización ayuda a reducir el sobreajuste. Intuitivamente la función elegirá el modelo con mejor predicción y con la más baja complejidad.

Dado que el modelo es entrenado de forma aditiva, la función objetivo se muestra de la siguiente manera:

$$L^t = \sum_{i=1}^n l(y_i, \hat{y}_i^{t-1} + f_t(\mathbf{x}_i)) + \Omega(f_t)$$

2.6. Red neuronal multicapas (*Multilayer Neural Network*)

Este tipo de red neuronal que contiene una capa de insumo, una o más capas escondidas, ponderaciones y una capa de resultado. Cada una de las capas tiene una o más neuronas artificiales.

La dinámica de este tipo de red neuronal comienza con los insumos que va a pasar por cada una de las capas ocultas, los cuales son previamente ponderados y sumados, en cada capa oculta se halla una función de activación que genera cambios en los valores.

Este procedimiento se repite tantas capas escondidas sean elegidas y finalmente se llega a una capa resultado.

$$h_{w,b}(x) = g(w \cdot x + b)$$

Donde:

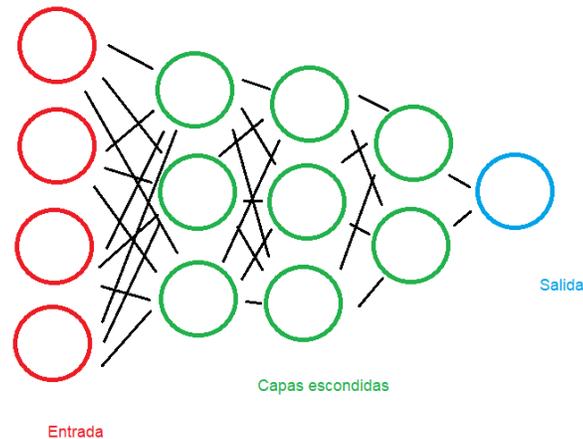
g = La función de activación.

w = Vector de todas las ponderaciones que llevan a la neurona.

x = Vector de las variables de entrada/insumo para el entrenamiento de la neurona.

b = El sesgo

Gráfico 2.1. Estructura de red neuronal de cuatro características de entrada, tres capas escondidas y una capa de salida.



Elaboración propia

2.6.1. Proceso de aprendizaje

El proceso de aprendizaje para cualquier algoritmo que use ponderaciones es el proceso de reajustar las ponderaciones y los sesgos, haciendo cada vez algunos más pequeños y otros más grandes, así se distribuye significativamente para ciertos bits de información y se minimizan para otros.

Las relaciones que aprenden las redes neuronales lo hacen de forma ciega, haciendo suposiciones basadas en los insumos y en las ponderaciones y luego miden que tan precisos son los resultados.

2.6.2. Propagación hacia adelante y hacia atrás

La propagación hacia adelante es el método de estimación de las redes neuronales que se basa en pasar por varias etapas la transformación de los datos, la primera transformación de los datos es llamada la propagación hacia adelante que consiste en

darle al modelo una primera visión de los datos de los cuales luego debe aprender. Esta propagación hacia adelante consiste en pasar los datos de insumo por todas las capas de la red neuronal realizar un primer cálculo con los ponderadores y transformar los datos con las funciones de activación.

Por otro lado, la propagación hacia atrás es un algoritmo de aprendizaje que se basa en calcular los posibles resultados de una red neuronal, este algoritmo es una parte importante de la reducción de errores en los modelos de redes neuronales. Utilizando las ponderaciones y las funciones de activación correspondiente, se evalúa la función de error y se intenta minimizarla yendo capa por capa desde la red resultado hasta el vector de insumos, modificando las ponderaciones con el objetivo de mejorar los resultados.

2.6.3. Función de activación

Las funciones que define el comportamiento de la red neuronal son las llamadas funciones de activación. Las funciones de activación transforman la combinación de vectores insumo, ponderaciones y sesgos.

Los productos de esas transformaciones son insumos para el siguiente nodo de la siguiente capa, este proceso se repite hasta llegar al nodo del resultado. Se usan las funciones de activación para las neuronas escondidas de una red neuronal con el objetivo de introducir la no linealidad dentro de las capacidades de modelización de la red.

2.6.4. Capas de una red neuronal

- Capa de entrada: es el vector que alimenta la red neuronal. El número de neuronas en una capa de insumo es normalmente igual al número de características para la red neuronal.

- Capas escondidas: Las capas escondidas son las capas que aplican a los valores de las ponderaciones en las conexiones entre las capas son como las redes neuronales sean codificadas el aprendedor de información extraído desde el conjunto de data de entrenamiento crudo.
- Capa-resultado: es la respuesta o predicción desde el modelo de las interacciones no lineales entre las capas anteriores. Dependiendo de la configuración de la red neuronal, el resultado final puede ser un resultado revaluado (regresión) o una configuración de probabilidad (clasificación).

Capítulo 3: Análisis de resultados

3.1. Datos

En este capítulo se desarrollan los modelos presentados en el capítulo anterior, para su comparación y posteriores conclusiones. En la confección de los modelos fueron utilizados datos de precios de acciones financieras que cotizan en las bolsas de los Estados Unidos. El motivo de la elección de los Estados Unidos y de las acciones financieras de este país, se debe a al hecho de que ha sido y es actualmente la economía de mayor preponderancia a nivel mundial, así como también una de las economías que posee el mayor desarrollo de mercado financiero en el mundo y a la facilidad de acceso a la información sobre precios de acciones financieras de empresas públicas. El período de tiempo en el que extrajeron los datos corresponde a los años 2018 hasta enero de 2020 ya que es el un período más cercano donde se ha podido observar mayor estabilidad en los mercados y se encuentra antes del inicio de la pandemia de COVID-19 y, por lo tanto, la estimación de parámetros como la media y la varianza de las series no sufren sesgos por eventos adversos.

Con el objetivo de realizar el análisis sectorial los datos para el desarrollo de los modelos provienen de las acciones más representativas de los índices de sectoriales, de la economía estadounidense que desarrolla S&P Dow Jones de los siguientes sectores: energía y materias primas, financiero, tecnológico, servicios de comunicación y salud. Los nombres de los índices utilizados son: *Energy Select Sector Fund*, *Finacial Select Sector Index*, *Communication Service Select Sector Fund*, *Technology Select Sector Index*, *Health Care Select Sector*. Para el año 2022, las empresas que se encuentran como los mayores ponderadores de estos índices son:

- *Energy Select Sector Fund (XLE):*
 1. Exxon Mobil Corporation (XOM)
 2. Chevron Corporation (CVX)
 3. Conoco Phillips (COP)
- *Financial Select Sector Index (IXM)*
 1. Berkshire Hathaway B (BRK-B)
 2. JP Morgan Chase & Co (JPM)
 3. Bank of America Corporation (BAC)
- *Communication Services Select Sector (XLC)*
 1. Meta Platforms Inc A (META)
 2. Alphabet Inc A (GOOGL)
 3. Netflix Inc (NFLX)
- *Technology Select Sector (IXT)*
 1. Apple Inc. (AAPL)
 2. Microsoft Corporation (MSFT)
 3. Nvidia Corp (NVDA)
- *Health Care Select Sector (IXV)*
 1. Johnson & Johnson (JNJ)
 2. United Health Group Inc (UNH)
 3. Pfizer Inc (PFE)

3.2. Resumen estadístico

Para el armado de los parámetros para el modelo del Movimiento Browniano Generalizado, fueron utilizados los precios de cierre de 5 ruedas de todas las acciones anteriormente mencionadas.

Cuadro 3.1 Resumen estadístico

	AAPL	BAC	BRK-B	COP	CVX	META	GOOGL	JNJ	JPM	MSFT	NFLX	NVDA	PFE	UNH	XOM
Muestra	252	252	252	252	252	252	252	252	252	252	252	252	252	252	252
Media	50.86	27.61	208.07	55.18	103.53	181.64	59.56	124.67	103.65	126.56	328.87	43.44	33.89	238.31	61.52
Std	8.65	2.22	7.79	4.21	3.59	16.05	3.99	4.73	10.27	14.97	34.49	6.26	2.18	17.58	3.20
Mín	34.41	22.70	191.66	45.60	92.09	131.74	51.27	114.63	87.40	93.77	254.59	31.78	29.28	206.31	55.93
Máx	72.14	33.58	227.05	63.26	109.47	208.10	68.12	137.29	128.34	155.30	385.03	59.67	37.62	285.28	68.71

Fuente: Elaboración propia en base a *Yahoo Finance*

De los cuales fueron calculados los retornos diarios y posteriormente se procedió a calcular el promedio móvil, el desvío móvil (σ) y el *drift* o deriva ($\mu - \frac{\sigma^2}{2}$) con el objetivo de calcular las trayectorias posibles de los precios.

Cuadro 3.2 Parámetros anualizados

	AAPL	BAC	BRK-B	COP	CVX	META	GOOGL	JNJ	JPM	MSFT	NFLX	NVDA	PFE	UNH	XOM
Media	67.15%	40.67%	11.51%	10.30%	15.91%	48.71%	27.63%	16.40%	40.47%	47.46%	24.95%	65.39%	-5.34%	21.41%	8.66%
Std	26.14%	23.16%	15.02%	27.96%	18.41%	27.88%	23.70%	16.45%	18.72%	19.83%	34.62%	40.60%	19.01%	25.27%	18.35%
Drift	63.73%	37.99%	10.38%	6.39%	14.21%	44.83%	24.82%	15.04%	38.72%	45.50%	18.96%	57.15%	-7.15%	18.22%	6.98%

Fuente: Elaboración propia en base a *Yahoo Finance*.

3.3. Movimiento Browniano Generalizado (MBG)

Utilizando un el modelo del Movimiento Browniano Generalizado es posible modelar los precios de una acción bajo el supuesto de varianza y *deriva* constante. Siendo estos dos supuestos los de mayor cuestionamiento sobre este modelo. De forma que la variación de precio de la acción se modela como:

$$dS = \mu S dt + \sigma S dz$$

Donde dS representa la variación del precio de la acción S .

3.3.1. Propiedad lognormal

Utilizando el Lemma de Ito para derivar el proceso como el logaritmo del precio de la forma:

$$G = g(x, S) = \ln(S)$$

Ya que:

$$\frac{\partial G}{\partial S} = \frac{1}{S}, \quad \frac{\partial^2 G}{\partial S^2} = -\frac{1}{S^2}, \quad \frac{\partial G}{\partial t} = 0$$

Se puede expresar la variación del proceso G , como:

$$dG = \left(\mu - \frac{\sigma^2}{2} \right) dt + \sigma dz$$

Ya que como μ y σ son constantes, la ecuación indica que $G = \ln(S)$ sigue un Movimiento Browniano Generalizado, con una deriva igual a $\mu - \frac{\sigma^2}{2}$ y varianza constante σ^2 . Así, las variaciones en $\ln(S)$ en los períodos de tiempo desde 0 a T se distribuyen normalmente con media $\left(\mu - \frac{\sigma^2}{2} \right) T$ y varianza $\sigma^2 T$. Lo cual implica que:

$$\ln S_T - \ln S_0 \sim \phi \left[\left(\mu - \frac{\sigma^2}{2} \right) T, \sigma^2 T \right]$$

o

$$\ln S_T \sim \phi \left[\ln S_0 + \left(\mu - \frac{\sigma^2}{2} \right) T, \sigma^2 T \right]$$

Así por el Lema de Ito, se puede expresar la variación logarítmica del precio de la acción como:

$$d \ln S = \left(\mu - \frac{\sigma^2}{2} \right) dt + \sigma dW$$

Luego,

$$\ln S(t) = \ln S(0) + \left(\mu - \frac{\sigma^2}{2} \right) t + \sigma W(t)$$

Así, la solución a la ecuación es:

$$S(t) = S(0)e^{\left(\mu - \frac{\sigma^2}{2} \right) t + \sigma W(t)}$$

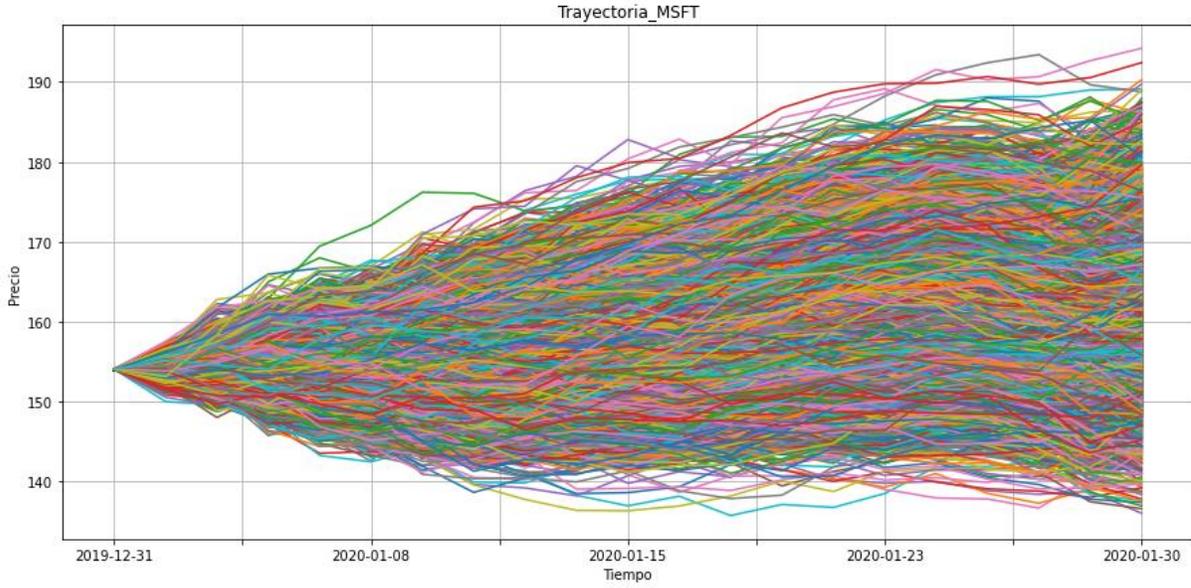
3.3.2. Simulación de Montecarlo

Una vez calculados los parámetros de cada una de las acciones se realizaron simulaciones de los precios de las acciones (100.000 simulaciones) logrando de esta forma estimar la trayectoria de los precios.

Luego de realizar las estimaciones de las curvas de cada una de las acciones se estimó un intervalo de confianza, del 95% con el objetivo de observar los límites de las trayectorias.

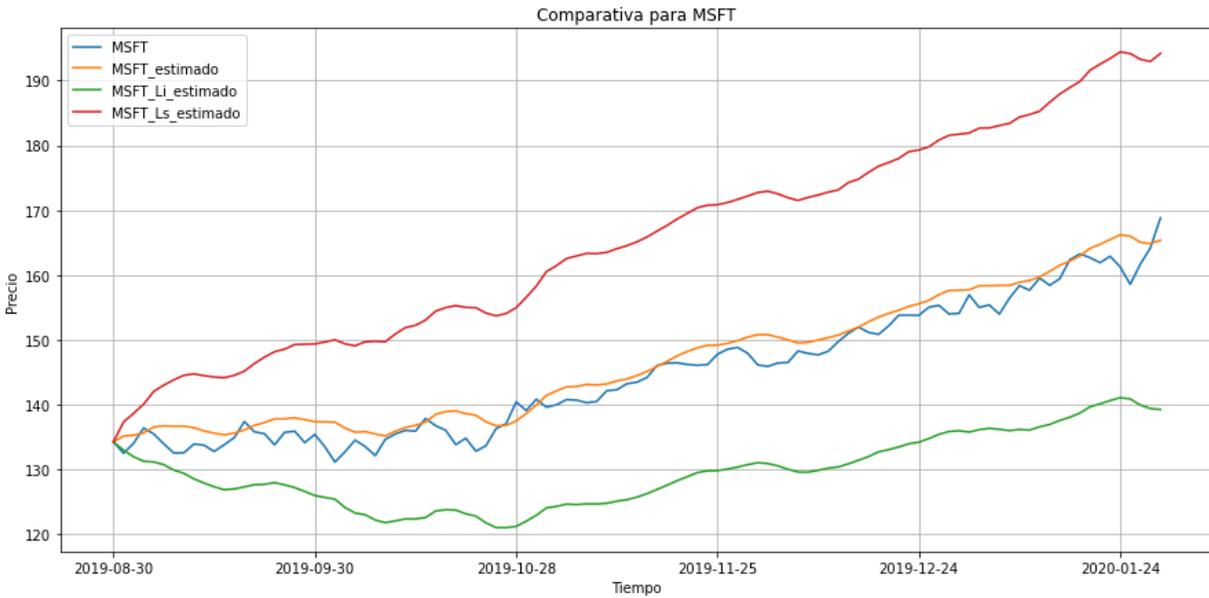
Finalmente, encontrado los límites se calculó el promedio las simulaciones obteniendo, así el precio estimado que se espera sea el valor más cercano de las estimaciones al verdadero precio.

Gráfico 3.1 Trayectorias estimadas para Microsoft (MSFT)



Fuente: Elaboración propia (Cálculos en Python)

Gráfico 3.2 Valores reales vs estimados - comparativa para Microsoft (MSFT)



Fuente: Elaboración propia (Cálculos en Python)

3.4. Modelos de Aprendizaje Automático

3.4.1. Características extraídas de los datos

Con el objetivo de entrenar los modelos predictivos de aprendizaje automático se calcularon los siguientes indicadores de tendencia (que son utilizados en la conformación de estrategias de trading) como característica de los precios de las acciones:

1. Índice de Fuerza Relativa (RSI)

El Índice de Fuerza Relativa (RSI) calcula la ratio entre la subida de precio más reciente y el movimiento del precio absoluto. El RSI se interpreta como un indicador de sobreventa o sobrecompra cuando los valores están sobre 70 o por debajo de 30.

$$RSI = 100 - \frac{Mayor\ precio\ promedio_t}{Mayor\ precio\ promedio_t + Menor\ precio\ promedio_t} * 100$$

Donde:

$$Mayor\ precio\ promedio_t = Precio_t - Precio_{t-1} \forall Precio_t - Precio_{t-1} > 0$$

$$Menor\ precio\ promedio_t = Precio_t - Precio_{t-1} \forall Precio_t - Precio_{t-1} < 0$$

2. Williams R (WILLR)

La tasa porcentual de Williams es un oscilador que basa en el momentum y es utilizado para identificar momento en los que exista sobrecompra o sobreventa.

$$WILLR = \frac{Mayor\ precio\ de\ cierre}{Menor\ precio\ de\ cierre} * 100$$

3. Momentum (MOM)

El momentum es una medición de la aceleración y desaceleración de los precios.

$$MOM = Precio_t - Precio_{t-n}$$

4. Índice de Canal de Materias Primas (CCI)

El CCI es utilizado para detectar el comienzo y final de las tendencias del mercado.

$$CCI = \frac{TP - ATP}{0.015 * MD}$$

Donde:

$$TP = \text{Precio típico} = \frac{\text{Mayor precio}_n + \text{Menor precio}_n + \text{precio de cierre}}{3}$$

$$ATP = \text{Media Móvil (TP)}$$

$$MD = \text{Desviación estándar (TP)}$$

5. Media Móvil Convergente/Divergente (MACD)

El MACD es un indicador que señala el cambio de tendencia e indica el inicio de una nueva tendencia.

$$MACD = \text{shortema} - \text{longema}$$

Donde:

$$\text{shortema} = 0.15 \text{ price} + 0.85 \text{ shortema}_{t-12}$$

$$\text{longema} = 0.075 \text{ price} + 0.925 \text{ longema}_{t-26}$$

3.4.2. Extreme Gradient Boosting (XGBoosting/XGB)

El Extreme Gradient Boosting es considerado como uno de los mejores algoritmos de la familia Gradient Boosting para el análisis predictivo a través de regresiones. Para su implementación fue utilizada la librería XGBoost para Python versión 3.9. En el entrenamiento se utilizaron los datos históricos de los índices mencionados en la sección 3.5. de los precios, para ellos fueron especificados los hiperparámetros.

Los índices fueron calculados sobre los retornos logarítmicos diarios de los precios. Los sets de datos constan de precios que se encuentran entre los años 2018 y 2020 los cuáles fueron divididos en grupos de entrenamiento y grupos de prueba, asignando un 80% de los datos al entrenamiento del modelo y el restante para la prueba de las

predicciones. Esto deja el modelo con un total de 418 datos para el entrenamiento y 105 datos para la prueba. Así, se intenta evitar posibles problemas de overfitting, en el modelo.

Luego, en la elección de la mejor combinación de hiperparámetros para el entrenamiento del modelo fue utilizada una técnica llamada Random-Search. Esta metodología se le aplicó para la optimización de cuatro hiperparámetros:

- Número de Estimadores
- Tasa de aprendizaje
- Regularización alpha L1
- Máxima profundidad del árbol

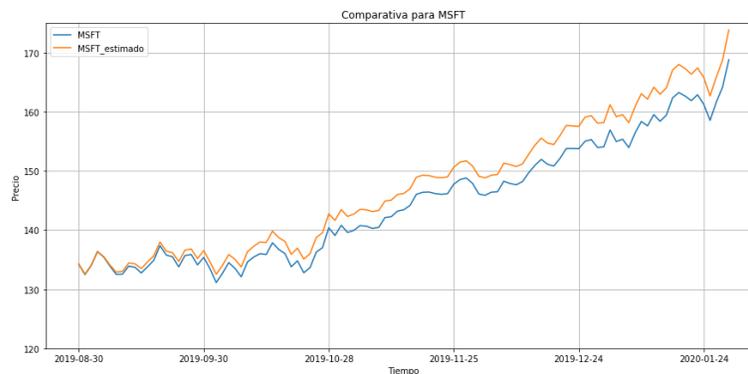
Este método fue elegido sobre el Grid-Search debido a que para aplicar esta metodología se debe tener un indicio sobre el mejor valor que se les pueden asignar a cada uno de los hiperparámetros, mientras que utilizando el Random-Search estos se eligen aleatoriamente de un conjunto de posibles valores luego, son calculados los resultados de usar las diferentes combinaciones y finalmente se elige la que menor diferencia genere con los datos originales. Los mismos al ser obtenidos de un conjunto de datos y probados entre ellos, son diferentes para cada una de las acciones:

Cuadro 3.3 Hiperparámetros XGBoosting

Acción	N° estimadores	Tasa de aprendizaje	Maxima profundidad	Regularización alpha L1
AAPL	200	0.05	20	0.0036
BAC	300	0.05	500	0.0017
BRK-B	300	0.05	5	0.0016
COP	1000	0.10	20	0.0025
CVX	500	0.05	5	0.0025
GOOGL	200	0.05	100	0.0030
JNJ	1000	0.10	20	0.0013
JPM	1000	0.10	5	0.0024
META	300	0.10	50	0.0011
MSFT	300	0.10	5	0.0011
NFLX	300	0.45	10	0.0026
NVDA	500	0.20	10	0.0015
PFE	200	0.05	500	0.0012
UNH	1000	0.05	5	0.0012
XOM	200	0.15	500	0.0011

Fuente: Elaboración propia (Cálculos en Python)

Gráfico 3.3 Valores reales vs estimados. comparativa XGBoosting para Microsoft (MSFT)



Fuente: Elaboración propia (Cálculos en Python)

3.4.3. Redes Neuronales (NN)

El segundo método aplicado de Machine Learning fue el de Redes Neuronales. Para su implementación fue utilizada la librería Sklearn, específicamente el apartado para redes neuronales para Python versión 3.9. Esta librería optimiza el cuadrado de la pérdida utilizando diferentes algoritmos de optimización.

Al igual que para el modelo de XGBoosting los índices fueron calculados utilizando los retornos diarios de los precios de las acciones, así como también el período de tiempo y la división entre los conjuntos de entrenamiento y prueba.

A diferencia del modelo de XGB en las redes neuronales es necesario utilizar algoritmo de optimización de la función de pérdida y funciones de activación, las cuales proveen al modelo herramientas para el aprendizaje de comportamientos no lineales de los datos. Tanto las metodologías de optimización como las funciones de activación varían según el tipo de datos utilizados por lo tanto en la elección de los hiperparámetros para el entrenamiento de las redes neuronales fue utilizada la metodología RandomSearch con el objetivo de probar las diferentes posibles combinaciones. Entre ellos se eligieron:

- Algoritmo de Optimización: [LBFGS, Stochastic Gradient Descent, ADAM]
- Función de Activación: [Logística, Tahn, Relu]
- Tasa de aprendizaje: [0.01,0.001,0.1]
- Capas escondidas: [300]
- Neuronas por capa: [10,25,30,35]
- Máxima iteración: [1.000,10.000]

Como resultado se obtuvo la siguiente tabla donde se muestran los parámetros elegidos para cada una de las acciones que minimizan el error cuadrático medio residual, entre los valores estimados y los reales.

Cuadro 3.4 Hiperparámetros Redes Neuronales

Acción	Algoritmo de Optimización	Función de Activación	Tasa de aprendizaje	Capas escondidas	Neuronas por capa	Máxima iteración
AAPL	adam	tanh	0,001	300	35	10.000
BAC	adam	tanh	0,001	300	35	1.000
BRK-B	adam	relu	0,01	300	35	1.000
COP	adam	tanh	0,001	300	30	1.000
CVX	adam	tanh	0,01	300	30	10.000
GOOGL	lbfgs	tanh	0,001	200	35	10.000
JNJ	adam	tanh	0,004	300	35	10.000
JPM	adam	tanh	0,001	300	35	10.000
META	adam	tanh	0,005	300	30	1.000
MSFT	adam	tanh	0,001	300	35	10.000
NFLX	adam	logistic	0,001	300	20	1.000
NVDA	adam	logistic	0,01	300	10	1.000
PFE	adam	logistic	0,001	300	10	1.000
UNH	adam	tanh	0,001	300	35	10.000
XOM	adam	tanh	0,001	300	35	1.000

Fuente: Elaboración propia (Cálculos en Python)

Gráfico 3.4 Valores reales vs estimados - comparativa Redes Neuronales para Microsoft (MSFT)



Fuente: Elaboración propia (Cálculos en Python)

3.5. Comparación entre modelos

Una vez finalizadas las predicciones tanto del modelo del Movimiento Browniano Generalizado como de los modelos de Machine Learning, fue calculada para cada una de las acciones y por cada uno de los modelos la raíz cuadrada del error cuadrático medio residual como medida de evaluación del rendimiento de los modelos. Esta comparación se realizó para distintos períodos de tiempo.

En una primera instancia se llevó a cabo el análisis para una ventana temporal de 15 días, luego para la ventana de un mes, de 3 meses y finalmente para la ventana de 5 meses. Donde se pudo evidenciar que, en la ventana de 5 meses, y solamente en 3 acciones, el modelo del Movimiento Generalizado Browniano superó la capacidad de predicción de los modelos de aprendizaje automático.

Entre ellas se encuentran AAPL, JNJ y NFLX, siendo el modelo de MBG para NFLX el que tuvo mayor distancia en la medida de rendimiento con respecto a los otros dos, el cual fue 8,49 (MBG NFLX) contra 9,05 (RN NFLX).

Cuadro 3.5 Comparativa de raíz cuadrada del error cuadrático medio entre modelos (ventana 5 meses)

		MBG	XGBoosting	Redes Neuronales
<i>Energy Select Sector</i>	XOM	1.39	0.45	0.54
	CVX	2.44	1.87	1.70
	COP	1.89	2.52	0.59
<i>Financial Select Sector</i>	BRK-B	4.36	2.92	0.77
	JPM	3.59	2.53	1.72
	BAC	1.00	0.36	0.45
<i>Communication Servicies Select Sector</i>	META	5.54	1.11	3.33
	GOOGL	1.18	0.69	1.12
	NFLX	8.49	35.65	9.05
<i>Technology Select Sector</i>	AAPL	1.02	2.35	1.07
	MSFT	2.54	2.16	1.32
	NVDA	1.90	1.88	1.28
<i>Health Care Select Sector</i>	JNJ	1.58	1.59	1.66
	UNH	8.62	3.96	3.37
	PFE	0.87	1.20	0.42

Fuente: Elaboración propia (Cálculos en Python)

Por otro lado, el mejor modelo para el cálculo de la ventana de 5 meses resultó ser el de las Redes Neuronales obteniendo un total de 8 acciones para las cuales generó el mejor rendimiento al momento de predecir precios futuros. Donde se destacan COP y PFE las cuales tuvieron tasas de errores menores a 1, 0,59 y 0.42 respectivamente. Así mismo, para esta ventana temporal, el XGB obtuvo el segundo puesto como mejor pronosticador al generar mejores predicciones entre todos los modelos. También se evidencia que, para esta ventana temporal, en 4 de 5 sectores las redes neuronales conforman el 66,6% de mejores capacidades predictivas, exceptuando el sector de comunicaciones donde esta tendencia cambia a favor del XGB.

Esta tendencia de supremacía del modelo de redes neuronales en comparación contra el modelo de XGB se mantuvo incluso para una ventana temporal a 3M.

Cuadro 3.6 Comparativa de raíz cuadrada del error cuadrático medio entre modelos (ventana 3 meses)

		MBG	XGBoosting	Redes Neuronales
<i>Energy Select Sector</i>	XOM	1.26	0.38	0.58
	CVX	2.26	1.68	1.63
	COP	1.67	2.29	0.58
<i>Financial Select Sector</i>	BRK-B	4.08	2.65	0.74
	JPM	3.26	2.18	1.38
	BAC	0.89	0.27	0.26
<i>Communication Services Select Sector</i>	META	5.20	0.67	3.10
	GOOGL	1.12	0.49	0.87
	NFLX	8.69	31.17	9.16
<i>Technology Select Sector</i>	AAPL	0.92	1.67	0.70
	MSFT	2.48	2.02	0.92
	NVDA	1.74	1.68	1.21
<i>Health Care Select Sector</i>	JNJ	1.67	1.67	1.78
	UNH	7.42	4.22	3.52
	PFE	0.76	1.08	0.45

Fuente: Elaboración propia (Cálculos en Python)

Luego, al evaluar los modelos utilizando ventanas temporales más cortas se obtuvo resultados diferentes. Se pudo evidenciar que en la medida en que la ventana temporal disminuía, la capacidad predictiva del modelo browniano disminuía y por el contrario el modelo del gradiente extremo ganó preponderancia.

En la ventana temporal de 5 meses los pesos se distribuyeron de la siguiente forma: un 20% (3 acciones) de mejores rendimientos para el MBG, un 26,67% (4 acciones) para el XGB y un 53% (8 acciones) para las RN. Sin embargo, como se mencionó anteriormente, esta tendencia fue disminuyendo conforme se redujo la ventana temporal del análisis, hasta observarse que, para ventanas temporales menores a los 3 meses, el modelo MBG no vuelve a ser el mejor modelo para predecir.

Cuadro 3.7 Evolución por ventana temporal del porcentaje de mejores predicciones por modelo

t	MBG	XGBoosting	Redes Neuronales
5M	20,00%	26,67%	53,33%
3M	6,67%	26,67%	66,67%
1M	0,00%	53,33%	46,67%
15D	0,00%	73,33%	26,67%

Fuente: Elaboración propia (Cálculos en Python)

También vale destacar que, así como el poder predictivo del MBG fue decreciendo en la medida en que se acortaban las ventanas temporales, en el caso del modelo XGB esta capacidad al contrario fue en aumento. Comenzando un porcentaje de acierto del 26,67%, a un 73,33% para una ventana temporal de 15 días.

De esta forma se pudo evidenciar que los modelos de aprendizaje automático tuvieron una mejor performance para predecir los precios de las acciones incluso entre distintos sectores de la economía y con diferentes niveles de volatilidad. Además, que no solo fueron superiores sino que para ventanas temporales cortas existe diferencia en la capacidad predictiva de los modelos de ML.

Finalmente, se realizó un promedio general de la capacidad predictiva de los modelos para todas las ventanas temporales y se encontró que el mejor modelo de todos es el de las redes neuronales. El cuál obtuvo un 48,33% (29 veces mejor modelo) de mejor capacidad predictiva a lo largo de todas las ventanas temporales contra un 45% (27 veces mejor modelo) para el XGB y un 6,7% (4 veces mejor modelo) para el MBG.

Capítulo IV: Conclusiones y futuras líneas de investigación

En el estudio del análisis de riesgo de mercado, existen diversas herramientas que potencian la capacidad de estimar el nivel de riesgo al que se enfrenta una institución financiera cuando decide conformar un portafolio de inversión. De tal relevancia es este tipo de riesgo que tanto las instituciones de regulación financiera domésticas como internacionales están constantemente en la búsqueda de mejoras en la normativa para la reducción de este tipo de riesgo. Es por ello que esta investigación analizó la capacidad predictiva del modelo más utilizado en el ámbito académico como lo es el Movimiento Browniano Generalizado, incluyendo esta maestría, en contra posición a dos modelos que subyacen en el espectro de una de las disciplinas más citadas en la actualidad como lo es el aprendizaje automático.

De tal forma se busca dar una vista sobre el panorama que la actualidad brinda en la implementación de modelos que no se sustentan en supuestos fuertes como lo es la hipótesis de los mercados eficientes y que aun así pueden convertirse probablemente en aliados indispensables del analista de riesgo en el futuro próximo, dándole así a este tipo de herramientas la relevancia necesaria para ocupar un puesto en los temarios de los estudios de posgrado de esta y otras casas de estudio.

A lo largo de este trabajo de investigación se realizó un análisis comparativo de distintos métodos de estimación que pueden ser utilizados para la predicción de precios de acciones financieras. Entre esos métodos, explicitados en el Capítulo II, se encuentran dos tipos de modelos, uno de corte paramétrico (el Movimiento Browniano Generalizado) y otros dos modelos de carácter no paramétrico, XGBoosting y Redes Neuronales. Estas metodologías fueron aplicadas a 15 acciones financieras de diferentes

rubros de la economía estadounidense, para estimar los últimos 5 meses de precios entre el año 2019 y 2020 antes del inicio de la pandemia por el COVID-19, entre ellos sector energía, financiero, comunicaciones y servicios, tecnología y salud. Como resultado se obtuvieron los precios estimados para cada uno de los modelos y se calculó el error cuadrático medio residual como indicador de exactitud en la predicción de diferentes cortes temporales para evaluar el poder de predictivo de cada modelo en distintos lapsos de tiempo.

De este análisis se obtuvo que en una ventana de estimación de 5M el mejor modelo para predecir los precios de las acciones fue el de las Redes Neuronales, siendo el más indicado para usarlo ya que para 15 acciones estimadas, las Redes Neuronales generaron la mejor predicción para 8 acciones. Así mismo cabe destacar que para esta ventana temporal, el modelo que generó el menor nivel de error cuadrático medio residual fue el XGB para la acción del Bank of America.

Luego, al evaluar los modelos en ventanas temporales más reducidas se encontraron cambios de tendencia. Se pudo notar que, para ventanas más cortas, la capacidad predictiva del modelo browniano fue cada vez menor en comparación con los modelos de aprendizaje automático, de tal manera que ya para los espacios temporales de 15 días y 1 mes no vuelve a ser el modelo con menor nivel de error cuadrático medio residual. Por el contrario, en la medida en que se reduce la ventana temporal de análisis el modelo del gradiente extremo ganó preponderancia.

Finalmente, en un promedio general se concluyó que las redes neuronales es el mejor modelo para realizar este tipo de análisis. Las redes neuronales fueron calificadas 29

veces mejor modelo el mejor modelo de mejor capacidad predictiva a lo largo de todas las ventanas temporales contra unas 27 veces para el XGB y apenas 4 para el MBG.

Así mismo se pudo evidenciar que entre no solo los modelos de aprendizaje automático superan al modelo de paramétrico, sino también que entre ellos para distintos momentos del tiempo existe una diferencia en las capacidades predictivas e inclusive entre sectores. Se pudo notar que, para períodos más cortos de tiempo, el XGB es el mejor modelo para utilizar en lugar de las redes neuronales y que para las acciones del rubro de telecomunicaciones el XGBoosting supera ampliamente a los demás modelos ya que fue el mejor predictor para 8 de cada 12 experimentos. Una situación similar es encontrada para el modelo de redes neuronales en los rubros de las finanzas y la salud (8 y 7 veces mejor predictor de cada 12 respectivamente).

Por otro lado, desde un punto de vista del costo computacional la calibración del modelo estocástico mucho más costoso que la de los modelos de aprendizaje automático. El modelo estocástico fue generado usando 100.000 simulaciones para cada una de las acciones, y cada simulación tuvo un tiempo promedio de 1 minuto 50 segundos. Por el lado de los modelos de aprendizaje automático este tiempo promedio no supera los 20 segundos por acción.

Los resultados obtenidos permiten concluir que la incorporación de métodos no paramétricos de la disciplina del aprendizaje automático al análisis y predicción de precios de acciones financieras son efectivamente una herramienta beneficiosa y enriquecedora para el analista de riesgo, inclusive si la tarea de este se encuentra en la construcción de carteras entre distintos sectores de la economía con diferentes niveles de volatilidad. Es tarea del analista de riesgo encontrar mejores modelos que le permitan

ser más eficientes y disminuir el riesgo hasta cuanto más sea posible en pro de mejorar la asignación de capital dedicándolo en menor cuantía para la cobertura ante un posible escenario adverso y así tener menos capital ocioso más allá de lo normado por los entes regulatorios.

Es importante destacar que esta investigación abre espacio a temáticas que requieren un análisis más exhaustivo en busca de desarrollar futuras líneas de investigación al respecto. Primeramente, se debe mencionar que los modelos de aprendizaje automático requieren la calibración de varios hiperparámetros así como el uso de algoritmos de optimización, los cuáles pueden mejorarse en la utilización de este tipo de metodología como el uso de algoritmos genéticos propuestos por Qiu y Song (2016), los cuales se encuentran totalmente fuera del alcance de esta investigación. Por otra parte, existen algoritmos de predicción más complejos que vale la pena proponer para futuras investigaciones como lo son los modelos de LSTM, Extreme Learning Machine.

Como última observación, una vez comprobado la efectividad de este tipo de análisis para acciones financieras se abre una ventana para el análisis de otras herramientas financieras como los son lo derivados financieros, la implementación de este tipo de técnicas al análisis de precios de futuros y opciones financieras generaría sin dudas una mejora factible en el análisis de riesgos financieros.

Referencias

- Amilon, H. (2003). A neural network versus Black-Scholes: a comparison of pricing and hedging performances. *Journal of Forecasting*, 22(4), 317–335. doi:10.1002/for.867
- Aria, Fideas (2006) “El proyecto de la investigación: introducción a la metodología científica” Editorial Episteme, 6° edición. Caracas, Venezuela.
- Basak, S., Kar, S., Saha, S., Khaidem, L., y Dey, S. R. (2019) Predicting the direction of stock market prices using tree-based classifiers. *The North American Journal of Economics and Finance*, vol 47, pág 552-567.
- Black, F., y Scholes M, (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81 637–59.
- Breiman Leo (2001). Random Forest. Statistics Department, University of California, Berkeley, CA 94720. *Machine Learning*. vol 45. pág 5-32.
- Chen, T., y Guestrin, C. (2016) XGBoosting: A Scalable Tree Bosting System. The Association for Computing Machinery. KDD'16. pág 13-17.
- Davis Marc (2021, Julio 17). The 8 Most Volatile Sectors. Investopedia. Recuperado de: <https://www.investopedia.com/financial-edge/0712/the-8-most-volatile-sectors.aspx>
- Dojima Rice Exchange - World's First Futures Exchange. (s.f). Recuperado de: <https://www.jpx.co.jp/dojima/en/index.html>
- Fama, E. F. (1970). *Efficient Capital Markets: A Review of Theory and Empirical Work*. *The Journal of Finance*, 25(2), 383–417. <https://doi.org/10.2307/2325486>
- Fama, E. F. (1965). *The Behavior of Stock-Market Prices*. *The Journal of Business*, 38(1), 34–105. <http://www.jstor.org/stable/2350752>

FM Labs (Sin fecha). *Indicators Reference*. Disponible en: <https://www.fmlabs.com/reference/default.htm?url=CCI.htm> [Consulta: 2021, noviembre]

Géron Aurélien (2019) *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*. O'Reilly Media. 2da Edición. Canadá.

Hastie., T., Tibshirani., R., Friedman., J. (2009) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer. 2da Edición. Estados Unidos.

Hegazy. O, Soliman, O. y Mustafa, A. (2013) *Machine Learning Model for Stock Market Prediction*. *International Journal of Computer Science and Telecommunications*. Volumen 4. Emisión 12. Pág 17-23.

Hernández, R., Fernández, C. y Baptista, P. (2014) "Metodología de La investigación". MCGRAW HILL INTERAMERICANA EDITORES, S.A. de C.V. Sexta Edición.

Hull, J. C., (2018) *Options, Futures, and Other Derivative Securities*. 10ma Ed. Prentice-Hall. [Capítulo 11 y 14]. Englewood Cliffs, New Jersey.

Hutchinson, J. M., Lo, A. W., & Poggio, T. (1994). A nonparametric approach to pricing and hedging derivative securities via learning networks. *The Journal of Finance*, 49 (3), 851–889.

IBM Cloud Education (2020, Junio 3). *What is Artificial Intelligence (AI)?* IBM. Recuperado de: <https://www.ibm.com/cloud/learn/what-is-artificial-intelligence>

Ivaşcu Codruţ-Florin (2017). Predicting the direction of stock market prices using tree-based classifiers. *Expert Systems with Applications*. Volume 163, January 2021, 113799.

- Jansen Stefan (2018). *Hands-On Machine Learning Model for Stock Market Prediction: Design and implement investment strategies based on smart algorithms that learn from data using Python*. Packet Publishing Ltd. Brimingham-Mumbai.
- Khaidem. L., Saha. S., y Dey. R. S. (2016). Predicting the direction of stock market prices using random forest. *Applied Mathematical Finance* April 2016. Pág 1–20.
- Lin, Sheldon, X (2006). *Introductory Stochastic Analysis for Finance and Insurance*. Capítulo 5 (pp 131 - 144). Jonh Wikey & Sons. Estados Unidos.
- Liang, X., Zhang, H., Xiao, J., & Chen, Y. (2009). Improving option price forecasts with neural networks and support vector regressions. *Neurocomputing*, 72, 3055–3065.
- Loffler. Gunter y Peter N. Posch. (2007) *Credit risk modeling using Excel and VBA*. John Wiley & Son. [Capítulo 2]. Chichester. Inglaterra.
- Malkiel, B. G. (2003). The Efficient Market Hypothesis and Its Critics. *The Journal of Economic Perspectives*, 17(1), 59–82. <http://www.jstor.org/stable/3216840>
- Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, 7(1), 77–91. <https://doi.org/10.2307/2975974>
- Mondragón, E. (2021, Julio 15). Biografía de Jim Simons: ¿Cuál es su estilo de inversión? Recuperado de: <https://www.bbva.ch/noticia/biografia-de-jim-simons-cual-es-su-estilo-de-inversion/>
- Patterson. J., y Gibson. A. (2017) *Deep learning: A Practitioner's Approach*. O'Reilly Media. Primera Edición. Estados Unidos de Norte América.
- Philippe Jorion (2007) *Value at Risk: The new benchmark for managing financial risk*. Thrid Edition. The McGraw-Hill Companies, Inc. Estados Unidos de Norte América.

- Qiu., M y Song., Y. (2016) Predicting the Direction of Stock Market Index Movement Using an Optimized Artificial Neural Network Model. PLoS ONE 11(5): e0155133. <https://doi.org/10.1371/journal.pone.0155133>.
- Rebentrost. P, Gupt. B, y Bromley. T (2018) Quantum computational finance: Monte Carlo pricing of financial derivatives. Journal Phys. Rev. A. 98.
- Risk Metrics (1996) *Technical Document*. 4th Ed. Morgan Guaranty Trust Company of New York.
- Shreve, E, Steven (2004) Stochastic Calculus for Finance II – Continuous – Time Models. Capítulo 3: Brownian Motion (pp. 93-98). Springer. Nueva York. Estados Unidos.
- Stringham, P., Currott, N. (2015). Institutions and Organizations. Capítulo 14: On the Origins of Stock Markets (pp. 324–344). Nueva York. Estados Unidos.
- Yahoo Finance (Sin fecha). Energy Select Sector SPDR Fund (XLE). Yahoo Finance. Disponible en: <https://finance.yahoo.com/quote/XLE/holdings?p=XLE> [Consulta:2021, noviembre]
- Yahoo Finance (Sin fecha). *Communication Services Select Sector (IXC)*. Yahoo Finance. Disponible en: <https://finance.yahoo.com/quote/XLC/holdings?p=XLC> [Consulta:2021, noviembre]
- S&P Dow Jones índices (Sin fecha). *Energy Select Sector (IXE)*. S&P Global. Disponible en: <https://www.spglobal.com/spdji/en/indices/equity/energy-select-sector-index/#overview> [Consulta:2021, noviembre]
- S&P Dow Jones índices (Sin fecha). *Financial Select Sector (IXM)*. S&P Global. Disponible en: <https://www.spglobal.com/spdji/en/indices/equity/financial-select-sector-index/#data> [Consulta:2021, noviembre]

S&P Dow Jones índices (Sin fecha). *Communication Services Select Sector (IXC)*. S&P Global. Disponible en: <https://www.spglobal.com/spdji/en/indices/equity/communication-services-select-sector/#overview> [Consulta:2021, noviembre]

S&P Dow Jones índices (Sin fecha). *Health Care Select Sector Index (IXV)*. S&P Global. Disponible en: <https://www.spglobal.com/spdji/es/indices/equity/health-care-select-sector-index/#data> [Consulta: 2021, noviembre]

S&P Dow Jones índices (Sin fecha). *Technology Select Sector (IXTC)*. S&P Global. Disponible en: <https://www.spglobal.com/spdji/en/indices/equity/technology-select-sector-index/#data> [Consulta:2021, noviembre]

Anexos

- 1. Código Python – Movimiento Browniano Generalizado**
- 2. Código Python – Extreme Gradient Boosting**
- 3. Código Python – Redes Neuronales**