UNIVERSIDAD
NACIONAL
DE COLOMBIA

Pronóstico de la Volatilidad del Precio de las Acciones usando Deep Learning para Construcción de Portafolios

Brismar Alexander Osorio Gutierrez

Universidad Nacional de Colombia
Facultad de Minas, Área Curricular de Sistemas e Informática
Medellín, Colombia

2022

Pronóstico de la Volatilidad del Precio de las Acciones usando Deep Learning para Construcción de Portafolios

Brismar Alexander Osorio Gutierrez

Trabajo de investigación presentado como requisito parcial para optar al título de:

Magister en Ingeniería Analítica

Director:
Juan David Velásquez Henao, PhD

Línea de Investigación:
Analítica Financiera
Grupo de Investigación:
Big Data & Data Analytics

Universidad Nacional de Colombia
Facultad de Minas, Área Curricular de Sistemas e Informática
Medellín, Colombia
2022

(Dedicatoria o lema)

A mi mamá Lina Gutierrez, mi tío John Gutierrez, mi abuela Maria Mercedes, mi familia y mi novia Sofía Vega quienes siempre me han brindado el apoyo necesario para terminar este y otros proyectos.

Declaración de obra original

Yo declaro lo siguiente:

He leído el Acuerdo 035 de 2003 del Consejo Académico de la Universidad Nacional. «Reglamento sobre propiedad intelectual» y la Normatividad Nacional relacionada al respeto de los derechos de autor. Esta disertación representa mi trabajo original, excepto donde he reconocido las ideas, las palabras, o materiales de otros autores.

Cuando se han presentado ideas o palabras de otros autores en esta disertación, he realizado su respectivo reconocimiento aplicando correctamente los esquemas de citas y referencias bibliográficas en el estilo requerido.

He obtenido el permiso del autor o editor para incluir cualquier material con derechos de autor (por ejemplo, tablas, figuras, instrumentos de encuesta o grandes porciones de texto).

Por último, he sometido esta disertación a la herramienta de integridad académica, definida por la universidad.

Brismar Alexander Osorio Gutierrez

14/03/2023

Agradecimientos

A mi familia, por contar siempre con su apoyo.

A mi novia quien fue más que un apoyo emocional, también fue una guía en la elaboración de este trabajo.

A Juan David Velásquez Henao, PhD, director de este trabajo de grado, por su apoyo, guía y paciencia en todo este proceso.

Y a mis compañeros de trabajo quienes me brindaron su apoyo y conocimiento en diferentes etapas de este proyecto.

Resumen

Pronóstico de la volatilidad del precio de las acciones usando Deep Learning para construcción de portafolios.

En este trabajo, se presenta un conjunto de simulaciones para evaluar los diferentes portafolios financieros creados con acciones del mercado financiero colombiano, bajo diferentes técnicas de modelado de la varianza; esto con el fin de evaluar si un portafolio financiero generado con técnicas de *Deep Learning*, genera mejores rendimientos que un portafolio calculado de la misma forma, con técnicas más sencillas de *Machine Learning* y tradicionales de series de tiempo. En la construcción de estos portafolios, se crearon modelos ARIMA para el pronóstico de la media de cada activo y de modelos ARCH/GARCH, XGBoost y redes neuronales con diferentes arquitecturas, para la estimación de la varianza. Cada portafolio, varía uno a otro en el modelo que utiliza para el pronóstico de la varianza, sin embargo, todos los portafolios se realizan bajo el modelo de Markowitz de mínima varianza; esto genera 3 escenarios diferentes, sobre los cuales se realiza la comparación de rendimientos al final del periodo en el cual se realizaron las simulaciones. El rendimiento del portafolio evaluado, es decir el portafolio con *Deep Learning* en el modelo de varianza, genera una rentabilidad del 21.24%, demostrando un mejor resultado que los portafolios generados con modelos ARCH/GARCH con un 7.36% y con XGBoost con una rentabilidad del 11.61%. Lo anterior, indica que estos modelos pueden ser útiles para el pronóstico de la varianza, pues su aplicación mejora los rendimientos de un portafolio financiero construido con acciones del mercado financiero colombiano.

Palabras clave: *Deep Learning*, Markowitz, varianza, portafolios financieros.

Abstract

Volatility forecast of the price of the stocks using Deep Learning for portfolio construction.

In this work, a set of simulations is presented to evaluate different financial portfolios created with shares of the Colombian financial market, under different variance modeling techniques, this, in order to evaluate whether a financial portfolio generated with *Deep Learning* techniques, generates better returns than a portfolio calculated in the same way, with simpler techniques of *Machine Learning* and traditional time series. In the construction of these portfolios, ARIMA models were created to forecast the mean of each asset and ARCH/GARCH models, XGBoost and neural networks with different architectures for variance estimation. Each portfolio varies from one to another in the model used for variance forecasting, however, all the portfolios are carried out under the Markowitz minimum variance model; this generates 3 different scenarios, on which the comparison of returns at the end of the period in which the simulations were performed is made. The performance of the evaluated portfolio, i.e. the portfolio with *Deep Learning* in the variance model, generates a return of 21.24%, showing a better result than the portfolios generated with ARCH/GARCH models with 7.36% and with XGBoost with a return of 11.61%. The above, indicates that these models can be useful for variance forecasting, since their application improves the returns of a financial portfolio built with shares of the Colombian financial market.

Keywords: *Deep Learning*, Markowitz, variance, financial portfolios.

Contenido

Resumen **9**
Abstract **10**
Lista de Figuras **13**
Lista de Tablas **14**
1. Introducción **15**
1.1 Definición del problema real 15
1.2 Definición del problema de analítica 17
1.3 Revisión de la Literatura 18
1.4 Problema por abordar en este trabajo final 21
1.5 Hipótesis 22
1.6 Objetivos 22
1.6.1 Objetivo general 22
1.6.2 Objetivos específicos 22
1.7 Información disponible 22
1.7.1 Selección de acciones 23
1.8 Mapa del documento 24
2 Metodología empleada **25**
2.1 Escenarios planteados 28
2.2 Modelo de pronóstico de la media 28
2.3 Modelo de pronóstico de la varianza 29
2.4 Modelo de conformación de portafolios de Markowitz 30
2.5 Balanceo y rebalanceo del portafolio 30
2.6 Simulaciones 30
3 Resultados y discusión **33**
3.1 Selección de periodos de entrenamiento y validación 33
3.2 Escenarios y simulación 34
3.3 Pronóstico de media en la simulación 34
3.3.1 Entendimiento de los datos de retorno 34
3.3.2 Variables independiente y dependiente para el modelo de media 37
3.3.3 Modelo y residuales 37
3.4 Pronóstico de varianza en la simulación 38

3.4.1 Entendimiento de los datos de varianza 38
3.4.2 Medida de desempeño 43
3.4.3 Pronostico ARCH/GARCH 43
3.4.4 Pronostico *Machine Learning* 45
3.4.5 Pronostico *Deep Learning* 48
3.4.5.1 Acotación 48
3.4.5.2 Arquitecturas 48
3.4.5.3 Escala de datos 51
3.4.5.4 Variables independientes para el modelo de varianza 51
3.4.5.5 Experimentos y resultados 51
3.5 Compra y venta de acciones según el portafolio optimo 58
3.6 Resultado de escenarios 59
4 Conclusiones 61
4.1 Pregunta de investigación 61
4.2 Cumplimiento de objetivos 62
4.2.1 Objetivo 1: Estimar la rentabilidad de un portafolio en el tiempo usando la metodología de Markowitz 62
4.2.2 Objetivo 2: Desarrollar modelos de pronóstico de la volatilidad de las acciones del portafolio utilizando *Deep Learning* 62
4.2.3 Objetivo 3: Estimar la rentabilidad del portafolio generado con el *Deep Learning* pronosticado 62
4.2.4 Objetivo 4: Comparar resultados de los portafolios 63
4.2.5 Objetivo general 63
5. Referencias 65

Lista de Figuras

Figura 1: Arquitectura interna de periodo de simulación (Pasos 2, 3 y 4). 26
Figura 2: Flujo macro de balanceo del portafolio (Paso 5). 27
Figura 3: Rebalanceo del portafolio (Paso 5). 27
Figura 4: Arquitectura interna de escenario. 31
Figura 5: Retornos logarítmicos ISA. 35
Figura 6: Retornos logarítmicos ECOPETROL. 35
Figura 7: Autocorrelación y autocorrelación parcial de las rentabilidades ISA. 36
Figura 8: Autocorrelación y autocorrelación parcial de las rentabilidades ECOPETROL. 36
Figura 9: Residuales del modelo de media de ISA. 38
Figura 10: Residuales del modelo de media de ECOPETROL. 38
Figura 11: Autocorrelación y autocorrelación parcial de los residuales de ISA. 39
Figura 12: Autocorrelación y autocorrelación parcial de los residuales de ECOPETROL.39
Figura 13: Residuales al cuadrado del modelo de media de ISA. 40
Figura 14: Residuales al cuadrado del modelo de media de ECOPETROL. 40
Figura 15: Autocorrelación parcial de los residuales al cuadrado del modelo de media de ISA. 41
Figura 16: Autocorrelación parcial de los residuales al cuadrado del modelo de media de ECOPETROL. 41
Figura 17: Comparación valores absolutos de los residuales del modelo de medía, frente a los residuales al cuadrado. 42
Figura 18: Red Neuronal Recurrente Simple/LSTM. 49
Figura 19: Stacked LSTM. 50
Figura 20: Histórico de rendimientos de los portafolios. 60

Lista de Tablas

Tabla 1: Acciones que pertenecieron al índice COLCAP en el año 2019. 23
Tabla 2: Escenarios planteados para la comparación y revisión de rentabilidades..... 28
Tabla 3: Resultados de los diferentes modelos ARCH/GARCH de ISA. 44
Tabla 4: Mejores modelos ARCH/GARCH. 45
Tabla 5: Espacio de hiperparametros para XGBoost. 46
Tabla 6: Mejores Modelos XGBoost. 47
Tabla 7: Experimentos iniciales. 52
Tabla 8: Top 10, mejores y peores modelos ISA. 53
Tabla 9: Top 10, mejores y peores modelos ECOPETROL. 53
Tabla 10: Experimentos secundarios. 54
Tabla 11: Resultados experimentos secundarios ISA. 55
Tabla 12: Resultados experimentos secundarios ECOPETROL. 56
Tabla 13: *Iteraciones en los modelos de Deep Learning.* 56
Tabla 14: Resultados LSTM. 57
Tabla 15: Comparación de modelos de varianza. 58
Tabla 16: Simulación de muestra para un periodo con operaciones de compra y venta.. 59
Tabla 17: Comparación de resultados de los portafolios. 60

1. Introducción

En este documento se presenta el Trabajo Final de Maestría titulado “Pronóstico de la volatilidad del precio de las acciones usando *Deep Learning*, para construcción de portafolios”, el cual, es desarrollado a partir de una simplificación de la metodología CRISP-DM.

El resto de este capítulo, está organizado de la siguiente manera: En la primera sección se hablará de la definición del problema real, en la que se abordara el modelo de selección de portafolios, las diferentes aproximaciones que este ha tenido desde lo tradicional y el por qué es importante revisar modelos de *Machine Learning* como una alternativa a la conformación de portafolios. En la segunda sección, se relacionará el problema real con un problema de analítica, mencionando la forma en que se han pronosticado las volatilidades de los activos financieros. La tercera sección, se usará para la revisión de la literatura, revisando en esta, la forma en la que se ha calculado la volatilidad y cuál ha sido el uso de *Machine Learning* en la conformación de portafolios. La cuarta sección se usará para la definición del problema a abordar, siguiendo con la definición de hipótesis y objetivos en la quinta y sexta sección. Continuando con la séptima sección, en la cual se definirá la información con la que se contará y finalizando con la octava sección, donde se revisará el mapa del documento para este Trabajo Final de Maestría.

1.1 Definición del problema real

En el mundo se ha estudiado en gran medida la conformación de portafolios financieros desde diferentes metodologías con el fin de encontrar portafolios óptimos. Estas metodologías, usan, principalmente, variables como la rentabilidad esperada y la volatilidad de los activos que conforman el portafolio, buscando aumentar la rentabilidad de los inversionistas frente a variaciones en los criterios de selección de activos y participación de estos en el portafolio.

Los inversionistas, hacen uso de modelos y metodologías para el cálculo de las volatilidades, pérdidas y ganancias de los activos financieros, buscando encontrar aquellos activos que en conjunto maximicen su rentabilidad. Dado lo anterior, se parte de la selección de activos financieros, con los cuales se desea trabajar (acciones, divisas, derivados…) y a partir de esto, y el conocimiento experto del inversionista, se decide la metodología para

la conformación de portafolios, teniendo la metodología seleccionada, se implementa y se simula un comportamiento del portafolio seleccionado para la estimación de pérdidas y ganancias en un período de tiempo.

Los inversionistas cuentan con diferentes metodologías para la conformación del portafolio óptimo. Markowitz (1959) propone crear portafolios óptimos a partir de la minimización del riesgo para una determinada rentabilidad. En esta metodología, se tiene en cuenta la rentabilidad esperada y la volatilidad, en ella, se parte de la conformación de la frontera eficiente con el fin de encontrar el punto de mínima varianza para una rentabilidad definida, dado que, al encontrar este punto, no existirá portafolio con mayor rentabilidad para esa varianza.

Sharpe (1964) propuso modificar la metodología de portafolios de Markowitz, al agregar un activo libre de riesgo a los activos ya abarcados en la conformación del portafolio, y al generar una línea de riesgo de mercado tangente a la frontera eficiente de Markowitz para la maximización del portafolio. Esta línea de mercado, es conformada por activos financieros representativos del mercado y consta de una relación lineal entre la rentabilidad del mercado y la volatilidad de este mismo.

El modelo de Black & Litterman (1991), surge como una mejora de los modelos anteriores, principalmente a la metodología de portafolios de Markowitz, y tiene como base importante la ley de Bayes. Este modelo, a diferencia del modelo de Markowitz, tiene en cuenta las expectativas de la rentabilidad de los activos que conformaran el portafolio; permitiendo agregar como variable, la subjetividad que se tiene sobre el comportamiento futuro de los diferentes activos financieros.

Las diferentes metodologías revisadas, hacen uso de la rentabilidad esperada y la volatilidad de cada activo financiero para la conformación de los diferentes portafolios. El cálculo de estas variables se hace de forma clásica utilizando la estadística tradicional y los supuestos relacionados según la metodología empleada. Estos cálculos han servido para la conformación de los portafolios, pero estas metodologías, al seguir una serie de supuestos como requisito para su uso, pueden dejar de lado opciones que podrían obtener mejores resultados, además, suelen ser metodologías cuyo supuesto es que la rentabilidad y la volatilidad de los diferentes activos es constante; este caso no pasa en un mercado

donde las diferentes noticias y sucesos pueden afectar considerablemente el comportamiento de los activos financieros.

Eugene Fama (1991) propuso la hipótesis de los mercados eficientes (EMH) la cual indica que, el precio de los activos, refleja toda la información existente de la cual se puede concluir que, como lo reviso Ante (2022), algo tan simple como un tweet, puede provocar un sinfín de reacciones en estos mercados, causando así, cambios en los comportamientos de los diferentes activos que lo conforman. Este cambio constante, dificulta la estimación de las diferentes variables necesarias para la conformación de portafolios, dado que el comportamiento de un activo hoy, podría ser muy diferente al comportamiento del mismo activo mañana. Este problema, es difícil de captar con modelos tradicionales de estadística, dado que, estos, en algunas ocasiones, no captan con rapidez estos cambios y dependen mucho del experto y su creatividad a la hora de realizar la extracción y selección de características.

La evaluación de los portafolios financieros con diferentes técnicas tradicionales para el cálculo de la volatilidad y la rentabilidad esperada, ha sido un tema muy estudiado a lo largo del tiempo, sin embargo, hallar un modelo simple que se ajuste a todos los activos que pueden componer un portafolio financiero, puede ser una tarea complicada y extensa, dado que, para cada activo, se debe realizar el análisis estadístico y hallar patrones y condiciones con el fin de lograr un buen modelo. Este problema, puede solventarse utilizando modelos de Deep Learning, pues estos, con una buena cantidad de información, logran captar de forma automática las variables y relaciones de estas, así como la identificación de nuevos patrones provocados por cambios en el mercado, por lo anterior, y al ser modelos que han funcionado bien con problemas no lineales complejos, la implementación de estos, podrían permitir la generación de mejores modelos con un fácil reentrenamiento.

En resumen, el problema real abordado en este Trabajo Final de Maestría, es determinar si el uso de técnicas de *Deep Learning* para el pronóstico de las volatilidades puede mejorar la rentabilidad de un portafolio financiero conformado a partir del mercado colombiano.

1.2 Definición del problema de analítica

El problema de conformación de portafolios, puede definirse matemáticamente como un problema de optimización, puesto que, para conformar el portafolio optimo se busca

minimizar la varianza del portafolio a determinada rentabilidad. Una de las formas en las que se puede lograr esto, es iterar sobre los diferentes pesos de los activos que conforman el portafolio, hallando así la combinación que minimiza la varianza.

Markowitz (1959), Sharpe (1964) y Merton (1971) proponen diferentes formas para abordar este problema, sin embargo, coinciden en que es necesario el uso de variables que puedan definir la rentabilidad, el riesgo del portafolio y sus activos. Dado lo anterior, para desarrollar la optimización de un portafolio, se debe estimar la medía y la varianza de cada acción.

Para realizar el cálculo de estas variables, se puede hacer uso de diferentes técnicas. Tradicionalmente, se han usado modelos de series de tiempo para el pronóstico de ambas, utilizando modelos ARIMA, para el pronóstico de la medía y modelos de la familia ARCH, para el pronóstico de la varianza, sin embargo, pueden ser considerados como modelos de regresión y, por ende, se pueden utilizar modelos de analítica predictiva que puedan generar mejores resultados.

En resumen, el problema de analítica a abordar en este Trabajo Final de Maestría, es el desarrollo de un modelo de simulación que permita la conformación de portafolios en el tiempo, los cuales utilizan técnicas de *Deep Learning* para el pronóstico de la volatilidad de los activos. Esto tiene como fin, poder determinar si, el uso de estas técnicas de pronóstico, permite obtener portafolios más rentables.

1.3 Revisión de la Literatura

En la literatura, existen diferentes tipos de volatilidades calculadas: la volatilidad histórica, cuyo cálculo se hace a partir de la varianza histórica de los activos y está formulada de la siguiente forma:

$$\sigma^2 = \frac{\sum_1^n (x_i - \bar{x})^2}{n}$$

Donde:

σ^2 : Volatilidad del activo x

x_i : Observación número i de la rentabilidad del activo x

$\bar{x}$: Promedio de rentabilidad del activo x

n : Numero de observaciones del activo x

Volatilidades calculadas con modelos tradicionales o estándar, como lo son los modelos con Heteroscedasticidad condicional autorregresiva (ARCH), los cuales tienen en cuenta que la varianza no es constante en el tiempo y dependen de su historia para generar el pronóstico Tsay (2001).

La principal idea de estos modelos, es que los retornos a de los activos financieros no están correlacionados, pero son dependientes. Esta dependencia, puede ser descrita como una función cuadrática simple de la forma:

$a_t = \sigma_t \epsilon_t$, $\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \cdots + \alpha_m a_{t-m}^2$;

En el cual para el caso de un ARCH(1)

$a_t = \sigma_t \epsilon_t$; $\sigma_t^2 = \alpha_0 + \alpha_1 \alpha_{t-1}^2$;

Donde $\alpha_0 > 0$ y $\alpha_1 \geq 0$

Con:

$E(a_t) = E[E(a_t|F_{t-1})] = E[\sigma_t E(\epsilon_t)] = 0$;

$Var(a_t) = E(a_t^2) = E[E(a_t^2|F_{t-1})] = E(\alpha_0 + \alpha_1 a_{t-1}^2) = \alpha_0 + \alpha_1 E(a_{t-1}^2)$;

Como evolución de los modelos anteriores, se presenta la generalización de los modelos con heterocedasticidad condicional autorregresiva (GARCH), en los cuales, el pronóstico no sólo depende de los errores pasados, sino también, de sus varianzas anteriores Tsay (2001)

Con una serie de retornos logarítmicos, se tiene que: $a_t = r_t - \mu_t$;

$a_t = \sigma_t \epsilon_t$; $\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-1}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2$;

$E(a_t^2) = \frac{\alpha_0}{1 - \sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i)}$;

Igualmente, se han probado técnicas de *Machine Learning* para el cálculo de la volatilidad. Estas técnicas, buscan solventar los problemas que se puedan encontrar en las técnicas tradicionales y se enfocan en el resultado del modelo. Christensen, Siggaard & Veliyev (2021), proponen calcular la volatilidad a partir de un modelo de redes neuronales

artificiales, dejando la mayoría de hiperparámetros por defecto, a como se sugiere desde la literatura; tomando como referencia la arquitectura definida como “pirámide geométrica” con función de activación L-ReLU y, en la cual, se usa como input del modelo, las varianzas realizadas de cada activo basado en los logaritmos de los retornos cada 5 minutos.

La ecuación general para la capa l-ésima de una red neuronal, se describe como:

$$a_t^{\theta_{l+1},b_{l+1}} = g_l(\sum_{j=1}^{N_l} \theta_j^{(l)} a_t^{\theta_l,b_l}); 1 \leq l \leq L$$

Donde:

g_l: Función de activación.

$\theta^{(l)}$: Matriz de pesos.

$b^{(l)}$: Error. Sirve como un umbral de activación para la siguiente neurona.

N_l: Numero de neuronas ocultas.

$a_t^{\theta_{l+1},b_{l+1}}$: Predicción.

Aguilar (2010), propone un modelo basado en la relación a partir de la media y varianza del modelo ARCH, con un modelo de redes neuronales con perceptrón multicapa, el cual es utilizado para modelar la volatilidad, en ocasiones en que la serie presenta heterocedasticidad, por lo que el modelo busca capturar la no linealidad de la serie.

Para esto, define un modelo MLP-ARCH de la forma:

$y_t = \beta_* + \sum_{h=1}^{H} \beta_h \mu\,(\alpha_{*,h} + \sum_{i=1}^{P} \alpha_{i,h} y_{t-i}) + a_t;$

Donde $a_t = \sigma_t \varepsilon_t$ con:

$\sigma_t^2 = w_0 + \sum_{i=1}^{q} w_i a_{t-i}^2;$

Carr, Wi & Zhang (2020), proponen hacer uso de algoritmos de regresión de *Machine Learning* sobre el precio de las opciones sobre los activos, para calcular la volatilidad futura de estos. En estas regresiones, se tuvieron en cuenta características, como: fecha de trade, fecha de expiración, precio spot (Precio actual), tasa de interés a término corto, tipo de opción financiera, bid, ask y se hizo uso específicamente de la Regresión Lineal, Ridge

Regression, feedforward neural network regression con función de activación ReLU y random forest regression demostrando que, si la función es lineal a trozos, se puede llegar a mejores resultados con estos algoritmos para el cálculo de la volatilidad.

En el caso colombiano, el uso de *Machine Learning* en el sector financiero, se ha empleado en diferentes ámbitos. Echeverri (2021), estudia la conformación de portafolios financieros, usando modelos de *Machine Learning* para pronosticar el movimiento de las acciones y así conformar el portafolio de inversión. Por otro lado, siguiendo con la aplicación en el mercado de valores, Villada, Muñoz & García (2012), hacen uso de redes neuronales para el pronóstico de los precios de dos acciones importantes para el mercado de valores colombiano. Carmona (2022), utiliza redes neuronales para el pronóstico de la inflación en Colombia.

1.4 Problema por abordar en este trabajo final

En los trabajos anteriores, se han revisado diferentes modelos para el pronóstico de los rendimientos y la varianza de los diferentes activos que pueden conformar un portafolio, variando entre modelos, técnicas y características utilizadas. La mayoría de estos, se han estudiado en mercados financieros desarrollados, como lo es la bolsa de valores de Nueva York.

Los estudios realizados para el mercado colombiano, se han enfocado, igualmente, en el pronóstico de las variables que influyen en la conformación de portafolios, sin embargo, la conformación de estos y validación de rendimientos no ha sido el tema más desarrollado.

Dado lo anterior, en este trabajo se plantea evaluar si las técnicas de *Deep Learning* para el pronóstico de la varianza, permiten portafolios más rentables, utilizando el modelo de mínima varianza de Markowitz. Esta idea se diferencia de las otras al revisar el impacto que tiene este tipo de pronóstico en el rendimiento final de un portafolio, sin variar la metodología propuesta por Markowitz.

Este trabajo, da continuidad a la tesis de Laura Cristina Echeverri Sánchez, abordando la conformación de estos portafolios financieros con el uso de técnicas de *Deep Learning* para el pronóstico de la varianza, problema que no se revisó en la tesis mencionada, pues en esta la conformación de portafolios financieros, se abordó desde el pronóstico de la media.

1.5 Hipótesis

La hipótesis de este trabajo es la siguiente: ¿Es posible mejorar la rentabilidad de un portafolio de activos construido utilizando la metodología de Markowitz cuando la volatilidad es pronosticada usando *Deep Learning*?

1.6 Objetivos

1.6.1 Objetivo general

El objetivo final de este Trabajo Final de Maestría es comparar la rentabilidad de un portafolio tradicional conformado con el modelo de Markowitz cuyas variables insumo se calculan de forma tradicional, con un portafolio que utiliza *Deep Learning* para pronosticar la volatilidad de sus activos y conformar el portafolio con el modelo de Markowitz.

1.6.2 Objetivos específicos

Los objetivos específicos de este Trabajo Final de Maestría son los siguientes:

1. Estimar la rentabilidad de un portafolio en el tiempo usando la metodología de Markowitz
2. Desarrollar modelos de pronóstico de la volatilidad de las acciones del portafolio utilizando *Deep Learning*
3. Estimar la rentabilidad del portafolio generado con el *Deep Learning* pronosticado
4. Comparar resultados de los portafolios

1.7 Información disponible

Para verificar y comparar los rendimientos de los diferentes portafolios creados con técnicas tradicionales y técnicas de *Deep Learning*, se hará uso de la información del mercado de valores colombiano administrada por la BVC (Bolsa de Valores de Colombia). Dado que, esta información está disponible en la página de la BVC, se implementarán técnicas de Webscraping para extraer la información de cada uno de los activos seleccionados.

Este proceso, se hizo con ayuda de la librería Selenium, generando un algoritmo de Webscraping en Python, usado para la extracción de la información histórica de cada uno de los activos a utilizar en los diferentes portafolios.

1.7.1 Selección de acciones

Para la selección de acciones, se usó como referencia el índice bursátil COLCAP, dado que, este refleja las acciones más liquidas del mercado, conformándose por mínimo 20 acciones, de 20 emisores diferentes. De este índice, se seleccionaron las acciones que se utilizaran en este trabajo y se pueden observar en la Tabla 1. Dado lo anterior, se hizo uso de la mayoría de los activos que conformaron este índice a finales del año 2019, ya que, por la pandemia, se acotó este periodo de entrenamiento y simulación a los años 2013-2019.

Tabla 1: Acciones que pertenecieron al índice COLCAP en el año 2019.

Acción	Decisión
ECOPETROL	Seleccionada
PFBCOLOM	Seleccionada
ISA	Seleccionada
GRUPOSURA	Seleccionada
BCOLOMBIA	Seleccionada
GEB	Se descarto por falta de información en el año 2013
NUTRESA	Seleccionada
PFAVAL	Seleccionada
GRUPOARGOS	Seleccionada
PFDAVVNDA	Seleccionada
CEMARGOS	Seleccionada
BOGOTA	Seleccionada
PFGRUPSURA	Seleccionada
PFGRUPOARG	Seleccionada
CORFICOLCF	Seleccionada
CELSIA	Seleccionada
PROMIGAS	Seleccionada
EXITO	Seleccionada
CNEC	Seleccionada
PFCEMARGOS	Seleccionada
CLH	Seleccionada
PFAVH	Seleccionada
MINEROS	Seleccionada
BVC	Seleccionada

1.8 Mapa del documento

El resto de este documento está organizado como se muestra a continuación. En el Capítulo 2, se presentará la metodología empleada para el desarrollo de los objetivos e hipótesis del trabajo. Siguiendo con el Capítulo 3, en el cual se revisarán los resultados del desarrollo de la metodología propuesta. Finalizando con el Capítulo 4, donde se harán las comparaciones y las conclusiones del desarrollo de la investigación.

2 Metodología empleada

En este capítulo, se describe la metodología empleada con la que se plantea resolver la pregunta de investigación. Para esto, se seguirán los siguientes pasos:

- **Paso 1:** Buscando comparar los portafolios con diferentes tipos de modelos para el pronóstico de su varianza, se definen escenarios a ejecutar. Estos escenarios, contarán con un periodo de simulación en el cual se realizarán diferentes actividades.
- **Paso 2:** Luego de definir los escenarios, se comienza con la generación del mejor modelo ARMA de series de tiempo, para el pronóstico de la rentabilidad; este modelo servirá como entrada al modelo de portafolio óptimo y, además, ayudará a generar los residuales necesarios para el pronóstico de la varianza.
- **Paso 3:** Siguiendo con el flujo, se desarrollarán los diferentes tipos de modelos para el pronóstico de la varianza. Estos modelos, serán modelos de la familia ARCH, *Machine Learning* o *Deep Learning*, según el escenario que se esté ejecutando.
- **Paso 4:** Al tener la media y varianza de cada activo, se procederá a conformar el portafolio de mínima varianza de Markowitz, como se puede observar en la Figura 1.
- **Paso 5:** Teniendo el portafolio de mínima varianza, se continua con el balanceo del portafolio, observado de forma gráfica en la Figura 2, o rebalanceo en la Figura 3, dependiendo de la simulación que se esté ejecutando.
- **Paso 6:** Para finalizar, se realizarán simulaciones diarias por un periodo de tiempo, en las cuales se ejecutarán los pasos 2, 3, 4 y 5 en cada simulación. Las diferentes simulaciones generaran la rentabilidad o pérdida del portafolio.

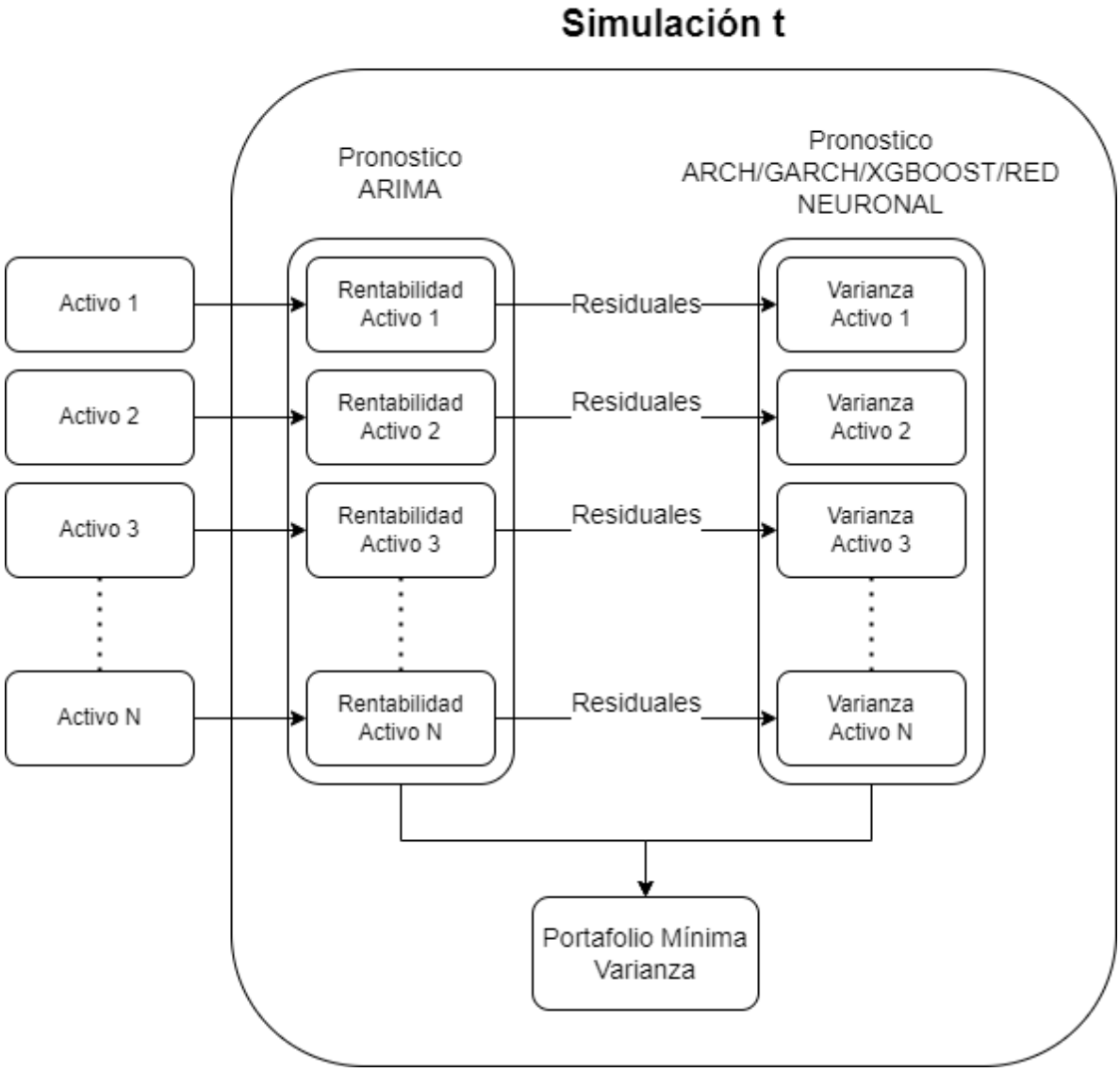

Figura 1: Arquitectura interna de periodo de simulación (Pasos 2, 3 y 4).

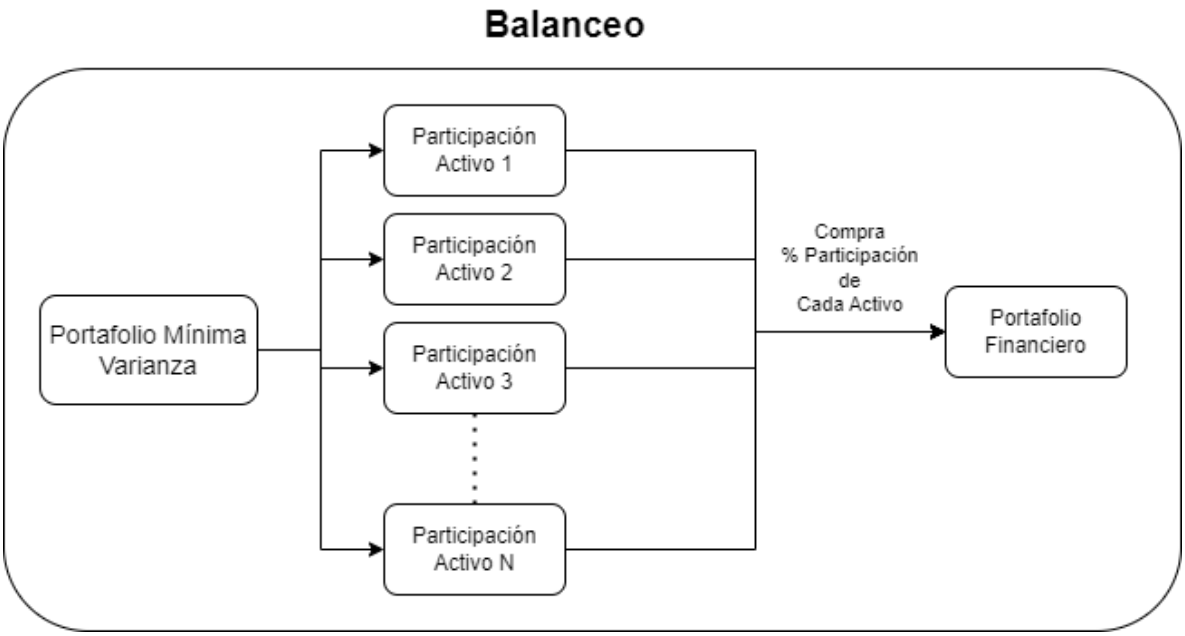

Figura 2: Flujo macro de balanceo del portafolio (Paso 5).

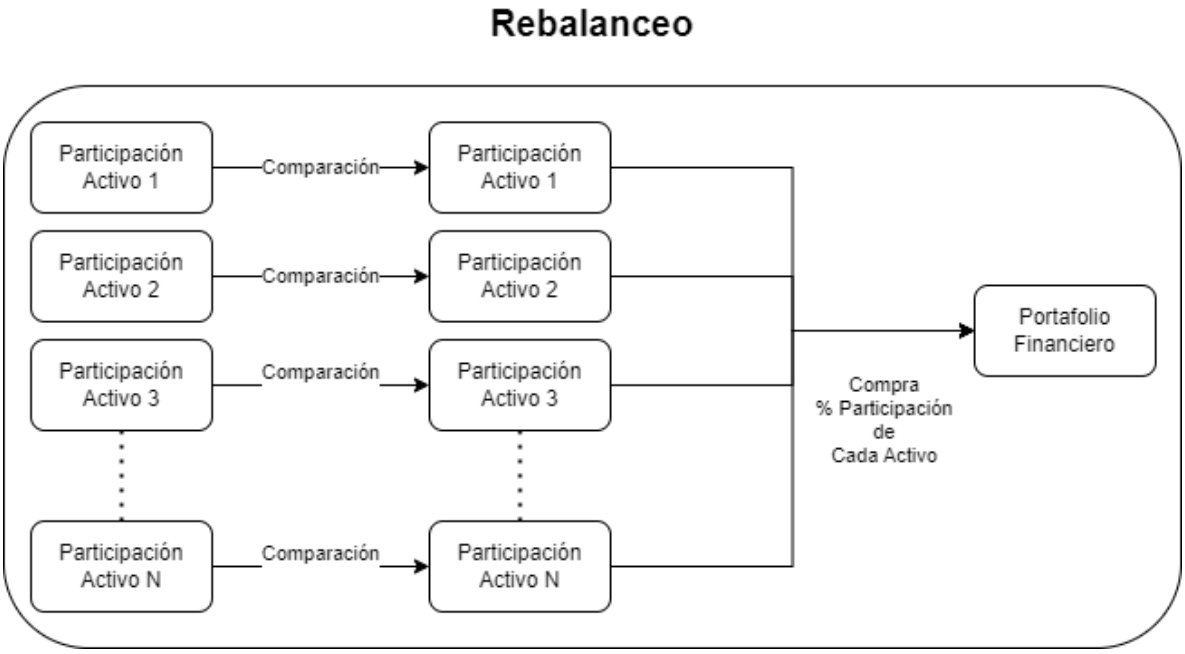

Figura 3: Rebalanceo del portafolio (Paso 5).

2.1 Escenarios planteados

Con el fin de analizar la rentabilidad generada por portafolios óptimos, utilizando modelos de *Deep Learning*, se planteó generar diferentes escenarios que se muestran en la Tabla 2, los cuales varían unos a otros en los modelos utilizados para el cálculo de la varianza de los diferentes activos.

Tabla 2: Escenarios planteados para la comparación y revisión de rentabilidades.

Escenario	Modelo Medía	Modelo Varianza
1	ARMA	GARCH
2	ARMA	ML
3	ARMA	DL

Dado que, con este desarrollo se espera tener portafolios equiparables, los cuales permitan validar que el pronóstico de la varianza con modelos de *Deep Learning* mejoren los rendimientos del portafolio. Se hará uso del mismo modelo ARMA de cada activo, para el pronóstico de la media en los 3 escenarios.

2.2 Modelo de pronóstico de la media

En el caso del modelo de media, se procederá al cálculo de la rentabilidad logarítmica de cada uno de los activos que podrían conformar el portafolio financiero.

$$R_t = \ln(\frac{P_t}{P_{t-1}})$$

La rentabilidad de cada uno de estos activos en los diferentes periodos de tiempo seleccionados para el análisis, generaran las diferentes series de tiempo que se usaran como entrada para el modelo tradicional ARIMA. En este sentido, para este modelo, sólo se tendrán en cuenta dos tipos de variables independientes, los valores rezagados y las medias móviles de la serie. De estos dos tipos, saldrán diferentes combinaciones de orden de rezago y medias móviles para la conformación del modelo.

Ya que se hará uso de diferentes combinaciones de orden autoregresivo y de medias móviles, se plantea utilizar el criterio de información de akaike (AIC), dado que, este permite comparar los diferentes modelos generados e indicar cual de estos es mejor sin revisar si el modelo de por sí es bueno.

$$AIC = 2k - 2\ln(L)$$

2.3 Modelo de pronóstico de la varianza

Para el pronóstico de la varianza, se realizarán 3 tipos de modelos (tradicionales, *Machine Learning* y *Deep Learning*), cada uno de estos, tendrá la misma variable a predecir, pero su forma de cálculo de las variables independientes, cambiará según el caso.

Estos modelos, harán uso de una de las salidas generadas en el modelo de pronóstico de media, el error calculado como el valor predicho, sobre el periodo de entrenamiento menos el valor real.

$$e_i = y_{pred} - y_{real}$$

Uno de estos modelos, seguirá el enfoque tradicional para el pronóstico de la varianza, como lo son, los modelos ARCH y GARCH.

En cuanto al modelo de *Machine Learning*, la dinámica se mantendrá, haciendo una pequeña modificación en las variables a tener en cuenta en este tipo de modelos. Por lo anterior, se generarán los rezagos y medias móviles de las diferentes combinaciones de modelos a probar, y se tomarán como variables independientes en el modelo de regresión resuelto con el modelo XGBoost.

$$F_m(X) = F_{m-1}(X) + \alpha_m h_m(X, r_{m-1})$$

$$\arg\ min_\alpha = \sum_{i=1}^{m} L(Y_i\,, F_{i-1}(X_i) + \alpha h_i(X_i, r_{i-1}))$$

Dónde:

$L\big(Y, F(X)\big)$: Función de perdida diferenciable.

α_i: Parámetro de regularización.

r_i: Residual del i^{er}árbol.

h_i: Función para predecir residuales usando X para el i^{er}arbol.

Para el caso de los modelos de *Deep Learning*, la metodología variara significativamente. Dado que, este tipo de modelos tiene un costo computacional muy alto. Se seleccionarán algunas de las acciones con mayor importancia para la elaboración de experimentos y acotación de parámetros e hiperparámetros a probar sobre todo el conjunto de series.

En este sentido, se plantea seleccionar alguna de estas series y modelar con diferentes arquitecturas de redes neuronales, variables independientes, hiperparámetros, escalas de estandarización y funciones de pérdida para evaluar cuáles de estos impactan considerablemente en el resultado de los pronósticos.

2.4 Modelo de conformación de portafolios de Markowitz

La conformación de portafolios óptima, se realizará siguiendo el modelo de portafolio de mínima varianza de Markowitz. Este modelo, será el mismo para todas las simulaciones, la única diferencia de una simulación a otra, serán los valores de la medía y varianza de cada activo en el momento en que se está simulando.

2.5 Balanceo y rebalanceo del portafolio

El balanceo del portafolio, consiste en realizar las operaciones de compra de cada acción, buscando tener la participación generada por el portafolio de mínima varianza. En el caso de contar con un portafolio, es decir, que la simulación que se esté ejecutando no sea la primera, se realizarán operaciones de compra y venta de cada acción con el mismo fin.

2.6 Simulaciones

Buscando evaluar el rendimiento de cada escenario a ejecutar, se propone realizar la simulación de cada escenario de forma diaria por un periodo de tiempo, realizando el rebalanceo del portafolio con cada conformación óptima de este, según el pronóstico diario de la media y varianza de cada uno de los activos financieros considerados. El rebalanceo

del portafolio, generará pérdidas o ganancias, las cuales se reinvertirán en el período de simulación y se evaluarán los portafolios al final de este periodo, revisando así, si este generó ganancias o pérdidas en toda su vida, permitiendo comparar estas frente a los diferentes escenarios a ejecutar.

Dado que, cada escenario contiene un número de simulaciones, en la Figura 4 se puede observar el flujo interno de cada uno de estos.

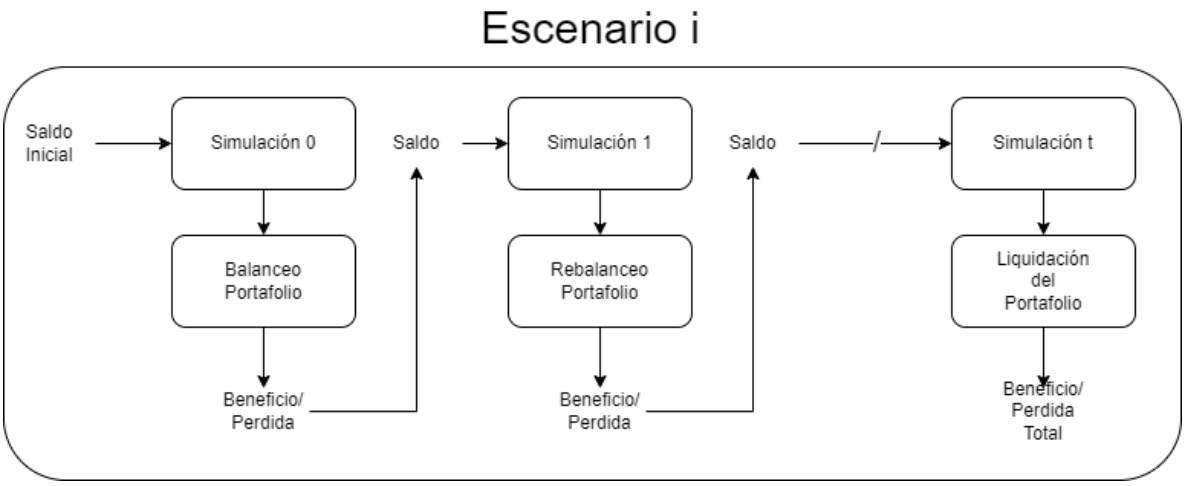

Figura 4: Arquitectura interna de escenario.

3 Resultados y discusión

En este capítulo, se hará uso de la metodología propuesta anteriormente junto con la metodología CRISP-DM para el completo desarrollo de los diferentes objetivos definidos al inicio. La metodología CRISP-DM, ayudará principalmente en la elaboración de los diferentes modelos de pronóstico de la varianza, que se usarán en cada simulación.

Inicialmente, se hablará del flujo general que se seguirá en cada simulación, para así seguir detallando el proceso de pronóstico de media requerido como parámetro en el modelo de conformación de portafolios de Markowitz y cómo se hará el cáculo de este para cada periodo a simular.

Luego de esto, se entrará a revisar la forma en que se acotaron los diferentes factores que pueden influenciar en la elaboración de los modelos de pronóstico de varianza, en todo el conjunto de series de tiempo generados para la conformación del portafolio y su posterior simulación. Usando esta acotación, se procederá a hablar del proceso creado para la búsqueda de los mejores modelos para todos los activos que se tienen en consideración.

Por último, se revisará cómo se realizó la conformación del portafolio siguiendo los modelos de portafolios de mínima varianza de Markowitz y evaluando la rentabilidad o pérdida generada por el portafolio en cada período de simulación y, al finalizar, todo el proceso.

3.1 Selección de periodos de entrenamiento y validación

Dado que, en 2020 se tuvo una pandemia mundial que afectó considerablemente los mercados financieros, se optó por tomar un período de observación para el total del análisis del año 2013 al año 2019, dejando el periodo conformado del año 2013 al 2018 como período de entrenamiento de los modelos y el año 2019, por completo, para la simulación con los modelos encontrados.

Desde el entrenamiento de los modelos, se contó con otra separación para la optimización de hiperparámetros y rezagos seleccionados para el modelo, dejando un conjunto de entrenamiento, conjunto de prueba y por último, de validación.

3.2 Escenarios y simulación

Cada uno de los escenarios ejecutados, constará de 250 simulaciones, aproximadamente; una por cada día del año. Siendo una simulación conformada por un flujo completo, en el cual se pronosticará la media, varianza y el portafolio de mínima varianza del día en que se esté ejecutando la simulación. La ejecución de cada simulación, generará los beneficios o pérdidas de cada día, reinvirtiendo los beneficios para el portafolio de la siguiente simulación. Al finalizar el periodo de simulaciones, se liquidará el portafolio, dejando así, el beneficio total en un año del escenario ejecutado.

El año que se tendrá en cuenta para las simulaciones será el 2019, año antes de la pandemia, que afectó el comportamiento regular de los mercados financieros.

3.3 Pronóstico de media en la simulación

3.3.1 Entendimiento de los datos de retorno

Con el fin de analizar el posible comportamiento de las 23 acciones a considerar en el modelo, se acota el análisis exploratorio a sólo dos de estas, siendo ISA y ECOPETROL las acciones seleccionadas, dado que, son dos de las acciones con mayor importancia en el índice COLCAP de referencia sobre el que se tomaron las diferentes acciones para el desarrollo.

Como se puede observar en las Figuras 5 y 6, en las que se muestra la rentabilidad de ambas acciones, no hay una relación gráfica que pueda permitir asegurar que ambas series comparten patrones o comportamientos; sus valores son muy diferentes y su desviación estándar es mayor para ECOPETROL.

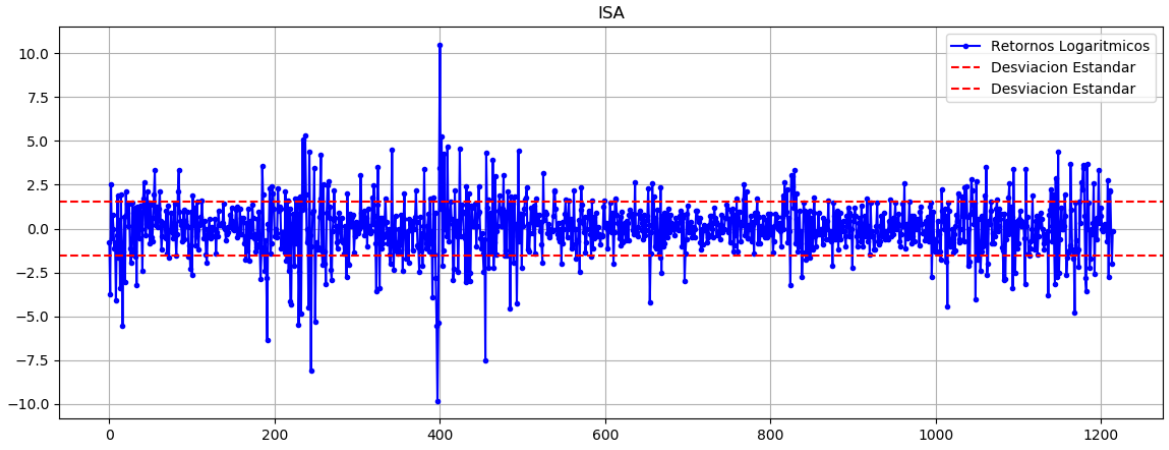

Figura 5: Retornos logarítmicos ISA.

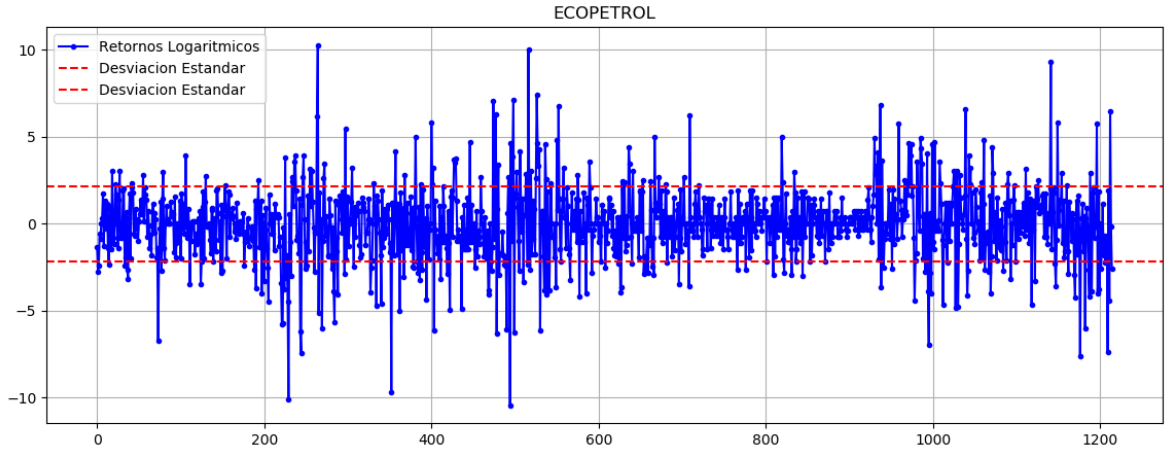

Figura 6: Retornos logarítmicos ECOPETROL.

Al revisar las Figuras 7 y 8 de autocorrelación y autocorrelación parcial de ambas acciones, se puede concluir que, existe una diferencia en el comportamiento de estas, puesto que, cada acción cuenta con un orden de valores rezagados y un orden de medias móviles diferente. De lo anterior, podemos concluir que, ambas acciones no pueden ser modeladas con el mismo orden de valores autorregresivos y de medias móviles.

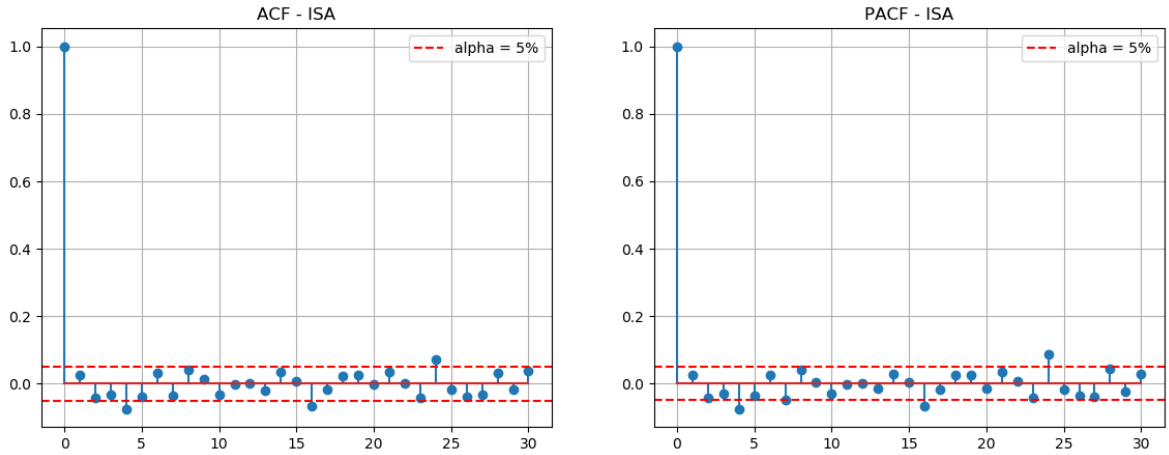

Figura 7: Autocorrelación y autocorrelación parcial de las rentabilidades ISA.

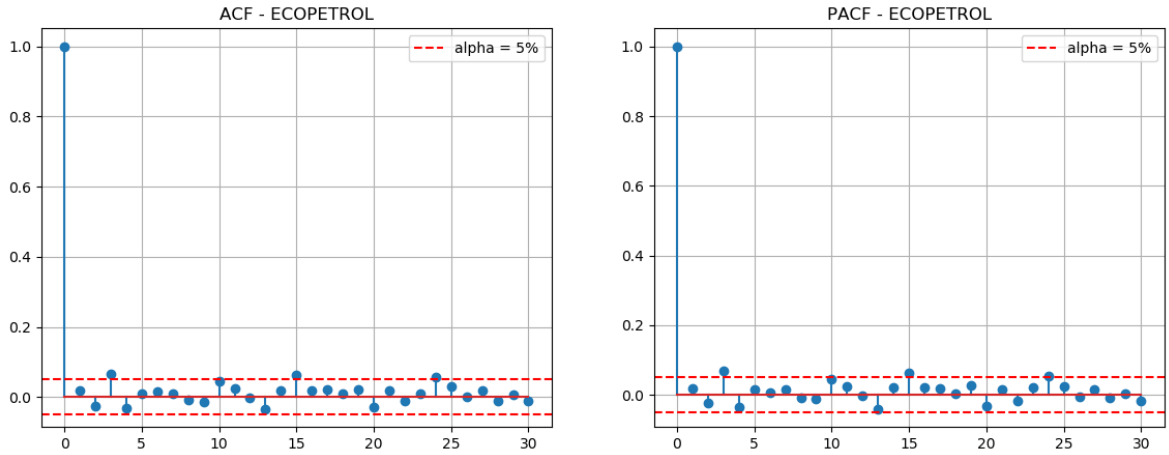

Figura 8: Autocorrelación y autocorrelación parcial de las rentabilidades ECOPETROL.

Dado el análisis anterior, para el caso de la media se deberá generar un modelo independiente para cada uno de los activos, que podrían conformar el portafolio financiero. Por esto, se buscarán formas automáticas para hallar el modelo que mejor se ajuste y genere un mejor desempeño.

3.3.2 Variables independiente y dependiente para el modelo de media

Siendo un problema de series de tiempo, la variable a pronosticar o dependiente será la diferencia logarítmica de los precios, es decir, el retorno logarítmico de estos. En este sentido, la construcción de esta variable se realizará como se había mencionado anteriormente.

En el caso de las variables independientes, se construirán rezagando la variable dependiente n veces, dependiendo de la iteración del modelo que se esté ejecutando. Este mismo ejercicio, se realizará para el cálculo de las medias móviles y así se construirá la tabla base que servirá para la elaboración de los modelos.

3.3.3 Modelo y residuales

Para la construcción de forma ágil de los diferentes modelos para cada una de las series, se hizo uso de la función auto_arima, localizada en la librería pmdarima de Python. Esta función, realiza la búsqueda automática del modelo correspondiente para cada uno de estos, iterando sobre el diferente orden de los rezagos y medias móviles de la serie, evaluando también la existencia de estacionalidad y modelándola, si es el caso de que existiese. Este proceso, lo hace a partir de una búsqueda inteligente, creando 4 modelos iniciales que varían en el orden de sus rezagos y medias móviles. Comparándolos con el criterio AIC, se selecciona el mejor y se procede a realizar diferentes modelos adicionales, con variaciones en el mejor modelo seleccionado previamente, comparando sus AIC y manteniendo el modelo cuyo AIC sea menor.

De este proceso, se obtuvieron 23 modelos; uno por cada serie a pronosticar y se calculó el pronóstico del día siguiente. Este pronóstico, se considera como la rentabilidad esperada que entrará al modelo de portafolio optimo, de igual forma, se generará el pronóstico sobre el conjunto de entrenamiento de esta serie. De este último paso, se obtuvieron los residuales del modelo, cuya salida servirá como entrada para el modelo de varianza que se desarrollará en los siguientes pasos.

3.4 Pronóstico de varianza en la simulación

3.4.1 Entendimiento de los datos de varianza

Siguiendo la misma dinámica llevada a cabo en el pronóstico de media, se realiza el análisis exploratorio a sólo dos acciones de los 23 totales, siendo ISA y ECOPETROL cómo se realizó en el proceso de pronóstico de media.

Al comparar las Figuras 9 y 10, se identifica una diferencia en el comportamiento de los residuales en ambas series, con ISA, teniendo una desviación estándar aproximada al 1.5 ,mientras que la de ECOPETROL se aproxima al 2.5.

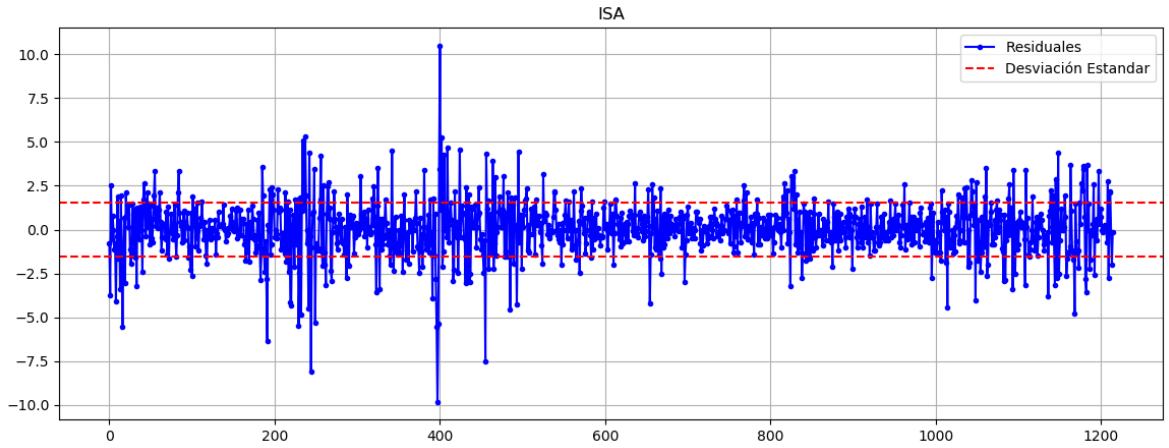

Figura 9: Residuales del modelo de media de ISA.

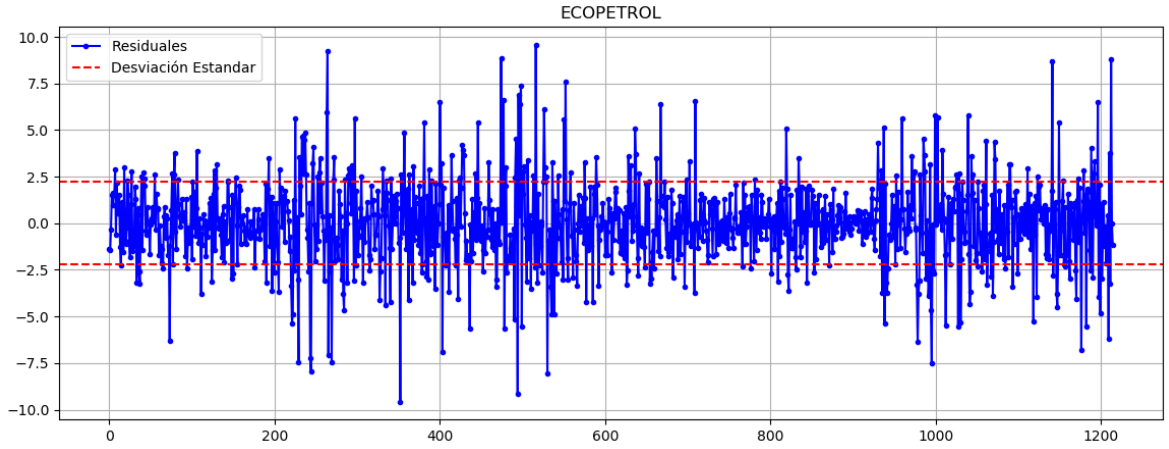

Figura 10: Residuales del modelo de media de ECOPETROL.

El análisis anterior, se puede complementar revisando las Figuras 11 y 12 de autocorrelación y autocorrelación parcial, de ambos activos, en los cuales, los gráficos en conjunto indican qué relación tienen los valores actuales con su valor rezagado n veces. Por lo anterior, se puede concluir que, los valores autorregresivos y de medias móviles que se podrían considerar para ISA, son diferentes de los que se deberían considerar para ECOPETROL.

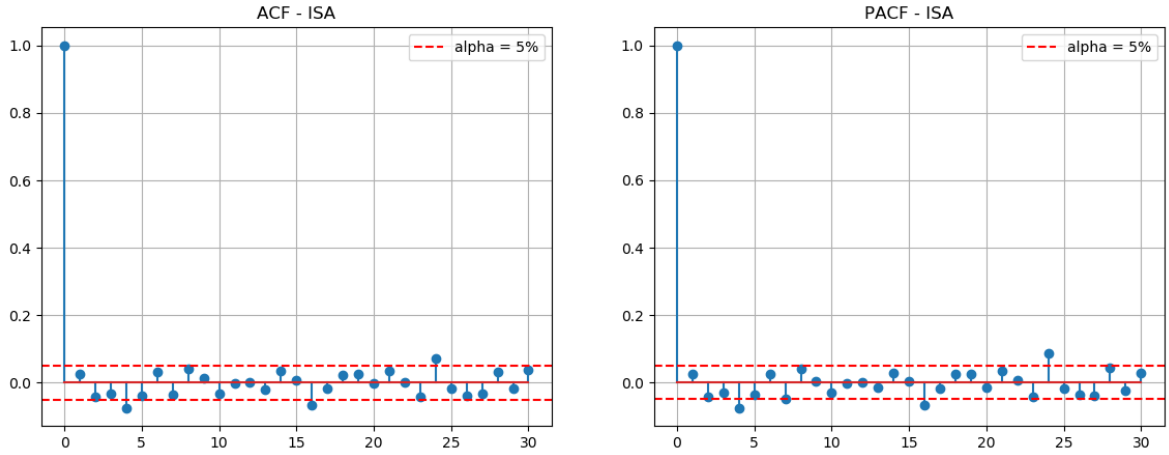

Figura 11: Autocorrelación y autocorrelación parcial de los residuales de ISA.

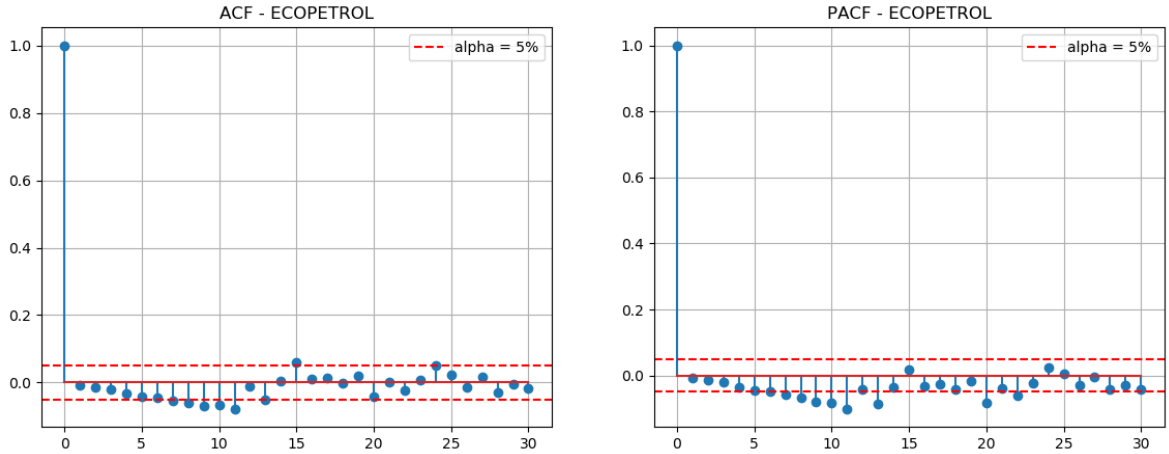

Figura 12: Autocorrelación y autocorrelación parcial de los residuales de ECOPETROL.

Puesto que se desea realizar el pronóstico de la varianza, es necesario calcular el cuadrado de los residuales, generando una nueva serie que se puede considerar como la varianza en cada instante.

Realizando el análisis de las Figuras 13 y 14 de los residuales al cuadrado, se puede seguir observando un comportamiento diferencial en las series, teniendo mayores valores y una desviación estándar más elevada en la serie de ECOPETROL.

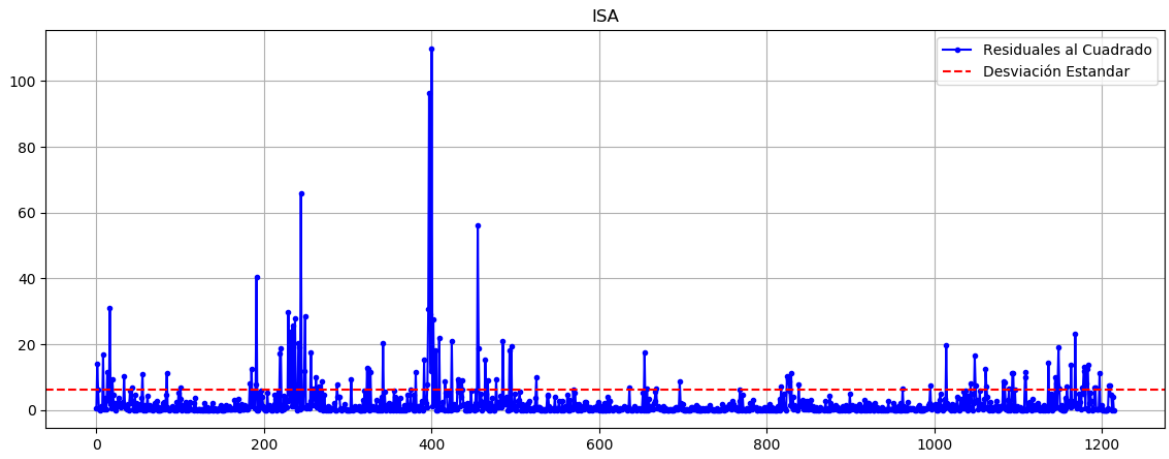

Figura 13: Residuales al cuadrado del modelo de media de ISA.

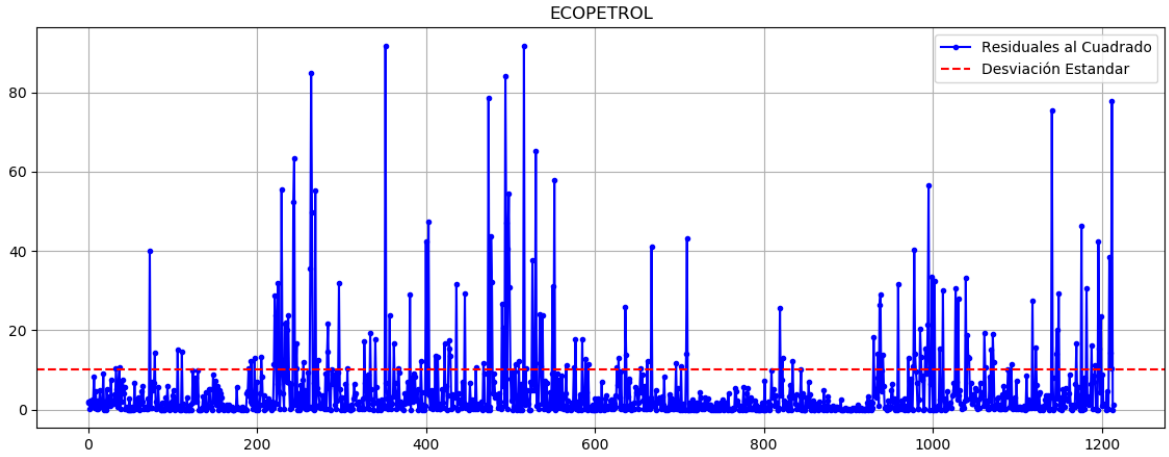

Figura 14: Residuales al cuadrado del modelo de media de ECOPETROL.

Igualmente, como se realizó en el paso anterior, se hace el análisis de los gráficos de auto correlación y auto correlación parcial con el fin de comprobar si ambas series tienen comportamiento similar.

De las Figuras 15 y 16, se puede observar que, el comportamiento de los residuales al cuadrado, se asemeja al comportamiento definido desde los residuales, por lo tanto, ambas series deben ser tratadas por separado, desde la selección de las variables independientes a utilizar como los modelos a generar en cada una de estas.

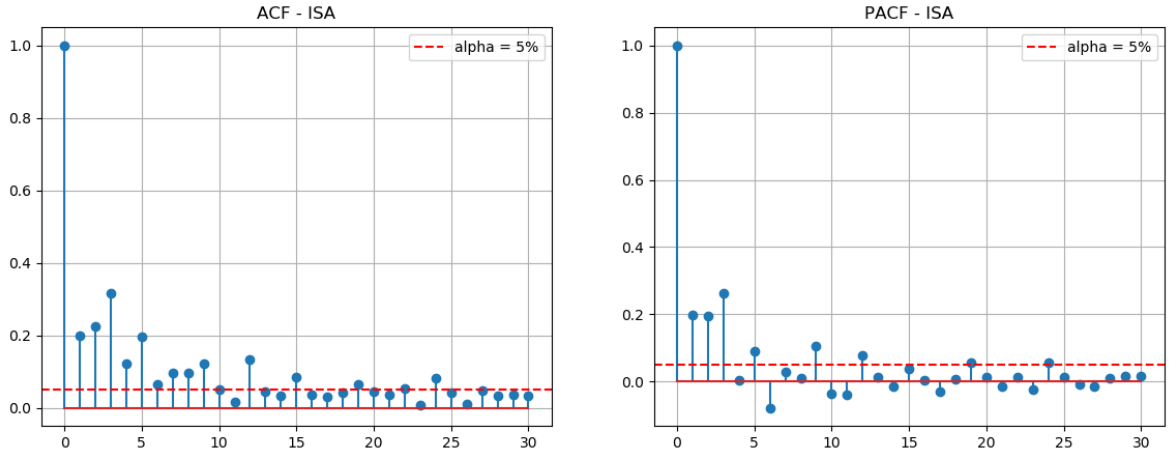

Figura 15: Autocorrelación parcial de los residuales al cuadrado del modelo de media de ISA.

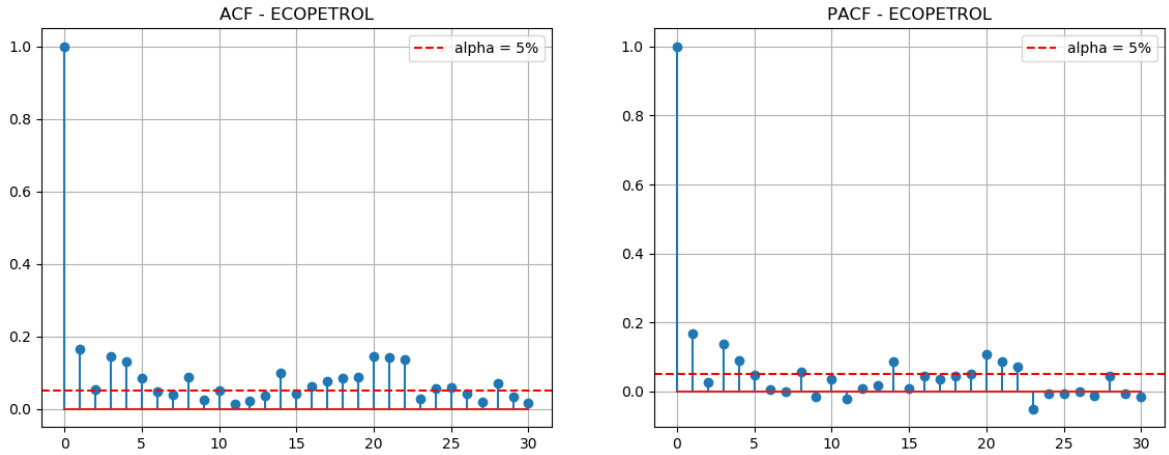

Figura 16: Autocorrelación parcial de los residuales al cuadrado del modelo de media de ECOPETROL.

Al revisar los residuales al cuadrado, se puede observar que se tienen valores que podrían ser considerados outliers y que, por ende, podrían afectar el rendimiento de los modelos que se generen para esto. Buscando reducir la escala de estos y los valores atípicos, se considera analizar el valor absoluto de los residuales, siendo una opción el generar modelos de pronóstico de la desviación, que elevados al cuadrado, generaran la varianza, la cual es el parámetro de entrada que se necesita para los modelos.

Desde la Figura 17, se puede ver cómo este acercamiento puede cumplir con lo que se desea. Los valores en los que se mueven ambas acciones, se reducen, y los valores atípicos podrían tener menos importancia que con el acercamiento de la varianza.

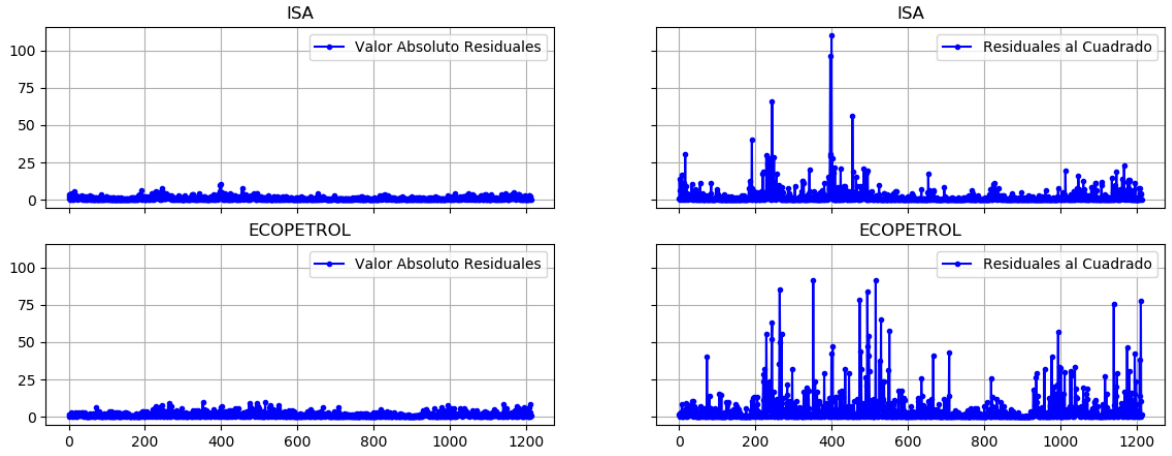

Figura 17: Comparación valores absolutos de los residuales del modelo de medía, frente a los residuales al cuadrado.

Al tener comportamientos tan diferentes, demostrados desde los diferentes análisis gráficos realizados anteriormente, se construirá un modelo propio para cada una de las series, tal como se realizó en el pronóstico de la media. Adicionalmente, para los casos de modelos de *Machine Learning* y *Deep Learning*, se generarán modelos del pronóstico de la desviación, buscando solventar los valores atípicos que se puedan considerar, de estos pronósticos se calculará su valor cuadrado, para obtener el parámetro tal cuál se necesita en los modelos de mínima varianza de Markowitz.

3.4.2 Medida de desempeño

Como ya que se pudo observar, las series a pronosticar tienen algunos valores atípicos, estos valores deben ser modelados, pues indican una alta varianza en determinados momentos del tiempo. Este análisis, ayudó a seleccionar el error cuadrático medio (MSE), como medida para comparar, tanto los diferentes modelos del mismo tipo para la misma serie con variación en sus variables independientes, como los diferentes tipos de modelos ejecutados en los diferentes escenarios.

Aunque lo común es revisar diferentes métricas de desempeño, en este caso no se tuvo en cuenta el Error Porcentual Absoluto Medio (MAPE), dado que, las series pueden contener muchos ceros y, en este caso, esta métrica no es posible calcularla. Del mismo modo, se descartó el Error Absoluto Medio (MAE) pues esta medida, penaliza los valores atípicos, no permitiendo que el modelo trate de capturarlos y, por lo tanto, dando un peor ajuste.

3.4.3 Pronostico ARCH/GARCH

Igual que como en el pronóstico de media, en el pronóstico de varianza también es necesaria la construcción de un modelo para cada acción, para esto, se define un espacio en el cual cada acción iterará sobre las variables a utilizar en el modelo, con el fin de encontrar de manera automática, cuál es el orden autoregresivo y de medias móviles que mejor resultado podría dar en cada una de estos.

Como se muestra en la Tabla 3, el espacio mencionado anteriormente, cuenta con un rango de orden para los valores autoregresivos de 1 a 4, mientras que, para el orden de las medias móviles, será de 0 a 4. Este espacio cuenta con 20 diferentes combinaciones de modelos, sobre el cual se itera y se obtiene el modelo con mejor medida de desempeño.

Tabla 3: Resultados de los diferentes modelos ARCH/GARCH de ISA.

Nemotecnico	Orden Rezago	Orden Media Movil	MSE	Modelo
ISA	1	4	3,8470	GARCH
ISA	1	3	3,8656	GARCH
ISA	1	1	3,8781	GARCH
ISA	1	2	3,8815	GARCH
ISA	3	1	3,8996	GARCH
ISA	4	1	3,8998	GARCH
ISA	2	1	3,9000	GARCH
ISA	3	4	3,9084	GARCH
ISA	4	4	3,9084	GARCH
ISA	2	4	3,9128	GARCH
ISA	4	2	3,9153	GARCH
ISA	3	2	3,9166	GARCH
ISA	2	2	3,9166	GARCH
ISA	2	3	3,9221	GARCH
ISA	3	3	3,9221	GARCH
ISA	4	3	3,9232	GARCH
ISA	3	0	3,9979	ARCH
ISA	1	0	4,0270	ARCH
ISA	4	0	4,0385	ARCH
ISA	2	0	4,0461	ARCH

Luego de seguir este proceso con las 23 series, se generó el resultado mostrado en la Tabla 4. Estos órdenes autorregresivos y de medias móviles, serán los utilizados por cada uno de los modelos que se ajustarán y pronosticarán el siguiente período de simulación en cada iteración.

Tabla 4: Mejores modelos ARCH/GARCH.

Nemotecnico	Orden Rezago	Orden Media Movil	MSE	Modelo
BCOLOMBIA	2	0	4,6575	ARCH
BOGOTA	2	2	2,4476	GARCH
BVC	1	0	1224,0121	ARCH
CELSIA	1	4	2,5357	GARCH
CEMARGOS	2	1	4,3090	GARCH
CLH	1	0	9,5407	ARCH
CNEC	3	4	5,8512	GARCH
CORFICOLCF	2	0	3,7549	ARCH
ECOPETROL	1	0	10,0950	ARCH
EXITO	4	2	3,4618	GARCH
GRUPOARGOS	1	4	3,5166	GARCH
GRUPOSURA	3	3	2,4507	GARCH
ISA	1	4	3,8470	GARCH
MINEROS	4	3	5,3584	GARCH
NUTRESA	1	4	1,5750	GARCH
PFAVAL	3	4	2,9748	GARCH
PFAVH	1	1	8,3805	GARCH
PFBCOLOM	1	2	4,1520	GARCH
PFCEMARGOS	2	0	5,2808	ARCH
PFDAVVNDA	4	4	3,0489	GARCH
PFGRUPOARG	2	4	4,7119	GARCH
PFGRUPSURA	1	4	3,5198	GARCH
PROMIGAS	1	0	6,7203	ARCH

3.4.4 Pronostico *Machine Learning*

Ya que los algoritmos de *Machine Learning* están enfocados en problemas de regresión y clasificación, es necesario hacer un ajuste y generar las variables independientes a partir de los rezagos y medias móviles de las acciones, esto, dado que, se piensan tratar como un problema de series de tiempo, sin variables exógenas. Igualmente, esta búsqueda se hará de forma automática, iterando sobre los mismos valores revisados en el pronóstico de modelos ARCH/GARCH con una diferencia, el rango de orden de los valores autorregresivos esta vez será de 0 a 4.

En este caso, se ejecutarán los pronósticos, haciendo uso del algoritmo XGBoost. Este mismo, se ajustará de forma iterativa por cada una de las series, cambiando sus variables

independientes y calculando su MSE como medida de desempeño. Al ser un algoritmo de *Machine Learning*, además de contar con sus variables independientes, que influirán en el ajuste del modelo, cuenta con hiperparámetros, que pueden ser optimizados para encontrar un mejor resultado.

Dado que, las iteraciones tienen un costo computacional muy elevado, se procede a hallar, inicialmente, la mejor combinación de variables independientes para cada serie, comparando su medida de desempeño.

Luego de hallar el mejor modelo a nivel de variables exógenas, se continúa con la optimización de sus hiperparámetros, buscando mejorar el rendimiento ya dado. Esta optimización se hará a partir de algoritmos de optimización bayesiana; este tipo de optimización se enfoca en minimizar la función objetivo, haciendo uso de técnicas de Bayesianas de *Machine Learning* y de la metodología Optimización Basada en Modelos Secuenciales (SMBO) por sus siglas en inglés, para hallar en menor tiempo posible los hiperparámetros, que mejor ajusten el modelo generado.

Continuando con lo anterior, se hace uso de la técnica Bayesiana de *Machine Learning* los Árboles de Estimación de Parzen (TPE) por sus siglas en inglés, modelando $p(x|y)$ transformando el proceso de la optimización Bayesiana al reemplazar la distribución de la configuración anterior, con densidades no paramétricas.

Siguiendo esta estrategia de optimización, se define el espacio sobre el cual buscar los mejores hiperparámetros, como se muestra en la Tabla 5, en la cual se puede observar el rango sobre el cuál se buscarán los mejores hiperparámetros de los diferentes modelos.

Tabla 5: Espacio de hiperparámetros para XGBoost.

Hiperparametros	**Rango**
Profundidad Maxima	1 a 18
Minimo Peso de las Ramas	0 a 18
Gamma	1 a 9
Alpha (Regularización)	0 a 180
Lambda (Regularización)	0 a 180

Por la forma en la que se buscarán los mejores hiperparámetros, este algoritmo no buscará en todas las posibles combinaciones del espacio, por el contrario, iterará 50 veces, haciendo una búsqueda de los hiperparámetros que podrían generar una mejora en el rendimiento en el modelo creado, teniendo el MSE como medida de desempeño como se habría mencionado anteriormente.

Por lo anterior, el flujo para hallar los modelos óptimos de *Machine Learning* para cada una de las acciones, se resume de la siguiente forma:

1) Generación de variables rezagadas y medias móviles.
2) Iteración del modelo con cambio en variables independientes.
3) Optimización de hiperparámetros de los mejores modelos.
4) Almacenamiento y predicción del modelo de cada acción.

El resultado de lo anterior, se puede ver en la Tabla 6, la cual muestra el orden de rezago y medias móviles que ingresaron como variables independientes al modelo de XGBoost de cada acción.

Tabla 6: Mejores Modelos XGBoost.

Nemotecnico	Orden Rezago	Orden Media Movil	MSE
BCOLOMBIA	3	0	2,457356224
BOGOTA	3	1	1,135324597
BVC	3	0	1073,472775
CELSIA	1	2	1,193239528
CEMARGOS	1	0	1,803506993
CLH	0	3	4,877910528
CNEC	0	4	2,797355401
CORFICOLCF	4	2	1,734126407
ECOPETROL	0	4	5,027716339
EXITO	2	3	1,472862153
GRUPOARGOS	4	3	1,646083935
GRUPOSURA	0	4	1,022312542
ISA	1	4	1,893554084
MINEROS	0	2	2,620710641
NUTRESA	0	3	0,791040979
PFAVAL	0	4	1,447828049
PFAVH	0	4	3,900370799
PFBCOLOM	4	4	2,118966901
PFCEMARGOS	4	4	2,736349839
PFDAVVNDA	3	2	1,547896392
PFGRUPOARG	0	2	2,341135556
PFGRUPSURA	4	0	1,699691143
PROMIGAS	2	0	4,463577425

3.4.5 Pronostico *Deep Learning*

3.4.5.1 Acotación

En este tipo de modelos, se encuentran diferentes factores que pueden afectar el rendimiento de estos, los factores pueden ser: la cantidad de variables independientes a tener en cuenta, el tipo de arquitectura, la escala de los valores y los hiperparámetros; como el número de neuronas y las funciones de pérdida. Ya que estos factores pueden ser bastantes y este tipo de modelos es de los más costosos a nivel computacional, se realizó una dinámica similar a la realizada en el entendimiento de los datos, es decir, se crearon diferentes experimentos variando entre algunos de los factores antes mencionados, para hallar cuales de estos pueden tener un mayor impacto en el rendimiento de los modelos. Para esto, igualmente, se tomaron las acciones de ISA y ECOPETROL.

3.4.5.2 Arquitecturas

En esta acotación se tuvieron en cuenta 3 tipos de arquitecturas, sobre las cuales se validaría con cuál de estas se obtiene un mejor resultado. Dado que se trata de un problema en el que la temporalidad es importante, es decir, el valor futuro dependerá del valor presente o pasados, es necesario hacer uso de redes neuronales recurrentes, las cuales son apropiadas en este tipo de problemas.

Sobre lo anterior, se optó por crear 3 variaciones de la red neuronal recurrente, es decir, 3 arquitecturas.

1) **Red neuronal recurrente simple (RNN):** inicialmente se prueba con una arquitectura RNN simple, la cual consta de una capa oculta de n neuronas y una capa densa de 1 neurona, ya que la salida de este modelo solo será un valor, el pronóstico de la desviación.

2) **Red neuronal de memoria de corto y largo plazo simple (LSTM):** como una mejora a la RNN, se presenta la LSTM. Esta red neuronal, mejora frente a la anterior al tener en cuenta una arquitectura que le permite tener memoria a corto y largo plazo, permitiéndole tener buenos resultados, tanto como con una cantidad alta de variables independientes, como teniendo una cantidad baja de variables independientes.

3) **Red neuronal de memoria de corto y largo plazo apilada (STACKED LSTM):** con el fin de encontrar la mejor arquitectura para los diferentes activos, también se hizo uso de la STACKED LSTM, arquitectura propuesta por Assaf, Di Fatta, & Nicosia (2022), los cuales consiguieron mejores resultados con este tipo de arquitectura, que con un modelo ARCH/GARCH para el pronóstico de la varianza.
A diferencia de la LSTM simple mencionada anteriormente, esta arquitectura cuenta con dos capas LSTM y una capa densa.

Dado lo anterior, la Figura 18 representa la arquitectura RNN y LSTM planteadas, de las cuales, sólo cuentan con una capa oculta que varía en el tipo de capa que se utilizará, mientras que en la Figura 19, podemos observar la arquitectura STACKED LSTM que cuenta con dos capas ocultas de tipo LSTM.

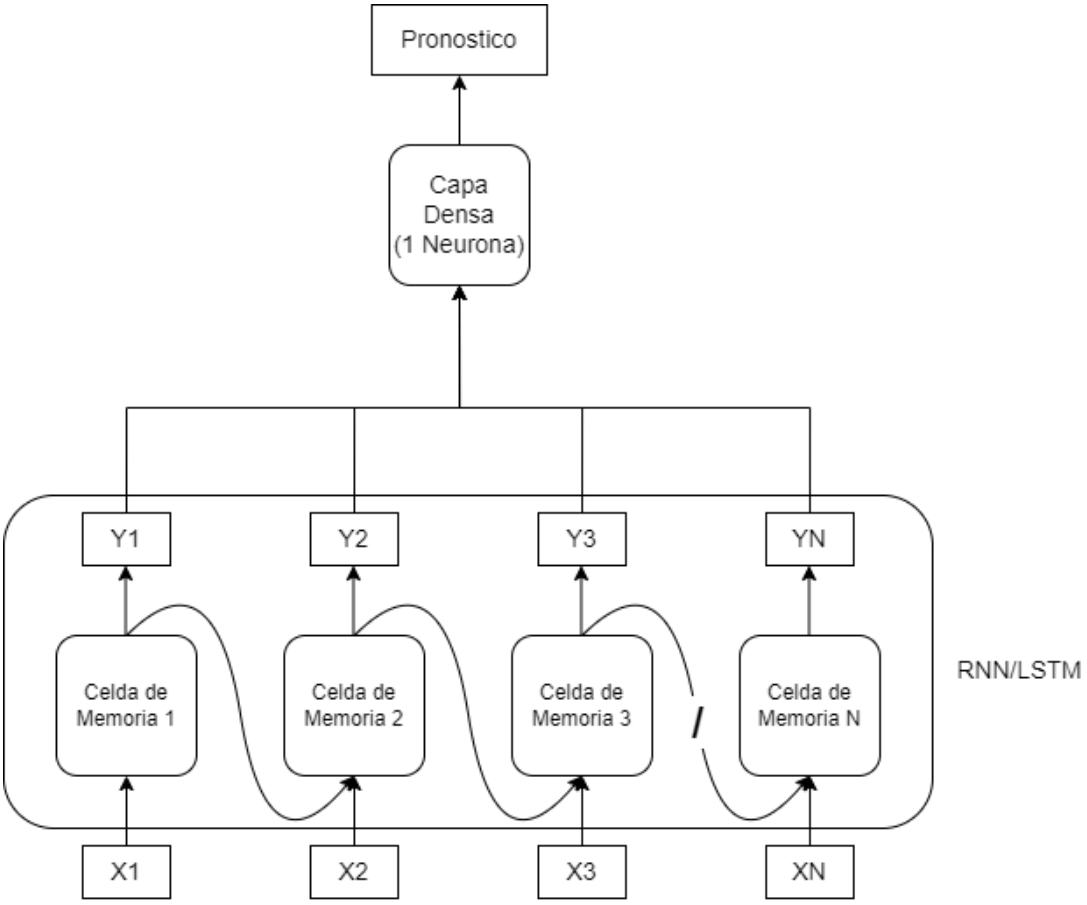

Figura 18: Red Neuronal Recurrente Simple/LSTM.

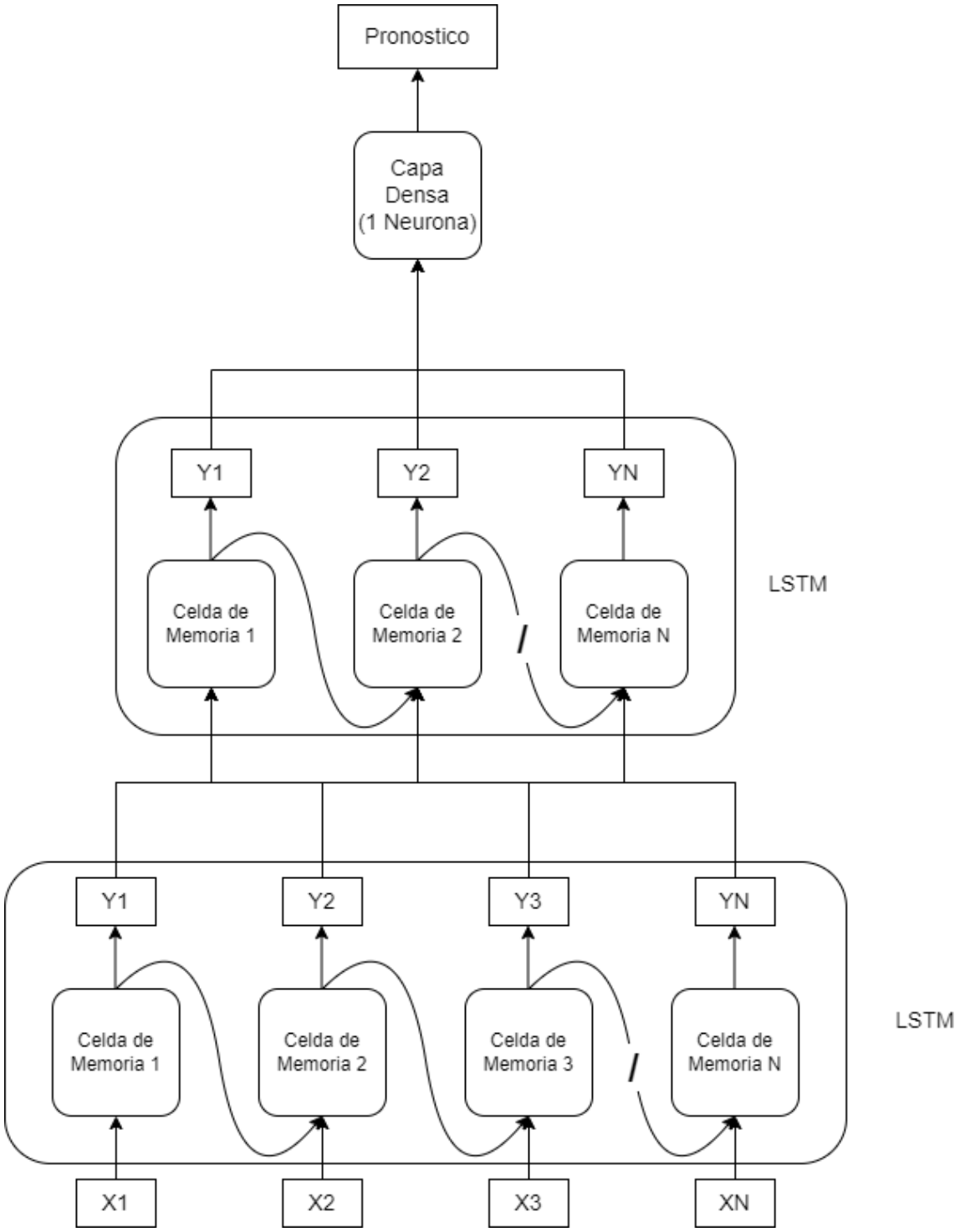

Figura 19: Stacked LSTM.

3.4.5.3 Escala de datos

Como se revisó anteriormente en la sección 3.3.1 Entendimiento de los datos, estas series cuentan con valores atípicos muy elevados, lo que podría afectar considerablemente el entrenamiento al hacer uso de modelos de redes neuronales. Dado lo anterior, se plantean 3 diferentes escalas sobre las cuales iterar para hallar el modelo que mejor se ajuste, estas escalas serán: MinMax y standar, las cuales se encuentran en la librería sklearn.

3.4.5.4 Variables independientes para el modelo de varianza

El tercer ítem sobre el que se itera para la acotación de los modelos, serán las variables independientes a utilizar. En este caso, como se ha hecho en las secciones anteriores, se cuenta con un problema de series de tiempo, por lo tanto, las variables independientes serán consideradas como la cantidad de periodos atrás que se tengan en cuenta para la predicción del día siguiente.

3.4.5.5 Experimentos y resultados

Dado los ítems mencionados anteriormente para la acotación de los modelos, se obtienen los siguientes experimentos a probar con las series de ISA y ECOPETROL. Esta búsqueda de mejores parámetros, se hace siguiendo la técnica de grid search, la cual se puede ver en la Tabla 7 junto a los resultados de los primeros experimentos ejecutados, dando un total de 54 modelos por cada acción.

Tabla 7: Experimentos iniciales.

Arquitectura	Variable	Escala	Numero de Neuronas	Numero de Rezagos
RNN	Valor absoluto residuales	Sin escala	32	1,3,5
	Valor absoluto residuales	MinMax	32	1,3,5
	Valor absoluto residuales	Estándar	32	1,3,5
	Valor absoluto residuales	Sin escala	64	1,3,5
	Valor absoluto residuales	MinMax	64	1,3,5
	Valor absoluto residuales	Estándar	64	1,3,5
LSTM	Valor absoluto residuales	Sin escala	32	60,100,140
	Valor absoluto residuales	MinMax	32	60,100,140
	Valor absoluto residuales	Estándar	32	60,100,140
	Valor absoluto residuales	Sin escala	64	60,100,140
	Valor absoluto residuales	MinMax	64	60,100,140
	Valor absoluto residuales	Estándar	64	60,100,140
STACKED LSTM	Valor absoluto residuales	Sin escala	32	60,100,140
	Valor absoluto residuales	MinMax	32	60,100,140
	Valor absoluto residuales	Estándar	32	60,100,140
	Valor absoluto residuales	Sin escala	64	60,100,140
	Valor absoluto residuales	MinMax	64	60,100,140
	Valor absoluto residuales	Estándar	64	60,100,140

Como se puede ver en la Tabla 8 y 9, del top 10 de los mejores y peores modelos de ISA y ECOPETROL, la arquitectura cuyo mejor resultado arroja es la RNN, con un número de rezagos bajos, por el contrario, la arquitectura STACKED LSTM, genera los peores resultados, con un MSE muy diferente del MSE de la RNN.

Tabla 8: Top 10, mejores y peores modelos ISA.

ISA					
Top 10 Mejores					
Arquitectura	**Variable**	**Escala**	**Numero de Neuronas**	**Numero de Rezagos**	**MSE**
RNN	Valor absoluto residuales	Sin Escala	64	1	0,4861
RNN	Valor absoluto residuales	MinMax	32	1	0,4879
RNN	Valor absoluto residuales	MinMax	64	1	0,4896
RNN	Valor absoluto residuales	Standar	64	1	0,4899
RNN	Valor absoluto residuales	Standar	32	1	0,4901
RNN	Valor absoluto residuales	Sin Escala	32	1	0,4934
RNN	Valor absoluto residuales	MinMax	32	3	0,5294
RNN	Valor absoluto residuales	MinMax	64	3	0,5336
LSTM	Valor absoluto residuales	Sin Escala	32	140	0,5817
RNN	Valor absoluto residuales	Sin Escala	32	3	0,5973
Top 10 Peores					
STACKED LSTM	Valor absoluto residuales	Sin Escala	32	100	1,2004
LSTM	Valor absoluto residuales	Sin Escala	64	140	1,2268
LSTM	Valor absoluto residuales	Sin Escala	64	100	1,2369
LSTM	Valor absoluto residuales	Standar	64	60	1,2412
STACKED LSTM	Valor absoluto residuales	Sin Escala	64	60	1,2562
LSTM	Valor absoluto residuales	Sin Escala	64	60	1,2580
STACKED LSTM	Valor absoluto residuales	Sin Escala	32	60	1,2668
STACKED LSTM	Valor absoluto residuales	Standar	64	140	1,2801
STACKED LSTM	Valor absoluto residuales	Standar	64	100	1,2848
STACKED LSTM	Valor absoluto residuales	Standar	64	60	1,2867

Tabla 9: Top 10, mejores y peores modelos ECOPETROL.

Ecopetrol					
Top 10 Mejores					
Arquitectura	**Variable**	**Escala**	**Numero de Neuronas**	**Numero de Rezagos**	**MSE**
RNN	Valor absoluto residuales	Standar	64	1	1,5742
RNN	Valor absoluto residuales	Standar	32	1	1,5844
RNN	Valor absoluto residuales	Sin Escala	32	1	1,5862
RNN	Valor absoluto residuales	Sin Escala	64	1	1,5969
RNN	Valor absoluto residuales	MinMax	64	3	1,6055
RNN	Valor absoluto residuales	MinMax	32	1	1,6111
RNN	Valor absoluto residuales	MinMax	64	1	1,6176
RNN	Valor absoluto residuales	MinMax	32	3	1,6597
RNN	Valor absoluto residuales	MinMax	32	5	1,8424
RNN	Valor absoluto residuales	Sin Escala	32	3	1,8618
Top 10 Peores					
STACKED LSTM	Valor absoluto residuales	Standar	64	140	3,6485
STACKED LSTM	Valor absoluto residuales	Sin Escala	32	60	3,6507
STACKED LSTM	Valor absoluto residuales	Sin Escala	64	100	3,6533
LSTM	Valor absoluto residuales	Sin Escala	64	140	3,6720
STACKED LSTM	Valor absoluto residuales	Sin Escala	32	100	3,6825
STACKED LSTM	Valor absoluto residuales	Standar	64	60	3,7578
LSTM	Valor absoluto residuales	Standar	64	60	3,7651
LSTM	Valor absoluto residuales	Sin Escala	64	60	3,8029
STACKED LSTM	Valor absoluto residuales	Sin Escala	64	60	3,8450
STACKED LSTM	Valor absoluto residuales	Sin Escala	64	140	3,8697

Dado que se puede observar que la arquitectura a utilizar y valores altos en el número de rezagos con estas arquitecturas afectan considerablemente el rendimiento del modelo, se descarta la STACKED LSTM, y se ejecuta un experimento nuevo, teniendo en cuenta sólo la LSTM y RNN. Cómo se puede ver en la Tabla 10, en los nuevos experimentos, sólo se probarán las arquitecturas RNN y LSTM con valores bajos en el número de rezagos y, además, no se itera sobre el número de neuronas, dejando cada arquitectura con 64.

Tabla 10: Experimentos secundarios.

Arquitectura	Variable	Escala	Numero de Rezagos
RNN	Valor absoluto residuales	Sin Escala	1,2,3,4,5
	Valor absoluto residuales	MinMax	1,2,3,4,5
	Valor absoluto residuales	Standar	1,2,3,4,5
	Valor absoluto residuales	Sin Escala	1,2,3,4,5
	Valor absoluto residuales	MinMax	1,2,3,4,5
	Valor absoluto residuales	Standar	1,2,3,4,5
LSTM	Valor absoluto residuales	Sin escala	1,2,3,4,5
	Valor absoluto residuales	MinMax	1,2,3,4,5
	Valor absoluto residuales	Standar	1,2,3,4,5
	Valor absoluto residuales	Sin Escala	1,2,3,4,5
	Valor absoluto residuales	MinMax	1,2,3,4,5
	Valor absoluto residuales	Standar	1,2,3,4,5

Revisando los resultados de las Tablas 11 y 12, se puede observar que, en esta ocasión, la LSTM con rezagos bajos, logró mejores resultados frente a la RNN.

Tabla 11: Resultados experimentos secundarios ISA.

ISA				
Top 10 Mejores				
Arquitectura	**Variable**	**Escala**	**Numero de Rezagos**	**MSE**
LSTM	Valor absoluto residuales	Sin Escala	1	0,2834
LSTM	Valor absoluto residuales	Sin Escala	1	0,3311
RNN	Valor absoluto residuales	Sin Escala	1	0,3986
RNN	Valor absoluto residuales	Sin Escala	1	0,4144
RNN	Valor absoluto residuales	Standar	1	0,4734
RNN	Valor absoluto residuales	Standar	1	0,4785
LSTM	Valor absoluto residuales	Standar	1	0,4880
RNN	Valor absoluto residuales	MinMax	1	0,4897
LSTM	Valor absoluto residuales	Standar	1	0,4906
RNN	Valor absoluto residuales	MinMax	3	0,4916
Top 10 Peores				
RNN	Valor absoluto residuales	MinMax	5	0,6016
RNN	Valor absoluto residuales	MinMax	5	0,6026
RNN	Valor absoluto residuales	MinMax	4	0,6027
LSTM	Valor absoluto residuales	Standar	4	0,6047
RNN	Valor absoluto residuales	Standar	4	0,6069
LSTM	Valor absoluto residuales	MinMax	5	0,6100
LSTM	Valor absoluto residuales	MinMax	5	0,6170
LSTM	Valor absoluto residuales	Standar	5	0,6206
RNN	Valor absoluto residuales	Standar	4	0,6250
RNN	Valor absoluto residuales	Standar	5	0,6352

Tabla 12: Resultados experimentos secundarios ECOPETROL.

Ecopetrol				
Top 10 Mejores				
Arquitectura	**Variable**	**Escala**	**Numero de Rezagos**	**MSE**
LSTM	Valor absoluto residuales	Sin Escala	1	0,9115
LSTM	Valor absoluto residuales	Sin Escala	1	1,0296
RNN	Valor absoluto residuales	Sin Escala	1	1,3936
RNN	Valor absoluto residuales	Sin Escala	1	1,4590
LSTM	Valor absoluto residuales	Standar	2	1,5791
LSTM	Valor absoluto residuales	Standar	2	1,5865
LSTM	Valor absoluto residuales	Sin Escala	2	1,5873
RNN	Valor absoluto residuales	Standar	2	1,6070
LSTM	Valor absoluto residuales	MinMax	1	1,6073
RNN	Valor absoluto residuales	MinMax	2	1,6092
Top 10 Peores				
LSTM	Valor absoluto residuales	MinMax	3	1,7717
LSTM	Valor absoluto residuales	MinMax	4	1,7717
LSTM	Valor absoluto residuales	Standar	5	1,7781
LSTM	Valor absoluto residuales	Standar	5	1,7797
RNN	Valor absoluto residuales	Standar	5	1,7875
LSTM	Valor absoluto residuales	MinMax	5	1,7878
LSTM	Valor absoluto residuales	MinMax	3	1,8011
LSTM	Valor absoluto residuales	MinMax	4	1,8019
LSTM	Valor absoluto residuales	MinMax	5	1,8188
RNN	Valor absoluto residuales	Standar	5	1,8228

Dado el análisis anterior, se plantean los modelos mostrados en la Tabla 13, sobre estos se iterará cada una de las acciones a considerar en el portafolio.

Tabla 13: *Iteraciones en los modelos de Deep Learning.*

Arquitectura	Variable	Escala	Numero de Rezagos
LSTM	Valor absoluto residuales	Sin escala	1,2,3,4,5,6
	Valor absoluto residuales	MinMax	1,2,3,4,5,6
	Valor absoluto residuales	Standar	1,2,3,4,5,6
	Valor absoluto residuales	Sin Escala	1,2,3,4,5,6
	Valor absoluto residuales	MinMax	1,2,3,4,5,6
	Valor absoluto residuales	Standar	1,2,3,4,5,6

Cómo se puede observar en la Tabla 14, referente a la configuración de los modelos finales para las 23 acciones, la mayoría de estos modelos obtuvo como mejor resultado, una configuración con sólo el rezago 1 y sin escala.

Tabla 14: Resultados LSTM.

Nemotecnico	Arquitectura	Variable	Escala	Numero de Rezagos	MSE
BCOLOMBIA	LSTM	Sin escala	32	1	0,4799
BOGOTA	LSTM	Sin escala	32	1	0,3644
BVC	LSTM	MinMax	64	2	926,4003
CELSIA	LSTM	Sin escala	32	1	0,2387
CEMARGOS	LSTM	Sin escala	32	1	0,3822
CLH	LSTM	Sin escala	32	1	1,9917
CNEC	LSTM	Sin escala	32	1	0,2443
CORFICOLCF	LSTM	Sin escala	32	1	0,3768
ECOPETROL	LSTM	Sin escala	32	1	0,9189
EXITO	LSTM	Sin escala	32	1	0,2148
GRUPOARGOS	LSTM	Sin escala	32	1	0,2628
GRUPOSURA	LSTM	Sin escala	32	1	0,1770
ISA	LSTM	Sin escala	32	1	0,2845
MINEROS	LSTM	Sin escala	32	1	1,0971
NUTRESA	LSTM	Sin escala	32	1	0,1649
PFAVAL	LSTM	Sin escala	32	1	0,3125
PFAVH	LSTM	Sin escala	32	1	0,6551
PFBCOLOM	LSTM	Sin escala	32	1	0,3354
PFCEMARGOS	LSTM	Sin escala	32	1	0,7704
PFDAVVNDA	LSTM	Sin escala	32	1	0,2536
PFGRUPOARG	LSTM	Sin escala	32	1	0,4423
PFGRUPSURA	LSTM	Sin escala	32	1	0,3104
PROMIGAS	LSTM	MinMax	32	1	2,7182

Para finalizar, en la Tabla 15 se puede observar la comparación de los modelos de cada acción, evidenciando una mejora significativa en el MSE de los modelos de redes neuronales.

Tabla 15: Comparación de modelos de varianza.

MSE			
Nemotecnico	**LSTM**	**XGBOOST**	**ARCH/GARCH**
NUTRESA	0,1649	0,7910	1,5750
GRUPOSURA	0,1770	1,0223	2,4507
EXITO	0,2148	1,4729	3,4618
CELSIA	0,2387	1,1932	2,5357
CNEC	0,2443	2,7974	5,8512
PFDAVVNDA	0,2536	1,5479	3,0489
GRUPOARGOS	0,2628	1,6461	3,5166
ISA	0,2845	1,8936	3,8470
PFGRUPSURA	0,3104	1,6997	3,5198
PFAVAL	0,3125	1,4478	2,9748
PFBCOLOM	0,3354	2,1190	4,1520
BOGOTA	0,3644	1,1353	2,4476
CORFICOLCF	0,3768	1,7341	3,7549
CEMARGOS	0,3822	1,8035	4,3090
PFGRUPOARG	0,4423	2,3411	4,7119
BCOLOMBIA	0,4799	2,4574	4,6575
PFAVH	0,6551	3,9004	8,3805
PFCEMARGOS	0,7704	2,7363	5,2808
ECOPETROL	0,9189	5,0277	10,0950
MINEROS	1,0971	2,6207	5,3584
CLH	1,9917	4,8779	9,5407
PROMIGAS	2,7182	4,4636	6,7203
BVC	926,4003	1073,4728	1224,0121

3.5 Compra y venta de acciones según el portafolio optimo

Habiendo generado las diferentes entradas para el modelo de portafolio óptimo, se procede a la ejecución de este, con el fin de encontrar las participaciones óptimas de cada activo para la simulación que se está ejecutando.

Hallada la nueva participación, se compara el nuevo portafolio con el actual, obteniendo un beneficio o pérdida diario. Los beneficios generados por este modelo se reinvierten en él mismo, hasta el fin de la simulación, en cuyo caso se liquidará por completo. Esta comparación se realiza teniendo en cuenta la participación que se tiene en el momento t-1 y t.

- Operación de Compra: La operación de compra se dará siempre que la participación del nuevo portafolio sea mayor a la del portafolio anterior. Se realiza la compra de nuevas acciones, con el fin de obtener la participación óptima del nuevo portafolio.
- Operación de Venta: Esta operación se realizará cuando la participación del nuevo portafolio sea menor a la del portafolio anterior. Se buscará vender acciones del portafolio actual con el fin de obtener la participación óptima del nuevo portafolio.

Como se muestra en la Tabla 16, para cada simulación se comparan las participaciones actuales, contra las nuevas, generando así, la operación a realizar para cada activo.

Tabla 16: Simulación de muestra para un período con operaciones de compra y venta.

Simulación 2/01/20219 - 3/01/2019						
Fecha	**Descripción**	**BCOLOMBIA**	**BOGOTA**	**CORFICOLCF**	**ECOPETROL**	**EXITO**
Portafolio Actual						
2/01/2019	Cantidad de Acciones Portafolio Actual	8	1	1	1	58
2/01/2019	Precio de Cierre Portafolio Actual	$ 30.400,00	$ 55.800,00	$ 16.400,00	$ 2.645,00	$ 12.400,00
2/01/2019	Participación Monetaria Portafolio Actual	$ 243.200,00	$ 55.800,00	$ 16.400,00	$ 2.645,00	$ 719.200,00
Nuevo Portafolio						
3/01/2019	Cantidad Acciones Nuevo Portafolio	1	10	3	1	4
3/01/2019	Precio de Cierre Nuevo Portafolio	$ 31.600,00	$ 54.060,00	$ 16.000,00	$ 2.650,00	$ 12.360,00
3/01/2019	Participación Monetaria Nuevo Portafolio	$ 31.600,00	$ 540.600,00	$ 48.000,00	$ 2.650,00	$ 49.440,00
Operación de Compra						
3/01/2019	Cantidad de Acciones a Comprar		9	2		
3/01/2019	Monto de la Operación de Compra		$ 486.540,00	$ 32.000,00		
Operación de Venta						
3/01/2019	Cantidad de Operaciones a Vender	7				54
3/01/2019	Monto de la Operación de Venta	$ 221.200,00				$ 667.440,00
3/01/2019	Benefició o Perdida de la operación	$ 8.400,00				-$ 2.160,00

3.6 Resultado de escenarios

Ya definidos los modelos sobre los cuales se harán los pronósticos de media y varianza según cada escenario, se procede a realizar la ejecución de los escenarios propuestos en la sección 2.1.

Con el fin de evaluar el desempeño de estos portafolios, se realiza la comparación sobre el saldo inicial frente al saldo final más la liquidación del portafolio final, en este caso, se tuvo un saldo inicial de $10.000.000 de pesos colombianos, de los cuales el 75% se usó para la compra de acciones, dejando el 25% restante como fondo para solventar las pérdidas y nuevas operaciones que se necesiten realizar.

Como se puede observar en la Figura 20, el portafolio con el modelo LSTM, genera unos beneficios y pérdidas mayores a los demás portafolios, mientras que el portafolio de XGBoost tiene un comportamiento similar gráficamente al portafolio con ARCH/GARCH, sin embargo, este tiene unas pérdidas y ganancias menos elevadas que el XGBoost.

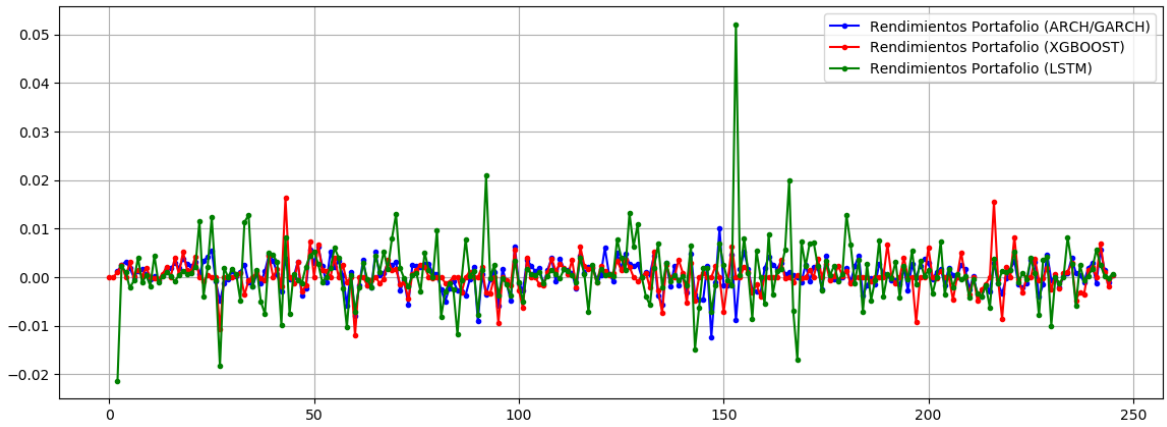

Figura 20: Histórico de rendimientos de los portafolios.

Revisando la Tabla 17, se puede observar que, al final del periodo de simulación, se logran mejores rendimientos con el portafolio creado con el modelo LSTM, con un rendimiento en 252 días del 21,24% terminando con ganancias de $2.123.603 pesos colombianos.

Tabla 17: Comparación de resultados de los portafolios.

Portafolio	Valor Inicial	Valor Final	Rendimiento
Portafolio (XGBOOST)	$ 10.000.000	$ 10.736.048	7,36%
Portafolio (ARCH/GARCH)	$ 10.000.000	$ 11.160.651	11,61%
Portafolio (LSTM)	$ 10.000.000	$ 12.123.603	21,24%

4 Conclusiones

4.1 Pregunta de investigación

¿Es posible mejorar la rentabilidad de un portafolio de activos construido utilizando la metodología de Markowitz cuando la volatilidad es pronosticada usando *Deep Learning*?

Respondiendo a esta pregunta, se encontró que sí es posible. A continuación, se justifica esta respuesta.

En los capítulos anteriores, se desarrolló el flujo necesario para la construcción de los diferentes portafolios; este flujo inicio con la construcción de modelos tradicionales ARIMA para el pronóstico de la rentabilidad de cada uno de los activos y continuó con la iteración de diferentes técnicas de modelado para el pronóstico de la varianza. En este sentido, se realizaron modelos Tradicionales (ARCH/GARCH), modelos de *Machine Learning* (XGBoost) y modelos de *Deep Learning* (LSTM) sobre los cuales se ejecutaron las diferentes simulaciones, correspondientes a los escenarios y por ende al resultado de un portafolio generado con el modelo de mínima varianza de Markowitz.

De este pronóstico de la varianza, se logra concluir que las técnicas de *Deep Learning*, logran tener mejores resultados en sus pronósticos, y sirven como entrada en los modelos de portafolios óptimos, generando una mejor rentabilidad. Este mismo resultado, no se cumple exactamente con los modelos de *Machine Learning*, puesto que, aunque sus pronósticos sean mejores que los modelos tradicionales, no logran obtener una mejor rentabilidad que estos. Esto se puede dar, dado que, no siempre genera pronósticos admisibles para el modelo de portafolio óptimo. Por lo anterior, se cuenta con simulaciones en las que el portafolio no cambia y, por ende, su rentabilidad o pérdida tampoco, problema que no ocurrió al hacer uso de los modelos tradicionales.

Adicionalmente, de los modelos de pronóstico de la varianza con modelos de *Deep Learning*, se puede decir que no se vuelve tan necesaria la iteración en un espacio de búsqueda muy grande para los factores como: escalas, neuronas y numero de rezagos, puesto que, como resultado, se encontró que al hallar una configuración de un activo, esta funciona generalmente bien para los demás. Por otro lado, la elección de la arquitectura a usar, tuvo mucho impacto en el rendimiento de los modelos, por esto, es uno de los puntos en donde se debería hacer un mayor enfoque a la hora de realizar este ejercicio.

4.2 Cumplimiento de objetivos

4.2.1 Objetivo 1: Estimar la rentabilidad de un portafolio en el tiempo usando la metodología de Markowitz

Para cumplir el objetivo 1, se realizaron diferentes tareas necesarias, para la creación de modelo de mínima varianza de Markowitz, y así, calcular la rentabilidad de este en el tiempo. En ese sentido, en la sección 3.2 se realizaron los modelos de pronóstico de la media, siguiendo con la sección 3.3.3 y 3.3.4 en los cuales se construyeron los diferentes modelos de pronóstico de la varianza, haciendo uso de modelos Tradicionales y ML. Este flujo, continuó con la sección 6.1, en la cual se realiza la compra y venta de acciones para el portafolio óptimo y se calculan las rentabilidades generadas por este.

Para finalizar, en la sección 6.2, se pueden observar los rendimientos totales, generados por los diferentes portafolios creados con la metodología de Markowitz.

4.2.2 Objetivo 2: Desarrollar modelos de pronóstico de la volatilidad de las acciones del portafolio utilizando *Deep Learning*

En la sección 3.3.5, se desarrollaron los diferentes modelos de *Deep Learning*, que tienen como fin, realizar el pronóstico de la varianza. Para esto, en la sección 3.3.5.5 se probaron las arquitecturas RNN, LSTM y STACKED LSTM, de los cuales se encontró como mejor resultado la arquitectura LSTM. Sobre esta, se realizaron diferentes iteraciones sobre los parámetros e hiperparámetros a probar en cada acción, con el fin de hallar el modelo que mejor se ajuste, de lo anterior, se crearon los modelos correspondientes a cada acción, de los cuales, el resultado común fueron modelos con 1 rezago como variable independiente, datos sin escalar y 32 neuronas en su capa LSTM.

4.2.3 Objetivo 3: Estimar la rentabilidad del portafolio generado con el *Deep Learning* pronosticado

El cumplimiento de este objetivo, se realizó siguiendo el flujo definido para el cumplimiento del objetivo 1. Es decir, su proceso inicia con el pronóstico de la media realizado en la sección 3.2, siguiendo con el pronóstico de la varianza con modelos de *Deep Learning*, desarrollados en la sección 3.3.5 y mencionados más detalladamente en el cumplimiento

del objetivo 2 en la sección 7.2.2. Continuando su flujo, la generación de operaciones de compra y venta del portafolio desarrollado en la sección 6.1 y finalizando con la revisión de rendimientos generados por el portafolio de *Deep Learning* al final del escenario, desarrollados en la sección 6.2.

4.2.4 Objetivo 4: Comparar resultados de los portafolios

Este objetivo, se realiza con el apoyo en el cumplimiento de los objetivos 1 y 3, en los cuales se desarrollan los diferentes portafolios óptimos, utilizando el modelo de portafolio óptimo de Markowitz. En la sección 6.2, se puede encontrar el completo desarrollo de este objetivo, generando las comparaciones y conclusiones, de los diferentes portafolios creados.

4.2.5 Objetivo general

El objetivo final de este Trabajo Final de Maestría, es comparar la rentabilidad de un portafolio tradicional, conformado por el modelo de Markowitz, cuyas variables insumo se calculan de forma Tradicional, con un portafolio que utiliza *Deep Learning* para pronosticar la volatilidad de sus activos y conformar el portafolio con el modelo de Markowitz.

Este objetivo, fue logrado mediante el cumplimiento de los demás objetivos, realizando el pronóstico de la rentabilidad o media, desarrollado en la sección 3.2, además, se utilizan los residuales de estos modelos como insumo para el desarrollo del pronóstico de varianza, realizado en la sección 3.3. Luego de la construcción de ambas predicciones, se procede a ejecutar las diferentes operaciones en el portafolio, este desarrollo se ve reflejado en la sección 6.1. Con todo lo anterior, se realizan las 250 simulaciones, aproximadamente, para los 3 escenarios definidos, generando una rentabilidad anual que se puede observar y comparar en la sección 6.2

5. Referencias

Aguilar, S. (2010). Red neuronal artificial con heterocedasticidad condicional. . *Rionegro: Universidad Católica de Oriente.*

Amazon Web Services. (12 de 12 de 2022). *https://docs.aws.amazon.com/.* Obtenido de https://docs.aws.amazon.com/sagemaker/latest/dg/xgboost-HowItWorks.html

Ante, L. (2022). How Elon Musk's Twitter activity moves cryptocurrency markets. *Blockchain Research Lab*.

Assaf, O., Di Fatta, G., & Nicosia, G. (2022). Multivariate LSTM for Stock Market Volatility Prediction.

Bergstra, J., Bardenet, R., Bengio, Y., & Kégl, B. (s.f.).

Black, F., & Litterman, R. (1991). Asset Allocation. *The Journal of Fixed Income*, 1(2), 7–18.

Carmona, N. (2022). Redes neuronales regularizadas para la predicción de la inflación Colombiana. Tesis de Maestría en Ingeniería - Analítica. *Medellín: Universidad Nacional de Colombia. Facultad de Minas. Área Curricular de Sistemas e Informática*.

Carr, P., Wu, L., & Zhang, Z. (2020). Using Machine Learning to Predict Rrealized Variance. *Nueva York: University of New York. Department of Finance and Risk Engineering.*

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *University of Washington*.

Christensen, K., Siggaard, M., & Veliyev, B. (2021). A machine learning approach to volatility forecasting. *Aarhus: Aarhus University. Department of Economics and Business Economics*.

Echeverri, L. (2021). Conformación Automática de Portafolios de Inversión Usando Analítica Financiera. Tesis de Maestría en Ingeniería - Analítica. *Medellín: Universidad Nacional de Colombia. Facultad de Minas. Área Curricular de Sistemas e Informática*.

Fama, E. F. (1991). Efficient Capital Markets: II. *The Journal of Finance*.

Gandini, G. (14 de 01 de 2020). 2019, el año del Colcap. *Semana*.

Hutter, F., H. Hoos, H., & Leyton-Brown, K. (2011). Sequential Model-Based Optimization for General Algorithm Configuration. *University of British Columbia*.

I. Frazier, P. (2018). A Tutorial on Bayesian Optimization.

Liu, Y. (2019). Novel volatility forecasting using deep learning–Long Short Term.

Markowitz, H. (1959). Portafolio Selection. *Journal Finance*, 7(1), 77–91.

Mercado accionario. (22 de 12 de 2022).

Merton, R. C. (1971). Optimum Consumption and Portfolio Rules in a Continuous-Time Model. *Journal of Economic Theory*.

Sharpe, W. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *The Journal of Finance*, 19(3), 425.

TSAY, R. S. (2001). *Analysis of Financial Time Series.* Wiley.

Velásquez, J. D., Gutiérrez, S., & Franco, C. J. (2013). USING A DYNAMIC ARTIFICIAL NEURAL NETWORK FOR FORECASTING THE VOLATILITY OF A FINANCIAL TIME SERIES.

Villada, F., Muñoz, N., & Garcia, E. (2012). Aplicación de las Redes Neuronales al Pronóstico de Precios en el Mercado de Valores. *Medellín: Universidad de Antioquia. Facultad de Ingeniería. Departamento de Ingenieria Eléctrica*.