



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



ESCUELA TÉCNICA
SUPERIOR INGENIEROS
INDUSTRIALES VALENCIA

Análisis y desarrollo de modelos cuantitativos para la predicción de mercados y la optimización de estrategias de inversión.

Ricardo Garcés Hidalgo.

Tutor: José Ramón Navarro Cerdán.

Cotutor: Juan Carlos García Díaz.



Agradecimientos.

Quiero comenzar expresando mi más sincero agradecimiento a José Ramón y a Juan Carlos, por su apoyo constante, su dedicación y por brindarme la oportunidad de adentrarme en el tema de este Trabajo de Fin de Grado. Gracias por haber estado siempre pendientes, por cada sugerencia y por guiarme con paciencia en este proceso.

Más allá del TFG, me gustaría enfocarme en estos cuatro años lleno de retos, aprendizajes y momentos intensos, no habría llegado hasta aquí sin las personas que me han acompañado.

A mi familia: gracias a mi madre, a mi padre, a Sergio y a mis hermanas, por haberme soportado en los peores días, cuando sentía que no podía más, cuando me preguntaba una y otra vez si tanto esfuerzo tendría su recompensa. Vuestro apoyo constante ha logrado que no me rindiese nunca y que me focalizase en los momentos más importantes.

A mis amigos de la universidad, gracias por guiarme siempre hacia el buen camino y por hacer que cada madrugón para ir a clase mereciese la pena. A mis compañeros de siempre, gracias por seguir ahí como el primer día, por obligarme a salir, a respirar, y poner un poco de pausa con esos paseos tan necesarios.

Y, por último, a esas personas que más han estado conmigo en estos últimos meses, gracias por acompañarme en este tramo final del camino, por compartir mis últimos momentos de la carrera con risas y momentos inolvidables, por estar ahí con esa constancia tranquila que ha hecho que esta etapa sea inolvidable.

“Tomorrow, our seeds will grow. All we need is dedication”

-Lauryn Hill, Everything is Everything



Resumen.

La aplicación de modelos estadísticos abarca todos los sectores imaginables, y el mercado financiero no es una excepción. Hoy en día, estos modelos son fundamentales para la dinámica y el éxito de las finanzas modernas.

Este Trabajo Final de Grado (TFG) tiene como objetivo el desarrollo y análisis de modelos cuantitativos aplicados al ámbito financiero, centrados en la predicción del comportamiento del mercado bursátil y la optimización de carteras de inversión. La investigación se sustenta en dos ejes fundamentales: Modelización predictiva mediante técnicas estadísticas y de aprendizaje automático y la teoría de la optimización de inversiones.

La combinación de modelos predictivos avanzados con técnicas de optimización permite ofrecer herramientas innovadoras que mejoran la toma de decisiones en entornos de alta incertidumbre. Este enfoque resulta altamente aplicable, especialmente en áreas como la gestión de activos o el análisis de riesgos y Fintech, representando un avance frente a modelos tradicionales de análisis e inversión.

En el ámbito de la predicción de mercados, se aplicarán modelos estadísticos clásicos como Holt-Winters, ARIMA etc. Así como métodos avanzados basados en redes neuronales recurrentes, específicamente LSTM (Long Short-Term Memory), con el objetivo de anticipar tendencias en series temporales financieras.

Por otro lado, se abordará la optimización de carteras de inversión, referenciándose en modelos teóricos consolidados. Buscando un estudio que permita encontrar una relación entre rentabilidad y riesgo y un posterior diseño de estrategias eficientes.

Además, todo lo nombrado anteriormente será apoyado mediante el uso de indicadores técnicos específicos en los mercados financieros, que servirán como apoyo en la toma de decisiones y de gestión.

El trabajo incluirá además el diseño de una interfaz dinámica con el fin de integrar de manera visual y funcional todos los componentes de análisis.

En conclusión, este TFG aspira a demostrar y a aportar un cambio en el entorno bursátil, proporcionando métodos precisos y automatizados para el análisis y la gestión financiera, adaptándose a la creciente digitalización y a la exigencia en el que se encuentra este sector actualmente.

Palabras Clave: Predicción Bursátil, Modelos estadísticos, Redes Neuronales LSTM, Series Temporales, Optimización de Carteras, Indicadores técnicos, Análisis Cuantitativos, Interfaz Dinámico.



Resum.

L'aplicació de models estadístics abasta tots els sectors imaginables, i el mercat financer no és una excepció. Hui en dia, aquests models són fonamentals per a la dinàmica i l'èxit de les finances modernes.

Aquest Treball Final de Grau (TFG) té com a objectiu el desenvolupament i anàlisi de models quantitatius aplicats a l'àmbit financer, centrats en la predicció del comportament del mercat borsari i l'optimització de carteres d'inversió. La investigació es fonamenta en dos eixos principals: la modelització predictiva mitjançant tècniques estadístiques i d'aprenentatge automàtic, i la teoria de l'optimització d'inversions.

La combinació de models predictius avançats amb tècniques d'optimització permet oferir eines innovadores que milloren la presa de decisions en entorns d'alta incertesa. Aquest enfocament resulta altament aplicable, especialment en àrees com la gestió d'actius, l'anàlisi de riscos i les Fintech, representant un avanç respecte als models tradicionals d'anàlisi i inversió.

En l'àmbit de la predicció de mercats, s'aplicaran models estadístics clàssics com Holt-Winters, ARIMA, etc., així com mètodes avançats basats en xarxes neuronals recurrents, específicament LSTM (Long Short-Term Memory), amb l'objectiu d'anticipar tendències en sèries temporals financeres.

D'altra banda, s'abordarà l'optimització de carteres d'inversió, referenciant-se en models teòrics consolidats. Es buscarà un estudi que permeta trobar una relació entre rendibilitat i risc, així com un posterior disseny d'estratègies eficients.

A més, tot allò esmentat anteriorment serà recolzat mitjançant l'ús d'indicadors tècnics específics en els mercats financers, els quals serviran com a suport en la presa de decisions i la gestió. El treball inclourà, a més, el disseny d'una interfície dinàmica amb la finalitat d'integrar de manera visual i funcional tots els components d'anàlisi.

En conclusió, aquest TFG aspira a demostrar i aportar un canvi en l'entorn borsari, proporcionant mètodes precisos i automatitzats per a l'anàlisi i la gestió financera, adaptant-se a la creixent digitalització i a l'exigència en què es troba aquest sector actualment.

Paraules clau: Predicció Borsària, Models Estadístics, Xarxes Neuronals LSTM, Sèries Temporals, Optimització de Carteres, Indicadors Tècnics, Anàlisi Quantitatiu, Interfície Dinàmica.



Abstract.

The application of statistical models spans all imaginable sectors, and the financial market is no exception. Nowadays, these models are fundamental to the dynamics and success of modern finance.

This Final Degree Project (TFG) aims to develop and analyse quantitative models applied to the financial field, focusing on predicting stock market behaviour and optimizing investment portfolios. The research is based on two key pillars: predictive modelling through statistical and machine learning techniques, and the theory of investment optimization.

The combination of advanced predictive models with optimization techniques offers innovative tools that enhance decision-making in highly uncertain environments. This approach is highly applicable, especially in areas such as asset management, risk analysis, and Fintech, representing progress compared to traditional analysis and investment models.

In the field of market prediction, classical statistical models such as Holt-Winters and ARIMA will be applied, as well as advanced methods based on recurrent neural networks, specifically LSTM (Long Short-Term Memory), with the aim of anticipating trends in financial time series. On the other hand, portfolio optimization will be addressed, based on established theoretical models. The study will seek to identify a relationship between return and risk, followed by the design of efficient investment strategies.

Additionally, all the work will be supported using specific technical indicators in financial markets, which will serve as a basis for decision-making and management. The project will also include the design of a dynamic interface to visually and functionally integrate all analytical components.

In conclusion, this TFG aspires to demonstrate and contribute to change in the stock market environment, providing precise and automated methods for financial analysis and management, adapting to the growing digitalization and current demands of the sector.

Keywords: Stock Market Prediction, Statistical Models, LSTM Neural Networks, Time Series, Portfolio Optimization, Technical Indicators, Quantitative Analysis, Dynamic Interface.



Índice.

1. INTRODUCCIÓN.....	1
1.1 Resumen.....	1
1.2 Objetivo.....	1
1.3 Motivación.....	2
1.4 Estructura documental.....	3
2. MARCO TEÓRICO.....	3
2.1 Mercados Financieros: Clasificación y Características.....	3
2.1.1 Mercado de Acciones.....	5
2.1.2 Índices Bursátiles.....	7
2.1.3 Renta fija: Mercado de Tipos de Interés.....	9
2.1.4 Fondos Indexados.....	12
2.1.5 ETF: Fondo Cotizado en Bolsa.....	13
2.1.6 Commodities.....	15
2.2 La Predicción en el Mercado Financiero.....	17
2.3 Series Temporales.....	18
2.4 Modelos Predictivos de Base Econométrica.....	21
2.4.1 Holt-Winters: Modelo basado en Suavizamiento Exponencial.....	21
2.4.2 Modelos ARIMA y Extensiones.....	26
2.4.3 Random Forest: Modelo de Ensamble basado en Árboles de Decisión.....	33
2.4.4 Redes Neuronales LSTM.....	35
2.5 Indicadores Técnicos y su función en la Toma de Decisiones.....	40
2.5.1 Relative Strength Index (RSI).....	40
2.5.2 Bandas de Bollinger.....	41
2.5.3 Moving Average Convergence Divergence (MACD).....	42
2.5.4 Value at Risk (VAR).....	43
2.6. Teoría de la Optimización de Carteras.....	45
2.6.1. Modelo de Markowitz: La Teoría Moderna de Carteras.....	45
2.6.2. Modelo Black-Litterman: Avance en la Optimización de Carteras.....	47
3. MARCO METODOLÓGICO.....	50
3.1 Lenguaje de Programación.....	50
3.2 Análisis de Datos.....	51
3.2.1 Obtención de Datos.....	51
3.2.2 Limpieza de Datos.....	52



3.2.3 Procesamiento de Datos.	52
3.3 Implementación Computacional.....	53
3.3.1 Modelos Predictivos.....	53
3.3.2. Implementación de Indicadores Técnicos	59
3.3.3. Optimización de las Carteras.....	60
3.4. Criterios de Evaluación.	64
3.4.1. Evaluación Predictiva.....	64
3.4.2. Evaluación del Rendimiento de las Carteras.	65
4. ANÁLISIS Y DISCUSIÓN DE LOS RESULTADOS.....	68
4.1. Evaluación de Modelos de Predicción.	68
4.2. Evaluación de Métodos de Optimización de Carteras.....	72
5. IMPLEMENTACIÓN DE LA INTERFAZ DINÁMICA.....	73
5.1. Lenguaje y Modelado.....	73
5.2. Visualización.	76
7. CONCLUSIÓN.	78
8. PERSPECTIVAS FUTURAS.	78
9. REFERENCIAS.	79
ANEXO I. PRESUPUESTO.....	82
1. Introducción.	82
2. Presupuesto Parcial.	82
2.1 Análisis y Diseño.....	82
2.2 Instalación y Desarrollo.	82
2.3 Validación, Calidad y Mantenimiento.	83
2.4 Licencias de Software y Equipo.....	83
2.5 Documentación Técnica	84
3. Presupuesto Total.	85
4. Conclusión.....	85
ANEXO II. OBJETIVOS DE DESARROLLO SOSTENIBLE.....	86



Índice de Figuras.

Figura 1. Tipos y clasificación de mercados financieros.....	4
Figura 2. Distribución Geográfica de los 10 Índices más representativos del mundo	9
Figura 3. Distribución Matemática de los Índices Bursátiles.....	9
Figura 4. Número de Empresas Españolas No Financieras Emisoras de Renta Fija.....	10
Figura 5. Ejemplificación de los Tipos de Interés al Contado.....	10
Figura 6. Ejemplificación de los Tipos de Interés forward.	11
Figura 7. Rentabilidad Neta de los Fondos Indexados vs Rentabilidad Neta de los Fondos de gestión Activa.....	12
Figura 8. Mejores 10 Fondos Cotizados en Bolsa de Renta Variable en el primer semestre de 2025.....	14
Figura 9. Variación Porcentual del Petróleo VS Inflación en EE.UU	16
Figura 10. Representación de los Componentes de las Series Temporales	19
Figura 11. Diferencias Representativas entre Esquema Aditivo y Multiplicativo.....	20
Figura 12. Representación Simulada de Modelo Holt-Winters.....	24
Figura 13. Representación del Ruido Blanco.....	27
Figura 14. ACF simulado de una Serie NO estacionaria.....	29
Figura 15. PACF con Componente AR.	29
Figura 16. Circulo de la Unidad en el Plano Complejo de un modelo ARIMA.....	30
Figura 17. ACF de Autocorrelación Simple Figura 18. PACF de Autocorrelación Simple. 30	
Figura 19. Representación Simulada del modelo ARIMA.....	32
Figura 20. División de los Nodos de un Random Forest.....	33
Figura 21. Ejemplo Simulado de Un Random Forest.	35
Figura 22. Red Neuronal Biológica VS Artificial	36
Figura 23. Representación multidimensional del Descenso del Gradiente	37
Figura 24. Funcionamiento de las Redes Neuronales Recurrentes	38
Figura 25. Funcionamiento de un LSTM.	38
Figura 26. Representación Simulada de las Bandas de Bollinger	42
Figura 27. Ejemplo Simulado del Indicador MACD	43
Figura 28. Visualización Distribución Normal del VaR.....	44
Figura 29. Frontera Eficiente Rentabilidad-Riesgo.....	46
Figura 30. Frontera Eficiente Rentabilidad-Riesgo + CAL.	47
Figura 31. Distribución Normal Ejemplo Black-Litterman	49
Figura 32. Representación Data Frame en Python.....	51



Figura 33. Correlación entre Variables Exógenas del Modelo ARIMAX.	57
Figura 34. Ejemplo Simulado del Valor de una Cartera de Inversión y su Drawdown.	68
Figura 35. Representación del Modelo predictivo LSTM estudiado para 2024.	71
Figura 36. Objetos Dinámicos en la Interfaz Implementada.	76
Figura 37. Ejemplo Simulado de una Acción en los Objetos Dinámicos de la interfaz.	77
Figura 38. Gráficos de la Interfaz Dinámica Generada.	77



Índice de Ecuaciones.

Ecuación 1. Modelo de Dividendos Descontados.	6
Ecuación 2. Modelo de Dividendos Descontados con Crecimiento h Constante.....	6
Ecuación 3. Simplificación del MDD.	6
Ecuación 4. Precio de las Acciones Fundamento con el MDD.	7
Ecuación 5. Valor del Índice para Market-Cap Weighted.	8
Ecuación 6. Valor del Índice para Price-Weighted.	8
Ecuación 7. Valor de Índice para Equal-Weighted.	8
Ecuación 8. Tipo Spot.	10
Ecuación 9. Tipo Forward Figura 6. Ejemplificación de los Tipos de Interés forward.	11
Ecuación 10. Tasa Interna de Rendimiento.	11
Ecuación 11. Serie temporal con Esquema Aditivo.	19
Ecuación 12. Serie Temporal con Esquema Multiplicativo.	20
Ecuación 13. Serie Temporal con Esquema Mixto.	20
Ecuación 14. Suavizamiento Exponencial Simple.	22
Ecuación 15. Suavizamiento Exponencial con término Sumatorio.....	22
Ecuación 16. Método de Doble Suavizamiento Exponencial.	22
Ecuación 17. Método De Holt-Winters con Estacionalidad Aditiva.	23
Ecuación 18. Método Holt-Winters con Estacionalidad Multiplicativa.	23
Ecuación 19. Proceso Autorregresivo del Modelo ARIMA.	28
Ecuación 20. Operador de Retardo Autorregresivo.....	28
Ecuación 21. Comprobación del Modelo Autorregresivo.	28
Ecuación 22. Obtención de la Varianza del proceso mediante Yule-Walker.	28
Ecuación 23. Proceso Media Móvil del Modelo ARIMA.	29
Ecuación 24. Operador de Retardo de la Media Móvil del Modelo ARIMA.....	29
Ecuación 25. Función de Autocorrelación de la Media Móvil.	30
Ecuación 26. Equivalencias de Operadores de Retardo para el proceso ARMA.	31
Ecuación 27. Proceso Integrativo del modelo ARIMA.	31
Ecuación 28. Operadores de Retardo para la extensión ARIMA.	32
Ecuación 29. Proceso SARIMA.....	32
Ecuación 30. Polinomios de Grado P y Q.	32
Ecuación 31. Media Móvil Simple.....	34
Ecuación 32. Retorno del Valor.....	34
Ecuación 33. Volatilidad.....	34



Ecuación 34. GINI del modelo Random Forest.	34
Ecuación 35. Regresión Lineal de una Neuron Artificial.	36
Ecuación 36. Gradiente del modelo Neuronal Artificial.	37
Ecuación 37. Sucesión para el Descenso del Gradiente.	37
Ecuación 38. Forget Gate Layer LSTM.	39
Ecuación 39. Entrada de Activación Sigmoidal del Input Gate Layer para LSTM.	39
Ecuación 40. Entrada de Activación tangencial del Input Gate Layer para LSTM.	39
Ecuación 41. Cell State para LSTM.	39
Ecuación 42. Activación del Output Layer.	39
Ecuación 43. Output Layer para LSTM.	40
Ecuación 44. Media Móvil Exponencial.	40
Ecuación 45. Relative Strength.	40
Ecuación 46. Relative Strength Index.	41
Ecuación 47. Desviación Estándar para las Bandas de Bollinger.	41
Ecuación 48. Límites Superiores e Inferiores de las Bandas de Bollinger.	41
Ecuación 49. Proceso MACD.	42
Ecuación 50. Proceso del Value at Risk.	43
Ecuación 51. Rentabilidad Esperada de una Cartera.	46
Ecuación 52. Varianza de la Cartera.	46
Ecuación 53. Rentabilidad Esperada de una Cartera + CAL.	47
Ecuación 54. Proceso Black - Litterman.	48
Ecuación 55. Vector Equilibrio del Mercado.	48
Ecuación 56. Vector de Retorno Sujeto a Opiniones del Inversor.	48
Ecuación 57. Distribución Normal del vector de Retornos Esperados.	48
Ecuación 58. Distribución Normal de errores asociados a Valores Subjetivos.	48
Ecuación 59. Matriz de Confianza Black-Litterman.	49
Ecuación 60. Riesgo de Aversión.	62
Ecuación 61. Error Medio Absoluto.	64
Ecuación 62. Error Cuadrático Medio.	65
Ecuación 63. Sharpe Ratio.	66
Ecuación 64. Maximum Drawdown.	67



Índice de Tablas.

Tabla 1. 10 índices Bursátiles más consolidados en el Mercado Financiero.....	8
Tabla 2. Comisiones Aplicables en los Fondos Indexados.....	13
Tabla 3. Mejores 10 ETF's de Renta Fija en el Primer Semestre de 2025 según su crecimiento porcentual.....	15
Tabla 4. Pronóstico Simulado de Holt-Winters.....	25
Tabla 5. Nivel, Tendencia y Estacionalidad Simulada.	25
Tabla 6. Valores de Distribuciones Normales.....	44
Tabla 7. Nombre de Activo VS Ticker.....	54
Tabla 8. Mercados Seleccionados para la Evaluación de Modelos.	69
Tabla 9. RMSE y MAE Obtenidos por Holt-Winters.....	69
Tabla 10. RMSE y MAE Obtenidos por SARIMA.....	69
Tabla 11. RMSE y MAE Obtenidos por ARIMAX.....	69
Tabla 12. RMSE y MAE Obtenidos por Random Forest.	70
Tabla 13. RMSE y MAE Obtenidos por LSTM.....	70
Tabla 14. Pesos por Modelo Markowitz.....	72
Tabla 15. Pesos por Modelo Black-Litterman.....	73



Índice de Estructura del Código.

Estructura del Código 1. Importación de la librería Yahoo! Finance y de la librería World Bank.	51
Estructura del Código 2. Eliminación de Datos Nulos.....	52
Estructura del Código 3. Eliminación de Columnas Irrelevantes.....	52
Estructura del Código 4. Evitar Datos Duplicados.....	52
Estructura del Código 5. Importación de Librerías.	53
Estructura del Código 6. Implementación del Ticker.	54
Estructura del Código 7. Descarga de Datos de Yahoo! Finance.	54
Estructura del Código 8. Separación de los Datos que se van a Entrenar.	54
Estructura del Código 9. Modelo Holt-Winters.	54
Estructura del Código 10. Modelo SARIMA.	55
Estructura del Código 11. Implementación de datos Yahoo Finance! y World Bank.....	56
Estructura del Código 12. Modelo ARIMAX.	56
Estructura del Código 13. Correlación de Variables Exógenas del modelo ARIMAX.	57
Estructura del Código 14. Implementación Variables del Random Forest.....	58
Estructura del Código 15. Implementación del Intervalo de Entrenamiento del Random Forest.	58
Estructura del Código 16. Modelo Random Forest.....	58
Estructura del Código 17. Precios Normalizados Para LSTM.....	58
Estructura del Código 18. Periodo de Entrenamiento y de Pronóstico para un LSTM.....	58
Estructura del Código 19. Modelo LSTM.....	59
Estructura del Código 20. Desnormalización del Modelo LSTM.....	59
Estructura del Código 21. Indicador Relative Strength Index.....	60
Estructura del Código 22. Indicador de las Bandas de Bollinger.....	60
Estructura del Código 23. Indicador MACD.....	60
Estructura del Código 24. Indicador Value at Risk.	60
Estructura del Código 25. Conjunto de Tickers descargados por Yahoo! Finance.....	61
Estructura del Código 26. Método Markowitz.....	61
Estructura del Código 27. Implementación de la Variables para Método Markowitz.....	61
Estructura del Código 28. Resultado Diario Markowitz.	62
Estructura del Código 29. Resultado Anual Markowitz.....	62
Estructura del Código 30. Variables Modelo Black-Litterman.	62
Estructura del Código 31. Variables Subjetivas Método Black-Litterman.....	63



Estructura del Código 32. Modelo Black-Litterman.	63
Estructura del Código 33. Resultados Óptimos Black-Litterman.	64
Estructura del Código 34. Cálculo de MAE.	65
Estructura del Código 35. MAE y RMSE.	65
Estructura del Código 36. Sharpe Ratio.	66
Estructura del Código 37. Maximum Drawdown.	67
Estructura del Código 38. Objetos Dinámicos del Interfaz.	74
Estructura del Código 39. Interpretación del Intervalo de Fechas a Pronosticar.	75
Estructura del Código 40. Predicción del Pronóstico para el Interfaz Dinámico.	75
Estructura del Código 41. Estilización del Gráfico Dinámico.	75
Estructura del Código 42. Implementación del Gráfico Valor de Mercado + Pronóstico.	76
Estructura del Código 43. Implementación del Gráfico "ZOOM".	76



1. Introducción.

1.1 Resumen.

El presente trabajo se presenta como una ayuda a la evolución de la transformación digital de los mercados financieros. En este se trata de complementar, e incluso discutir, los métodos tradicionales de inversión y predicción bursátil. Para ello, se abordan nuevas herramientas fundamentadas en la utilización de datos y la interpretación predictiva.

Este análisis combina dos de los grandes pilares en la gestión económica: la predicción de mercados mediante bases estadísticas, y la optimización de carteras de inversión sujeto a fundamentos basados en los criterios de rentabilidad-riesgo. Dicha dualidad refleja una tendencia actual del sector financiero, donde se pretende generar modelos con capacidades de alta precisión en el cambiante terreno de la bolsa de valores.

En lo que se refiere al ámbito predictivo, se realiza una comparación tanto teórica como experimental de sistemas más conservadores como Holt-Winters o ARIMA, con modelos más complejos y sofisticados como el Random Forest o las Redes Neuronales Artificiales. Todo ello se ve enriquecido por indicadores técnicos que permiten un mejor entendimiento de la situación en la que se encuentra el mercado.

En paralelo, se aborda la creación óptima de un portafolio de inversión. Para ello, se desarrolla métodos que pretenden buscar una eficacia mediante los fundamentos del modelo de Markowitz y de Black-Litterman, resaltando una diferencia entre estos procedimientos enfocada en la flexibilidad y la subjetividad del agente financiero.

Además, para contemplar estas estructuras, se implementa en un entorno global como es Google Colab, un lenguaje computacional totalmente aplicable como es Python. De esta manera se puede observar la funcionalidad real de este trabajo. También se presenta una interfaz dinámica como medio para facilitar el acceso a este análisis.

Sintetizando todas las ideas, este trabajo pretende que la ingeniería, la estadística y las finanzas converjan en un nuevo paradigma estratégico basado en datos, combatiendo los desafíos que presentan la economía digital y ofreciendo nuevos puntos de vista.

1.2 Objetivo.

La finalidad se basa en diseñar, desarrollar y aplicar procesos cuantitativos que faciliten la anticipación de la incertidumbre en los mercados y la optimización de estrategias de inversión. El propósito es mejorar la toma de decisiones en un contexto financiero complejo, ofreciendo



soluciones precisas y automatizadas, ajustadas a las necesidades evolutivas de la transformación digital. Además, se ofrece como base a un proyecto futuro donde se puede llegar a ensamblar los sistemas que se definen en este trabajo y que sustentan los mejores rendimientos, presentándolo a un público interesado en la bursatilidad y cotización de la bolsa.

1.3 Motivación.

En un mundo donde la toma de decisiones se ve cada vez más influenciada por la digitalización, los algoritmos y los datos, el ámbito financiero no es ninguna excepción. Hoy en día, existe una autentica necesidad en entender el funcionamiento interno de la bolsa. La tendencia a proponer cambios tecnológicos más sofisticados ha transformado por completo los métodos de gestión y negocio. En este punto, es donde surge el impulso y la motivación de utilizar análisis numéricos como herramientas fundamentales para aquellos que buscan solucionar dicho problema, anticipándose y actuando sobre los movimientos del mercado.

Como señala (O'Neil, 2016), los modelos cuantitativos no son otra cosa más que opiniones incrustadas en matemáticas, diseñados con el objetivo de convertirse en el aliado más poderoso para las adopciones de decisiones. La capacidad de poder pronosticar tendencias de mercado mediante bases estadísticas no solo presenta un avance técnico a nivel global, sino una transformación de los cánones focalizados en la forma de abordar una inversión.

Por otra parte, en la actualidad los portafolios estratégicos también han sufrido una rápida evolución digitalizada. Hoy en día, las decisiones ya no residen principalmente en la intuición del gestor, las últimas actualizaciones rigen una obligatoriedad de procesar grandes volúmenes de información, permitiendo identificar patrones para crear estrategias diversificadas y adaptables. Defendiendo las palabras sujetas a (Dalio, 2017), los inversores cometen el grave error de creer que lo que ha ocurrido en un pasado seguirá sucediendo en el futuro. Por eso, fundamentarse en herramientas numéricas ayuda a reducir el sesgo humano y a cuantificar la incertidumbre.

Pero más allá de los modelos, las gráficas y las estrategias, este trabajo surge de una convicción personal. En un sistema financiero normalmente inaccesible e injusto, la tecnología y la ciencia de datos permite una extensión de oportunidades, eliminando barreras. Esta motivación no nace únicamente gracias a la fascinación técnica de las finanzas, sino ser un grano de arena más en la montaña que permita un cambio estructural. Porque como dice (Yunus, 2007), no se puede afrontar el desafío de un negocio pensando que es un problema, sino una solución.

En definitiva, este trabajo pretende focalizar un punto en común entre la estadística, el aprendizaje automático y las finanzas, demostrando como su aplicación permite construir tanto estrategias más robustas como facilidades para la introducción a la bolsa bursátil.



1.4 Estructura documental.

Este documento se dividirá en tres grupos principales, primero se expone un marco teórico con el fin de definir y entender los diferentes componentes que se presentan. Por otro lado, se representa un marco metodológico con el fin de exponer de manera práctica el funcionamiento de los distintos modelos destacados en este trabajo. El último punto principal sugiere una evaluación y discusión de los resultados obtenidos para formalizar una respuesta a la duda sobre que método.

Además, se añade otros tres grupos adicionales. Uno centrado en la implementación de una representación dinámica, con el fin de simplificar una visualización de los resultados. Los otros dos analizan el comportamiento de un proceso real simulado y ofrecen una posibilidad futura para mejorar lo que se expone.

2. Marco Teórico.

2.1 Mercados Financieros: Clasificación y Características.

Enfocándose en el contexto histórico, los Mercados Financieros se consolidaron de manera oficial durante el siglo XVII, específicamente en 1602, con la creación de la Bolsa de Ámsterdam, en los Países Bajos. En este año, la Compañía Neerlandesa de las Indias Orientales (VOC) fue la primera corporación en emitir acciones de manera pública, estableciendo una base fundamental en el sistema financiero moderno. Desde entonces, estos mercados han evolucionado exponencialmente, permitiendo un desarrollo extraordinario a las nuevas economías.

Los Mercados Financieros se podrían definir como estructuras complejas, de constante evolución, creadas con el principal objetivo de permitir intercambios de activos financieros, así como la fijación de sus precios.

Dentro de este campo, un Mercado Bursátil es un tipo de específico de mercado financiero donde se negocian e intercambian acciones, bonos y otros tipos de activos financieros. Su principal objetivo es poner en contacto a agentes económicos como inversores, intermediarios financieros y otros actores del sistema. Son estructuras fundamentales en el ámbito de las finanzas modernas, ya que proporcionan liquidez, transparencia y eficiencia a las transacciones de activos, contribuyendo de manera óptima a las economías globales.

A la hora de clasificar los mercados financieros, cabe la posibilidad de agruparlos según infinidad de normas debido a la complejidad que estos ofrecen. No obstante, una de las maneras más comunes y sencillas para entender los diferentes tipos de mercados es mediante su grado de gestión y rentabilidad. Esta clasificación considera como se estudian las transacciones, los activos negociados y las monedas utilizadas mediante una principal diferenciación entre los mercados con moneda nacional o los mercados con divisa.

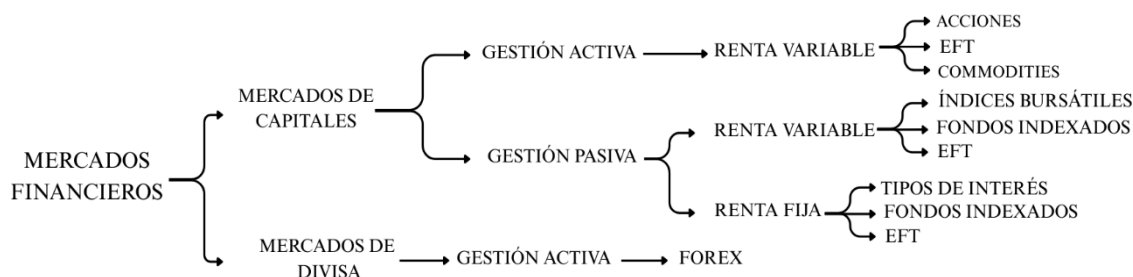


Figura 1. Tipos y clasificación de mercados financieros.

Fuente: De elaboración propia.

A partir de esta clasificación, es posible enfocar el presente trabajo hacia un análisis preciso de los mercados financieros dentro del ámbito del mercado de capitales. Esta estructura servirá para abordar de manera detallada el funcionamiento y las aplicaciones de estos tipos de instrumentos económicos, esenciales para la gestión de riesgo, rentabilidad y formación de precios en los mercados modernos.

Antes de profundizar en los diferentes tipos de mercados financieros, es fundamental comprender algunos términos básicos que suelen aparecer en el contexto. Estas definiciones permitirán interpretar adecuadamente el lenguaje técnico utilizado en el ámbito económico y financiero:

- **Activo financiero**, Se trata de un instrumento que otorga a su titular un derecho económico sobre otra entidad. Los activos financieros derivan su valor en función de diferentes variables, como la oferta, la demanda o diferentes expectativas y especulaciones de mercado.
- **Liquidez**, se refiere a la capacidad de un activo para transformarse fácilmente en dinero en efectivo sin que se vea una alteración significativa de su valor. Es una propiedad esencial en los mercados financieros, ya que permite a los inversores comprar o vender activos sin afectar demasiado su precio.
- **Riesgo financiero**, es la posibilidad de que los resultados reales de una decisión financiera difieran de los esperados. En términos más simples, implica un cierto nivel de incertidumbre, y puede suponer tanto pérdidas como ganancias imprevistas. El riesgo puede derivarse de diversos factores: cambios en los tipos de interés, variaciones de precios en el mercado, incumplimientos crediticios, movimientos geopolíticos, o incluso eventos imprevistos como crisis sanitarias.
- **Rentabilidad**, es el rendimiento económico que genera una inversión respecto al valor capital invertido, en un período temporal determinado. Se expresa generalmente en porcentaje y permite medir la eficiencia o el desempeño de una inversión.
- **Interés**, es el coste del dinero en el tiempo. En términos financieros, se refiere a la retribución que se obtiene por prestar dinero, generalmente expresado en forma de porcentaje sobre el capital inicial.



- **Volumen**, representa la cantidad de activos financieros que se han negociado en un periodo determinado (por ejemplo, número de acciones intercambiadas en una sesión bursátil).

2.1.1 Mercado de Acciones.

“Una acción es un título que da a un individuo la propiedad de una parte pequeña de una empresa” (Gómez Pineda, Javier G., 2010, p.183).

El mercado de acciones se enmarca en la clasificación de los valores negociados en el mercado de renta variable, dado que la rentabilidad periódica que genera no es fija ni predecible. Esto contrasta con los instrumentos de renta fija, como los bonos, que ofrecen pagos periódicos preestablecidos y con menor tendencia a la volatilidad.

El mercado financiero permite que el patrimonio de las grandes corporaciones esté dividido por multitud de partes, permitiendo que los inversores individuales o institucionales puedan ser dueños de pequeñas fracciones de estas, convirtiéndose en copropietarios de la entidad. Cada fracción en la que se dividen estos activos se le denomina acción.

En una sociedad es posible encontrar dos grandes grupos de accionistas:

- **Accionistas de referencia**, aquellos que debido a su elevado porcentaje de participación tienen una significativa influencia en la gestión y la dirección de la empresa.
- **Accionistas minoritarios**, poseen una menor proporción del capital social y, por lo tanto, no tienen capacidad de intervención, engloba a la mayoría de los individuos.

Refiriéndose al primer grupo nombrado anteriormente, al adquirir acciones, el representante de estas obtiene ciertos derechos tanto económicos como políticos según la cantidad de acciones que posee.

Dentro de los derechos económicos, se puede encontrar diferentes puntos de vista.

- **Derecho al dividendo**, es el derecho al beneficio que la junta general de accionistas decide repartir a propuesta del consejo de administración.
- **Derecho de suscripción**, con el fin de evitar la disolución de la participación de los accionistas, cuando ciertas entidades llevan a cabo ampliaciones de capital, sus participantes se benefician de la preferencia de obtención de nuevas acciones sobre el resto de los individuos.
- **Derecho a la cuota de liquidación**, en caso de que la sociedad llegase a disolverse, los accionistas tendrían derecho a recibir una proporción de los activos remanentes.

En lo que se refiere a los derechos políticos, algunos subapartados que contempla este apartado son los siguientes.

- **Derecho de información**, permite un seguimiento actualizado de las acciones de la entidad
- **Derecho de asistencia y voto** en las juntas de accionistas.

- **Derecho a impugnar los acuerdos adoptados** por la junta general de accionistas.

Las acciones normalmente son transadas en mercados secundarios, como las bolsas de valores, permitiendo dos grandes impactos: por una parte, ofrece la posibilidad de que una parte de la empresa pueda ser trasladada de un individuo a otro fácilmente, dando lugar a lo que se conoce hoy en día como compraventa de acciones o *trading*. Por otra parte, y a consecuencia de la primera, otorga al propietario una gran liquidez y facilita la diversificación de las carteras de inversión.

Para todos los individuos involucrados en el mercado bursátil, siempre existe una incertidumbre, también conocido como riesgo, sobre el resultado futuro. En el caso de la inversión en acciones, el riesgo para el propietario recae en la pérdida o en la subida de su capital, como consecuencia de la alteración del precio de las mismas acciones. Generalmente, el riesgo de invertir en acciones suele ser mucho mayor al de intervenir en otras actividades financieras como por ejemplo los bonos o los índices. Sin embargo, las acciones también ofrecen una mayor rentabilidad potencial a largo plazo, lo que se conoce como prima por riesgo.

A la hora de poder calcular el precio de una acción, el método principal para poder conocer su valor es el Modelo de Dividendos Descontados (DDM). Este nombre se da ya que hace referencia a que el precio de una acción es igual a los dividendos futuros descontados a valor presente.

La ecuación en su expresión más sencilla y básica sería la siguiente.

$$P_0 = \sum_{i=1}^n \frac{D_i}{(1 + j_i)^i}$$

Ecuación 1. Modelo de Dividendos Descontados.

Donde P_0 es el valor de la acción en el periodo $t = 0$, la variable D_i se atribuye a los dividendos que la entidad distribuye en el periodo i y j_i es la tasa de interés en el mismo periodo.

La expresión anterior se puede simplificar en el hipotético caso de que los dividendos crecieran a una tasa constante h , y que, además, la tasa de interés permaneciese constante.

$$P_0 = \sum_{i=1}^n \frac{(1 + h)^i D_0}{(1 + j)^i}$$

Ecuación 2. Modelo de Dividendos Descontados con Crecimiento h Constante.

Además, se podría simplificar de manera algebraica quedando de la siguiente manera.

$$P_0 = \sum_{i=1}^n \frac{D_1}{(i - h)}$$

Ecuación 3. Simplificación del MDD.

Por último, enfocándose en la característica definida anteriormente como riesgo, es interesante conocer que la tasa de interés total está compuesta por la suma de la propia tasa de interés libre de riesgo (i_F) más la prima por el riesgo (ϕ_E) de las acciones. La tasa de interés libre de riesgo es



aquella impuesta por los rendimientos de los bonos del gobierno, es igual a la tasa de interés de inflación más la inflación esperada.

Atendiendo a la evolución de las ecuaciones anteriores, finalmente la ecuación que determinaría el precio de las acciones sería la siguiente.

$$P_0 = \sum_{i=1}^n \frac{D_1}{i_F + \phi_E - h}$$

Ecuación 4. Precio de las Acciones Fundamento con el MDD.

2.1.2 Índices Bursátiles.

Un Índice Bursátil se define como un indicador financiero que refleja la evolución del valor de un grupo de activos que cotizan en un mercado común y específico. Dichos activos, como acciones o bonos, comparten características parecidas, como pertenecer a la misma región geográfica o sector financiero. La función principal de este instrumento económico se centra en medir el rendimiento de la selección representativa de estos valores financieros, permitiendo una evaluación rápida y sencilla del comportamiento real del mercado al que se le asocia.

En términos generales, los índices bursátiles resumen y sintetizan el comportamiento de los precios de un grupo de activos, permitiendo una visualización clara de las tendencias del mercado. Permitiendo a los diferentes agentes económicos tomar decisiones informadas basadas en datos precisos.

Para clarificar este suceso con un ejemplo, si se dice que el Ibex-35, el cual está compuesto por los 35 valores más líquidos de las cuatro Bolsas españolas del Mercado Continuo, ha subido un 2% con respecto la sesión anterior, esto indica que, en promedio, las 35 empresas que componen dicho índice bursátil han incrementado su valor en un 2% durante este periodo.

Atendiendo a la clasificación de este tipo de mercado, los índices bursátiles pueden estructurarse de distintas maneras, siendo una de las más comunes la agrupación según la región geográfica de los activos que la componen. Esta clasificación es especialmente útil para evaluar el diferente desempeño de los mercados a nivel mundial, ya que cada geolocalización está determinada por características políticas, económicas y financieras muy distintas que influyen en sus mercados de manera particular y única. Además, esto permite un análisis comparativo de las diferentes economías mucho más sencillo, permitiendo a los diferentes agentes financieros diversificar su portafolios de una manera más eficiente y óptima.

Otra forma bastante reconocida de clasificarlos es basándose en el método de cálculo del valor del índice que representan. El valor de un índice bursátil es una medida clave para el seguimiento del desempeño de este en un mercado específico. Refleja las condiciones económicas y permite identificar tendencias alcistas o bajistas. Según el método se pueden clasificar de la siguiente manera.

- **Índices Ponderados por Capitalización de Mercado (Market-Cap Weighted):** Las empresas con mayor capitalización tienen un peso proporcionalmente más alto en el índice.

$$\text{Valor del Índice}(\%) = \frac{\sum_{i=1}^n (P_i \times N_i \times IWF_i)}{\text{Divisor del Índice}}$$

Ecuación 5. Valor del Índice para Market-Cap Weighted.

- **Índices Ponderados por Precio (Price-Weighted):** No depende del tamaño de capitalización de la empresa, solo del precio accionario de esta.

$$\text{Valor del Índice}(\%) = \frac{\sum_{i=1}^n P_i}{\text{Divisor del Índice}}$$

Ecuación 6. Valor del Índice para Price-Weighted.

- **Índices Ponderados por Igual (Equal-Weighted):** Todas las acciones tienen el mismo peso.

$$\text{Valor del Índice}(\%) = \frac{1}{n} \sum_{i=1}^n R_i$$

Ecuación 7. Valor de Índice para Equal-Weighted.

Donde P_i es el precio de cada acción y se refiere al precio de mercado de cada activo en un momento dado, el N_i es el número de acciones en circulación de una entidad que están disponibles para ser negociadas, el IWF_i es el factor de flotación libre de cada empresa, refiriéndose al porcentaje que refleja la porción de las acciones disponibles para ser negociada libremente. Por otra parte, el *Divisor del Índice* es un factor de ajuste utilizado con la finalidad de mantener consistencia en el índice a lo largo del tiempo. Por último, el R_i hace referencia al rendimiento de la acción i . Con el fin de ejemplificar la distribución geográfica y metodología de cálculo de los índices bursátiles, a continuación, se presenta una clasificación de los 10 más reconocidos a nivel mundial.

Índice	Código	Activos Principales
MSCI- World	W1DOW	1500 Acciones Globales Desarrolladas.
FTSE All-World	AWORLDS	3900 Acciones Globales Desarrolladas.
EURO STOXX 50	SX5E	50 Mayores Empresas de la Eurozona.
Down Jones Stoxx 600	SXXP	600 Empresas Europeas.
IBEX-35	IBEX	35 más Líquidas de la Bolsa Española.
S&P 500	GSPC	500 Mayores Empresas de EE.UU.
Down Jones Industrial Average	DJI	30 Mayores Empresas Industriales de EE.UU.
NASDAQ Composite	IXIC	Más de 3000 Acciones Tecnológicas.
DAX 40	GDAXI	40 Mayores Empresas Alemanas
Nikkei 225	N225	225 Grandes Empresas Japonesas

Tabla 1. 10 índices Bursátiles más consolidados en el Mercado Financiero.

Fuente: De Elaboración Propia.

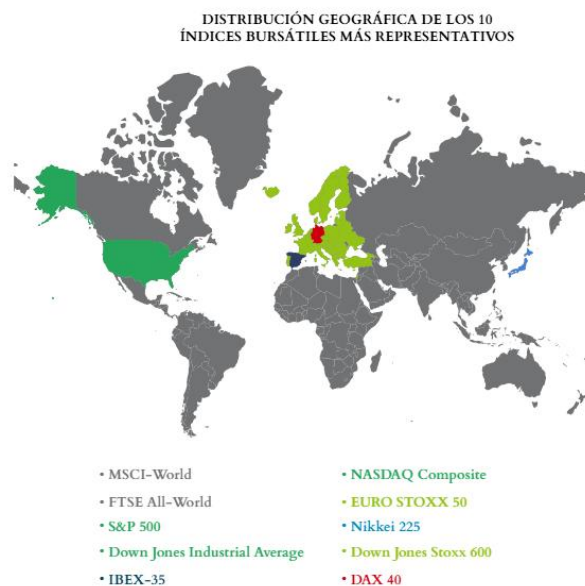


Figura 2. Distribución Geográfica de los 10 Índices más representativos del mundo
Fuente: De elaboración propia



Figura 3. Distribución Matemática de los Índices Bursátiles.
Fuente: De elaboración propia.

2.1.3 Renta fija: Mercado de Tipos de Interés.

En el ámbito de la economía bursátil, el mercado enfocado a la renta fija y crédito desempeña una parte esencial en la estructura financiera global. De hecho, según un informe actualizado de la Organización para la Cooperación y el Desarrollo Económico (OCDE) citado por *El País* (2025), se estima que el flujo de los bonos soberanos y otros instrumentos de renta fija alcancen un récord de 17 billones de dólares en 2025. Estos activos son fundamentales para el funcionamiento económico, dado que “permiten asignar el precio al coste del dinero y del paso del tiempo” (Fernández Rodríguez et al., 2013, p.1).

En la Figura 3 se puede apreciar la tendencia de la participación de las corporaciones ibéricas en el mercado de renta fija, evidenciando una evolución alcista a lo largo de la última década del siglo XXI, con un incremento sostenido hasta 2021.

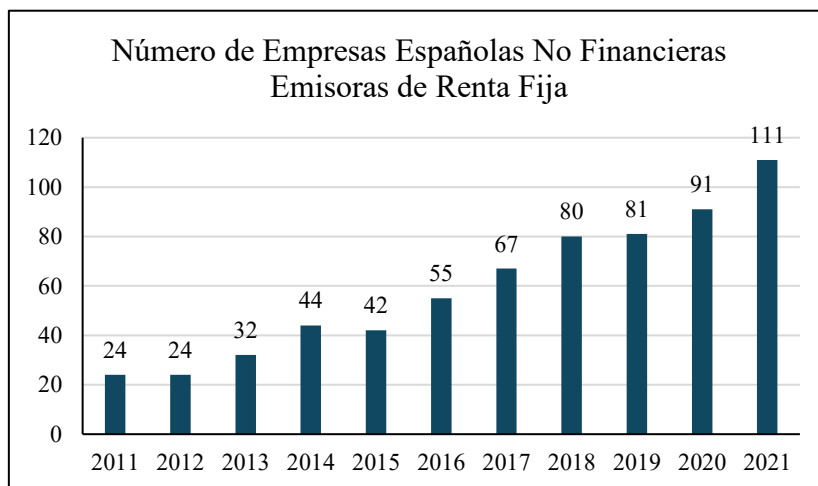


Figura 4. Número de Empresas Españolas No Financieras Emisoras de Renta Fija
Fuente: Elaboración propia a partir de datos del Banco de España.

El término de tipo de interés está asociado al costo de usar dinero durante un periodo de tiempo. Es, en esencia, la compensación que espera recibir quien presta un capital y el costo que debe pagar aquel quien lo recibe. "El interés es el precio que se paga por el uso del dinero durante un periodo de tiempo." (Keynes, 1936).

Aunque quepa la posibilidad de encontrarse una infinita diversidad de clasificaciones de tipos de interés, a la hora de estudiarlos con respecto a la evolución temporal de estos, se han de diferenciar tres grupos principales.

- **Tipos de Interés al contado (Tipo Spot)**, es la rentabilidad efectiva anual que se obtiene al comprar un activo de renta fija y mantenerlo hasta su vencimiento. Es un tipo de interés basado en un periodo fijo de tiempo, es decir, no incluye pagos intermedios, tan solo un único flujo de entrada y un único flujo de salida. Además, el rendimiento de este es perfectamente conocido desde el momento de la compra, por lo que no es necesario un estudio de previsiones futuras, ya que se elimina la incertidumbre asociada a fluctuaciones del mercado. El cálculo de la rentabilidad esperada R_n vendría dado por el precio nominal del activo en el mismo momento P_0 , el valor de reembolso P_n en el vencimiento y el mismo plazo determinado n .

$$R_n = \left(\frac{P_n}{P_0} \right)^{\frac{1}{n}} - 1$$

Ecuación 8. Tipo Spot.

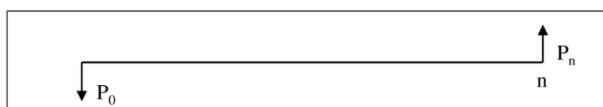


Figura 5. Ejemplificación de los Tipos de Interés al Contado.

Fuente: Mercados de Activos Financieros, JL Martín Marín.

- **Tipos de Interés a plazo implícito (Tipo *forward*)**, es la tasa de interés esperada para un periodo futuro, deriva de la combinación de dos tipos *spot* diferentes. Estima la rentabilidad del interés previsto en periodo específico en el futuro. Los Tipos de Interés a plazo implícito reflejan las expectativas de mercado sobre como evolucionarán los tipos de interés en un futuro. Para ello, se obtiene el tipo de interés al contado R_n en n años, otro R_m obtenido en m años, y se calcula en función de los mismos plazos temporales n y m , siendo $m > n$.

$$R_{n,m} = \left(\frac{1 + R_m}{1 + R_n} \right)^{\frac{1}{m-n}} - 1$$

Ecuación 9. Tipo Forward

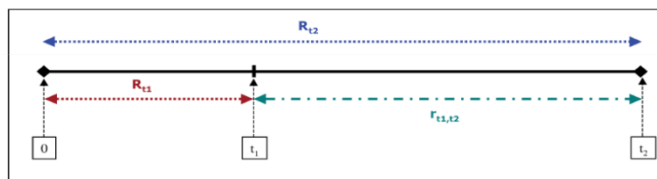


Figura 6. Ejemplificación de los Tipos de Interés forward.

Fuente: Mercados de Activos Financieros, JL Martín Marín.

- **Tasa Interna de Rendimiento (TIR) o Rentabilidad al vencimiento**, es uno de los instrumentos más utilizados en la estimación de activos de renta fija. Representa el rendimiento total de la rentabilidad de un interés si se mantiene hasta su vencimiento, considerando todos los flujos de efectivo que genera el título. El TIR asume que el bono se mantiene en la cartera hasta su amortización y, además, asume una reinversión constante con la misma tasa.

La equivalencia financiera que permite evaluar el rendimiento interno (TIR) se calcula mediante la Ecuación 10.

$$P_0 = \sum_{t=1}^n \frac{C}{(1 + TIR)^t} + \frac{P_n}{(1 + TIR)^n}$$

Ecuación 10. Tasa Interna de Rendimiento.

La P_0 denota el precio inicial del título, la C representa la reinversión periódica de intereses o cupón, la P_n el valor de reembolso o valor estimado de amortización y la n el periodo esperado de amortización.

Los instrumentos de inversión que se ven alterados por la actuación del mercado relacionado con los tipos de interés se pueden dividir en tres grandes grupos.

- **Bonos**, es un instrumento de deuda proporcionado por el gobierno o por corporaciones de medio o largo plazo. Un inversor presta dinero al emisor y recibe intereses periódicos, además del reembolso del capital una vez se cumple el vencimiento. Representa una relación inversa con los tipos de interés, cuando estos suben el valor de los bonos bajan, ya que ofrecen un mejor rendimiento.
- **Letras del Tesoro**, es un título de deuda emitido a corto plazo. No paga intereses, se trata de una compra por debajo de su valor nominal y, una vez haya vencido, se recibe el 100% del valor. Su rendimiento depende directamente de los tipos de interés.
- **Pagarés**, representa la promesa de pagar una cantidad fijada en un futuro, con intereses, los cuales varían según el tipo al que está sujeto.

2.1.4 Fondos Indexados.

Tal como se ha mencionado anteriormente, los fondos indexados son instrumentos de inversión colectiva que se enmarcan en la gestión pasiva, ya que no requiere supervisar sus inversiones tan de cerca como la forma activa.

Según se muestra en la Figura 7, los fondos de gestión pasiva han demostrado, en muchos casos, una rentabilidad neta superior a la de los fondos gestionados activamente.

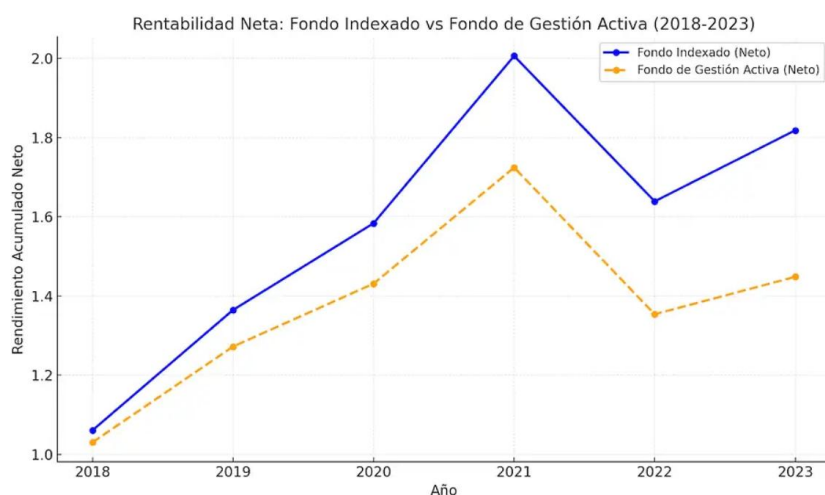


Figura 7. Rentabilidad Neta de los Fondos Indexados vs Rentabilidad Neta de los Fondos de gestión Activa

Fuente: *nosgustaintervenir.com*

Al igual que el resto de tipo de fondos, su principal objetivo es la réplica del comportamiento de índices, ya sea en la casilla de la renta variable o la renta fija. Además, están también gestionados por una sociedad gestora que es la encargada de gestión, la administración y la supervisión de la cartera.

Sin embargo, la diferencia fundamental con el resto radica en que su objetivo no es superar al mercado mediante selección de valores, sino igualar la renta del índice de referencia, reduciendo de esta forma la exposición a errores derivados de decisiones humanas. Como apunta (Feron, 2023) “Los fondos indexados te permiten obtener rendimientos potenciales con “barandillas” adicionales, ya que muchas personas no se sienten cómodas eligiendo o gestionando sus acciones individuales”.

Existen diferentes maneras para replicar un activo. Una de ellas es la réplica “parcial” del índice que se busque imitar, donde solo se seleccionan los valores con un gran peso y relevancia dentro de este, aunque siempre buscando una semejanza. Esto ocurre para evitar algunos inconvenientes que pueden llegar a surgir a la hora de invertir en este tipo de agentes económicos. Como, por

ejemplo, los grandes costes debidos a comisiones, surgiendo como consecuencia del gran volumen de activos que se gestiona. Otro problema aparece debido a la fluctuación de las empresas con menor capitalización en el índice, provocando problemas a nivel de gestión o de desembolso.

A continuación, se detalla un cuadro con todas las comisiones aplicables en los fondos indexados, lo que permite entender la magnitud de los costes que generan la administración de estos.

Tipos de Comisión	Cantidad	Aplicación
Por Suscripción	5%	En la Inversión
Por Reembolso	5%	En el Retorno
Por Gestión	2.25%	Anual
Por Éxito	18%	En el Beneficio
Por Gestión + Éxito	1.35% + 9%	Anual
Por Depósito	0.20%	Anual

Tabla 2. Comisiones Aplicables en los Fondos Indexados.

Fuente: Creación propia a partir de datos de Rankia.

Por último, es útil clasificar los fondos indexados según el tipo de activos financieros en los que invierten. Esta estructuración permite al inversor identificar la naturaleza de su exposición en función del perfil de riesgo y horizonte temporal:

- **Fondo Indexados de Renta Variable**, corporaciones colectivas que trabajan replicando índices de actividades variables, por ejemplo, índices bursátiles como *S&P500*, *IBEX35*. Están orientados a inversores que buscan crecimiento a largo plazo, aceptando una mayor volatilidad.
- **Fondos Indexados de Renta Fija**, engloba a todas las agrupaciones que tratan de imitar a índices de bonos gubernamentales o corporativos, como *Bloomberg Euro Government Bond Inde* o *iBoxx Euro Corporate Bond Index*. Suelen presentar una menor volatilidad, pero también una rentabilidad potencial más moderada.
- **Fondos Indexados Mixtos**, mezclan los dos tipos de rentabilidad. Ejemplos destacados son *Vanguard LifeStrategy 60% Equity* o *Amundi IS Diversified Moderate Allocation*.

2.1.5 ETF: Fondo Cotizado en Bolsa.

Los ETF (Exchange Traded Funds), conocidos en español como Fondos Cotizados en Bolsa, son instrumentos financieros relativamente recientes que han revolucionado el panorama de la inversión colectiva mediante la combinación de las ventajas de los fondos de inversión tradicionales con la flexibilidad operativa que proporcionan las acciones bursátiles. Tal y como defiende (Gastineau 2001) los Fondos de Cotización en Bolsa han perdurado a lo largo del tiempo, y continuarán ganando gran relevancia en el futuro, gracias a dos factores imprescindibles

e innovadores que los hace separarse de los Fondos de Inversión tradicionales: Su bajo coste y su capacidad de negociación intradía.

Un ETF es, en esencia, un producto que se forma mediante la composición de una cartera diversificada. Es decir, es un instrumento de inversión colectiva que está gestionado por entidades financieras que agrupa el capital de distintos inversores para invertirlo en una selección de activos financieros.

Aunque este enfoque puede parecer muy similar al del resto del mercado centrado en los fondos financieros colectivos, como los indexados, los ETF presentan unas características que los hacen destacar.

- **Flexibilidad de Cotización y Liquidez en Bolsa**, se negocia en el mercado financiero como una acción bursátil y es posible la compra y la venta de esta en cualquier periodo del día.
- **Fluctuación del precio**, varía durante toda la sesión según la oferta y demanda, similar al comportamiento de las acciones cotizadas.
- **Flexibilidad Operativa**, Los inversores pueden aplicar estrategias utilizando órdenes limitadas, *stop-loss*, o indicadores técnicos como medias móviles, lo que incrementa las posibilidades de gestión activa.
- **Transparencia**, los inversores son capaces de conocer en todo momento la composición subyacente de la cartera.
- **Naturaleza híbrida**, los principales ETF imitan índices de forma pasiva, pero existen otros que siguen estrategias activas o temáticas, lo que permite clasificarlos dentro de ambos estilos de gestión.

Estructurándolos de igual manera que se ha realizado anteriormente en los Fondos Indexados, es posible encontrar tres grupos diferenciados según la réplica de este.

- **ETF de Renta variable**, son aquellos que buscan replicar activos volátiles como es el caso de los índices bursátiles. En la Figura 8 aparecen los más reconocidos a nivel global, así como su crecimiento en la sesión de 2025.

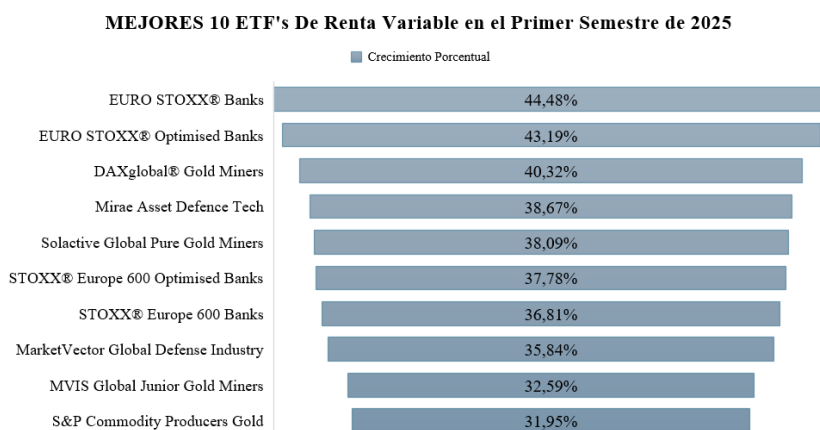


Figura 8. Mejores 10 Fondos Cotizados en Bolsa de Renta Variable en el primer semestre de 2025

Fuente: De creación propia a partir de datos de JustETF.

- **ETF de Renta fija**, son aquellos que imitan carteras de bonos, letras del tesoro u otros instrumentos que reflejan la actuación de los tipos de interés. En la Tabla 3 se recogen los más influyentes y su crecimiento porcentual durante este año 2025.

MEJORES 10 ETF's de Renta Fija en el Primer Semestre de 2025		
	Enfoque de Inversión	Crecimiento Porcentual
Bloomberg Emerging Markets Local Currency Liquid Government Bond (EUR Hedged)	Bonos Mercados emergentes Deuda pública	7,42%
FTSE Qualified Global Convertible (EUR Hedged)	Bonos Mundial Bonos convertibles	6,07%
FTSE Qualified Global Convertible (GBP Hedged)	Bonos Mundial Bonos convertibles	4,55%
iBoxx® MSCI ESG USD Asia ex-Japan High Yield Capped (EUR Hedged)	Bonos Asia Pacífico Deuda corporativa USD	4,04%
iBoxx® USD Liquid High Yield Capped (EUR Hedged)	Bonos Mundial Deuda corporativa USD	3,28%
Bloomberg US Enhanced Inflation (EUR Hedged)	Bonos Estados Unidos Ligados a la inflación USD	3,23%
JP Morgan USD Emerging Markets Diversified 3% capped 1-5 (EUR Hedged)	Bonos Mercados emergentes Agregado USD	3,15%
iBoxx® USD Contingent Convertible Liquid Developed Market AT1 (EUR Hedged)	Bonos Mundial Bonos convertibles USD	2,98%
Bloomberg Emerging Markets USD Sovereign & Agency 3% Country Capped (EUR Hedged)	Bonos Mercados emergentes Deuda pública USD	2,97%
JP Morgan ESG EMBI Global Diversified (EUR Hedged)	Bonos Mercados emergentes Deuda pública USD	2,93%

Tabla 3. Mejores 10 ETF's de Renta Fija en el Primer Semestre de 2025 según su crecimiento porcentual.

Fuente: Creación propia con datos de JustETF.

- **ETF de Renta mixta**, agrupan el capital invertido en carteras que combinan activos de renta fija y de renta variable. Estos fondos son ideales para quienes buscan un punto medio entre estabilidad y rentabilidad potencial.

2.1.6 Commodities.

En el término financiero, el concepto *commodity* hace referencia a materias primas tangibles que se negocian en los mercados, ya sea mediante venta, compra o intercambio. Desde el punto de vista bursátil, los *commodities* son bienes físicos sobre los que se pueden especular y obtener un rendimiento económico. Es posible adquirirlos con capital para transformarlos y aumentar así su valor añadido (Equipo de Expertos en Empresa de la Universidad Internacional de Valencia, 2023).

El mercado de materias primas presenta una gran distinción con el resto de los instrumentos financieros, como las acciones o índices. Esta diferencia fundamental radica en que no están expuestos a riesgos de obsolescencia, por lo que se reduce la probabilidad de perder la totalidad del valor en algún periodo futuro de tiempo. Así mismo, tal y como se definen los *commodities*, estos son elementos vitales y esenciales dentro del tejido económico global, ya que constituyen recursos básicos para el funcionamiento de cualquier civilización (Taulli, 2011).

Profundizando en sus ventajas dentro del mercado financiero, este tipo de agente bursátil es muy útil a la hora de realizar un análisis técnico en la predicción económica global, ya que sus bases

están relacionadas a la macroeconomía de un país, por lo que los *commodities* pueden ser vistos como indicadores adelantados de la situación financiera. Un claro ejemplo es la evolución de la inflación estadounidense a partir del valor del petróleo, ejemplificando como la variación del precio de esta materia prima puede alterar la economía de un país.

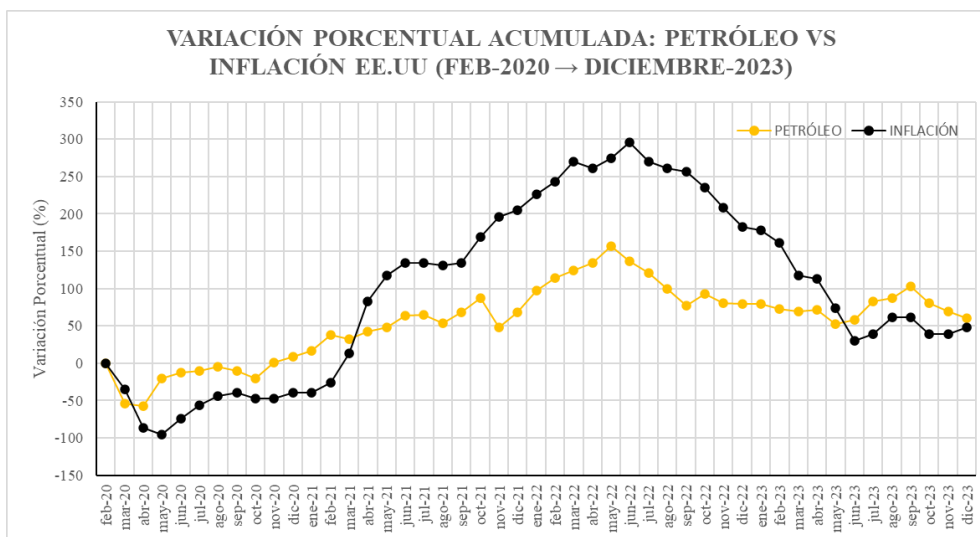


Figura 9. Variación Porcentual del Petróleo VS Inflación en EE.UU

Fuente: Elaboración Propia a partir de datos de Investing.com y Statista.

Una de las características que define a este tipo de inversión es la homogeneidad. Según (Fabozzi & Markowitz , 2006), “Las materias primas son bienes homogéneos, es decir, intercambiables entre sí sin importar su origen, siempre que cumplan con unos estándares de calidad definidos internacionalmente.” Esto implica que, por mucho que estas materias estén producidas o creadas por multitud de empresas o fabricantes, no existe una diferencia significativa entre los productos finales si cumplen con las especificaciones técnicas. Ejemplos como el oro, el gas natural o el petróleo Brent ilustran esta uniformidad. Este hecho produce multitud de virtudes como la facilidad de negociaciones entre mercados internacionales organizados, ya que elimina la necesidad de inspecciones físicas por parte de los agentes comerciales (Mochón, F., 2015).

Tradicionalmente, académicos como (Geman , 2005) y empíricos como (Taulli , 2011) estructuran las *commodities* en tres grupos principales, que se determinan según sus propiedades intrínsecas.

- **Agricultura**, engloba a todos aquellos recursos que derivan de organismos vivos y que son utilizados principalmente como alimento o como fibra natural. Los mercados más destacados dentro de este sector son *Chicago Mercantile Exchange (CME Group, EE. UU)*, *Intercontinental Exchange (ICE, Reino Unido)*, *Dalian Commodity Exchange (DCE, China)*, *Brasil Bolsa Balcão (B3, Brasil)* *Euronext (Europa)*.
- **Metales**, incluye a todos esos bienes que se extraen de la tierra y que se definen como elementos químicos sólidos a temperatura ordinaria, con un brillo característico y con grandes cualidades conductoras. Los comercios más reconocidos donde se negocia dicha materia son *Chicago Mercantile Exchange (CME Group, EE. UU)*, *London Metal*



Exchange (LME, Reino Unido), Tokio Commodity Exchange (TOCOM, Japón), Shanghai Futures Exchange (SHFE, China), MCX (Multi Commodity Exchange, India).

- **Energía y Combustible**, referido a todos aquellos materiales que son capaces de generar trabajo, ya sea de manera directa o mediante transformaciones. Los mercados más importantes y destacados son *New York Mercantile Exchange (NYMEX, EE. UU)*, *Tokio Commodity Exchange (TOCOM, Japón)*, *Intercontinental Exchange (ICE, Reino Unido)*.

2.2 La Predicción en el Mercado Financiero.

A lo largo del tiempo, hemos identificado anomalías estadísticas —como la reversión a la media y el momentum— que permiten, al menos temporalmente, anticipar movimientos futuros del mercado. Esto sugiere que, en lugar de ser completamente eficientes, los mercados financieros son sistemas adaptativos en los que las oportunidades de predicción emergen y desaparecen a medida que los inversores aprenden y modifican su comportamiento en respuesta a nuevas condiciones. (Lo & MacKinlay, 1999).

La anticipación en los mercados financieros ha representado durante años uno de los mayores desafíos dentro del campo económico. La predicción en bolsa es un terreno donde conviven tanto el potencial analítico como la gran incertidumbre inevitable de conocer el futuro. Sin embargo, existe una posibilidad de encontrar una relación entre el presente y los acontecimientos que ocurrirán en una serie temporal próxima, mediante el desarrollo de múltiples teorías, modelos estadísticos y algoritmos avanzados.

En esencia, predecir es formular hipótesis fundamentadas sobre eventos futuros, basándose en datos y patrones visibles en la actualidad. Esta tarea resulta extremadamente difícil en contextos como el bursátil, donde la volatilidad y la incertidumbre de múltiples variables imprevisibles dificultan los modelos de precisión (Timmermann & Granger, 2004). Además, uno de los pilares teóricos más reconocidos en este campo es la Hipótesis del Mercado Eficiente, planteado principalmente por (Fama, 1970). Esta postulación sugiere que los precios de los instrumentos financieros incorporan de una manera total la información relevante disponible en el mercado. Desde este punto de vista, ningún agente económico podría obtener una rentabilidad excesiva, ya que cualquier método analítico que produzca una ventaja competitiva con respecto al resto de inversores sería neutralizado por la reacción del mercado a nueva información (Malkiel, 2003). Bajo esta perspectiva, se plantea que el mercado actúa de forma eficiente y que cualquier serie localizada por inversores tenderá a desaparecer cuanto más explotada esté por un número significativo de individuos que participen con estos patrones.

Sin embargo, la teoría de la Hipótesis del Mercado eficiente ha sido debatida en multitud de ocasiones, y en ciertas ocasiones, refutada. Ciñéndose a la teoría defendida por (Atsalakis & Valavinis 2009), la clave no se basa en alcanzar una predicción con exactitud matemática, sino aportar ventajas competitivas al reducir la incertidumbre y permitir una mejor gestión del riesgo. De igual manera, economistas como (Fama & French, 1996), identifican otros factores para tener en cuenta los modelos predictivos, como el tamaño de la empresa, el efecto del calendario o la



relación valor-precio revelan que existen anomalías persistentes que no pueden arraigarse al azar ni se ven necesariamente eliminadas por la cantidad de participantes que operan en él.

(Kaufman, 2005), destaca la diferencia entre reconocer una tendencia ya en curso y anticipar un futuro. Aunque lo primero puede servir como ayuda ante una anticipación del mercado, esto puede lograrse con un análisis analítico, mientras tanto, lo segundo requiere de un entendimiento profundo de la dinámica bursátil y otros factores que intervienen en ella, como la macroeconomía, políticas monetarias o comportamiento de instituciones.

En la práctica, es imprescindible una combinación de ambas estrategias analíticas. Es por ello, que en las últimas décadas ha habido una gran relevancia del uso del análisis técnico apoyado por herramientas predictivas de base econométrica como es el caso de árboles de decisión, modelos estadísticos o redes neuronales artificiales, permitiendo manejar grandes volúmenes de datos y detectar patrones no lineales difíciles de identificar a simple vista.

En este trabajo se estudiará algunos modelos predictivos apoyados por algunos de los indicadores técnicos más usados en el terreno económico.

2.3 Series Temporales.

Una serie temporal, también denominada serie de tiempo, se define como un conjunto de observaciones cuantitativas registradas secuencialmente a lo largo del tiempo. Este análisis suele realizarse mediante intervalos periódicos regulares (días, meses, trimestres, etc.) y representan la evolución de una variable en el tiempo.

El término serie temporal se puede encontrar en multitud de campos no necesariamente ligados a la economía: Por ejemplo, en la meteorología, se estudian variantes como las precipitaciones o las temperaturas en periodos consecutivos; o en la demografía, mediante indicadores evolutivos de la población. (González Velasco & del Puerto García, 2009).

Todo análisis sobre la secuencia temporal se basa en la representación gráfica de esta, representando en el eje X el periodo temporal y en el eje Y los valores de una serie (K_t). Según (Box et al. 2008) el valor de la serie (K_t) se ha de descomponer en componentes como Tendencia (T), estacionalidad (S), ciclos (C) e Irregularidades (I), lo cual permitirá entender su estructura interna y aplicar técnicas para su pronóstico.

- **Ciclo (C_t)**, alude a oscilaciones repetitivas sin periodos fijos, irregulares en magnitud y duración, debido a algunas causas físicas relacionadas con las alteraciones económicas y sociales, dificultando una identificación precisa de esta.
- **Tendencia (T_t)**, representa la evolución general de la serie en un largo plazo de tiempo. Se manifiesta como la variación arraigada al valor medio de dicha serie a lo largo de un periodo, posibilitando un comportamiento alcista, decreciente o constante. La tendencia puede englobar en muchas ocasiones los efectos cíclicos de una extensa duración, motivo por el cual ambos elementos se pueden agrupar bajo la componente tendencia-ciclo (González Velasco & del Puerto García, 2009).

- **Estacionalidad (S_t)**, hace referencia a los comportamientos que se repiten con una cierta regularidad dentro de una misma estación. Estas fluctuaciones suelen tener una periodicidad fija y se deben a patrones relacionados con fenómenos externos basados en actividades humanas o naturales. Por ejemplo, en la demanda eléctrica, es posible ver una mayor tendencia consumista durante los meses de verano e invierno.
- **Irregularidad (I_t)**, comprenden a aquellos movimientos aleatorios e imprevistos que no pueden explicarse debido a la estacionalidad.

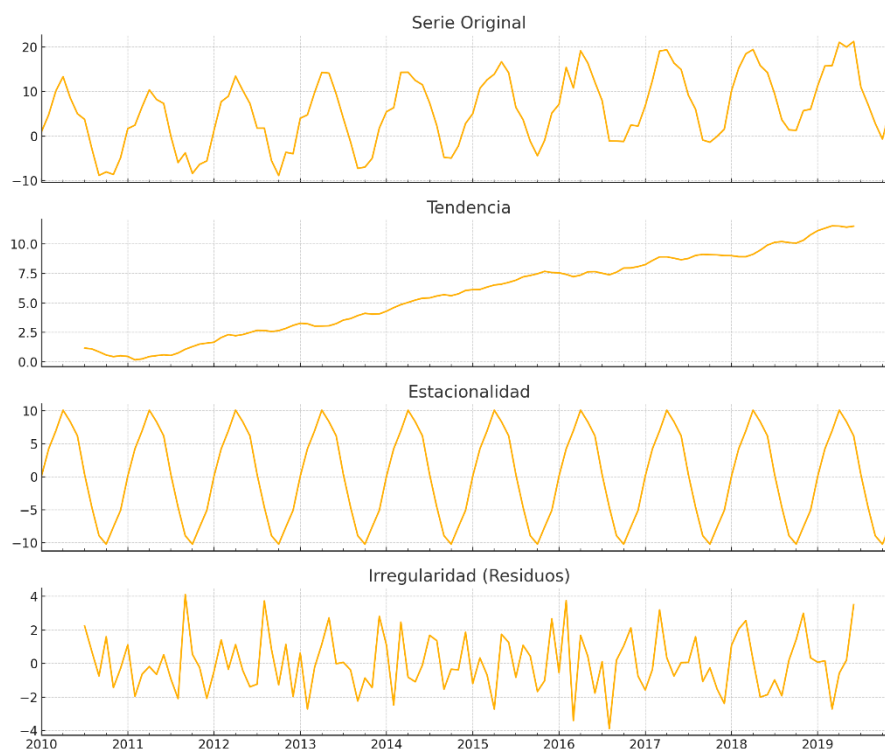


Figura 10. Representación de los Componentes de las Series Temporales

Fuente: De Elaboración propia.

Según la relación mutua entre las diferentes componentes fundamentales pueden definir el tipo de estructura subyacente de la serie. Según (García Díaz, 2016), estas relaciones permiten realizar un análisis de la serie temporal calificándola en un modelo de descomposición estructurado por tres esquemas: aditivo, multiplicativo y mixto.

- **Esquema aditivo**, la serie puede asumir su resultado de manera que cada componente actúa de manera independiente, por lo que la ecuación se expresa mediante la suma de los cuatro. Este modelo es adecuado cuando la fluctuación estacional y de tendencia mantiene una magnitud constante a lo largo del tiempo.

$$K_t = T_t + S_t + I_t$$

Ecuación 11. Serie temporal con Esquema Aditivo.

- **Esquema multiplicativo**, no cabe la posibilidad de tratar los componentes como casos independientes, ya que existe una interrelación entre ellos. Dicho modelo es útil cuando

la magnitud estacional depende directamente del nivel de la serie, aumentando o disminuyendo proporcionalmente según su tendencia.

$$K_t = T_t \times S_t \times I_t$$

Ecuación 12. Serie Temporal con Esquema Multiplicativo.

- **Esquema mixto**, a causa de la aleatoriedad producida por el componente irregular, lo más habitual es tratarla como un síntoma independiente, provocando un comportamiento mixto.

$$K_t = T_t \times S_t + I_t$$

Ecuación 13. Serie Temporal con Esquema Mixto.

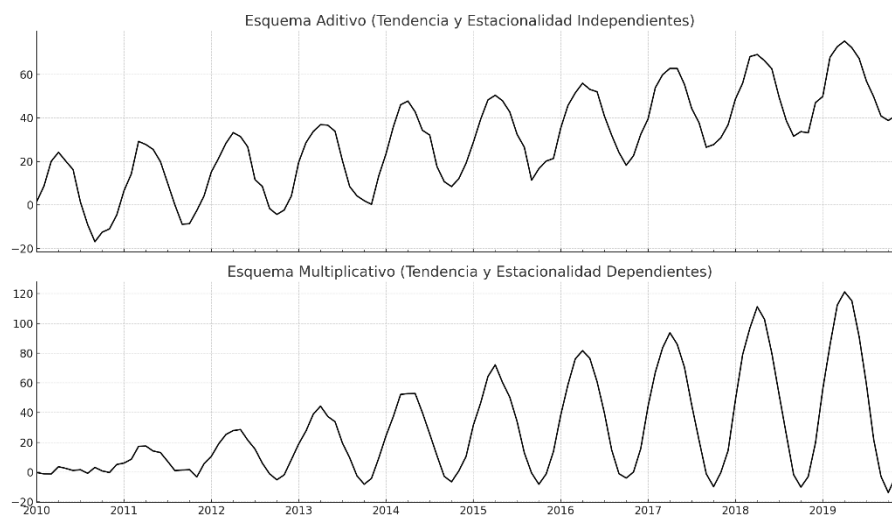


Figura 11. Diferencias Representativas entre Esquema Aditivo y Multiplicativo

Fuente: De elaboración propia.

Además, es esencial introducir el aspecto de dependencia temporal, es decir, es la correlación que existe entre los valores actuales y pasados de una variable observada en el tiempo. Esta dependencia puede aparecer a través de patrones como ciclos o tendencias, los cuales deben ser identificados correctamente para poder entender el comportamiento de una serie y estudiar las posibles predicciones útiles (Chatfield, 2004).

En este contexto, aparece el concepto de estacionariedad, lo que define una característica estadística de ciertas series temporales en las que esta dependencia se mantiene de manera constante a lo largo del tiempo. Una serie será estacionaria si sus valores se ven alterados alrededor de una media constante y si la varianza no experimenta cambios sistemáticos. En tales casos, la relación entre datos pasados y presentes se mantienen constantes y predecibles, lo que facilita una modelización estadística.

En definitiva, tal y como el análisis de las series temporales tiene como finalidad la estimación de valores futuros, lo que posiciona la predicción como piedra angular de este enfoque estadístico. Para ello, como se ha nombrado anteriormente, es necesario una correcta descripción y visualización de estas, permitiendo detectar tanto patrones como ciertas anomalías que pueden alterar el modelo. Paralelamente, el análisis busca construir un prototipo matemático que



represente adecuadamente los datos generados, permitiendo una interpretación fidedigna de su comportamiento para poder supervisar y mantener una estabilidad del proceso.

2.4 Modelos Predictivos de Base Econométrica.

En el campo de la predicción bursátil, los modelos cuantitativos de base econométrica establecen un papel fundamental, ya que permiten proyecciones futuras a partir de datos históricos. A diferencia de los modelos cualitativos, apoyados por la intuición, los modelos cuantitativos sugieren que las variables financieras contienen valores fundamentales para el pronóstico de eventos futuros. En este ámbito, el análisis bursátil se sostiene mediante la teoría de las series temporales estocásticas, las cuales imposibilitan un futuro con total certeza, pero sí apoyarse mediante distribuciones probabilísticas condicionadas por la recopilación de los precios, el volumen de negociación u otros indicadores. Como afirman (Campbell et al. 1997), la predicción en los mercados financieros está inevitablemente sujeta a incertidumbre debido a la naturaleza estocástica de los datos, pero los modelos econométricos ofrecen una estructura formal para cuantificar y gestionar dicha incertidumbre de manera efectiva. Así pues, el objetivo de los modelos predictivos basado en la econometría es maximizar el ajuste del pronóstico mediante un estudio adecuado de la dinámica estocástica de los datos.

2.4.1 Holt-Winters: Modelo basado en Suavizamiento Exponencial.

Una de las metodologías más extensamente utilizadas en la predicción de contextos financieros, económicos o industriales es el modelo basado en el suavizamiento exponencial. Su alta demanda es causada gracias a su eficiencia computacional, su facilidad de implementación y, sobre todo, por su capacidad de estimar pronósticos fieles en periodos de corto y medio plazo, pese a encontrarse en entornos de datos muy fluctuantes y limitados. Así pues, como resalta (Chatfield, 2000) una ventaja clave de este modelo es su aplicabilidad en situaciones con escasa información histórica, es por ello por lo que se convierte en una opción muy viable ante escenarios de incertidumbre o altamente evolutivos.

A diferencia de otros métodos técnicos más restringidos, el suavizamiento exponencial propone un estudio de datos con diferentes pesos, asignando un valor decreciente a las observaciones más antiguas con respecto a las más actuales. Este enfoque sugiere que el comportamiento del futuro se verá más afectado a causa de datos actuales, por lo que ajusta los pronósticos de forma dinámica a través de un sistema de retroalimentación constante. Tal como redacta (García Díaz, 2016), “el suavizado exponencial se basa en la actualización continua de la predicción anterior, corrigiéndola en función del error cometido, lo cual permite un proceso adaptativo de ajuste automático”.

Una de las mayores fortalezas del suavizamiento exponencial es su capacidad de adaptación y estabilidad en la predicción. Como define (Gardner, 1985), “el éxito del suavizamiento

exponencial radica en su capacidad de ofrecer una sorprendente precisión con un esfuerzo mínimo en la identificación del modelo”, convirtiéndolo en una herramienta espléndida en terrenos operativos donde se necesite rapidez y confiabilidad.

Dentro de esta metodología, cabe la posibilidad de distinguir tres estructuras principales, cada una adaptada a un contexto específico.

- **Suavizamiento Exponencial Simple**, requiere de un solo parámetro de suavizado “ α ”, encargado de regular el peso relativo entre la predicción anterior y la observación más reciente. Es adecuado para series sin tendencia ni estacionalidad y con una fluctuación baja:

$$P_{t+1} = P_t + \alpha(X_t - P_t)$$

Ecuación 14. Suavizamiento Exponencial Simple.

La predicción futura P_{t+1} se actualiza en función de dos componentes, la estimación previa P_t y el valor real recién observado X_t , los cuales se ajustan mediante un componente α que oscila entre 0 y 1, según el peso del dato observado. Ajustando la ecuación y añadiéndole un componente sumatorio sucesivo, el resultado dado es el de la ecuación 15:

$$P_{t+1} = \alpha X_t + \alpha(1 - \alpha)X_{t-1} + \alpha(1 - \alpha)^2 X_{t-2} + \dots + \alpha(1 - \alpha)^{t-1} X_{t-1} + (1 - \alpha)^t P_t$$

Ecuación 15. Suavizamiento Exponencial con término Sumatorio.

- **Método de Holt o de Doble Suavizamiento**, surge como una evolución del modelo de suavizamiento exponencial simple, con el objetivo de superar sus limitaciones en el análisis de series temporales más complejas, las cuales no pueden ser captadas de una manera precisa mediante una formulación básica. Introduce un segundo parámetro “ β ” que permite ajustar la tendencia lineal de la serie, siendo útil ante situaciones donde existe una evolución sucesiva de los datos. En el enfoque matemático, la predicción está compuesta por una suma entre dos componentes: N_t , donde se representa una estimación del nivel de la serie un momento t y S_t , que estima la pendiente de la serie en ese mismo instante:

$$N_t = \alpha X_t + (1 - \alpha)(N_{t-1} + s_{t-1})$$

$$s_t = \beta(N_t - N_{t-1}) + (1 - \beta)s_{t-1}$$

$$P_{t+m} = N_t + s_t \times m$$

Ecuación 16. Método de Doble Suavizamiento Exponencial.

- **Modelo de Holt-Winters o de Triple Suavizamiento**, parte de la necesidad de cuantificar la estacionalidad de una serie. Cuando una serie temporal no presenta indicios de tendencia ni estacionalidad es posible analizarla mediante métodos simples de suavizamiento exponencial. En situaciones donde la serie evidencia un patrón evolutivo creciente o decreciente de manera sistemática, es necesario recurrir a sistemas más complejos que estudien este dinamismo como el modelo propuesto por Charles Holt.

Sin embargo, siguiendo la explicación de los investigadores (Hyndman, R. J., & Athanasopoulos, G. 2021), cuando además de una tendencia se identifica una estacionalidad, es decir, un comportamiento fijo que se repite bajo una frecuencia temporal fija es necesario atender a esta complejidad mediante la aplicación de una teoría más efectiva. Por eso, Peter Winters amplió el modelo de Holt añadiendo una tercera variable destinada a plasmar un patrón estacional, originando así el modelo de Holt-Winters.

Este modelo se fundamenta en tres ecuaciones que actualizan reiterativamente el nivel, la tendencia y la estacionalidad de una serie, cada una ajustada y controlada por un parámetro de suavizado específico: “ α ”, “ β ” y “ γ ”. Esta combinación matemática permite un estudio predictivo ante situaciones de series temporales altamente complejas y recurrentes, como es el caso de las relacionadas con el término financiero.

Además, dicho sistema permite adaptarse a situaciones más simples. En el caso que alguna de las variables concretadas anteriormente no se encuentre presentes en la serie que se va a analizar solo haría falta anular el parámetro de suavizado a la que esta arraigada dicha variable.

Existen dos métodos diferentes Holt-Winters según la dependencia de los componentes:

- **Estacionalidad aditiva**, para componentes totalmente dependientes, no suele ser tan utilizada. Las ecuaciones básicas son las siguientes:

$$N_t = \alpha(X_t - C_{t-s}) + (1 - \alpha)(N_{t-1} + s_{t-1})$$

$$s_t = \beta(N_t - N_{t-1}) + (1 - \beta)s_{t-1}$$

$$C_t = \gamma(X_t - N_t) + (1 - \gamma)C_{t-s}$$

$$P_{t+m} = N_t + s_t \times m + C_{t-s+m}$$

Ecuación 17. Método De Holt-Winters con Estacionalidad Aditiva.

- **Estacionalidad multiplicativa**, para componentes interrelacionados entre ellos mismos. Las ecuaciones que engloban los componentes fundamentales son las siguientes:

$$N_t = \alpha \frac{X_t}{C_{t-s}} + (1 - \alpha)(N_{t-1} + s_{t-1})$$

$$s_t = \beta(N_t - N_{t-1}) + (1 - \beta)s_{t-1}$$

$$C_t = \gamma \frac{X_t}{N_t} + (1 - \gamma)C_{t-s}$$

$$P_{t+m} = (N_t + s_t \times m) \times C_{t-s+m}$$

Ecuación 18. Método Holt-Winters con Estacionalidad Multiplicativa.

Las ecuaciones fundamentadas en este modelo no difieren mucho a las del método Holt. Sin embargo, aparece una nueva componente s que se refiere al número de periodos que abarca la estacionalidad.

Con el objetivo de representar de forma sencilla los conceptos asociados al modelo de Holt-Winters, se ha recreado un ejemplo arbitrario utilizando el lenguaje de programación R. Este permite visualizar, tanto a nivel gráfico como cuantitativo, cómo se comporta el pronóstico de una serie temporal que presenta una estacionalidad aditiva. La elección de esta configuración responde a su simplicidad estructural, lo que facilita el entendimiento del funcionamiento metodológico y de sus componentes principales: nivel, tendencia y estacionalidad.



Figura 12. Representación Simulada de Modelo Holt-Winters

Fuente: De elaboración Propia.

A partir de la gráfica histórica cabe la posibilidad de obtener unos datos seleccionados mediante un periodo mensual. Más tarde se han ajustado estas observaciones para solucionar algunas anomalías observables y mediante el sistema Holt-Winters poder realizar un pronóstico de las siguientes 12 mensualidades.

Fecha	Observado	Ajustado	Pronóstico
enero-20	0.933	0.648	
febrero-20	5.223	3.996	
marzo-20	10.956	8.328	
abril-20	14.546	12.713	
mayo-20	10.192	10.552	
junio-20	7.032	7.926	
julio-20	6.158	3.423	
agosto-20	0.035	-0.600	
septiembre-20	-5.599	-4.462	
octubre-20	-4.415	-5.784	
noviembre-20	-4.587	-3.231	
diciembre-20	-0.431	-0.081	
...
enero-24			23.458
febrero-24			26.806
marzo-24			31.138
abril-24			35.523
mayo-24			33.362
junio-24			30.736
julio-24			26.223
agosto-24			22.209
septiembre-24			18.348
octubre-24			17.026
noviembre-24			19.579
diciembre-24			22.729

Tabla 4. Pronóstico Simulado de Holt-Winters.

Fuente: De Elaboración Propia.

Fecha	Nivel	Tendencia	Estacionalidad
enero-20	0.437	0.475	0.211
febrero-20	0.912	0.475	3.084
marzo-20	1.387	0.475	6.941
abril-20	1.863	0.475	10.850
mayo-20	2.338	0.475	8.214
junio-20	2.813	0.475	5.113
julio-20	3.288	0.475	0.135
agosto-20	3.763	0.475	-4.364
septiembre-20	4.239	0.475	-8.700
octubre-20	4.714	0.475	-10.498
noviembre-20	5.189	0.475	-8.420
diciembre-20	5.664	0.475	-5.745
...
enero-23	17.544	0.475	0.211
febrero-23	18.019	0.475	3.084
marzo-23	18.495	0.475	6.941
abril-23	18.970	0.475	10.850
mayo-23	19.445	0.475	8.214
junio-23	19.920	0.475	5.113
julio-23	20.395	0.475	0.135
agosto-23	20.871	0.475	-4.364
septiembre-23	21.346	0.475	-8.700
octubre-23	21.821	0.475	-10.498
noviembre-23	22.296	0.475	-8.420
diciembre-23	22.771	0.475	-5.745

Tabla 5. Nivel, Tendencia y Estacionalidad Simulada.

Fuente: De Elaboración Propia.

Los datos representados en la Tabla 4 surgen directamente de los cálculos realizados a partir de las ecuaciones previamente expuestas, en los que se han asignado valores arbitrarios a los parámetros de suavizado α , β y γ con el fin de ejemplificar el funcionamiento del modelo.

Ejemplificando una de las predicciones dentro de las 12 calculadas, aplicando la ecuación del modelo de estacionalidad aditiva. Se extraen los siguientes datos de la Tabla 5 para estimar el valor predictivo.

Nivel obtenido en el último periodo observado (Diciembre de 2023) = 22.771.

Tendencia estimada mediante ecuación sinusoidal = 0.475.

Estacionalidad referida a la mensualidad que se desea estimar (Enero) = 0.211.

$$P_{\text{Enero } 2024} = 22.771 + 0.475 + 0.211 \approx \mathbf{23.458}$$

Este proceso se realizará sucesivamente para conseguir el conjunto completo de pronósticos correspondientes al horizonte temporal deseado.

2.4.2 Modelos ARIMA y Extensiones.

El modelo ARIMA (Autoregressive Integrated Moving Average) surge como solución ante la problemática de analizar y predecir series estacionarias o no estacionarias sin un claro nivel de tendencia ni estacionalidad.

Los modelos ARIMA han sido ampliamente utilizados en el ámbito económico debido a su flexibilidad para modelar distintos comportamientos de series temporales y a su capacidad para generar pronósticos fiables a corto plazo de indicadores macroeconómicos como los tipos de cambio.
Pindyck, R. S., & Rubinfeld, D. L. (1998). Modelos econométricos y pronósticos económicos.

El mercado bursátil es, por naturaleza, una disciplina dinámica que investiga situaciones cuya evolución con el tiempo es esencial para su comprensión. En este punto es cuando aparece la metodología ARIMA, popularmente utilizada en este sector debido a la eficiencia que existe para capturar el dinamismo interno que aparece en las series económicas. Además, a causa de su estructura modular y compatibilidad con diferentes métodos técnicos, hacen que el modelo sea muy ajustable a las realidades que ofrece la predicción financiera. Esta ductilidad ha permitido que el método Autorregresivo Integrado y de Medias Móviles se haya consolidado como un referente entre todo el arsenal de herramientas basadas en la econometría.

Los analistas (Box & Jenkins 1970), fueron los partícipes de la teorización de este sistema. En particular, la metodología Box-Jenkins es un procedimiento de análisis estadístico que pretende un ajuste del modelo mediante ciclos iterativos. Cada iteración está compuesta por varias etapas consecutivas que permiten afinar o precisar la calidad del modelo.

En primer lugar, se define una clase general de modelos posibles que son considerados en el análisis, lo cual supone un entendimiento previo de la estructura de los datos y de los patrones probables presentes.

Posteriormente, se procede a la identificación del modelo más adecuado, utilizando herramientas y criterios estadísticos que permiten determinar los órdenes de los componentes autorregresivos (p), de diferenciación (d) y de media móvil (q).

Una vez realizados los pasos anteriores, se avanza a la fase de estimación de los parámetros, donde se deciden los coeficientes que mejor representan la dinámica observada según los datos históricos.

Finalmente, se realiza una verificación del modelo ajustado, permitiendo una evaluación para comprender si el modelo tiene en cuenta los residuos aleatorios y no tiene en cuenta tan solo una información sistemática, lo cual es esencial para garantizar las tareas predictivas.

Basándose en la metodología impartida originalmente por Box y Jenkins, y atendiendo a nuevas aportaciones como la de (Hyndman & Athanasopoulos, 2021), el término ARIMA está compuesto por tres componentes fundamentales: La parte autorregresiva (AR), la parte integrada (I), atendiendo a la diferenciación de los datos para lograr la estacionalidad; y la parte de media móvil (MA).

Con el fin de entender los diferentes procesos que conforman el modelo ARIMA, es necesario interpretar una serie de definiciones fundamentales que establecen los pilares fundamentales del funcionamiento sistemático de dicho método.

- **Estacionariedad**, hace referencia a la inferencia estadística sobre el comportamiento de un proceso estocástico. La idea fundamental de la estacionariedad se basa en demostrar que las distribuciones probabilísticas no varían con el tiempo. Se afirma que un proceso es estacionario si su propiedad probabilística es invariante ante traslaciones temporales sin importar que puntos específicos se elijan.
- **Ruido Blanco**, también denominado ruido aleatorio, es una secuencia de variables puramente aleatorias. Se utiliza principalmente en el análisis de series temporales para plasmar novedades que surgen en una serie, modelando la parte imprevisible de las variaciones aleatorias. Se define por las siguientes propiedades:
 - Media cero: $E(\varepsilon_t) = 0$
 - Varianza constante: $Var(\varepsilon_t) = \sigma$
 - No autocorrelación: $Cov(\varepsilon_t, \varepsilon_{t-k}) = 0$ para todo $k \neq 0$

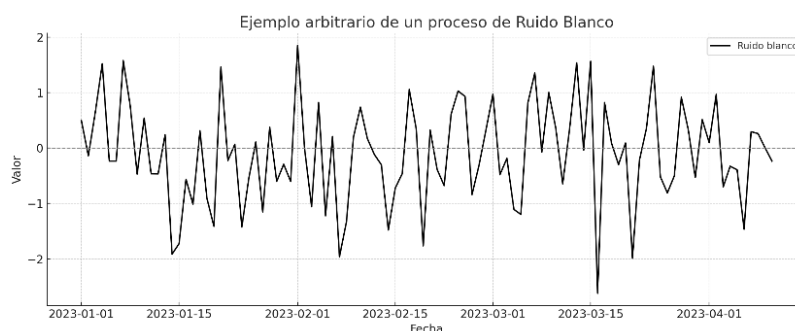


Figura 13. Representación del Ruido Blanco.
Fuente: De elaboración propia.

- **Autocorrelación o correlograma simple**, mide la relación lineal entre valores de la serie separados en distintos rezagos o *lags*. Detecta dependencia a corto y largo plazo con el objetivo de identificar si la serie es estacionaria. Si decrece lentamente suele indicar una tendencia no estacionaria. Por otra parte, si corta rápidamente identifica un proceso de Media Móvil. En este interfiere tanto la tendencia como el grado de la Media Móvil.
- **Autocorrelación parcial o correlograma parcial**, encargada de medir la correlación entre dos puntos de la serie, eliminando el efecto de los valores intermediarios. Es muy útil a la hora de identificar n orden en los procesos ARIMA. Este depende del grado del modelo autorregresivo.

Para comprender el proceso del modelo ARIMA, es esencial examinar los métodos por los que está compuesto, es decir, aquellos procesos autorregresivos, integrados y de media móvil.

- **Procesos Autorregresivos**, son aquellos que implican realizar una regresión de la variable sobre sus propios valores rezagados. Un proceso autorregresivo de orden p satisface a la ecuación general:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + e_t$$

Ecuación 19. Proceso Autorregresivo del Modelo ARIMA.

Donde X_t representa el valor actual, ϕ_p se refiere a los coeficientes que permiten ajustar el modelo autorregresivo y e_t plasma ese ruido blanco definido anteriormente como el valor aleatorio de la serie.

Los coeficientes que controlan el modelo se determinan durante un proceso de ajuste a través de métodos de estimación como los mínimos cuadrados o la máxima verosimilitud.

Además de los fundamentos estructurales, el principal objetivo es identificar si el modelo AR es estacionario o no. Para ello será requisito que las raíces que conforman el polinomio autorregresivo de orden p sea mayor a 1 en su valor absoluto, siendo B el operador de retardo.

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \quad \phi(B)X_t = X_{t-k}$$

Ecuación 20. Operador de Retardo Autorregresivo.

Una manera de comprobar su valor absoluto es mediante ciertas aproximaciones como las siguientes:

$$\phi_1 + \phi_2 + \dots + \phi_p < 1, \quad \phi_p - \dots - \phi_2 - \phi_1 < 1 \quad y \quad |\phi_p| < 1$$

Ecuación 21. Comprobación del Modelo Autorregresivo.

Identificar que una serie temporal es estacionaria implica, entre otras cosas, que su valor medio es constante.

$$E[X_t] = \mu \text{ para todo } t,$$

Una vez asumido una serie estacionaria y con una media constante, se pueden usar las llamadas ecuaciones de Yule-Walker para obtener la varianza del proceso y, de esta manera, calcular las antes definidas autocorrelaciones. Dichas ecuaciones se escriben en forma de relaciones sucesivas entre las autocorrelaciones y los coeficientes ϕ_p , obteniendo los valores numéricos de la autocorrelación para cualquier número de rezagos.

$$\gamma_0 = \phi_1 \gamma_1 + \phi_2 \gamma_2 + \dots + \phi_p \gamma_p + \sigma_e^2, \quad \rho_p = \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + \dots + \phi_p,$$

$$\text{teniendo en cuenta que } \rho_p = \frac{\gamma_p}{\gamma_0}, \quad \text{entonces: } \gamma_0 = \frac{\sigma_e^2}{1 - \phi_1 \gamma_1 - \phi_2 \gamma_2 - \dots - \phi_p \gamma_p}$$

Ecuación 22. Obtención de la Varianza del proceso mediante Yule-Walker.

Donde γ_0 hace referencia a la varianza de la serie, ρ_p representa el coeficiente de autocorrelación en el rezago p y es una medida estandarizada que indica cuanta relación existe entre el valor X_t y X_{t-p} . Por último, la γ_p define la autocovarianza en el rezago p .

En la práctica, los análisis gráficos presentados anteriormente como correlogramas simples y parciales también son fundamentales para el estudio de la dependencia de la serie y para identificar el orden del modelo. Dando como resultados una decadencia exponencial en la Figura 14. y unas ondas amortiguadas que van perdiendo amplitud, como se puede ver en la En la Imagen 15.



Figura 14. ACF simulado de una Serie NO estacionaria.

Fuente: de elaboración propia.

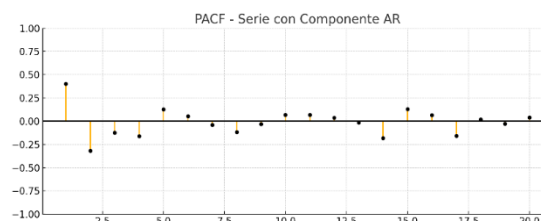


Figura 15. PACF con Componente AR

Fuente: de elaboración propia.

- **Proceso de Media Móvil**, es un método basado en el suavizamiento de las series temporales mediante la recopilación de datos para obtener una predicción cuando no se presenta ni tendencia ni estacionalidad. A diferencia del modelo autorregresivo, este es un tipo de método de series temporales que no se basa en los valores históricos de la serie, sino en los errores aleatorios o “ruidos” que ocurrieron en periodos anteriores (Box et al. 2008).

$$X_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q}$$

Ecuación 23. Proceso Media Móvil del Modelo ARIMA.

X_t hace referencia al valor actual de la serie, e_t representa lo denominado anteriormente como “ruido blanco” o error aleatorio en el periodo t y θ_q expresa los coeficientes de ajuste del modelo MA.

Para el estudio sobre la viabilidad analítica del modelo es necesario usar el polinomio de media móvil:

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \qquad X_t = \theta(B)e_t$$

Ecuación 24. Operador de Retardo de la Media Móvil del Modelo ARIMA.

Los modelos MA son estacionarios por normativa, lo que implica que su valor medio sea constante a lo largo del tiempo. Para comprobar si el proceso es válido, es necesario un estudio de su invertibilidad, lo cual permite reescribir el modelo MA de orden q en uno AR de orden infinito. Para ello, se examina si las raíces del polinomio se encuentran fuera del rango del círculo de la unidad, una herramienta geométrica que recoge todos los números complejos x con valor absoluto menor o igual a la unidad.

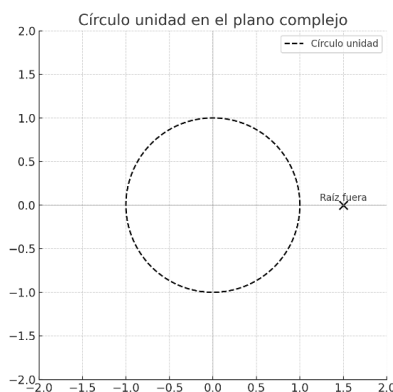


Figura 16. Círculo de la Unidad en el Plano Complejo de un modelo ARIMA

Fuente: De elaboración propia

Una vez estudiada la viabilidad del proceso, al igual que en el modelo autorregresivo, se obtiene la función de autocorrelación simple y parcial mediante la ecuación 25. Obteniendo los valores de los rezagos q para la interpretación de la dependencia y el orden del método.

$$\rho_k = \begin{cases} \frac{-\theta_k + \theta_1\theta_{k+1} + \theta_2\theta_{k+2} + \dots + \theta_{q-k}\theta_q}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2}, & \text{para } k = 1, 2, \dots, q \\ 0, & \text{para } k > q \end{cases}$$

$$\alpha_k = -\frac{\theta^k(1 - \theta^2)}{1 - \theta^{2(k+1)}}$$

Ecuación 25. Función de Autocorrelación de la Media Móvil.

Dando como resultado dos representaciones gráficas. La Figura 17 muestra cómo se corta directamente una vez el número de rezagos supera el orden q . Por otra parte, la Imagen 18, a diferencia del modelo autorregresivo, muestra un patrón de sinusoidales amortiguadas decrecientes hasta un valor infinito.

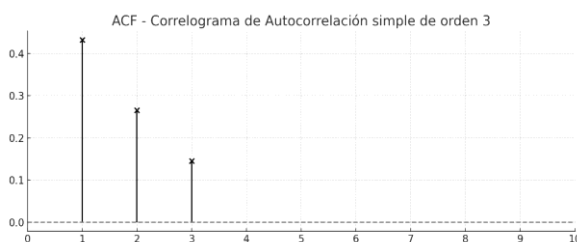


Figura 17. ACF de Autocorrelación Simple

Fuente: De elaboración propia

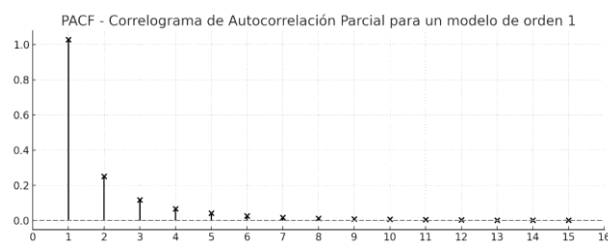


Figura 18. PACF de Autocorrelación Simple.

Fuente: De Elaboración propia.

- **Proceso Mixto (ARMA)**, como su propio nombre indica se trata de un modelo que combina en su estructura componentes autorregresivos (AR) y de media móvil (MA), incluyendo así atributos de ambos tipos de proceso. El término ARMA (p, q) indica que el modelo integra p términos autorregresivos y q términos de media móvil. Su relación aparece en sus polinomios característicos y en la ecuación general ARIMA dando como resultado las siguientes ecuaciones generales:

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q},$$

o escrito con el operador de retardo B ,

$$\phi(B) = \theta(B)$$

Ecuación 26. Equivalencias de Operadores de Retardo para el proceso ARMA.

En cuanto al análisis gráfico, el correlograma de autocorrección simple presentará un patrón de decaimiento poco exagerado, resultante de la combinación entre oscilaciones amortiguadas y exponenciales. Por otra parte, el correlograma de autocorrelación parcial mostrará un comportamiento parecido al de autocorrelación simple, con el detalle de que puede presentar ciertas desviaciones en los comportamientos de los $p-q+1$ rezagos, siempre que p sea mayor que $q+1$.

- **Proceso Integrativo (I)**, permite transformar una serie temporal no estacionaria en estacionaria dentro del modelo ARIMA, es decir, aquella cuyos valores tienen un valor medio y una varianza constante en el tiempo. Esto ocurre gracias a la llamada diferenciación, que permite eliminar tendencias o variaciones reiterativas.

El operador se representa mediante un operador de diferencia (∇). La forma general del operador de diferencia d se expresa de la siguiente forma:

$$\nabla^d X_t = (1 - x)^d X_t$$

Ecuación 27. Proceso Integrativo del modelo ARIMA.

La letra d indica el número de veces que se ha aplicado dicha diferenciación. Gracias a este proceso integrador, los modelos ARIMA son capaces de pronosticar series que no presentan una tendencia estacionaria o determinista.

Para entender el funcionamiento del método ARIMA, a continuación, se presenta un ejemplo muy simplificado de una serie temporal totalmente arbitraria que, tras una breve inspección, es fácil observar que la serie tiene tanto valor medio como varianza constante, por lo que no es necesaria una diferenciación. El pronóstico del gráfico N. se ha obtenido mediante la fórmula general del modelo ARMA, con orden (1,1). Escogiendo componentes autorregresivos, de media móvil y de ruido arbitrarios.

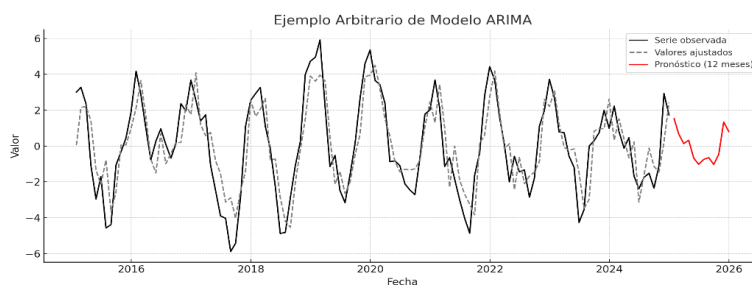


Figura 19. Representación Simulada del modelo ARIMA

Fuente: De elaboración propia.

Además del ya definido modelo ARIMA, existen diferentes extensiones que se ajustan idóneamente a situaciones específicas que fuerzan algunas series temporales.

- **ARIMAX**, es una ampliación del modelo ARIMA, conocido como modelo autorregresivo integrado de medias móviles con variables exógenas. Esta extensión permite incorporar múltiples variables adicionales, lo cual permite una mayor ventaja a la hora de tener en cuenta diferentes aspectos no correlacionados en una serie temporal. Cada variante añadida tendrá un componente propio:

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q} + \beta_1 Z_{1,t} + \dots + \beta_v Z_{v,t}$$

O mediante operadores de rezago:

$$\phi(B) = \beta_x \theta(B)$$

Ecuación 28. Operadores de Retardo para la extensión ARIMA.

Siendo β_t el coeficiente de ajuste correspondiente a cada variable exógena $Z_{v,t}$, donde el término t corresponde al periodo actual y v la cantidad de variables exógenas aplicadas en esa determinada serie temporal.

- **SARIMA**, también conocido como modelo de media móvil integrado autorregresivo estacional, es una extensión del modelo original ARIMA capaz de capturar de forma explícita patrones estacionales a partir de los datos obtenidos. A diferencia del modelo base, SARIMA incorpora estructuras cíclicas reiterativas, siendo fundamental para series con estacionalidad. Es decir, es capaz de representar una estructura temporal dinámica de una serie. Su formulación está compuesta por los polinomios característicos originarios del modelo ARIMA junto con tres adicionales referidos a los componentes de los datos estacionales (P, D, Q):

$$\phi(B)\Phi_p(B^S)(1-B)^d(1-B^S)^D X_t = \theta_q(B)\Theta_Q(B^S)e_t$$

Ecuación 29. Proceso SARIMA.

Siendo $\Phi_p(B^S)$ y $\Theta_Q(B^S)$ los polinomios de grado P y Q de la parte autorregresiva y de media móvil respectivamente de la parte estacional y S es la duración de la temporada:

$$\Phi_p(B^S) = 1 - \Phi_s(B^S) - \Phi_{s2}(B^{2S}) - \dots - \Phi_{sp}(B^{pS})$$

$$\Theta_Q(B^S) = 1 - \Theta_s(B^S) - \Theta_{s2}(B^{2S}) - \dots - \Theta_{sQ}(B^{QS})$$

Ecuación 30. Polinomios de Grado P y Q .

2.4.3 Random Forest: Modelo de Ensamble basado en Árboles de Decisión.

Los árboles de decisión constituyen una de las herramientas más visuales y eficaces tanto en el terreno de la toma de decisiones como de los pronósticos probabilísticos. Se entienden como diagramas de ramas estructurados que permiten entender un problema decisivo mediante bifurcaciones, conocidas como “nodos” y trayectorias, denominadas ramas y representadas mediante flechas, las cuales unen los nodos. Cada nodo refleja una elección correspondiente a eventos aleatorios cuya probabilidad está estimada anteriormente (Moreno Osorio, 2011). Los árboles de decisión siguen una lógica de izquierda a derecha o de arriba abajo, finalizando con los resultados de las distintas posibles combinaciones.



Figura 20. División de los Nodos de un Random Forest.
Fuente: Canós Darós, L. (2024). Árboles de decisión.

En el ámbito del estudio de datos, este método se ha consolidado como uno de los modelos más intuitivos y simples para las tareas predictivas. Como expresa el economista (Wang, 2020) este sistema funciona mediante la división iterativa de un conjunto de datos en subconjuntos más homogéneos, guiados por las estimaciones de los atributos analizados. Cada valor posible del atributo estudiado genera una divergencia y una subdivisión del conjunto de datos estudiados. El algoritmo del modelo emula este proceso reiterativamente hasta que se logra alcanzar una condición de terminación. Esta puede estar definida por diferentes criterios como la profundidad máxima del árbol, la cantidad mínima de nodos o de medidas de impureza aceptadas.

La estructuración de los árboles utiliza criterios cuantitativos que evalúan la calidad de las bifurcaciones. Entre las más comunes es posible encontrar:

- **Entropía**, sirve como medida del desorden.
- **Ganancia de información**, cuantifica la reducción de incertidumbre tras una división.
- **GINI**, cuantifica una métrica de impureza del nodo.

Estos criterios permiten seleccionar los atributos más explicativos en cada proceso.

No obstante, una de las desventajas más común en los árboles de decisión individuales es su propensión al sobreajuste, es decir, se ciñe demasiado a los datos entrenados, perdiendo capacidad de generalización ante nuevos datos. Para solucionar este problema, (Leo Breiman, 2001) propuso un método innovador llamado “Random Forest”, que actúa como ensamblador de multitud de árboles de decisión utilizando una serie de datos y atributos aleatorios, mejorando

significativamente la estabilidad y la precisión del modelo. Su aplicación se ha extendido a dominios tan diversos como la medicina, las ciencias biológicas, la ingeniería y las finanzas, donde se exige un análisis riguroso en entornos inciertos (Hasan & Horvat, 2024).

En el sector financiero, Random Forest es empleado tanto en la predicción de valores bursátiles como la evaluación de riesgos. Los atributos más utilizados en este sector son los siguientes:

- **Moving Average**, técnica muy utilizada en el análisis financiero para suavizar las fluctuaciones de precios en el mercado. Consiste en promediar los precios de una acción durante un número concreto de días. Generalmente se usa un periodo de 10 días (MA10) y 50 días (MA50):

$$MA_n = \frac{M_1 + M_2 + \dots + M_n}{n}$$

Ecuación 31. Media Móvil Simple.

M_n hace referencia al precio activo en el día que corresponde y n es el periodo que se quiere estudiar.

- **Retorno**, es un indicador directo que permite estimar la ganancia o la pérdida en un determinado periodo. Mide el cambio porcentual del cierre P_t de un recurso financiero con respecto al día anterior P_{t-1} :

$$Ret = \frac{P_t - P_{t-1}}{P_{t-1}}$$

Ecuación 32. Retorno del Valor.

- **Volatilidad**, media técnica que indica cuánto varía el precio de un activo financiero en un periodo determinado. La forma más común de estimar la volatilidad es mediante la desviación estándar del retorno:

$$Vol_n = \sqrt{\frac{1}{n} \sum_{i=1}^n (R_i - \bar{R})^2}$$

Ecuación 33. Volatilidad.

Siendo la n el número de días a considerar, R_i el retorno del día que se está estudiando y \bar{R} es el promedio de los retornos diarios.

El criterio más utilizado para establecer la calidad de una división en este sector es el llamado GINI. Como se ha explicado anteriormente es un método que evalúa la impureza midiendo la probabilidad de que un atributo seleccionado sea clasificado incorrectamente en referencia a su distribución. Es decir, cuanto más homogéneo sea un resultado previsto, más puro será y, por lo tanto, menor será su índice GINI. La formulación general está formada por la proporción de elementos p de la clase i de cada resultado:

$$GINI = 1 - \sum_{i=1}^k p_i^2$$

Ecuación 34. GINI del modelo Random Forest.

A la hora de modelizar un sistema de Random Forest, se analiza los atributos seleccionados previamente y los datos disponibles en el conjunto de muestras. Cada nodo representa una condición basada en los atributos elegidos, que dividen las muestras o *samples* de una forma reiterativa según se cumpla o no la condición estimada, clasificando de esta manera si sube o baja. El valor o *value* indica la distribución de las clases entre las muestras que se han analizado en dicho nodo. Aunque, como se puede observar en la Figura 21. la suma de valores de un nodo no coincide con la de muestras, esto se debe al fenómeno *Bootstrap*, o técnica de conteo por reemplazo. En esta no se usa todo el conjunto original de *samples*, en su lugar se escoge un subconjunto aleatorio del mismo tamaño, pero con reemplazo, por lo que algunas observaciones podrán repetirse y, por otro lado, otras podrán no aparecer. Por último, el índice Gini evaluará la pureza de la división, determinando que atributo separa mejor las clases en el siguiente nodo, hasta llegar a uno totalmente puro.

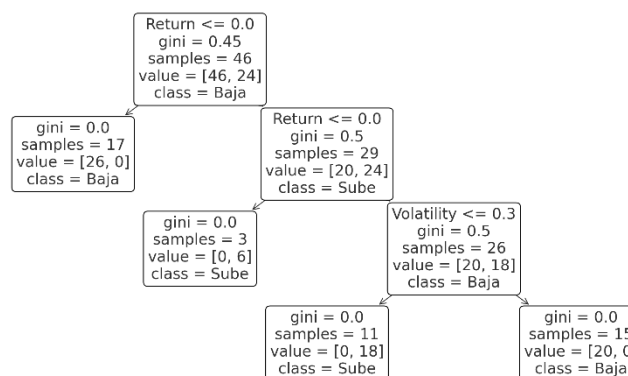


Figura 21. Ejemplo Simulado de Un Random Forest.
Fuente: De elaboración propia.

2.4.4 Redes Neuronales LSTM

Las Redes Neuronales son, actualmente, unas de las herramientas más recurrentes dentro del sector del *Machine Learning* y de la Inteligencia Artificial. Se definen como formulaciones informáticas de aprendizaje automático basadas en los fundamentos de las Redes Neuronales Biológicas. Sencillamente, replican la organización física del cerebro, reproduciendo de manera simplificada sus características computacionales (Buchanan, 2005).

De manera estructural, el funcionamiento de una neurona biológica y artificial es la misma. De igual manera que en las Redes Neuronales Biológicas, las RNA están creadas por multitud de unidades unidas entre sí mediante los llamados *inputs* y *outputs*, simulando las dendritas y los

axones respectivamente. Además, otra similitud reside en el interior de la neurona, donde se procesa toda la información recibida, llamado soma en el caso biológico.

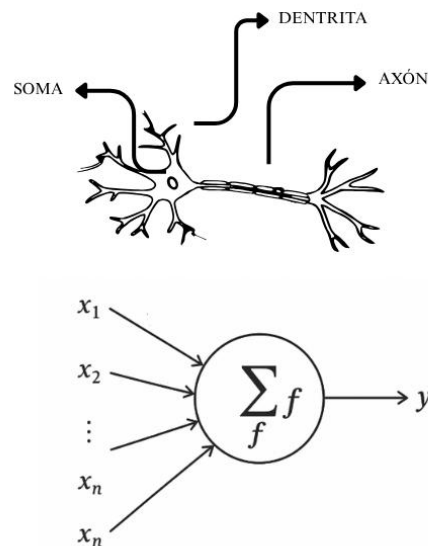


Figura 22. Red Neuronal Biológica VS Artificial
Fuente: De elaboración propia.

Una vez recibida la información por una serie de estímulos externos, o valores de entradas, esta se transfiere al interior de la neurona o perceptrón.

En el perceptrón, tal y como explica (Goodfellow, 2016) se calcula una combinación lineal de entradas, seguida de una función que determina una salida binaria. Este cálculo es la suma ponderada de todos los valores de entrada con una ponderación que asigna un peso sobre cada una de estas entradas. Dichos parámetros son los encargados de ajustar el programa de la RNA para permitir una predicción eficiente. Además, cada neurona contiene otro componente llamado sesgo, permitiendo una adecuación lo más fiel posible a los datos obtenidos. Resumidamente lo que hace cada neurona es un problema de regresión lineal como el siguiente:

$$Y_i = \sum_n P_{in} x X_n + s$$

Ecuación 35. Regresión Lineal de una Neurona Artificial.

Donde Y_i es la salida, P_n se refiere al peso de cada parámetro, X_n el valor de entrada o también llamado estímulo y s es el sesgo.

Para conocer en qué nivel de adaptación se encuentra el modelo de Red Neuronal que se está estudiando, es necesario formular una función llamada “log-likelihood” que plasme el error encontrado en cada una de las combinaciones de los parámetros de ajuste dentro de la función de coste, provocando un proceso de aprendizaje en el modelo. Dicha función tendrá que ser refinada mediante el “Descenso del gradiente”, el cual sigue las bases de un algoritmo de optimización

iterativa que busca el mínimo de una función con orden n , ajustando los parámetros paso a paso de una manera opuesta a la del gradiente de la función de coste (Géron, 2019).

Para entenderlo de una manera sencilla, lo que se pretende de primera estancia es representar una pendiente a partir del modelo de coste obtenido originalmente. Sin embargo, debido a la complejidad de esta formulación, la pendiente se representará dentro de un campo multidimensional con el mismo orden que el número de parámetros con los que se trabaje.

Para obtener dicha pendiente será necesario el cálculo de un gradiente que consiste en la derivada parcial de cada uno de los parámetros con los que se operen en el proceso, todo ello evaluado desde un punto x del dominio f , obteniendo una forma de cuantificar la dirección del campo que varía más rápidamente.

$$\nabla f(y) = \left(\frac{\partial f(y)}{\partial x_1}, \frac{\partial f(y)}{\partial x_2}, \frac{\partial f(y)}{\partial x_3}, \dots, \frac{\partial f(y)}{\partial x_i} \right)$$

Ecuación 36. Gradiente del modelo log-likelihood de Neuronal Artificial.

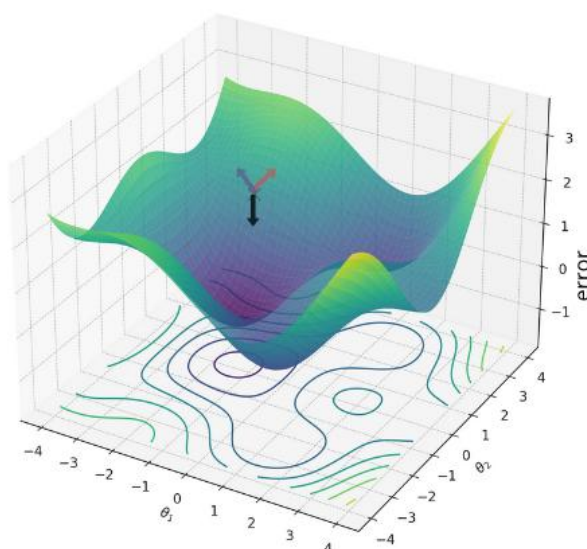


Figura 23. Representación multidimensional del Descenso del Gradiente

Fuente: De elaboración propia

Como se observa en la Figura 23, en la representación simulada de una función de orden 3, tras aplicar las derivadas parciales de cada componente se obtiene tres direcciones distintas, atendiendo al algoritmo, la dirección a tomar sería la de la flecha negra. Este proceso se realizaría sucesivamente hasta obtener un dominio con una pendiente próxima 0. La ecuación que plasma esta iteración es la siguiente:

$$\theta_i = \theta_{i-1} - \eta x \nabla f(y_{i-1})$$

Ecuación 37. Sucesión para el Descenso del Gradiente.

Donde θ_i se expresa como los parámetros del modelo en el momento i , θ_{i-1} son los referidos a un momento anterior, $\nabla f(y_{i-1})$ es el gradiente de la función de coste con respecto a los

parámetros de la iteración anterior y η es lo denominado como tasa de aprendizaje, que determina el tamaño del paso de la dirección del gradiente. Si es muy pequeño se necesitarán más sucesiones que si es de mayor tamaño.

El tipo de RNA más utilizado en el mundo predictivo son las conocidas como Redes Neuronales Recurrentes o RNN, estas están formadas por bucles que permiten una contención de la información generada anteriormente, permitiendo que las neuronas artificiales posean memoria. Esto se debe gracias a la estructura del perceptrón, la cual permite que sus salidas se reúsen como nuevos estímulos en etapas posteriores del procesamiento.

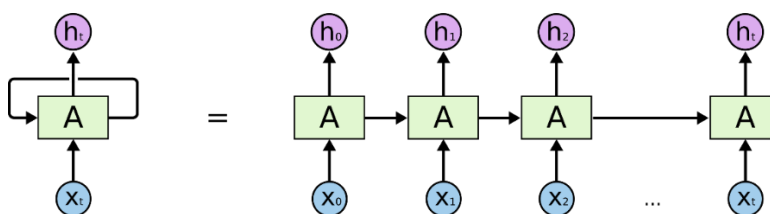


Figura 24. Funcionamiento de las Redes Neuronales Recurrentes
Fuente: Github.io Autor: alexis96.

Como se puede visualizar en la Figura 24, gracias a la composición en cadena, la RNA puede obtener una mayor eficacia y capacidad de representación en comparación con un esquema en el que no existe esa retroalimentación en los *outputs*.

Es posible encontrar una gran cantidad de Redes Neuronales Artificiales Recurrentes. Sin embargo, aquellas más con más aplicabilidad son las de tipos LSTM (Long-Short Term Memory). Propuestas por (Hochreiter & Schmidhuber 1997) con el fin de superar las limitaciones impuestas por las redes neuronales del momento, demostrando ser muy eficaces en situaciones donde se manejan una gran cantidad de datos secuenciales.

Las LSTM, a diferencia de las RNN básicas, están compuestas por cuatro etapas de redes neuronales que se conectan entre ellos mediante composiciones neuronales de tipo sigma o tangencial, finalizando con un operador, como es posible ver en la Imagen 24.

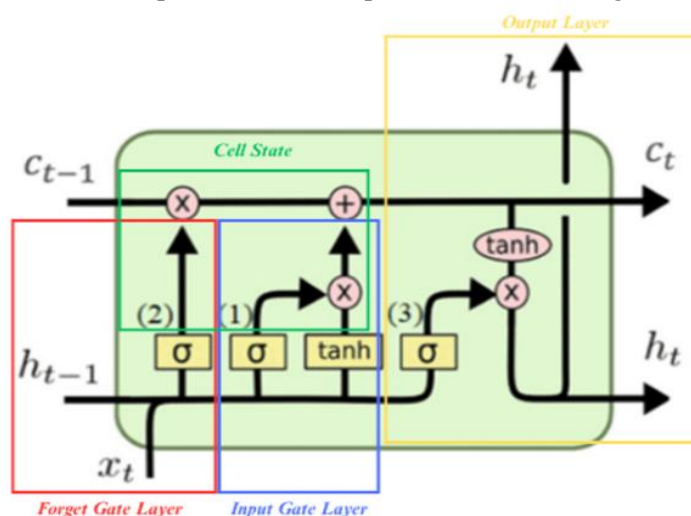


Figura 25. Funcionamiento de un LSTM.
Fuente: De Elaboración Propia.

Dentro de la arquitectura de una Red LSTM, cabe la posibilidad de encontrar distintos componentes o también llamado *compuertas*, que se encargan de regular el flujo de información dentro de cada celda de memoria:

- **Forget Gate Layer**, el modelo evalúa aquella información que ha de ser descartada de la celda de memoria (*Cell State*). Dicha decisión es tomada mediante una función de activación de tipo sigma σ , la cual da un valor entre 0 y 1 depende de si se desea olvidar la confirmación, dando un 0 como resultado, o mantenerla, obteniendo un 1 en la salida. Además, surge de la multiplicación entre la matriz de pesos P_{f2} con la concatenación del estado oculto anterior h_{t-1} y la entrada actual x_t . A ese mismo resultado se le aplica la suma con el término del vector sesgo s_{f2} :

$$f_2 = \sigma (P_{f2} \cdot [h_{t-1}, x_t] + s_{f2})$$

Ecuación 38. Forget Gate Layer LSTM.

- **Input Gate Layer y Generación de Candidatos**, compuertas encargadas de la información que se debe añadir al estado de la celda. La función de cada una es la siguiente:
 - o Una entrada de activación sigmoideal σ encargada de decidir qué elementos se actualizan.
 - o Una capa de candidatos que usa una activación tangencial \tanh con el objetivo de generar posibles valores restringidos entre -1 y 1.

$$f_1 = \sigma (P_{f1} \cdot [h_{t-1}, x_t] + s_{f1})$$

Ecuación 39. Entrada de Activación Sigmoideal del Input Gate Layer para LSTM.

$$f_{\tanh} = \tanh (P_{f_{\tanh}} \cdot [h_{t-1}, x_t] + s_{f_{\tanh}})$$

Ecuación 40. Entrada de Activación tangencial del Input Gate Layer para LSTM.

- **Cell State**, acción que confirma la actualización previa de la celda. Esta viene dada por la multiplicación del estado de las celdas anteriores C_{t-1} y el valor de la compuerta de olvido f_2 y la adición de las operaciones previamente multiplicadas de la compuerta de entrada f_1 y la capa generativa f_{\tanh} :

$$C_t = f_2 * C_{t-1} + f_1 * f_{\tanh}$$

Ecuación 41. Cell State para LSTM.

- **Output Layer**, determina la salida de la unidad en el estado de la celda actual. Primero se ha de calcular la activación de la compuerta de salida para decidir qué parte contiene la influencia principal de la salida, siendo esta de tipo sigma. Más tarde, se aplica la operación tangencial al estado de la celda actual C_t , y dicho resultado se multiplica con la salida de la puerta generada.

$$f_3 = \sigma (P_{f3} \cdot [h_{t-1}, x_t] + s_{f3})$$

Ecuación 42. Activación del Output Layer.

$$h_t = f_3 * \tanh(C_t)$$

Ecuación 43. Output Layer para LSTM.

2.5 Indicadores Técnicos y su función en la Toma de Decisiones.

Los indicadores técnicos en la bolsa son formulaciones fundamentales que permiten a los agentes económicos un análisis básico del comportamiento del mercado. Se basan en la creación de señales de compra o venta mediante la utilización de datos anteriores. Como define Murphy (1999), aunque no garantice un resultado seguro, estos análisis técnicos, medidos principalmente a través de gráficos, actúan de soporte esencial en la toma de decisiones dentro del entorno volátil y cambiante, complementado eficazmente los modelos predictivos de base econométrica.

Para conocer de una forma generalizada el funcionamiento de los indicadores técnicos, a continuación, se va a detallar algunos de los más populares en el ámbito financiero.

2.5.1 Relative Strength Index (RSI).

El Índice de Fuerza Relativa (RSI) es un indicador que se utiliza con el fin de medir la intensidad de los cambios más recientes del precio de un activo, mediante una comparativa de las alteraciones alcistas y bajistas en un periodo determinado. Este se representa mediante un valor entre 0 y 100, dando de esta forma una conclusión de sobrecompra o sobreventa.

El cálculo de este indicador se realiza mediante dos etapas distintas, aunque una de ellas necesita de un paso previo para conocer el valor intermedio.

Primero, se ha de calcular la Fuerza Relativa (*Relative Strength*) mediante el cociente de las medias móviles exponenciales alcistas y las de tendencia bajista en un periodo n determinado.

La Media Móvil Exponencial (EMA) consiste en asignar un peso a cada valor a estimar, teniendo en cuenta que los datos más actuales tendrán mayor relevancia ante los más lejanos.

$$EMA_t = P_t * \alpha + EMA_{t-1} * (1 - \alpha)$$

Ecuación 44. Media Móvil Exponencial.

Donde P_t se refiere al valor del periodo actual, α es el peso asignado a dicho valor y EMA_{t-1} se define como la media móvil exponencial del periodo anterior.

$$RS = \frac{EMA \text{ alcista de } n \text{ periodos}}{EMA \text{ bajista de } n \text{ periodos}}$$

Ecuación 45. Relative Strength.

Se dice que una Media Móvil Exponencial es alcista cuando la suma de todas las medias en los n periodos sea positiva. Por lo contrario, se denomina como bajista.

Una vez obtenido el valor RS se aplica la fórmula que permite obtener el Índice de Fuerza Relativa:

$$RSI = 100 - \left(\frac{100}{1 + RS} \right)$$

Ecuación 46. Relative Strength Index.

Según el valor obtenido, se puede interpretar de la siguiente manera:

- **RSI >70**, indica una tendencia de sobrecompra.
- **RSI <30**, señala una tendencia se sobreventa.

2.5.2 Bandas de Bollinger.

Este indicador se basa en la superposición de tres bandas dentro de los gráficos evolutivos referidos a los precios de los activos. Ciñéndose a las palabras de su creador John Bollinger (2001), al ofrecer una solución relativa a los valores del mercado reciente, permite comparar el comportamiento de los precios en el mercado, siendo especialmente útil para identificar cambios de tendencias o periodos de mucha o poca volatilidad.

Entre sus componentes, aparece una banda media que corresponde a una media móvil simple x , calculada por un periodo n de referencia. A partir de esta, derivan una banda superior y una inferior, expresadas como la suma y la resta, respectivamente, de un múltiplo K por la desviación estándar σ_n respecto a la media móvil. Este múltiplo es un valor subjetivo interpretado por cada individuo financiero, típicamente cercano a dos.

$$\sigma_n = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}$$

Ecuación 47. Desviación Estándar para las Bandas de Bollinger.

$$UB = X + K * \sigma_n$$

$$LB = X - K * \sigma_n$$

Ecuación 48. Límites Superiores e Inferiores de las Bandas de Bollinger.

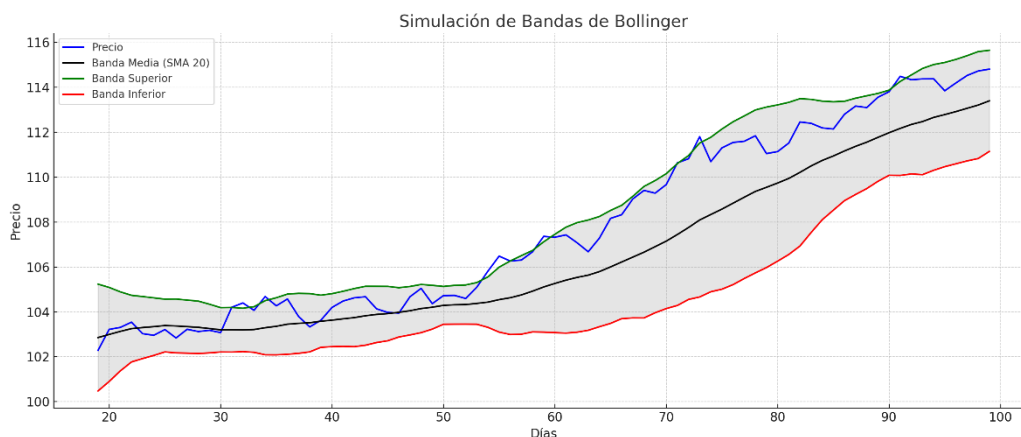


Figura 26. Representación Simulada de las Bandas de Bollinger
Fuente: De elaboración Propia

2.5.3 Moving Average Convergence Divergence (MACD).

La Media Móvil de Convergencia Divergencia (MACD, sus siglas en inglés), se utiliza principalmente para analizar cambios de volumen, de fuerza, de tendencia y de duración, dándole una gran versatilidad como indicador técnico.

Su formulación se basa en la diferencia entre dos Medias Móviles Exponenciales (EMA) con un periodo distinto. La primera referida a un corto plazo n_0 , usualmente de 12 periodos, y la segunda de medio plazo n_1 , normalmente de unos 26 periodos. La ecuación general de la MACD es la siguiente:

$$MACD = EMA(n_0) - EMA(n_1)$$

Ecuación 49. Proceso MACD.

Es necesario tener en cuenta que, a menor periodo, mayor será la sensibilidad de la EMA en las alteraciones de los precios. Permitiendo detectar más fácilmente los giros de tendencia.

Como se ejemplifica en el gráfico N, la parte inferior está compuesto por tres fundamentos:

- **Línea MACD**, obtenida mediante la ecuación MACD.
- **Línea de Señal**, es un EMA respecto de la línea MACD, usualmente de periodo 9.
- **Histograma**, es la diferencia entre la línea MACD y la de Señal. Cuando está es positiva, muestra una tendencia alcista; cuando es negativa, un impulso bajista.

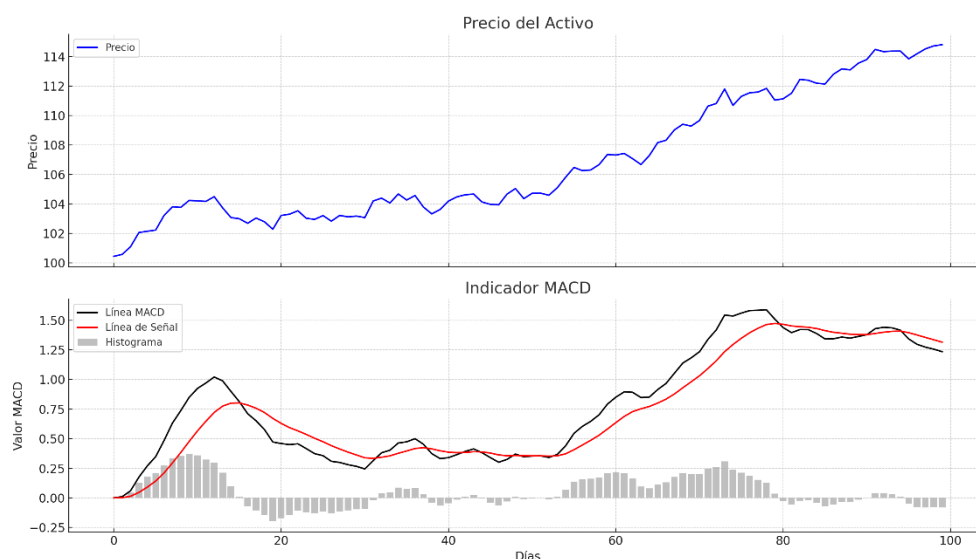


Figura 27. Ejemplo Simulado del Indicador MACD

Fuente: de Elaboración Propia.

2.5.4 Value at Risk (VAR)

El valor en riesgo de una cartera de inversión (VAR) se conoce como un instrumento estadístico encargado de cuantificar el riesgo de pérdida del valor de una inversión en un horizonte temporal. Es decir, como explica (Jorion, 2007), es una herramienta que pretende contestar a la pregunta de qué potencial existe de que ocurra una máxima pérdida relativa del valor de la cartera en condiciones normales del mercado.

Realmente este indicador se puede calcular desde diferentes puntos de vista, ya sea por datos históricos o por simulaciones. Sin embargo, una manera de entenderlo de una forma sencilla es mediante su método paramétrico de varianza – covarianza, donde se asume que los rendimientos actúan mediante una distribución normal. Su formulación viene expresada por el propio rendimiento esperado de la cartera μ , por el valor crítico de la distribución normal z , mostrado en la Tabla 6, y por la desviación estándar σ , la cual representa la volatilidad:

$$VaR = \mu - z * \sigma$$

Ecuación 50. Proceso del Value at Risk.

Para contemplar un ejemplo visual, en el caso de que existiese un 5% de posibilidades de entrar dentro de esa zona de riesgo extremo, la representación gráfica de la distribución normal se vería de la siguiente manera:

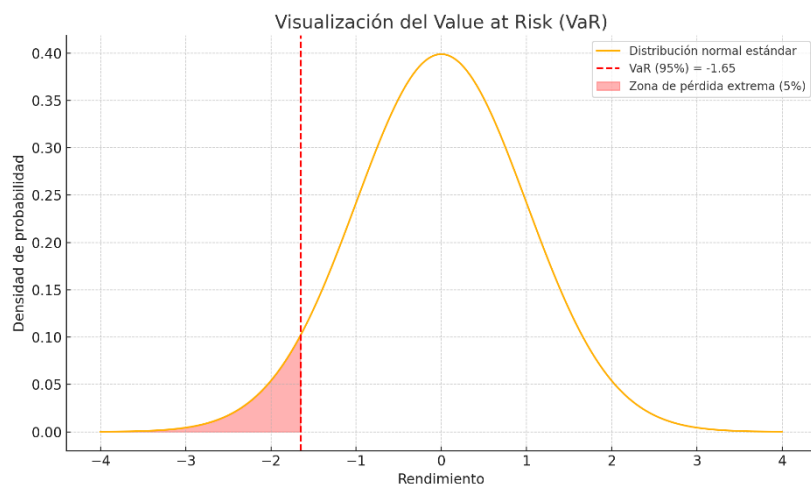


Figura 28. Visualización Distribución Normal del VaR

Fuente: De Elaboración Propia.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998

Tabla 6. Valores de Distribuciones Normales.

Fuente: De Elaboración Propia.

2.6. Teoría de la Optimización de Carteras.

Uno de los pilares fundamentales de la teoría cuantitativa en el ámbito financiero es la optimización de carteras de inversión. Su principal objetivo es determinar la relación más adecuada de activos financieros que permitan alcanzar un balance máximo entre el riesgo asumido y la rentabilidad esperada. Tal es su importancia que, como aseguran los economistas (Fabozzi et al. 2006), durante estas últimas décadas esta problemática ha sido abordada mediante multitud de técnicas estadísticas, matemáticas y económicas que permiten formalizar decisiones bajo condiciones de incertidumbre.

Con el desarrollo de los instrumentos computacionales y debido al constante desarrollo del mercado financiero, en la actualidad surgen nuevas técnicas de optimización más eficientes y robustas, buscando una constante adaptación a las condiciones reales de mercado. Entre estas destacan métodos como el *Deep Reinforcement Learning*.

Sin embargo, con el fin de posicionar las raíces en el ámbito inversivo, el presente trabajo se centrará en analizar las principales bases de la optimización de carteras, evaluando tanto sus fundamentos teóricos como su aplicabilidad práctica.

2.6.1. Modelo de Markowitz: La Teoría Moderna de Carteras.

Harry Markowitz es, sin discusión, el padre de la teoría moderna de carteras y su principal referente práctico. Su trabajo transformó radicalmente la profesión, la teoría, la legislación y los métodos de gestión de inversiones. Las innovaciones que introdujo han mejorado la vida de incontables personas que dependen, directa o indirectamente, de los rendimientos financieros para asegurar su futuro económico. (Goetzmann, 2023)

En épocas anteriores a Markowitz, los inversores se centraban en los activos que más rentabilidad pronosticasen recibir. Sin embargo, tras la irrupción del economista estadounidense, se demostró que la manera más eficiente de crear una cartera financiera se basaba en la idea de diversificar las carteras de inversión en diferentes activos que mantuviesen un equilibrio entre rentabilidad y riesgo, dando lugar al concepto actualmente conocido como Frontera Eficiente, donde se forman las carteras bajo dos teorías:

- 1- Dado un nivel de rentabilidad esperada, se busca el nivel más bajo de riesgo.
- 2- Dado un nivel de riesgo, se alcanza la mayor rentabilidad esperada.

Para entender el proceso de este modelo, lo primero será definir que activos del universo financiero son atractivos para el inversor ya que, a partir de estos, se conocerá las posibles combinaciones de carteras, añadiéndole un peso w_i a cada activo.

Como define (Markowitz, 1952), la rentabilidad de una cartera $E(R_p)$ es simplemente la media ponderada de cada una de las rentabilidades $E(R_i)$ de la totalidad de activos N que la conforman:

$$E(R_p) = \sum_{i=1}^N w_i * E(R_i)$$

Ecuación 51. Rentabilidad Esperada de una Cartera.

Por consiguiente, el siguiente paso será el cálculo del riesgo, definido como varianza de la cartera $Var(R_p)$. En esta no solo importa la media ponderada de los riesgos individuales $Var(R_i)$, sino que también es necesario analizar el comportamiento de los activos entre sí, mediante la covarianza $Cov(R_i, R_j)$. Formulándose de tal manera:

$$Var(R_p) = \sum_{i=1}^N w_i^2 * Var(R_i) + 2 \sum_{j=1}^N \sum_{i=1}^N w_i w_j * Cov(R_i, R_j)$$

Ecuación 52. Varianza de la Cartera.

En este caso, si los activos brindan una correlación baja o negativa, tal y como define el principio de diversificación, se podrá reducir en su totalidad, ya que los movimientos de un activo son contrarios a los del otro, generando una cobertura perfecta mediante pesos específicos que conducen a una inversión sin riesgo.

Una vez obtenido tanto el riesgo como el rendimiento, ya se puede representar gráficamente una frontera eficiente de una cartera mediante la contemplación de multitud de combinaciones de estas, asumiéndose como variables los pesos de cada activo. Como se puede observar en la Gráfica 29, el eje X señala el riesgo y el eje Y la rentabilidad esperada.

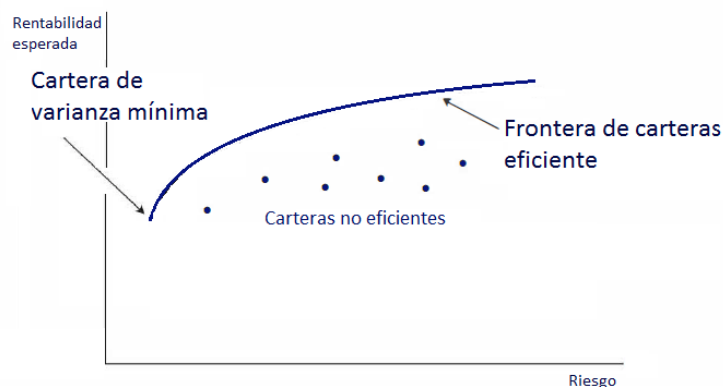


Figura 29. Frontera Eficiente Rentabilidad-Riesgo

Fuente: De Federico J. Caballero Ferrari en Economipedia.com

Todos aquellos puntos que no estén en la frontera eficiente se considerarán carteras ineficientemente optimizadas.

Sin embargo, en un modelo práctico, los inversores no solo tienen la posibilidad de elegir entre activos denominados como “riesgosos”, sino que también tienen la posibilidad de invertir en agentes con menor incertidumbre, como aquellos anexados a la rentabilidad fija. Esto permite nuevas oportunidades: combinar activos según su volatilidad en una cartera eficiente.

Esta combinación genera lo denominado como Línea de Asignación de Capital (CAL), representada dentro del gráfico de rentabilidad-riesgo como una línea recta tangencial a la frontera eficiente. Su ecuación es la siguiente:

$$E(R_p) = R_F + \left(\frac{E(R_i) - R_F}{\sigma_i} \right) * \sigma_p$$

Ecuación 53. Rentabilidad Esperada de una Cartera + CAL.

Donde la rentabilidad del activo financiero libre de riesgo es representada como R_F , la desviación estándar de cada componente es σ_i y de la cartera total se refiere a σ_p . De una manera representativa se vería como con la Figura 30, donde aparece una nueva variable representativa aparte de la ya conocida frontera eficiente. Esta nueva variable muestra las combinaciones más eficientes entre riesgo y rentabilidad al incluir un activo libre de riesgo. Su pendiente plasma la mejor tasa de compensación por unidad de riesgo.

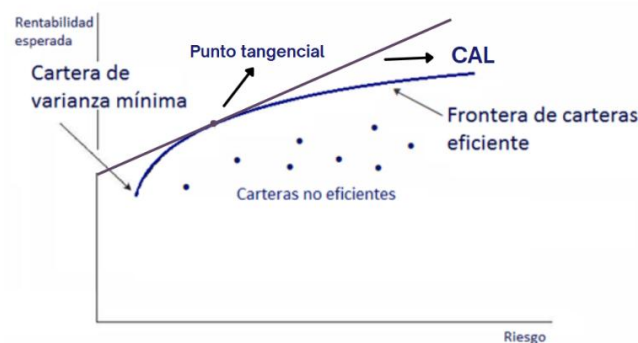


Figura 30. Frontera Eficiente Rentabilidad-Riesgo + CAL.

Fuente: De Elaboración Propia a partir de Federico J. Caballero Ferrari en Economipedia.com

2.6.2. Modelo Black-Litterman: Avance en la Optimización de Carteras.

Uno de los principales problemas de las carteras basadas en la teoría moderna trata de poder combinar los modelos cuantitativos con procedimientos subjetivos tratados por cada individuo. Siguiendo los argumentos de (Satchell & Scowcroft, 2000), los agentes económicos más conservadores se alejan de la gestión cuantitativa, ya que sienten que no capturan eficazmente el valor añadido que ellos aportan.

Como solución a este rompecabezas, (Black & Litterman, 1992) deciden crear un sistema capaz de aumentar su aplicabilidad en los casos prácticos de la optimización de carteras.

El modelo BLM (Black-Litterman) pretende compaginar opiniones subjetivas de un inversor con la cartera de equilibrio del mercado mediante operaciones basadas en el teorema de Bayes. Es decir, permite pasar de una creencia inicial a una más ajustada basada en evidencias. Además, el modelo BLM propone un análisis de mercado más ajustado a la realidad, donde no existe un equilibrio fijo, sino un punto dinámico y cambiante.

El resultado final de esta metodología es un vector de retornos esperados para cada activo, el cual se obtiene por otros dos retornos: los retornos esperados de equilibrio y los retornos esperados subjetivos. Por lo tanto, el resultado final respecto al retorno de equilibrio estará condicionado por las tendencias subjetivas del inversor. La ecuación principal de Black-Litterman es la representada en la fórmula N:

$$E(r) = \Pi + \tau \Sigma P^T (P \tau \Sigma P^T + \Omega)^{-1} (Q - P \Pi)$$

Ecuación 54. Proceso Black - Litterman.

Para poder entender dicha ecuación es necesario explicar sus dos componentes principales.

Por un lado, la parte de equilibrio del mercado Π es el vector que representa aquello que el mercado cree que será el retorno, obviando supuestas subjetividades de los inversores. Su formulación viene derivada del modelo de Valoración de Activos Financieros (CAPM), el cual, a su vez, constituye una extensión de la metodología de la Línea de Asignación de Capital, previamente definido en la página 47.

$$\Pi = \frac{\Sigma w (R_M - R_F)}{w^T \Sigma w}$$

Ecuación 55. Vector Equilibrio del Mercado.

El término Σ hace referencia a la matriz de varianzas-covarianzas de cada activo, los pesos de cada uno de estos se representan con w y, por último, R_M y R_F definen, respectivamente, los retornos esperados del mercado y los retornos de activos libres de riesgo.

Por el otro lado, es necesario entender el vector de retornos sujeto a las opiniones de los inversores Q , donde cada elemento corresponde al retorno esperado que el inversor cree que debería tener el activo que representa. Además, es necesario tener en cuenta que el vector Q siempre viene acompañado de un vector P que interpreta a que activos se refiere cada opinión, dando como resultado la siguiente ecuación:

$$Q = P * \mu + \varepsilon$$

Ecuación 56. Vector de Retorno Sujeto a Opiniones del Inversor.

Donde μ representa el vector de retornos esperados verdaderos de los activos y sigue una distribución normal con media Π y varianza $\tau \Sigma$.

$$\mu \sim N(\Pi, \tau \Sigma)$$

Ecuación 57. Distribución Normal del vector de Retornos Esperados.

Siendo τ un valor escalar que representa el grado de incertidumbre asociado al vector de retornos implícito.

ε define el término de error asociado a las valoraciones subjetivas del inversor y se modela estadísticamente como la siguiente distribución normal:

$$\varepsilon \sim N(0, \Omega)$$

Ecuación 58. Distribución Normal de errores asociados a Valores Subjetivos.

El término Ω refleja el nivel de confianza del individuo que invierte. Se define mediante una matriz que está compuesta por la varianza al cuadrado de cada activo.

$$\Omega = \begin{pmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \dots & 0 \\ 0 & 0 & \sigma_v^2 \end{pmatrix}$$

Ecuación 59. Matriz de Confianza Black-Litterman.

Para entender mejor la última parte de la ecuación de Black-Litterman, se dará un ejemplo práctico:

Supongamos que tenemos tres activos A, B, C y D donde:

- Existe un 100% esperado del retorno de A de un 12%. (Opinión Absoluta)
- Está 80% seguro de que el retorno esperado del activo B será mayor que el de C, y que la diferencia estará entre el 4% y el 6%. (Opinión Relativa)
- No tiene opinión subjetiva sobre el activo D, y por tanto no crea una opinión sobre D.

Siguiendo la ecuación 56:

$$\begin{pmatrix} 0.12 \\ 0.05 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \end{pmatrix} * \mu + \varepsilon$$

La obtención de Q se trata de un vector columna formado por los valores retorno esperados según el inversor. La primera fila de la matriz P corresponde a la opinión absoluta, ya que solo se refiere en este ejemplo al activo A, se asignará un 1 en la primera columna de dicha fila. Por otra parte, en la fila dos, referida a la opinión relativa, aquellos activos que tengan un mejor desempeño tendrán un valor positivo y los que reciban un desempeño peor esperado serán negativos.

Más tarde, se obtendría la distribución normal de los valores μ y ε . A partir del último, sería posible obtener la matriz de nivel de confianza Ω de la siguiente manera.

Suponiendo un caso simulado con una distribución normal que sigue la siguiente representación 31.

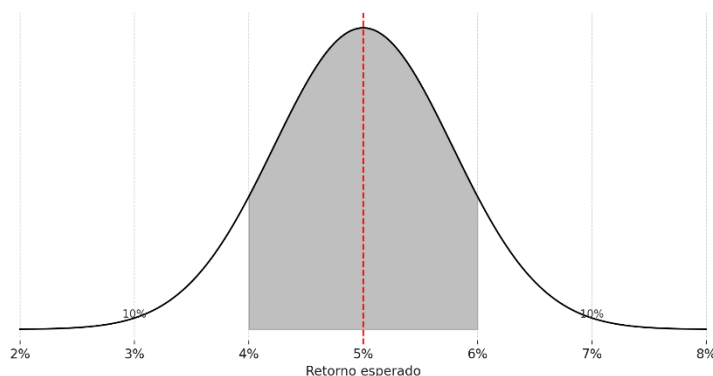


Figura 31. Distribución Normal Ejemplo Black-Litterman

Fuente: De elaboración Propia

$$P(-1,28 \leq Z \leq 1,28) = 0.80$$

$$Z = \frac{r - m}{\sigma} = \frac{0.06 - 0.05}{\sigma} = 1.28$$

$$\sigma = \frac{0,01}{1,28} \approx 0,0078$$

$$\Omega = \begin{pmatrix} 0.0078^2 & 0 \\ 0 & 0.0078^2 \end{pmatrix} = \begin{pmatrix} 6.1^{-5} & 0 \\ 0 & 6.1^{-5} \end{pmatrix}$$

Una vez obtenidas todas las matrices correspondientes a la opinión subjetiva del inversor, solo quedará obtener la matriz de retorno implícito del mercado para, de esta manera, conocer el resultado final del modelo BLM.

3. Marco Metodológico.

3.1 Lenguaje de Programación.

A la hora de realizar un estudio experimental sobre los diferentes modelos predictivos de base econométrica es necesario el uso de un lenguaje computacional que permita cierta complejidad operativa. Hoy en día existen multitud de programas suficientemente potentes como para poder realizar un análisis tan complejo como es el analizado en este trabajo, donde es necesario un tratamiento preciso de datos, capacidad para controlar grandes volúmenes de información y existe la oportunidad de comparar a nivel efectivo los diferentes sistemas.

Por todo ello, se ha optado finalmente por utilizar como lenguaje principal de programación Python, ya que brinda de multitud de beneficios con respecto a los demás, además, proporciona una gran versatilidad en el ámbito académico y profesional.

Python cuenta con multitud de usuarios y desarrolladores que permiten una constante evolución y actualización de recursos, documentación o bibliotecas especializadas. Entre las ventajas más destacadas encontramos las siguientes:

- **Claridad**, gracias a su sintaxis legible es posible implementar de manera sencilla modelos más complejos sin el requerimiento de un aprendizaje excesivo.
- **Especificación**, dispone de librerías orientadas principalmente al análisis de datos y estructuras de bases econométricas, como Pandas, Statsmodels, Scikit-learn o NumPy, permitiendo tanto la manipulación avanzada de datos como una estimación precisa de modelos preventivos.
- **Integridad**, permite una gran combinación de entornos y herramientas con diferentes utilidades, facilitando tareas de visualización de resultados mediante bibliotecas como Matplotlib o Seaborn.

- **Aplicabilidad**, existe una gran versatilidad en múltiples campos, dando la posibilidad de ejecutar el programa en multitud de compiladores.

En conjunto, Python es un instrumento ideal para abordar con eficacia y exactitud las necesidades que exigen este trabajo.

3.2 Análisis de Datos.

3.2.1 Obtención de Datos.

Para llevar a cabo un estudio exhaustivo de los datos históricos que se van a emplear en los distintos métodos preventivos sobre la evolución bursátil, es indispensable realizar una búsqueda profunda y fundamentada que permita la obtención de estos de una manera simple y versátil.

Es por ello, que, para la información financiera agregada a la microeconomía, como los precios históricos o el volumen de negociación, la librería que se ha utilizado ha sido Yahoo! Finanzas.

Por otro lado, para la información relacionada con la macroeconomía, como tasas de empleabilidad, la inflación o el producto interior bruto; se utilizará datos obtenidos de la fuente World Bank, donde existe una gran variedad de referencias a nivel global y a los que se pueden acceder de una manera sencilla.

La implementación de las siguientes librerías en un script se visualizar de la siguiente manera:

```
import yfinance as yf
import wldata
```

Estructura del Código 1. Importación de la librería Yahoo! Finance y de la librería World Bank.

Dicha llamada a esta fuente de datos devolverá una combinación de datos agrupados en pandas (DataFrame), como es posible ver en la siguiente Imagen, el cual se ha utilizado un ejemplo de la empresa tecnológica Apple:

	Close_AAPL	High_AAPL	Low_AAPL	Open_AAPL	Volume_AAPL
Date					
2020-01-02	72.620827	72.681274	71.373203	71.627077	135480400
2020-01-03	71.914803	72.676431	71.689942	71.847102	146322800
2020-01-06	72.487854	72.526541	70.783256	71.034717	118387200
2020-01-07	72.146942	72.753823	71.926915	72.497529	108872000
2020-01-08	73.307518	73.609752	71.849540	71.849540	132079200

Figura 32. Representación Data Frame en Python

Fuente: De Elaboración Propia.

3.2.2 Limpieza de Datos.

A la hora de utilizar Pandas, eliminar los datos que no vayan a ser relevantes en el análisis es fundamental para poder procesar de manera óptima toda la información que se requiere a nivel operativo. Por eso, hay que tener en cuenta las siguientes posibilidades que pueden surgir a la hora de tratar con fuentes externas:

- **Eliminar datos nulos**, dada la exportación de referencias históricas desde fuentes externas, pueden aparecer datos nulos representados como “NaN”, los cuales han de ser eliminados o intercambiados con otro valor como para poder realizar un procesamiento correcto. Este proceso se realiza mediante la siguiente codificación:

```
df = df.dropna()  
df = df.fillna(0)
```

Estructura del Código 2. Eliminación de Datos Nulos.

- **Eliminar Columnas irrelevantes**, a veces los datos descargados incluyen columnas innecesarias en el estudio, se eliminarán de la siguiente manera:

```
df = df.drop(columns=["ColumnName"])
```

Estructura del Código 3. Eliminación de Columnas Irrelevantes.

- **Eliminar datos duplicados**, debido a la posibilidad de encontrar filas repetidas, sobre todo al mezclar fuentes, es necesario tener en cuenta este punto mediante el siguiente código:

```
df = df.drop_duplicates()
```

Estructura del Código 4. Evitar Datos Duplicados.

3.2.3 Procesamiento de Datos.

A la hora de ejecutar el código que plasme el análisis de cada método estudiado durante este trabajo, es necesario que sea simple y legible, pero a la vez, ha de ser lo suficientemente potente como para soportar la complejidad cuantitativa de estos.

Después de un estudio profundo sobre las posibilidades que brinda este lenguaje de programación, se ha optado por utilizar Google Colaboratory, definido como una herramienta en la nube que permite escribir y ejecutar código Python directamente desde un navegador web. Entre sus beneficios podemos encontrar los siguientes:

- **Coste**, al ser gratuito permite una mayor participación de los usuarios y desarrolladores.
- **Disponición**, es posible encontrar una gran multitud de librerías ya preinstaladas, facilitando el acceso a estas a la hora de la escritura del código.

- **Accesibilidad**, al ser un instrumento situado en la nube, permite que se pueda utilizar un mismo archivo desde diferentes componentes, ya que este se encuentra guardado en Google Drive.

Por otro lado, para poder realizar un procesamiento correcto y compacto de los sistemas a analizar, es necesario la utilización de librerías como conjuntos de funciones, instrucciones o herramientas ya escritas con el objetivo de evitar una escritura total desde cero. Las utilizadas en esta investigación serán las siguientes:

- **Numpy**, útil para las operaciones numéricas y vectoriales con cierta complejidad.
- **Pandas**, manipulación de datos obtenidos mediante la representación en DataFrames.
- **Datetime**, utilizado con la finalidad de manejar fechas y tiempos. Permiten un cálculo preciso de intervalos y la definición de rangos.
- **Statsmodels**, implementación de matemática estadística, en este caso metodología predictiva.
- **Sklearn.metrics**, permite y facilita el cálculo de errores y evaluación (RMSE y MAE).

3.3 Implementación Computacional

3.3.1 Modelos Predictivos.

Cuando se pretende obtener un resultado predictivo de un sistema estadístico en Python, es necesario seguir unos pasos fundamentales en su código para implementarlos, de tal forma que el lenguaje sea capaz de calcular dicho pronóstico. En el presente análisis, con el objetivo de evaluar los modelos predictivos seleccionados, se realizará un *Training Test*, que constará de un entrenamiento del modelo utilizando datos históricos correspondientes al periodo comprendido entre 2020 y 2023. A partir de este, se generará una simulación del pronóstico para el 2024. La representación gráfica proporcionará tanto las predicciones generadas por los modelos como los valores reales observados, permitiendo así una comparación visual entre ambos conjuntos.

En este proceso, por mucho que se estudien métodos distintos, existen algunos puntos en común:

- 1- Importación de librerías, permitiendo su uso en dicho script o cuaderno.

```
import yfinance as yf
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from statsmodels.tsa.holtwinters import ExponentialSmoothing
from sklearn.metrics import mean_absolute_error, mean_squared_error
import warnings
warnings.filterwarnings('ignore')
```

Estructura del Código 5. Importación de Librerías.

- 2- Definir el símbolo, nombre o *ticker* de los productos de inversión financieros a estudiar.

Nombre del Producto	Ticker del producto
Microsoft	MSFT
S&P500	^GSPC
Invesco NASDAQ 100 ETF	QQQM

Tabla 7. Nombre de Activo VS Ticker.

Fuente: De elaboración propia.

```
ticker = 'AAPL'
```

Estructura del Código 6. Implementación del Ticker.

- 3- Descargar los datos que se quieran manipular con su debido intervalo temporal y realizar la limpieza de información no necesaria:

```
df = yf.download(ticker, start='2019-01-01', end='2025-01-01', interval='1wk')  
df = df[['close']].dropna()
```

Estructura del Código 7. Descarga de Datos de Yahoo! Finance.

- 4- Separar en intervalos de tiempo los datos con los que se van a entrenar el modelo analizado y los datos reales que se encuentran en las fechas a pronosticar:

```
df_train = df[(df.index >= '2020-01-01') & (df.index < '2024-01-01')]  
df_real_2024 = df[(df.index >= '2024-01-01') & (df.index < '2025-01-01')]
```

Estructura del Código 8. Separación de los Datos que se van a Entrenar.

Por otro lado, según el método predictivo que se quiera estudiar, habrá ciertos cambios en la programación que vayan acordes a las operaciones y teorías que utilizan estos sistemas econométricos:

- **Holt-Winters**, donde se implementa el código entrenado basado en el suavizamiento exponencial triple, donde se tiene en cuenta parámetros como la estacionalidad y la tendencia, En este caso, se analiza una estacionalidad semanal a lo largo de un periodo de 52 semanas (1 año).:

```
modelo = ExponentialSmoothing(df_train['close'], trend='add', seasonal='add', seasonal_periods=52)  
ajuste = modelo.fit()  
pronostico_2024 = ajuste.forecast(len(df_real_2024))  
pronostico_2024.index = df_real_2024.index
```

Estructura del Código 9. Modelo Holt-Winters.

- **SARIMA (ARIMA)**, se aplica siguiendo los fundamentos de la teoría Autorregresiva Integrada de Media Móvil con la capacidad de captar la estacionalidad de la serie temporal. Para ello, en la codificación, se interpretan combinación de parámetros arbitrarios que permiten evaluar distintos modelos candidatos. La mejor simulación se determina mediante el Criterio de Información de Akaike (AIC), una métrica que equilibra la calidad del ajuste con la complejidad del modelo, siendo el mejor el que menor valor represente:


```
param_combinaciones = [
    ((1,1,1), (0,1,1,52)),
    ((2,1,1), (1,1,1,52)),
    ((1,1,2), (1,1,0,52)),
    ((2,1,2), (1,1,1,52)),
    ((0,1,1), (0,1,1,52)),
    ((5,2,1), (1,1,1,52))
]

mejor_aic = float('inf')
mejor_order = None
mejor_seasonal_order = None
mejor_model = None

for order, seasonal_order in param_combinaciones:
    try:
        model = sm.tsa.statespace.SARIMAX(train,
                                           order=order,
                                           seasonal_order=seasonal_order,
                                           enforce_stationarity=False,
                                           enforce_invertibility=False)

        results = model.fit(dispatch=False)
        print(f'Tested SARIMA{order}x{seasonal_order} - AIC: {results.aic:.2f}')

        if results.aic < mejor_aic:
            mejor_aic = results.aic
            mejor_order = order
            mejor_seasonal_order = seasonal_order
            mejor_model = results

    except Exception as e:
        print(f'Model SARIMA{order}x{seasonal_order} failed: {e}')
pronostico_2024 = mejor_model.forecast(len(df_real_2024))
pronostico_2024.index = df_real_2024.index
```

Estructura del Código 10. Modelo SARIMA.

- **ARIMAX (ARIMA)**, esta extensión del sistema ARIMA contiene la particularidad de que se observan datos históricos diferentes a los valores de cierre de la bolsa. En este caso, se han considerado tres tipos de información adicional:
 - **Índice de Volatilidad (VIX)**, también conocido como índice del miedo, mide la incertidumbre o la volatilidad esperada del mercado en los próximos 30 días.
 - **Índice de Precios al Consumidor (IPC)**, cuantifica el valor de la inflación, es decir, el aumento de precios en bienes y servicios. Es un valor muy vigilado ya que puede marcar una subida o bajada del tipo de interés.
 - **Tasa de Interés Real**, refleja el rendimiento efectivo de una inversión, una vez descontado la inflación. Es un valor muypreciado en la bolsa de valores, ya que determina el atractivo de invertir, el crecimiento de las empresas y la actitud de los consumidores.

El primer paso para realizar es la descarga de los datos históricos pertinentes al estudio ARIMAX. En este caso, los valores de cierre y la magnitud VIX será obtenida mediante la librería de Yahoo! Finance, mientras que los valores de IPC y Tasa Real estarán sujetos a la fuente procedente del World Bank, importado gracias al llamamiento de la sintaxis computacional 2. Estos datos tendrán que ser procesados y limpiados de una manera muy precisa y unidos en un solo DataFrame para poder manipularlos en conjunto.

```
ticker_objetivo = 'AAPL'
vix_ticker = '^VIX'
fecha_inicio = '2019-01-01'
fecha_fin = '2025-01-01'

df_stock = yf.download(ticker_objetivo, start=fecha_inicio, end=fecha_fin, interval='1wk')
df_stock = df_stock[['Close']]
df_stock.columns = ['Stock']

df_vix = yf.download(vix_ticker, start=fecha_inicio, end=fecha_fin, interval='1wk')[['Close']]
df_vix.columns = ['VIX']

df = df_stock.join(df_vix, how='inner')
indicadores_exo = {
    'FR.INR.RINR': 'InterestRate',
    'FP.CPI.TOTL.ZG': 'CPI'
}

wb_df = wbdata.get_dataframe(indicadores_exo, country='USA')
wb_df.index = pd.to_datetime(wb_df.index)
wb_df = wb_df.sort_index()
wb_df = wb_df.reindex(df.index, method='ffill')

if isinstance(df.index, pd.MultiIndex):
    df.index = df.index.get_level_values(0)
if isinstance(wb_df.index, pd.MultiIndex):
    wb_df.index = wb_df.index.get_level_values(0)

df.index = pd.to_datetime(df.index)
wb_df.index = pd.to_datetime(wb_df.index)
df = df.join(wb_df[['InterestRate', 'CPI']], how='left')
df = df.fillna(method='ffill')
```

Estructura del Código 11. Implementación de datos Yahoo! Finance y World Bank.

A partir de este procesamiento, al igual que en la otra extensión enfocada en el modelo Autorregresivo Integrado de Media Móvil, se entrena el modelo desde 2020 hasta 2023. Más tarde, se proponen combinaciones para los parámetros que formalizan dicho método y, por último, con la mejor combinación se realiza la predicción del modelo analizado.

```
df_train = df[(df.index >= '2020-01-01') & (df.index < '2024-01-01')]
df_test = df[(df.index >= '2024-01-01') & (df.index < '2025-01-01')]

y_train = df_train['Stock']
X_train = df_train[['VIX_rolling', 'InterestRate_rolling', 'CPI_rolling']]
X_test = df_test[['VIX_rolling', 'InterestRate_rolling', 'CPI_rolling']]

# Prueba de combinaciones
combinaciones = [(1,1,1), (2,1,1), (1,1,2), (2,1,2), (0,1,1), (5,1,2)]
mejor_aic = float('inf')
mejor_order = None
mejor_modelo = None

for order in combinaciones:
    try:
        modelo = sm.tsa.SARIMAX(endog=y_train, exog=X_train, order=order,
                                enforce_stationarity=False, enforce_invertibility=False)
        resultado = modelo.fit(dispatch=False)
        print(f"ARIMAX{order} - AIC: {resultado.aic:.2f}")

        if resultado.aic < mejor_aic:
            mejor_aic = resultado.aic
            mejor_order = order
            mejor_modelo = resultado

    except Exception as e:
        print(f"Error con ARIMAX{order}: {e}")

# Predicción con el mejor modelo
forecast = mejor_modelo.predict(start=df_test.index[0], end=df_test.index[-1], exog=X_test)
forecast.index = df_test.index
```

Estructura del Código 12. Modelo ARIMAX.

Además, de una manera opcional, también es posible graficar mediante una matriz de correlaciones, la importancia de las diferentes variables expuestas en el modelo entre sí. Obteniendo información sobre la efectividad de las variables exógenas en la evaluación realizada.

```
df['Stock_pct_change'] = df['Stock'].pct_change()
df['VIX_rolling'] = df['VIX'].rolling(window=4).mean()
df['InterestRate_rolling'] = df['InterestRate'].rolling(window=4).mean()
df['CPI_rolling'] = df['CPI'].rolling(window=4).mean()
df = df.fillna(method='bfill')
columnas_filtradas = [col for col in df.columns if 'rolling' not in col and col != 'Stock_pct_change']
corr = df[columnas_filtradas].corr()
# Graficar la matriz de correlación
import seaborn as sns
import matplotlib.pyplot as plt
plt.figure(figsize=(10, 6))
sns.heatmap(corr, annot=True, cmap='coolwarm')
plt.title('Matriz de correlación (solo variables originales)')
plt.show()
```

Estructura del Código 13. Correlación de Variables Exógenas del modelo ARIMAX.

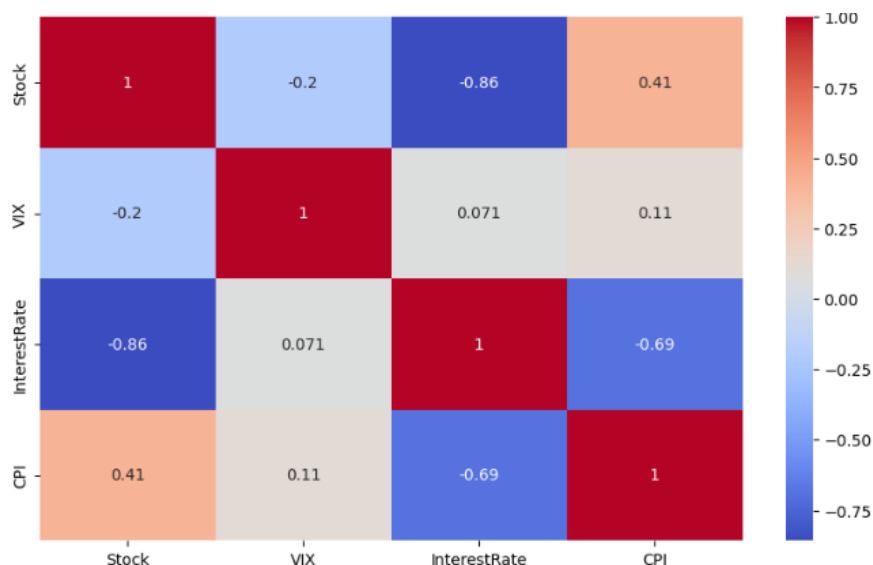


Figura 33. Correlación entre Variables Exógenas del Modelo ARIMAX.

Fuente: De Elaboración Propia.

En la Figura 34, los vínculos de variables positivas indican una mayor relación cuanto más grande sea la magnitud absoluta de esta. Por otro, el signo de dicha correlación indica si es directamente proporcional o inversamente proporcional.

- **Random Forest**, como se ha visto anteriormente en el marco metodológico, para realizar un análisis sobre este modelo es necesario seleccionar un conjunto de variables que determinen una frontera a la hora de tomar decisiones tanto en los valores como en la tendencia del pronóstico. Es por ello por lo que antes de implementar el código pertinente a la previsión, se ha de calcular los valores de estos atributos para todas las semanas sobre las que se vayan a predecir el precio:

```
df['Return'] = df['Close'].pct_change()  
df['Volatility'] = df['Close'].rolling(window=10).std()  
df['MA10'] = df['Close'].rolling(window=10).mean()  
df['MA50'] = df['Close'].rolling(window=50).mean()  
df.dropna(inplace=True)  
weekly_df = df.resample('W').last()  
weekly_df['Target'] = weekly_df['Close'].shift(-1)  
weekly_df.dropna(inplace=True)
```

Estructura del Código 14. Implementación Variables del Random Forest.

Por consiguiente, como en los códigos anteriores, se tendrá que entrenar al sistema mediante un intervalo de tiempo suficiente como para asegurarse que es capaz de realizar un estudio óptimo:

```
df_train = weekly_df[weekly_df.index.year < 2024]  
df_real_2024 = weekly_df[weekly_df.index.year == 2024]
```

Estructura del Código 15. Implementación del Intervalo de Entrenamiento del Random Forest.

Finalmente, se presenta el código capaz de pronosticar el intervalo que se desea. Para obtener estos resultados, se expone una semilla aleatoria fijada mediante el parámetro “random_state=45”:

```
features = ['Return', 'Volatility', 'MA10', 'MA50']  
X_train = df_train[features]  
y_train = df_train['Target']  
X_test = df_real_2024[features]  
y_test = df_real_2024['Target']  
rf = RandomForestRegressor(n_estimators=100, random_state=45)  
rf.fit(X_train, y_train)  
pronostico_2024 = rf.predict(X_test)
```

Estructura del Código 16. Modelo Random Forest.

- **LSTM**, a diferencia de la sintaxis del resto de métodos, esta codificación ofrece mayor complejidad a la hora de ejecutar el programa. En primer lugar, es necesario una normalización del precio y volumen a analizar, con el objetivo de tratar los datos con unidades similares, mejorando de esta manera su rendimiento:

```
normal = MinMaxScaler()  
normalizado_df = normal.fit_transform(df)
```

Estructura del Código 17. Precios Normalizados Para LSTM.

Por otro lado, se procede a separar según los intervalos de entrenamiento y de pronóstico:

```
df_normalizado = pd.DataFrame(df, index=df.index, columns=df.columns)  
train = df_normalizado.loc["2020":"2023"]  
test = df_normalizado.loc["2024"]
```

Estructura del Código 18. Periodo de Entrenamiento y de Pronóstico para un LSTM.

A continuación, se presenta el paso más complejo. A la hora de escribir el modelo, primero se ha de preparar una secuencia de entradas y salidas, explicadas en el marco teórico. Más

tarde, se simula una red neuronal de 2 capas, con el fin de obtener un resultado más fiel a la realidad, pero sin sobrecargar el modelo. La primera compuesta por 100 neuronas y la segunda por 50. A la red se le añade una capa Dropout para reducir el sobreajuste y una capa de salida, optimizada con error cuadrático medio:

```
def create_sequences(df, target_col="Price", seq_length=12):
    x, y = [], []
    for i in range(len(df) - seq_length):
        x.append(df.iloc[i:i+seq_length].values)
        y.append(df.iloc[i+seq_length][target_col])
    return np.array(x), np.array(y)

SEQ_LEN = 12
X_train, y_train = create_sequences(train)

modelo = Sequential()
modelo.add(LSTM(100, return_sequences=True, input_shape=(SEQ_LEN, X_train.shape[2])))
modelo.add(Dropout(0.2))
modelo.add(LSTM(50))
modelo.add(Dropout(0.2))
modelo.add(Dense(1))
modelo.compile(optimizer='adam', loss='mse')
modelo.fit(X_train, y_train, epochs=100, batch_size=8, verbose=0)

X_test, y_test = create_sequences(test)
prediccion_normalizada = modelo.predict(X_test)
```

Estructura del Código 19. Modelo LSTM.

Por último, será necesario volver a desnormalizar el precio predicho para conocer los valores reales a comparar:

```
normal_precio = MinMaxScaler()
normal_precio.min_, normal_precio.scale_ = normal.min_[0], normal.scale_[0]
prediccion_2024 = normal_precio.inverse_transform(prediccion_normalizada)
df_real_2024 = normal_precio.inverse_transform(y_test.reshape(-1, 1))
```

Estructura del Código 20. Desnormalización del Modelo LSTM.

3.3.2. Implementación de Indicadores Técnicos

Con el objetivo de ayudar a los modelos predictivos de base econométrica, como se ha mencionado anteriormente, existe una serie de parámetros operativos básicos que sirven como indicaciones de cómo actuar ante ciertas situaciones en el mercado bursátil. Para aplicarlas en un entorno de Google Colaboratory, se ha de importar la librería “.ta” (Technical Analysis), que sirve para calcular indicadores técnicos en el campo bursátil. Los códigos de cada indicador son los siguientes:

- **RSI (Relative Strength Index)**, con un estudio de los 14 días anteriores (window=14).

```
def compute_rsi(df, window=14):
    parametro = df['Close'].diff()
    ganancia = (parametro.where(parametro > 0, 0)).rolling(window=window).mean()
    perdidas = (-parametro.where(parametro < 0, 0)).rolling(window=window).mean()
    rs = gain / loss
    rsi = 100 - (100 / (1 + rs))
    return rsi

df['RSI'] = compute_rsi(df)
```

Estructura del Código 21. Indicador Relative Strength Index.

- **Bandas de Bollinger**, en este caso el window óptimo y más frecuente es de 20 días. Además, el parámetro K de desviación elegido es de 2:

```
rolling_mean = df['Close'].rolling(window=14).mean()
rolling_std = df['Close'].rolling(window=14).std()
df['Bollinger_Upper'] = rolling_mean + (rolling_std * 2)
df['Bollinger_Lower'] = rolling_mean - (rolling_std * 2)
```

Estructura del Código 22. Indicador de las Bandas de Bollinger.

- **MACD (Moving Average Convergence Divergence)**, siguiendo el fundamento teórico de este indicador técnico, en el desarrollo computacional se contemplará el cálculo de las medias exponenciales, una de señal rápida y otra lenta, definidas mediante el comando “span”:

```
exp1 = df['Close'].ewm(span=12, adjust=False).mean()
exp2 = df['Close'].ewm(span=26, adjust=False).mean()
df['MACD'] = exp1 - exp2
df['Signal_Line'] = df['MACD'].ewm(span=9, adjust=False).mean()
```

Estructura del Código 23. Indicador MACD.

- **VAR (Value at Risk)**, una vez estimado un retorno sobre la rentabilidad de la cartera gracias a los métodos de optimización de cartera, y de atribuir un valor subjetivo al nivel de confianza de dicho riesgo, es posible calcular el valor histórico del VAR mediante el siguiente líneas de código:

```
nivel_de_confianza = 0.95
percentil_VaR = np.percentile(retorno, (1 - nivel_de_confianza) * 100)
valor_VaR = VaR_percentil * df[-1]
```

Estructura del Código 24. Indicador Value at Risk.

3.3.3. Optimización de las Carteras.

El estudio metodológico sobre la teoría moderna que permite la optimización de carteras de inversión dará sentido al estudio previo de modelos predictivos, ya que, una vez visualizados los posibles resultados eficientes sobre el futuro de los instrumentos bursátiles analizados, este proceso permitirá conocer como invertir en ellos, teniendo en cuenta tanto rentabilidad esperada como la incertidumbre generada por la combinación de todos los activos añadidos en la cartera.

A diferencia del análisis seguido en los casos anteriores, para la optimización de carteras se presentará un estudio que recogerá datos desde el inicio de 2024 hasta el primer semestre de 2025, ya que será de mayor utilidad una información actualizada, con el fin de entender la situación actual en la que se encuentran los diferentes activos. A partir de este proceso, se aplicarán los dos modelos explicados en el marco teórico, dando como resultado el peso que necesita recibir cada título de inversión, la rentabilidad anual esperada y el riesgo al que se somete la cartera generada por el usuario.

Antes de adentrarse en los diferentes sistemas, es necesario mencionar que, en este caso, se necesitará descargar datos de todos aquellos activos que aparecerán en la cartera de inversión. Por ello, la forma de descargar la información asociada a estos se realizará de la siguiente manera:

```
tickers = ["^GSPC", "F", "AAPL", "GC=F"]  
df = yf.download(tickers, start="2024-01-01", end="2025-06-01")["Close"]  
retorno = data.pct_change().dropna()
```

Estructura del Código 25. Conjunto de Tickers descargados por Yahoo! Finance.

- **Método Markowitz**, una vez obtenido y analizado todos los datos pertinentes de esta exploración, se calcula las primeras variables iniciales de rentabilidad, volatilidad y la matriz de covarianza. Con estos parámetros se estima el Sharpe ratio explicado más adelante y evaluado a partir de pesos simulados en la cartera de inversión:

```
def portfolio_performance(pesos, media_de_retornos, matriz_covarianza):  
    retorno = np.dot(pesos, media_de_retornos)  
    std = np.sqrt(np.dot(pesos.T, np.dot(matriz_covarianza, pesos)))  
    return retorno, std  
  
def negative_sharpe_ratio(pesos, media_de_retornos, matriz_covarianza, tasa_libre_riesgo):  
    p_return, p_std = portfolio_performance(pesos, media_de_retornos, matriz_covarianza)  
    return -(p_return - tasa_libre_riesgo) / p_std
```

Estructura del Código 26. Método Markowitz.

A continuación, se define una tasa libre de riesgo diaria, habitualmente estimada en 0,02 anual debido a valores históricamente conservadores y con rendimientos realistas, la cual es dividida entre los 252 hábiles que tiene la bolsa en un año. El objetivo es formular una primera solución inicial a partir de los pesos anteriores, a los que se les establece unos límites y unas restricciones, donde cada peso tan solo podrá valer de 0 a 1 y la suma de todos no podrán ser mayores a la unidad:

```
tasa_libre_riesgo = 0.02 / 252 # tasa libre de riesgo diaria  
init_guess = cantidad * [1. / cantidad]  
bounds = tuple((0, 1) for _ in range(cantidad))  
constraints = ({'type': 'eq', 'fun': lambda x: np.sum(x) - 1})
```

Estructura del Código 27. Implementación de la Variables para Método Markowitz.

Seguidamente, se procesa un método iterativo de optimización con el objetivo de encontrar el mejor valor de Sharpe Ratio mediante el método de Programación Secuencial por Mínimos Cuadrados (SLSQP), algoritmo no lineal que permite la maximización o minimización de una función objetivo con restricciones, lineales o no lineales. A partir de

la obtención optimizada del Sharpe ratio, se podrá obtener tanto la rentabilidad, el riesgo y los pesos más eficientes:

```
sharpe_result = minimize(negative_sharpe_ratio, init_guess, args=(media_de_retornos, matriz_covarianza, tasa_libre_riesgo),  
                        method='SLSQP', bounds=bounds, constraints=constraints)  
opt_pesos = sharpe_result.x  
opt_retorno, opt_riesgo = portfolio_performance(opt_pesos, media_de_retornos, matriz_covarianza)  
sharpe_ratio = (opt_retorno - tasa_libre_riesgo) / opt_riesgo
```

Estructura del Código 28. Resultado Diario Markowitz.

Por último, se anualizan los valores calculados anteriormente con el fin de obtener resultados basados en un intervalo de los 252 días hábiles de la bolsa:

```
retorno_anual = opt_retorno * 252  
riesgo_anual = opt_riesgo * np.sqrt(252)  
riesgolibre_anual = tasa_libre_riesgo * 252  
sharpe_ratio_anual = (retorno_anual - riesgolibre_anual) / riesgo_anual
```

Estructura del Código 29. Resultado Anual Markowitz.

- **Método Black-Litterman**, aunque parezca que este código se pueda complicar debido a la aparición de matrices en su formulación, realmente los pasos a seguir son sencillos. En primer lugar, es esencial el cálculo de la matriz de equilibrio del mercado, representada en la Ecuación 55 (*Página 48*), con el símbolo Π . Para ello, será necesario la matriz de covarianza y el riesgo de aversión λ , parámetro subjetivo que refleja la importancia del inversor al riesgo frente al retorno. El fundamento mencionado se puede obtener de la siguiente manera:

$$\lambda = \frac{(R_M - R_F)}{w^T \Sigma w n}$$

Ecuación 60. Riesgo de Aversión.

Como se puede observar, es realmente una equivalencia a una de las partes que aparecen en la Ecuación 56 (*Página 48*), que se refiere al peso de los retornos esperados del mercado R_M y los retornos de activos libres de riesgo R_F .

El llamamiento de “risk.models” aparece mediante la implementación de la librería “pypfopt”, la cual facilita la optimización de carteras de inversión mediante la implicación de rendimientos esperados y matrices de riesgo y covarianza. Además, el riesgo de aversión establecido es de 2,5, ya que es un valor que refleja una aversión ciertamente moderada al riesgo, ni muy conservadora ni muy arriesgada.

```
matriz_covarianza = risk_models.sample_cov(df)  
capital_mercado = pd.Series([1/len(tickers)] * len(tickers), index=tickers)  
riesgo_aversion = 2.5  
rendimientos_mercado = market_implied_prior_returns(capital_mercado, riesgo_aversion, matriz_covarianza)
```

Estructura del Código 30. Variables Modelo Black-Litterman.

A continuación, se presenta la parte subjetiva del modelo, donde el individuo ha de seleccionar los valores relativos y absolutos sobre los retornos esperados, guiándose en su experiencia o cálculos de índole histórico. En la Estructura del Código 31, aparece un

ejemplo simulado, donde la agrupación de vectores “relaciones” se refiere a la relación que existe entre los diferentes instrumentos activos analizados en la cartera de inversión.

Por un lado, en la relación relativa, se asigna con un 1 al activo sobre el que se espera un mayor rendimiento y un -1 a aquel sobre el que recae la comparación.

Por otra parte, en cuanto a los vectores de retorno absoluto, tan solo se verá reflejada la unidad en aquellos activos que intervengan.

Además, la agrupación de vectores “valores”, determinará la magnitud de retorno asignada a cada relación, en el caso de las opiniones relativas, el valor reflejará la diferencia esperada de retornos entre los activos comparados. En el caso de las absolutas, el valor será la magnitud de retorno esperado de ese mismo instrumento bursátil.

```
relaciones = np.array([
    [0, -1, 0, 1],
    [0, 0, 1, 0],
])

valores = np.array([
    0.09,
    0.05,
])

# Activo D tendrá un retorno esperado de 0.09 sobre el activo B
# Activo C tendrá un retorno esperado de 0.05
```

Estructura del Código 31. Variables Subjetivas Método Black-Litterman.

Más tarde, se aplicará la llamada a la función que se encargue de calcular el riesgo y la rentabilidad esperada de una forma ajustada, pero no optimizada. En esta parte del código se obtendrá todos los parámetros que aparecen en la Ecuación 56 (Página 48). Cabe destacar que el término de grado de incertidumbre “tau” tiene un valor de 0,05 por ser un valor común en el campo financiero. Sin embargo, este es totalmente optativo. Para introducir la llamada al modelo BLM es necesario importar la librería “pypfopt.black_litterman” que permite optimizar procesos de cartera de inversión enfocados al modelo Black-Litterman.

```
bl = BlackLittermanModel(
    cov_matrix=matriz_covarianza,
    pi=rendimientos_mercado,
    P=relaciones,
    Q=valores,
    tau=0.05
)

bl_returns = bl.bl_returns()
```

Estructura del Código 32. Modelo Black-Litterman.

Por último, se obtiene los valores más eficientes de retorno y riesgo esperado, y los pesos sujetos a cada activo mediante un proceso iterativo de optimización.

```
eficiencia = EfficientFrontier(bl_returns, matriz_covarianza)
pesos = eficiencia.max_sharpe()
opt_pesos = eficiencia.clean_weights()
eficiencia.portfolio_performance(verbose=True)

print("Ponderaciones óptimas de la Cartera de Inversión BLACK-LITTERMAN:")
for ticker, peso in opt_pesos.items():
    print(f"{ticker}: {peso:.2%}")
```

Estructura del Código 33. Resultados Óptimos Black-Litterman.

3.4. Criterios de Evaluación.

3.4.1. Evaluación Predictiva.

Para poder decidir si un método es ciertamente preciso comparado con la evolución real del mercado, es necesario un estudio que cuantifique como de fiel es dicha proyección estimada.

En este punto aparecen diferentes medidores de error basados en comparaciones entre los valores reales y los pronosticados. Ofreciendo una visión completa del rendimiento del modelo, lo cual es indispensable a la hora de la toma de decisiones. Los más comunes y, por tanto, los que se van a exponer, son los denominados como Error Absoluto Medio (MAE) y Error Cuadrático Medio (RMSE).

3.4.1.1. MAE (Mean Absolute Error).

Es un proceso técnico que se fundamenta en la desviación absoluta entre todos los valores promediados \hat{v}_i y los valores reales v_i , sin tener en cuenta la dirección del error ya sea positivo o negativo. Ciertamente, es una técnica muy robusta para la detección de valores atípicos, ya que el impacto de un error grande tendrá la misma importancia que el de uno pequeño. La ecuación asociada al valor del indicador MAE es la siguiente:

$$MAE = \frac{1}{n} \sum_{i=1}^n |v_i - \hat{v}_i|$$

Ecuación 61. Error Medio Absoluto.

Siendo n el número de observaciones estudiadas.

A la hora de implementarlo en un entorno Python, el código se vería de esta manera:

```
mae = mean_absolute_error(df_real_2024['Close'], pronostico_2024)
```

Estructura del Código 34. Cálculo de MAE.

Importado directamente de la librería “sklearn.metrics”.

3.4.1.2. RSME (Root Mean Squared Error).

A diferencia de otros medidores erráticos, el Error Cuadrático Medio (RMSE) se basa en la distancia euclidiana, calculando la desviación más directa (cuadrática) entre los valores promediados y los reales, como en el Teorema de Pitágoras. Por ello, este proceso penaliza más acentuadamente los grandes errores, proporcionándole consecuencias más graves. En comparación con el error cuadrático absoluto, este permite una observación más sensible. Su formulación es la que aparece en la Ecuación 64:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (v_i - \hat{v}_i)^2}$$

Ecuación 62. Error Cuadrático Medio.

El resultado operacional en un lenguaje computacional se obtendría con la siguiente escritura:

```
mse = mean_squared_error(df_real_2024['Close'], pronostico_2024)
rmse = np.sqrt(mse)
```

Estructura del Código 35. MAE y RMSE.

En primera estancia se calcularía la media de los errores al cuadrado y más tarde se aplicaría la raíz cuadrada con el fin de corregir la unidad representada, ya que el mse la expresaría al cuadrado, imposibilitando su interpretación.

3.4.2. Evaluación del Rendimiento de las Carteras.

A la hora de seleccionar el mejor método para encontrar una cartera de inversión que proponga valores reales y eficientes, es necesario un modelo evaluativo que permita conocer una comparativa de riesgo con respecto a la rentabilidad de cada combinación de títulos financieros.

Como defienden (Fabozzi et al. 2002), modelos como Sharpe Ratio y el Drawdown son unos de los más usados en el campo bursátil, ya que muestran de una manera muy representativa e intuitiva la exposición del inversor a una cierta volatilidad y el potencial de rendimiento de una estrategia de inversión.

3.4.2.1 Sharpe Ratio.

El Sharpe Ratio es uno de los primeros procesos críticos de las carteras de inversión, propuesto por William F. Sharpe en 1966, aparece con el principal objetivo de unificar el estudio de beneficio e incertidumbre al que se somete un individuo a la hora de crear un plan de inversión.

Indica cuánto rendimiento adicional obtiene un inversor por cada unidad de riesgo que asume. En su formulación aparecen términos como la rentabilidad media de la cartera R_C y aquella libre de riesgo R_F . Además, se presenta también el parámetro de la desviación estándar de los rendimientos σ_C , manifestando la volatilidad de la cartera.

$$\text{Sharpe Ratio} = \frac{R_C - R_F}{\sigma_C}$$

Ecuación 63. Sharpe Ratio.

A la hora de implementarlo de una manera computacional, dependiendo del modelo que se quiera usar se realizará de una manera o de otra.

Por una parte, en la teoría de Markowitz se estima en primera instancia una simulación poco ajustada y con signo negativo del Sharpe Ratio, mediante los fundamentos que intervienen en la ecuación de este proceso, es decir, la Ecuación 63. Después, mediante un proceso iterativo, se minimiza ese valor negativo del primer resultado del Sharpe Ratio, encontrando el valor más óptimo. Aunque parezca extraño que se busque minimizar la función, realmente es una especie de maximización de esta misma, ya que está disminuyendo un valor con signo negativo. Por último, se obtienen de una manera inversa los valores que intervienen en la ecuación del Sharpe Ratio, y con ellos se calcula de nuevo el Sharpe Ratio que servirá como criterio evaluativo.

```
def negative_sharpe_ratio(pesos, media_de_retornos, matriz_covarianza, tasa_libre_riesgo):
    p_return, p_std = portfolio_performance(pesos, media_de_retornos, matriz_covarianza)
    return -(p_return - tasa_libre_riesgo) / p_std
    tasa_libre_riesgo = 0.02 / 252 # tasa libre de riesgo diaria
    init_guess = cantidad * [1. / cantidad]
    bounds = tuple((0, 1) for _ in range(cantidad))
    constraints = ({'type': 'eq', 'fun': lambda x: np.sum(x) - 1})

    sharpe_result = minimize(negative_sharpe_ratio, init_guess, args=(media_de_retornos, matriz_covarianza, tasa_libre_riesgo),
                             method='SLSQP', bounds=bounds, constraints=constraints)
    opt_pesos = sharpe_result.x
    opt_retorno, opt_riesgo = portfolio_performance(opt_pesos, media_de_retornos, matriz_covarianza)
    sharpe_ratio = (opt_retorno - tasa_libre_riesgo) / opt_riesgo
```

Estructura del Código 36. Sharpe Ratio.

Por otro lado, en el caso del sistema Black-Litterman, su implementación será mucho más sencilla, ya que el cálculo de este ya viene implícito en la misma función de llamada al propio modelo, por lo que no habría que realizar ningún código adicional.

3.4.2.2. Drawdown.

El Drawdown es una evaluación financiera que consiste en darle una magnitud a la posible pérdida esperada de una inversión desde que alcanza su punto máximo hasta el siguiente pico mínimo que sufre antes de que remonte. Es decir, este método evaluativo permite representar la diferencia máxima relativa del valor de una estrategia de inversión.

Dentro de este campo, el instrumento más común es el llamado Maximum Drawdown, es un indicador que registra dentro de un periodo decidido por el inversor todas las caídas relativas que aparecen en el valor de la cartera, seleccionando aquella más pronunciada. Es realmente útil para identificar la pérdida efectiva que sufre la cartera de un inversor, ayudándole a realizar un análisis estratégico con el fin de evitar maximizar las pérdidas, a diferencia del Sharpe Ratio, que busca únicamente aumentar la rentabilidad ajustada al riesgo.

Su formulación está compuesta por el valor más alto de la inversión o *peak* V_{Max} y por el valor más bajo posterior al pico V_B , representándose de la siguiente manera:

$$\text{Maximum Drawdown} = \frac{V_B - V_{Max}}{V_{Max}}$$

Ecuación 64. Maximum Drawdown.

A la hora de plasmarlo en un entorno de programación, los pasos a seguir son el cálculo del rendimiento histórico de la cartera, mediante la multiplicación de los retornos diarios por los pesos de la cartera. Más tarde, con la estimación obtenida del rendimiento se consigue el valor de la cartera, donde se representa el precio acumulativo de esta a lo largo de un periodo de tiempo, cabe destacar que el primer término de esta ecuación representa la inversión inicial, por lo que es totalmente subjetivo. Por último, se obtendría el resultado mediante una ecuación homologa a la mostrada anteriormente, donde se utiliza el pico máximo acumulado y el valor de la cartera total.

```
rendimientos_cartera = retorno.dot(opt_pesos)
valor_cartera = 1*(1 + rendimientos_cartera).cumprod()
pico_maximo = valor_cartera.cummax()
drawdown = (valor_cartera - pico_maximo) / pico_maximo
max_drawdown = drawdown.min()
```

Estructura del Código 37. Maximum Drawdown.

Conviene resaltar que el resultado siempre será negativo, ya que se está calculando la máxima pérdida relativa de la estrategia de inversión creada por un individuo. Además, sus unidades serán en porcentajes, ya que muestra como la magnitud de la pérdida de inversión con respecto al valor de la cartera.

A continuación, se muestra un gráfico simulado muy intuitivo para ejemplificar como se visualizaría el término del Drawdown en una cartera, donde los valores esperados de esta se representan con una línea azul y los picos máximos con las líneas discontinuas. A partir de estos se podrá obtener cuantificado el Drawdown, representado como el área que aparece entre la diferencia del pico máximo y los valores inferiores que siguen a este punto.

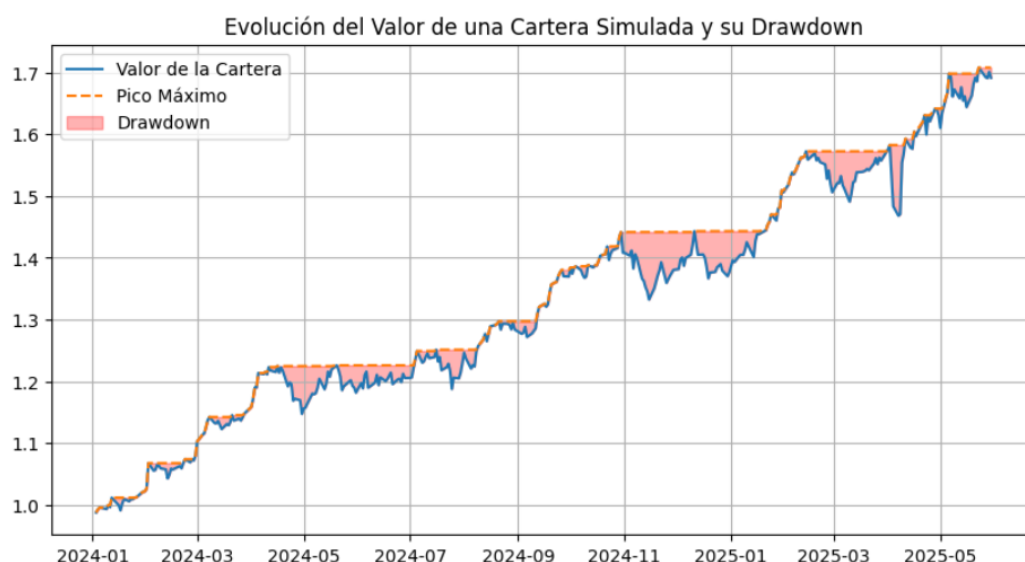


Figura 34. Ejemplo Simulado del Valor de una Cartera de Inversión y su Drawdown.

Fuente: De elaboración propia.

4. Análisis y Discusión de los Resultados.

4.1. Evaluación de Modelos de Predicción.

En lo referido a la discusión de los diferentes sistemas implementados, aunque todos plasmen un resultado ciertamente similar, es necesario resaltar uno sobre todos los demás, con el objetivo de mostrar que el modelo cumple el problema propuesto de la manera más fiel y precisa.

Para ello, el método a seguir se centrará en la elección aleatoria de 2 activos bursátiles, cada uno perteneciente a un tipo distinto según la clasificación definida en el marco teórico del proyecto, Todos ellos serán analizados con todos los modelos predictivos presentados. Más tarde, se obtendrá el Error Cuadrático Medio (RMSE) y el Error Medio Absoluto (MAE), se hará una media de todos aquellos y se ponderará un 50% a cada uno de los criterios evaluativos. De este modo, seleccionando aquel que menor magnitud tenga, se obtendrá un valor final que permitirá, de una forma sencilla, conocer que método es el más preciso y, por lo tanto, el más indicado para poder conocer el comportamiento futuro del mercado.

Los activos seleccionados son los visualizados en la Tabla 8:

MERCADOS	NOMBRES	SÍMBOLOS
MERCADO DE ACCIONES	Apple	AAPL
	Microsoft	MSFT
ÍNDICES BURSÁTILES	IBEX35	^IBEX
	S&P500	^GSPC
MERCADO DE TIPOS DE INTERÉS	Bono del Tesoro de EE.UU a 30 años	^TYX
	Gilts Británicos	INXG.L
FONDOS INDEXADOS	Myinvestor NASDAQ 100	0P0001MRGW.F
	Vanguard Global Stock INDEX	0P00011OE3
ETF	iShares Silver Trust	SLV
	Invesco S&P500 Equal Weight ETF	RSP
COMMODITIES	Oro	GC=F
	Petroleo Crudo	CL=F

Tabla 8. Mercados Seleccionados para la Evaluación de Modelos.

Fuente: De Elaboración Propia.

Siendo el RMSE y el MAE de cada activo por modelo los siguientes:

- **Holt-Winters:**

MODELO	SÍMBOLOS	RMSE	MAE
HOLT-WINTERS	AAPL	18.03	15.1
	MSFT	30.03	25.48
	^IBEX	861.01	753.68
	^GSPC	377.85	333.3
	^TYX	0.35	0.31
	INXG.L	0.66	0.55
	0P0001MRGW.F	0.21	0.19
	0P00011OE3	28.04	25.39
	SLV	5.35	4.68
	RSP	12.87	10.91
	GC=F	243.34	200.07
	CL=F	7.29	5.75
	MEDIAS	132.09	114.62
	RESULTADO	123.35	

Tabla 9. RMSE y MAE Obtenidos por Holt-Winters.

Fuente: De Elaboración Propia.

- **SARIMA:**

MODELO	SÍMBOLOS	RMSE	MAE
SARIMA	AAPL	41.99	51.21
	MSFT	51.84	47.94
	^IBEX	549.98	449.88
	^GSPC	1001.01	910.31
	^TYX	0.31	0.24
	INXG.L	1.28	1.06
	0P0001MRGW.F	0.36	0.33
	0P00011OE3	46.4	42.61
	SLV	5.5	4.79
	RSP	31.79	28.06
	GC=F	328.9	278.71
	CL=F	10.59	8.6
	MEDIAS	172.50	151.98
	RESULTADO	162.24	

Tabla 10. RMSE y MAE Obtenidos por SARIMA.

Fuente: De Elaboración Propia.

- ARIMAX:

MODELO	SÍMBOLOS	RMSE	MAE
ARIMAX	AAPL	31.16	26.51
	MSFT	41.36	38.29
	^IBEX	521.87	418.76
	^GSPC	829.33	731.82
	^TYX	0.52	0.48
	INXG.L	0.72	0.66
	0P0001MRGW.F	0.33	0.21
	0P0001IOE3	35.31	31.18
	SLV	4.91	4.3
	RSP	18.06	15.01
	GC=F	395.79	340.3
	CL=F	7.25	5.9
	MEDIAS	157.22	134.45
	RESULTADO	145.83	

Tabla 11. RMSE y MAE Obtenidos por ARIMAX.

Fuente: De Elaboración Propia.

- Random Forest:

MODELO	SÍMBOLOS	RMSE	MAE
RANDOM FOREST	AAPL	36.16	29.25
	MSFT	52.3	49.86
	^IBEX	1390.13	1246.325
	^GSPC	991.13	925.67
	^TYX	0.19	0.16
	INXG.L	0.25	0.21
	0P0001MRGW.F	0.13	0.1
	0P0001IOE3	38.66	35.57
	SLV	2.88	2.41
	RSP	21.65	19.74
	GC=F	423.56	366.48
	CL=F	4.63	3.92
	MEDIAS	246.81	223.31
	RESULTADO	235.06	

Tabla 12. RMSE y MAE Obtenidos por Random Forest.

Fuente: De Elaboración Propia.

- LSTM:

MODELO	SÍMBOLOS	RMSE	MAE
LSTM	AAPL	18.16	15.95
	MSFT	17.63	13.27
	^IBEX	434.72	386.27
	^GSPC	210.55	188.75
	^TYX	0.13	0.1
	INXG.L	0.22	0.18
	0P0001MRGW.F	0.03	0.03
	0P0001IOE3	11.55	10.6
	SLV	1.67	1.32
	RSP	12.03	11.06
	GC=F	130.49	121.35
	CL=F	5.04	4.28
	MEDIAS	70.19	62.76
	RESULTADO	66.47	

Tabla 13. RMSE y MAE Obtenidos por LSTM.

Fuente: De Elaboración Propia.

Tras obtener y analizar todos los cálculos pertinentes, es posible extraer una serie de conclusiones destacadas que son necesarias comentar.

Por un lado, cabe destacar que los modelos cuantitativos basados en el suavizamiento exponencial y aquellos fundamentados en la autorregresión con medias móviles sugieren una cierta tendencia de afinidad en el comportamiento futuro del mercado. Sin embargo, en dicha relación se exceptúa los activos de los mercados de acciones y los índices bursátiles, donde las diferencias son más destacables. Estas discrepancias permiten clarificar una mayor eficacia del modelo Holt-Winters sobre las extensiones ARIMA.

Por otro lado, es fácil visualizar que el modelo que mejores respuestas a ofrecido ha sido aquel que ejemplifica el funcionamiento de las Redes Neuronales Long-Short Term Memory (LSTM), dando pronósticos altamente precisos del futuro bursátil de los diferentes títulos financieros analizados. De hecho, obviando algunas irregularidades provocadas por agentes externos al propio mercado, la línea de valores que representa los precios de cierre pronosticados es realmente afín a la línea que muestra el valor real de dicho precio de cierre real, como se observa en la representación 35.



Figura 35. Representación del Modelo predictivo LSTM estudiado para 2024

Fuente: De Elaboración Propia.

A decir verdad, los pronósticos son tan similares que cabe una mínima posibilidad de que haya surgido un problema en la implementación del código relacionado con el término *Data Lackage*. Dicho concepto se utiliza cuando ha habido una fuga de datos durante el proceso de análisis y procesamiento de estos. Se trata de un error constructivo que introduce, de manera no intencionada, información útil para la estructuración de un futuro estimado en el entrenamiento de un modelo predictivo.

4.2. Evaluación de Métodos de Optimización de Carteras.

En lo que se refiere a las carteras de inversión, realmente no existe ningún criterio que decida que método es el correcto, pues los diferentes modelos de estrategia vienen diferenciados por un factor clave que no puede medirse en una magnitud errática, la subjetividad del inversor.

Es por ello por lo que los fundamentos evaluativos que se han explicado durante este análisis estratégico de carteras tan solo permiten conocer algunas características sobre que rendimiento tendrán los activos que las forman después de una total optimización.

Para saber si un portafolio financiero cumple con sus expectativas se ha decidido imponer unos valores que sirven como evaluación representativa del rendimiento y la eficacia de las diferentes posibles combinaciones de instrumentos bursátiles.

Por un lado, en lo referido al Sharpe Ratio, sus resultados se interpretarán de la siguiente manera:

- **< 0**, Mal rendimiento.
- **0 – 1**, Bajo rendimiento.
- **1 – 2**, Rendimiento aceptable.
- **2 – 3**, Cartera de inversión con un rendimiento muy bueno.
- **> 3**, Excelente rendimiento. No obstante, es improbable que se mantenga en un largo plazo.

Por otro lado, enfocándose en el Drawdown, las magnitudes porcentuales se entienden como se detalla a continuación:

- **0%**, La cartera no sufre caídas, es un caso altamente irreal.
- **≈ -5%**, El valor de la cartera sufre caídas muy poco pronunciadas y con mucha facilidad de recuperación.
- **≈ -20%**, Se estima un portafolio de inversión con una serie de pérdidas serias.
- **< -50%**, Valor de la cartera colapsado.

Para analizar las diferencias entre los diferentes métodos expuestos. A continuación, se presenta un ejemplo simulado que contempla los resultados obtenidos por la aplicación de la teoría de Markowitz y por la implementación del modelo de Black-Litterman, donde se ejemplifica la subjetividad del inversor de la siguiente manera:

- El retorno relativo esperado de Apple sobre el oro es aproximadamente de 2%.

- El retorno absoluto esperado del índice S&P500 es un 5% y el de los Bonos del Tesoro de Estados Unidos es del 3%.

MARKOWITZ	PESOS	
	APPLE	0.00%
	S&P500	20.17%
	ORO	22.59%
	Bono del Tesoro de EE.UU	57.24%
	SHARPE RATIO ANUAL	2.21
	MAXIMUM DRAWDOWN	-5.43%

Tabla 14. Pesos por Modelo Markowitz.

Fuente: De Elaboración Propia.

BLACK-LITTERMAN	PESOS	
	APPLE	17.93%
	S&P500	32.91%
	ORO	24.97%
	Bono del Tesoro de EE.UU	24.19%
	SHARPE RATIO ANUAL	0.39
	MAXIMUM DRAWDOWN	-11.42%

Tabla 15. Pesos por Modelo Black-Litterman.

Fuente: De Elaboración Propia.

A simple vista puede parecer que el ejemplo visualizado con el sistema del economista estadounidense Harry Markowitz da un beneficio mucho mayor al comparado con el propuesto por los financieros Robert Litterman y Fischer Black. No obstante, sería un error afirmar dicha suposición, ya que realmente el modelo BL lo que ofrece es una versión más compleja de las optimizaciones de carteras, entregándole una gran responsabilidad interpretativa al creador de dicha estrategia, permitiendo un resultado más real y afín a la visión del participante, aunque a veces esta percepción pueda ser errónea.

5. Implementación de la Interfaz Dinámica.

5.1. Lenguaje y Modelado.

Con el fin de lograr un entendimiento claro, sencillo y visual sobre el comportamiento bursátil, resulta fundamental implementar lo que se denomina interfaz dinámica. Es un tipo de entorno que permite representar de manera intuitiva e interactiva el comportamiento de diferentes modelos computacionales. En este caso, debido al mejor resultado obtenido en los análisis procesados, se realizará una implementación de interfaces dinámicas del método predictivo de las Redes Neuronales LSTM.

Para seguir estableciendo un objetivo centrado en un enfoque accesible y replicable, a la vez que útil, se utilizará como desarrollo Google Colaboratory. A pesar de que existen diferentes plataformas con una especialización y robustez para la visualización de datos, “Colab” sigue siendo un método práctico, gratuito y ampliamente disponible para todos los usuarios que pretendan adentrarse en el mundo computacional aplicado a las finanzas.

No obstante, aunque el campo de ejecución sea el mismo, la aplicación de este entorno requiere la instalación de diferentes librerías específicas, las cuales serán fundamentales para la ejecución y manipulación de información obtenida.

- **matplotlib.pyplot**, permite generar gráficos de una manera estática, interactiva o personalizada. Facilita de una manera considerable el análisis visual de los datos manipulados y procesados.
- **ipywidgets**, utilizado para crear controles interactivos, ya sean deslizadores, menús o botones. Es muy útil para explorar resultados o “inputs” sin la necesidad de reescribir el código.
- **IPython.display**, simplemente se presenta como una manera de estilizar las salidas de todos los análisis dinámicos que aparecen en la codificación, ya sean introducidos mediante **matplotlib.pyplot** o **ipywidgets**. Son los encargados de darle forma a los modelos interactivos.

Con respecto al modelado. Lo primero de todo se trata en hacer una llamada a las entradas que se desean implementar en este apartado dinámico, para ello se realiza lo siguiente:

```
ticker_input = widgets.Text(  
    value='',  
    description='Escribir Activo a Pronosticar:',  
    layout=widgets.Layout(width='400px'),  
    style={'description_width': 'initial'}  
)  
days_slider = widgets.IntSlider(  
    value=1,  
    min=1,  
    max=30,  
    step=1,  
    description='Días a predecir:',  
    layout=widgets.Layout(width='400px'),  
    style={'description_width': 'initial'}  
)  
run_button = widgets.Button(  
    description='Ejecutar',  
    layout=widgets.Layout(width='150px')  
)  
display(ticker_input, days_slider, run_button)
```

Estructura del Código 38. Objetos Dinámicos del Interfaz.

En este, se pide el título bursátil que se desea predecir y los días sobre los que se quiere hacer dicha estimación. También aparece la implementación de un botón que es el que permite acceder a los datos y al modelo LSTM del activo elegido por el usuario. Como se observa, cada entrada dispone de un tamaño o *layout* y un estilo o *style* que permiten refinar estos objetos dinámicos. Además, el deslizador que permite seleccionar el intervalo de pronóstico que ha de tener un valor mínimo y máximo. En este caso, se ha seleccionado un rango mensual, ya que permite una investigación futura a corto o medio plazo.

A continuación, se aplican los valores obtenidos por los objetos de entrada a las variables correspondientes, como se puede observar en el Estructura del Código 39. Más tarde, con la pertinente información se generará el modelo computacional asociado a las Redes Neuronales LSTM (como se muestra en el Estructura del Código 40 de la página 59), teniendo en cuenta que el modelo se entrenará con la información obtenida desde 2022 hasta 2025 y que ya no pronosticará el año de 2024, sino que realizará una previsión de los valores de cierre de las fechas futuras seleccionadas por el participante.

```
def run_forecast(b):  
    clear_output(wait=True)  
    display(ticker_input, days_slider, run_button)  
  
    TICKER = ticker_input.value.upper()  
    fechas_pronostico = days_slider.value  
    hoy = datetime.today()  
    ayer = pd.to_datetime((hoy - timedelta(days=1)).date())
```

Estructura del Código 41. Interpretación del Intervalo de Fechas a Pronosticar.

El modelo dará como resultado una predicción normalizada, que habrá que invertirla para obtener sus precios reales, y las fechas que se desean pronosticar, como se ve a continuación.

```
prediccion = np.array(prediccion)  
prediccion_real = scaler.inverse_transform(prediccion)  
fechas_pronostico = pd.date_range(start=df.index[-1] + timedelta(days=1), periods=fechas_pronostico, freq='B')
```

Estructura del Código 42. Predicción del Pronóstico para el Interfaz Dinámico.

Una vez obtenido todo lo anterior, se podrá empezar con la realización de las gráficas dinámicas. Primeramente, se establecen unos parámetros a nivel estéticos, donde se definen los colores de fondo del gráfico y aquellos correspondientes a las líneas de valores reales y las referidas a los valores futuros.

```
plt.style.use("dark_background") # Fondo negro  
color_real = '#FFFFFF' # Color Blanco atribuido a los valores reales  
# Gráfico completo estilizado  
fig, ax1 = plt.subplots(figsize=(14, 6))  
fig.patch.set_facecolor('#121212')  
# Elegir color del pronóstico según tendencia general  
if prediccion_real[-1, 0] - prediccion_real[0, 0] > 0.001:  
    color_pronostico = 'green'  
elif prediccion_real[-1, 0] - prediccion_real[0, 0] < -0.001:  
    color_pronostico = 'red'  
else:  
    color_pronostico = 'orange'
```

Estructura del Código 43. Estilización del Gráfico Dinámico.

Con el objetivo de mostrar de manera más visual la rentabilidad del activo futuro, se ha propuesto un sistema de codificación por colores según su tendencia. Asignando un color verde si la proyección indicada tiene una tendencia alcista, rojo si es bajista y naranja si es relativamente permanente.

Por último, se modelan los dos gráficos que intervienen en este entorno, uno referido a todos los valores desde 2020 hasta la actualidad y otro, a modo de *zoom*, enfocado en la representación de los valores de cierre de las últimas tres semanas de un activo y en su previsión futura. Este último, además contará con unas marcas día a día donde se muestren los precios y la fecha correspondiente.

```
# Gráfico de Valores Reales desde 2022 + Pronóstico
ax1.plot(df.index, df["Price"], label="Histórico", color=color_real, linewidth=2)
ax1.plot(fechas_pronostico, prediccion_real[:, 0], label="Pronóstico", color=color_pronostico, linestyle='--', linewidth=2)
ax1.set_title(f"{TICKER} - Valor de la Bolsa + Pronóstico Deseado.", fontsize=18, fontweight='bold', color='white', pad=20)
ax1.set_xlabel("Fecha", fontsize=12)
ax1.set_ylabel("Precio", fontsize=12)
ax1.legend(loc='upper left', fontsize=12, frameon=False)
ax1.xaxis.set_major_formatter(mdates.DateFormatter('%b %d'))
ax1.tick_params(axis='x', rotation=45)
ax1.tick_params(axis='both', labelsize=10)
plt.tight_layout()
plt.show()
```

Estructura del Código 44. Implementación del Gráfico Valor de Mercado + Pronóstico.

```
# Gráfico ZOOM : Últimas 3 semanas + Pronóstico.
zoom_df = df.iloc[-15:]
fig, ax2 = plt.subplots(figsize=(14, 5))
fig.patch.set_facecolor('#121212')
# Correctly indented plotting code for ax2
ax2.plot(zoom_df.index, zoom_df["Price"], label="Últimas 3 semanas", color=color_real, linewidth=2)
ax2.plot(fechas_pronostico, prediccion_real[:, 0], label="Pronóstico", color=color_pronostico, marker='x', linestyle='--', linewidth=2)
# Añadir texto con fecha y precio sobre los puntos del pronóstico
for date, price in zip(fechas_pronostico, prediccion_real[:, 0]):
    ax2.text(date, price, f"{price:.2f}\n{date.strftime('%d-%b')}",
             color='white', fontsize=6, ha='center', va='bottom', rotation=0)
ax2.set_title(f"{TICKER} - Últimas 3 semanas + Pronóstico Deseado.", fontsize=16, color='white', pad=15)
ax2.set_xlabel("Fecha", fontsize=12)
ax2.set_ylabel("Precio", fontsize=12)
ax2.legend(loc='upper left', fontsize=11, frameon=False)
ax2.xaxis.set_major_formatter(mdates.DateFormatter('%b %d'))
ax2.tick_params(axis='x', rotation=45)
ax2.tick_params(axis='both', labelsize=10)
plt.tight_layout()
plt.show()
```

Estructura del Código 45. Implementación del Gráfico "ZOOM".

5.2. Visualización.

Los objetos interactivos codificados en el apartado anterior se representarían de una forma visual de la siguiente manera:

Escribir Activo a Pronosticar (Formato Ticker):

Días a predecir: 1

Figura 36. Objetos Dinámicos en la Interfaz Implementada.

Fuente: DE Elaboración Propia.

En la imagen 36, se puede observar cómo hay una casilla para escribir el instrumento bursátil a analizar, los días que se quieren pronosticar y el botón de ejecución que daría como resultado las gráficas dinámicas asociadas al futuro de dicho activo.

Para entender el funcionamiento de este proceso se propone un ejemplo a partir de un título bursátil aleatorio. En este caso, Ford Motor Company, analizado en un futuro de 19 días.

Escribir Activo a Pronosticar (Formato Ticker):

Días a predecir: 19

Figura 37. Ejemplo Simulado de una Acción en los Objetos Dinámicos de la interfaz.

Fuente: De Elaboración Propia.

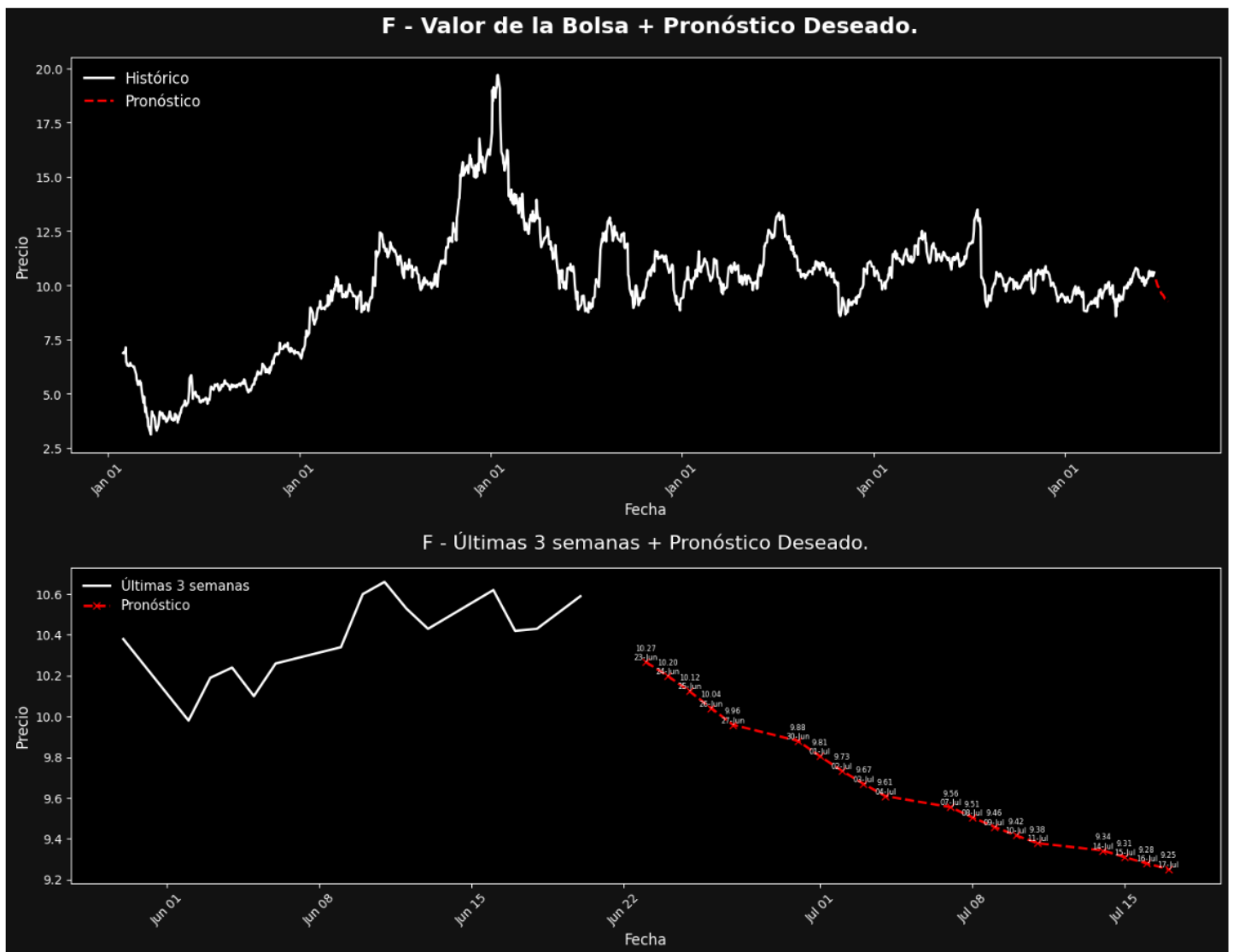


Figura 38. Gráficos de la Interfaz Dinámica Generada.

Fuente: De Elaboración Propia.

En este caso, al tratarse de un activo cuya proyección futura es potencialmente descendente con respecto al valor actual, la línea correspondiente al pronóstico se representa de un color rojo. Asimismo, se visualizan los valores estimados junto a las fechas correspondientes, facilitando de esta forma la lectura temporal de la evolución prevista.



7. Conclusión.

Este trabajo ha procurado en todo momento complementar el análisis estadístico mediante la combinación de la modelización de series temporales con la economía volátil, demostrando que existe una probabilidad de mejorar la interpretación de entornos dinámicos y variables. Los recursos basados en datos, algoritmos y decisiones han permitido ofrecer una claridad en medio de un mar lleno de incertidumbre.

Entre las aproximaciones definidas, se ha podido observar un claro sistema destacado sobre los demás, las Redes Neuronales Long-Short Term Memory (LSTM) que han mostrado una sorprendente capacidad de adaptación para aprender patrones temporales, ajustándose de una manera sobresaliente a los valores futuros de los diferentes activos bursátiles. Su estructura, basada en una máquina de aprendizaje automática con gestión eficiente de su memoria supera con creces al resto de técnicas detalladas. Esta tecnología, al igual que en otros terrenos, representa un paso adelante hacia la aplicación de la inteligencia artificial en nuestra sociedad, emergiendo como una potencial herramienta de ayuda en multitud de desafíos a los que se enfrenta el ser humano.

Por otro lado, la propuesta de cuantificar una eficiencia en las diferentes estrategias de carteras manifiesta una dualidad entre predicción y optimización que permite crear un plan de inversión donde un individuo es capaz de tomar decisiones a partir de pronósticos y pesos ajustados.

Sin embargo, más allá de lograr los objetivos técnicos que se proponen, lo que pretende este análisis de modelos bursátiles cuantitativos es facilitar al usuario, de una manera funcional, adentrarse en el mundo financiero. Por ello, se ha presentado al final de este trabajo una serie de implementaciones visuales y procesos simulados que pretenden confirmar que los datos son una fuente altamente eficaz en este campo, siempre que se sepan usar de una manera correcta.

8. Perspectivas Futuras.

La combinación de previsiones futuras y refinamientos de los pesos de los activos en los portafolios de inversión pueden tener una gran aplicación en la sociedad siempre que se asocien de una manera apropiada.

Una propuesta futura podría ser la posibilidad de mejorar el actual algoritmo con una sintaxis enfocada en la reducción del tiempo de compilación y en el desarrollo de procesamiento, ya que con el lenguaje implementado el número de redes neuronales suponen una alta carga de parámetros los cuales no pueden ser procesados por dispositivos con *hardware* limitados. De este modo serán posibles procedimientos mucho más complejos y ajustados a las posibles noticias externas que provocan ciertas irregularidades en el mercado.

A partir de la introducción de estas mejoras, también se puede proponer un ensamblaje de los modelos predictivos con los métodos de optimización de carteras mediante interfaces mucho más



complejas, creando un prototipo de aplicación o web que actúe como un simulador simple y educativo para aquellos usuarios que quieran adentrarse en el mundo financiero.

9. Referencias.

Atsalakis, G. S., & Valavanis, K. P. (2009). Surveying stock market forecasting techniques – Part II: Soft computing methods. *Expert Systems with Applications*, 36(3), 5932–5941.

Black, F., & Litterman, R. (1992). Global portfolio optimization. *Financial Analysts Journal*, 48(5), 28–43. Recuperado el 29 de mayo de 2025, de <https://doi.org/10.2469/faj.v48.n5.28>

Bollinger, J. (2001). *Bollinger on Bollinger bands*. McGraw-Hill.

Box, G. E. P., & Jenkins, G. M. (1970). *Time series analysis: Forecasting and control*. Holden-Day.

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2008). *Time series analysis: Forecasting and control* (4th ed.). John Wiley & Sons.

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. Recuperado el 23 de mayo de 2025, de <https://doi.org/10.1023/A:1010933404324>

Buchanan, B. G. (2005). A (very) brief history of artificial intelligence. *AI Magazine*, 26(4), 53–60. <https://doi.org/10.1609/aimag.v26i4.1848>

Campbell, J. Y., Lo, A. W., & MacKinlay, A. C. (1997). *The econometrics of financial markets*. Princeton University Press.

Chatfield, C. (2000). *Time-series forecasting*. Chapman & Hall/CRC.

Chatfield, C. (2004). *The analysis of time series: An introduction* (6th ed.). CRC Press.

Dalio, R. (2017). *Principles: Life and work*. Simon & Schuster.

El País. (2025, marzo 20). *La deuda global entra en zona de riesgo: crece, es más cara y hay presión para elevar la inversión*. Recuperado el 15 de mayo de 2025, de <https://elpais.com/economia/2025-03-20/la-deuda-global-entra-en-zona-de-riesgo-crece-es-mas-cara-y-hay-presion-para-elevar-la-inversion.html>

Equipo de Expertos en Empresa de la Universidad Internacional de Valencia. (2023). *Principios de economía y gestión empresarial*. Universidad Internacional de Valencia.

Fabozzi, F. J., & Markowitz, H. M. (2006). *The theory and practice of investment management*. John Wiley & Sons.

Fabozzi, F. J., Focardi, S. M., & Kolm, P. N. (2006). *Quantitative equity investing: Techniques and strategies*. Wiley.



- Fabozzi, F. J., Gupta, F., & Markowitz, H. M. (2002). The legacy of modern portfolio theory. *The Journal of Investing*, 11(3), 7–22. Recuperado el 6 de junio de 2025, de <https://doi.org/10.3905/joi.2002.319510>
- Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2), 383–417.
- Fama, E. F., & French, K. R. (1996). Multifactor explanations of asset pricing anomalies. *The Journal of Finance*, 51(1), 55–84.
- Fernández Rodríguez, J. M., Arguedas Calle, E., & De la Peña Velasco, C. (2013). *Renta fija: fundamentos y estrategias*. ESIC Editorial.
- Feron, S. (2023). *Fondos indexados y gestión pasiva para principiantes*.
- García Díaz, J. C. (2016). *Predicción en el dominio del tiempo: Análisis de series temporales para ingenieros*. Editorial de la Universidad Politécnica de Valencia.
- Gardner, E. S. (1985). Exponential smoothing: The state of the art. *Journal of Forecasting*, 4(1), 1–28. Recuperado el 20 de mayo de 2025, de <https://doi.org/10.1002/for.3980040103>
- Gastineau, G. L. (2001). *The exchange-traded funds manual*. John Wiley & Sons.
- Geman, H. (2005). *Commodities and commodity derivatives: Modelling and pricing for agriculturals, metals and energy*. John Wiley & Sons.
- Goetzmann, W. N. (2023). *El legado de la teoría moderna de carteras*. EBSCO.
- González Velasco, C., & del Puerto García, P. (2009). *Series temporales*. Universidad de Extremadura. Recuperado el 10 de mayo de 2025, de https://publicauex.unex.es/libro/series-temporales_135473/
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
- Géron, A. (2019). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems* (2nd ed.). O'Reilly Media.
- Gómez Pineda, J. G. (2010). *Mercados financieros: Instrumentos, estructura y funcionamiento*. Ecoe Ediciones.
- Hasan, M., & Horvat, T. (2024). *Machine learning applications in uncertain environments*. Springer.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780.
- Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3^a ed.). OTexts. <https://otexts.com/fpp3/>
- Jorion, P. (2007). *Value at Risk: The New Benchmark for Managing Financial Risk* (3rd ed.). McGraw-Hill.



Kaufman, P. J. (2005). *Alpha trading: Profitable strategies that remove directional risk*. John Wiley & Sons.

Keynes, J. M. (1936). *The general theory of employment, interest and money*. Macmillan.

Lo, A. W., & MacKinlay, A. C. (1999). *A non-random walk down Wall Street*. Princeton University Press.

Malkiel, B. G. (2003). The efficient market hypothesis and its critics. *Journal of Economic Perspectives*, 17(1), 59–82.

Mochón, F. (2015). *Fundamentos de economía*. McGraw-Hill Education.

Moreno Osorio, L. G. (2011). *Teoría de la decisión*. Editorial Universidad Nacional de Colombia.

Murphy, J. J. (1999). *Technical analysis of the financial markets: A comprehensive guide to trading methods and applications*. New York Institute of Finance.

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown Publishing Group.

Pindyck, R. S., & Rubinfeld, D. L. (1998). *Modelos econométricos y pronósticos económicos* (4.^a ed.). Pearson Educación.

Satchell, S., & Scowcroft, A. (2000). A demystification of the Black–Litterman model: Managing quantitative and traditional portfolio views. *Journal of Asset Management*, 1(2), 138–150. Recuperado el 6 de junio de 2025, de <https://doi.org/10.1057/palgrave.jam.2240007>

Taulli, T. (2011). *All about commodities*. McGraw-Hill Education.

Timmermann, A., & Granger, C. W. J. (2004). Efficient market hypothesis and forecasting. *International Journal of Forecasting*, 20(1), 15–27.

Wang, J., Zhao, X., & Liu, Y. (2020). *Artificial intelligence: Methods and applications*. Springer.



Anexo I. Presupuesto.

1. Introducción.

El objetivo principal de este documento es presentar un presupuesto del coste de elaboración del presente proyecto sobre la creación de un software web profesional a partir del estudio pertinente de los modelos analizados. Para ello, se presentarán unos precios desglosados según su función. Por último, se expondrá un precio final teniendo en cuenta diferentes aspectos como gastos generales, beneficios e impuestos aplicados.

2. Presupuesto Parcial.

2.1 Análisis y Diseño.

Recoge todas las definiciones de requerimientos, la convocatoria de reuniones, objetivos y restricciones del proyecto y los detalles deseados por el cliente.

<i>Recurso</i>	<i>Perfil Profesional</i>	<i>Horas</i>	<i>Precio/Hora (€)</i>	<i>Coste (€)</i>
Análisis del Proyecto	Ingeniero Principal	80	50	€ 4.000,00
Costes Indirectos (1,00%)				€ 40,00
Total Análisis y Diseño				€ 4.040,00

2.2 Instalación y Desarrollo.

Engloba aquellas tareas dedicadas a la implementación de modelos y construcción de códigos para la ejecución e instalación del programa. Presenta los principales perfiles que actuarán en el proyecto de desarrollo para la construcción del software.

<i>Recurso</i>	<i>Perfil Profesional</i>	<i>Horas</i>	<i>Precio/Hora (€)</i>	<i>Coste (€)</i>
Desarrollo de API's	Ingeniero de Software	120	50	€ 6.000,00
Montaje de la Parte Gráfica	Ingeniero de Software	80	50	€ 4.000,00
Desarrollo de Modelos	Data Scientist	150	60	€ 9.000,00
Montaje técnico en la Nube	Ingeniero de Cloud Computing	40	55	€ 2.200,00
Costes Indirectos (1,00%)				€ 212,00
Total Instalación y Desarrollo				€ 21.412,00

2.3 Validación, Calidad y Mantenimiento.

Dicho apartado recoge todo aquello relacionado con la comprobación del estado del software, ya sea la calidad de este como la operatividad y la funcionalidad.

<i>Recurso</i>	<i>Perfil Profesional</i>	<i>Horas</i>	<i>Precio/Hora (€)</i>	<i>Coste (€)</i>
Training and Test	Analista de Calidad	40	35	€ 1.400,00
Revisión Técnica Externa	Consultor Externo	10	100	€ 1.000,00
Actualizaciones y Soporte Técnico a 6 Meses Vista	Equipo Técnico	80	45	€ 3.600,00
Costes Indirectos (1,00%)				€ 60,00
Total Validación, Calidad y Mantenimiento.				€ 6.060,00

2.4 Licencias de Software y Equipo.

Este bloque indica todos los gastos relacionados con los diferentes softwares centrados en librerías de API's privadas y softwares más profesionalizados, ya sea de entorno para el lenguaje o para la visualización dinámica del modelo.



Licencias de Softwares y API's externas	€ 1.200,00
Equipo Informático	€ 2.000,00
Costes Indirectos (1,00%)	€ 32,00
Total Licencias y Equipo	€ 3.232,00

2.5 Documentación Técnica

La documentación se refiere a todo aquello relacionado con manuales técnicos que permiten un mayor entendimiento del funcionamiento del software. Además, puede recoger otro tipo de documentación como modificaciones del sistema o del propio software.

Recurso	Perfil Profesional	Horas	Precio/Hora (€)	Coste (€)
Documentación y Manual Técnico	Ingeniero de Datos	20	45	€ 900,00
Costes Indirectos (1,00%)				€ 9,00
Total Documentación Técnica				€ 909,00



3. Presupuesto Total.

El presupuesto total desarrollará la suma de todos los aspectos parciales mencionados anteriormente más la suma de los Gastos Generales y los Beneficios Industriales, que se refieren respectivamente a aquellos costes a los que se somete la empresa para llevar a cabo el proyecto y el margen de ganancia que se aplica para asegurar que dicho trabajo sea rentable. Además, una vez obtenido el Presupuesto de Ejecución por Contrata, se le añadirá el Impuesto de Valor Añadido (IVA) para poder prestar de manera oficial un servicio profesional.

Presupuesto Total

1. Análisis y Diseño.	€ 4.040,00
2. Instalación y Desarrollo.	€ 21.412,00
3. Validación, Calidad y Mantenimiento.	€ 6.060,00
4. Licencias de Software y Equipo.	€ 3.232,00
5. Documentación Técnica.	€ 909,00
Gastos Generales (13%)	€ 4.634,89
Beneficio Industrial (6%)	€ 2.139,18
Presupuesto de Ejecución por Contrata (PEC = Costes + Gastos Gen. + Beneficio Industrial)	€ 42.427,07
IVA 21%	€ 8.909,68
Presupuesto Base de Ilitación (PBL = PEC +IVA)	€ 51.336,75

4. Conclusión.

Este presupuesto tan solo es una estimación orientada a un desarrollo futuro de un software más profesionalizado de los modelos y técnicas presentadas. Cabe destacar que no corresponde a la dificultad presentada en el trabajo actual, cuyo objetivo, como se ha dicho anteriormente, es introducir a los interesados en la bolsa en métodos de inversión cuantificables que permitan una mayor facilidad, sirviendo como base conceptual para posibles aplicaciones más avanzadas.

Anexo II. Objetivos de Desarrollo Sostenible.

Atendiendo a la tabla adjuntada sobre los diferentes objetivos de Desarrollo Sostenible de la Agenda 2030, es posible encontrar algunos niveles de vinculación entre el contenido desarrollado en el presente trabajo y los ODS.

Grado de relación del trabajo con los Objetivos de Desarrollo Sostenible (ODS).

Objetivos de Desarrollo Sostenibles	Alto	Medio	Bajo	No Procede
ODS 1. Fin de la pobreza.				X
ODS 2. Hambre cero.				X
ODS 3. Salud y bienestar.				X
ODS 4. Educación de calidad.		X		
ODS 5. Igualdad de género.				X
ODS 6. Agua limpia y saneamiento.				X
ODS 7. Energía asequible y no contaminante.				X
ODS 8. Trabajo decente y crecimiento económico.	X			
ODS 9. Industria, innovación e infraestructuras.			X	
ODS 10. Reducción de las desigualdades.				X
ODS 11. Ciudades y comunidades sostenibles.				X
ODS 12. Producción y consumo responsables.			X	
ODS 13. Acción por el clima.				X
ODS 14. Vida submarina.				X
ODS 15. Vida de ecosistemas terrestres.				X
ODS 16. Paz, justicia e instituciones sólidas.				X
ODS 17. Alianzas para lograr objetivos.			X	

Aunque no aparezca una relación directa, a lo largo del proyecto se abordan diferentes temas que, de manera implícita, contribuyen a diferentes características de los objetivos.

1. ODS 4 – EDUCACIÓN DE CALIDAD.

Este trabajo tiene un enfoque tanto académico como divulgativo, acercando a los estudiantes e inversores no expertos a modelos de inversión cuantitativa y herramientas de predicción de base econométrica, promoviendo el acceso a un conocimiento técnico.

2. ODS 8 – TRABAJO DECENTE Y CRECIMIENTO ECONÓMICO.

El proyecto sugiere herramientas que apoyan al análisis financiero enfocado en un rendimiento responsable y en un crecimiento económico, facilitando decisiones de inversión más informadas, eficientes y sostenibles.

3. ODS 9 – INDUSTRIA, INNOVACIÓN E INFRAESTRUCTURAS.

Se incluyen desarrollos basados en los fundamentos de la Inteligencia Artificial y modelos econométricos, promoviendo una gran innovación tecnológica del sector financiero.



4. ODS 12 – PRODUCCIÓN Y CONSUMO RESPONSABLE.

El proyecto contribuye a una forma de producción y consumo de inversión más sostenible y consciente, ya que se incentiva la toma de decisiones de una manera más racional y cuantificada.

5. ODS 17 – ALIANZAS PARA LOGRAR LOS OBJETIVOS.

Gracias a la interpretación basada en futuras colaboraciones con inversores o empresas de ámbito financiero, se promueven ciertas alianzas orientadas al desarrollo económico y educativo.

En conclusión, este trabajo ofrece un desarrollo de soluciones tecnológicas aplicadas al ámbito financiero que promueven la innovación, la optimización y el acceso al conocimiento y aprendizaje. Fomenta una inversión totalmente razonada e informada, basada en datos, impulsando un modelo económico más sostenible e inclusivo. Además, exponer estos fundamentos de manera pública favorece a la educación, la concienciación y las colaboraciones con diferentes agentes que proponen un objetivo académico, social y profesional.