MÁSTER EN INGENIERÍA DE TELECOMUNICACIONES

TRABAJO FIN DE MÁSTER

NUEVOS ENFOQUES EN LA GESTIÓN BÁSICA DE CARTERAS DE VALORES BASADOS EN EL RANKING Y EL RIESGO. EL USO DE DATOS DE INTERVALOS

Autor: Teresa Domingo Górriz

Director: Carlos Maté Jiménez

Madrid

NUEVOS ENFOQUES EN LA GESTIÓN BÁSICA DE CARTERAS DE VALORES BASADOS EN EL RANKING Y EL RIESGO. EL USO DE DATOS DE INTERVALOS

Autor: Domingo Górriz, Teresa
Director: Maté Jiménez, Carlos
Entidad Colaboradora: ICAI – Universidad Pontificia Comillas

RESUMEN DEL PROYECTO

En los últimos años se ha dado un incremento exponencial en la cantidad de datos recogidos, tratados y analizados gracias a los avances de la tecnología. Con ello, el análisis de datos puntuales es habitual en muchos ámbitos, sin embargo, el análisis de intervalos (*Interval Analysis,* IA) no se ha estudiado con tanta profundidad y puede ser de gran ayuda para representar la incertidumbre de una forma más precisa. Gracias al uso de estos datos es posible obtener una comprensión más completa de la información disponible, lo cual permite tomar decisiones y evaluar riesgos de una forma mejorada.

Uno de los ámbitos en los que el uso de IA puede ser de gran utilidad es en el mundo de los mercados financieros y es por esta razón por la que este proyecto se va a enfocar a ellos. La combinación de IA con modelos de optimización y predicción puede ofrecer un punto de vista interesante en la gestión, optimización y formación de carteras. Por ello, uno de los objetivos que se persiguen en el curso de este proyecto es el desarrollo de un sistema original, de utilización sencilla y visual que sirva de herramienta a los inversores para realizar el análisis del comportamiento de los precios de las acciones en el pasado, presente y futuro, para poder formar su cartera óptima de valores basada en el rendimiento y riesgo que deseen obtener. Así, se trabajará con información de diversos mercados financieros como el IBEX35 o el S&P100, en diferentes periodos para lograr una serie de resultados y obtener de ellos las conclusiones pertinentes.

Palabras clave: Análisis de Intervalos, Optimización de Carteras, Mercados Financieros, Riesgo, Retorno, Ranking, Métodos de Predicción, Matlab.

1. Introducción

Los mercados financieros constituyen un elemento imprescindible para las economías de los diferentes países, así como una ayuda para el desarrollo y potenciación de su salud económica.

A nivel particular, son muchas las personas para las cuales estos mercados representan una forma de ahorro o negocio, buscando obtener cierta rentabilidad de su dinero a partir de la compra de acciones, bonos u otros valores con los que forman una cartera.

La optimización de carteras consiste en la generación de diferentes carteras de valores con el objetivo de elegir aquella que mejor satisfaga los objetivos de retorno y riesgo de cada inversor. A la hora de crear estas carteras existen diferentes herramientas que se pueden utilizar. Desde la primera Teoría Moderna de Carteras (*Modern Portfolio Theory,* MPT) (Kolm, Tütüncü, & Fabozzi, 2014) se ha continuado investigando sobre este tema y se han desarrollado otras herramientas para ello (Cosín López-Medel, 2019).

Por otro lado, los datos de intervalo han sido subestimados durante mucho tiempo en comparación con los datos puntuales o clásicos. Sin embargo, mientras que los datos clásicos pueden presentar desafíos en términos de precisión, ya que no capturan toda la información inherente a variables complejas, los datos de intervalo ofrecen una perspectiva más amplia y precisa, lo cual resulta valioso en la toma de decisiones y la comprensión precisa de la información.

A lo largo de este proyecto se estudiarán diversos métodos de análisis de datos históricos, así como modelos de predicción de series temporales con datos de intervalo (Maté C. , 2022) de cara a enfocar estos análisis a la optimización de carteras. Para ello, se estudiarán registros pasados de diferentes mercados financieros, como el IBEX35 o el S&P100, con el propósito de desarrollar una herramienta que brinde asistencia al inversor en la toma de decisiones.

2. Definición del proyecto

El proyecto va a constar de dos partes principales. La primera de ellas consistirá en el estudio de la documentación relativa a los conceptos más importantes de la gestión y optimización de carteras, rendimiento y riesgo, métodos de predicción y análisis de datos de intervalo. Todo ello será recogido en los primeros capítulos y servirá de base para el entendimiento y desarrollo posterior de la herramienta.

La segunda rama del proyecto consiste en la creación de la herramienta. Es un hecho conocido que el mundo de los mercados financieros es un ámbito impredecible ya que estos son afectados por un gran número de factores. Por ello, se busca desarrollar una herramienta que aúne el estudio de algunos de estos factores y métodos estadísticos para intentar estimar futuros rendimientos de acciones y así formar una cartera optimizada para cada inversor.

Así, el proyecto buscará facilitar a los inversores la toma de decisiones de inversión ya que no solo les va a ofrecer la posibilidad de analizar los comportamientos pasados de diferentes compañías operando en un índice bursátil y compararlas entre ellas, sino que también les ofrecerá diversos métodos de creación de carteras a partir del uso de modelos de predicción. Todo ello será ofrecido al usuario por medio de diferentes gráficas y tablas que le ayudarán a comprender los resultados de una manera intuitiva y visual, aportando información de gran valor.

De cara a desarrollar la herramienta se hará uso de la funcionalidad de App Designer de Matlab, un entorno de desarrollo interactivo que permite diseñar una aplicación y programar su comportamiento (Matlab, s.f.). Así, el programa desarrollado será aplicado a valores reales de datos obtenidos de distintos índices bursátiles para obtener un modelo real y más tarde, poder comprobar la precisión y efectividad del mismo.

3. Descripción de la herramienta

En la elaboración de la herramienta se parte de la base de los conceptos teóricos estudiados en profundidad previamente y expuestos a lo largo del proyecto. La herramienta contará con dos funciones principales de análisis de valores que difieren en su situación temporal.

La primera de ellas trata del análisis de valores históricos, es decir, de precios del pasado. Dentro de esta opción el usuario podrá escoger entre observar la información de una

empresa o de dos, en caso de que esté interesado en comparar el comportamiento de dos compañías concretas. A continuación, el usuario deberá escoger el tipo de análisis que desee obtener y así el programa elaborará un gráfico en función de una serie de parámetros de entrada que el usuario habrá seleccionado previamente.

La segunda de las funcionalidades que ofrece la herramienta es la predicción de valores futuros a partir del análisis de datos pasados. Para ello, la herramienta escoge una serie de valores históricos de la compañía o compañías de interés para el usuario y según el método de predicción escogido por él mismo, elaborará una estimación del valor futuro de las acciones. Así, será capaz de generar un ranking que las ordene en función de su retorno esperado para que el usuario pueda tener una idea de los resultados que se esperan de cada compañía y así formar su propia cartera de acuerdo a sus niveles de retorno y riesgo deseados.

Además, cabe destacar que el usuario tendrá libertad para decidir desde la herramienta si los datos que desea estudiar son semanales, mensuales o trimestrales.

A continuación, se muestra un esquema de las funcionalidades que ofrece la herramienta para facilitar su comprensión.

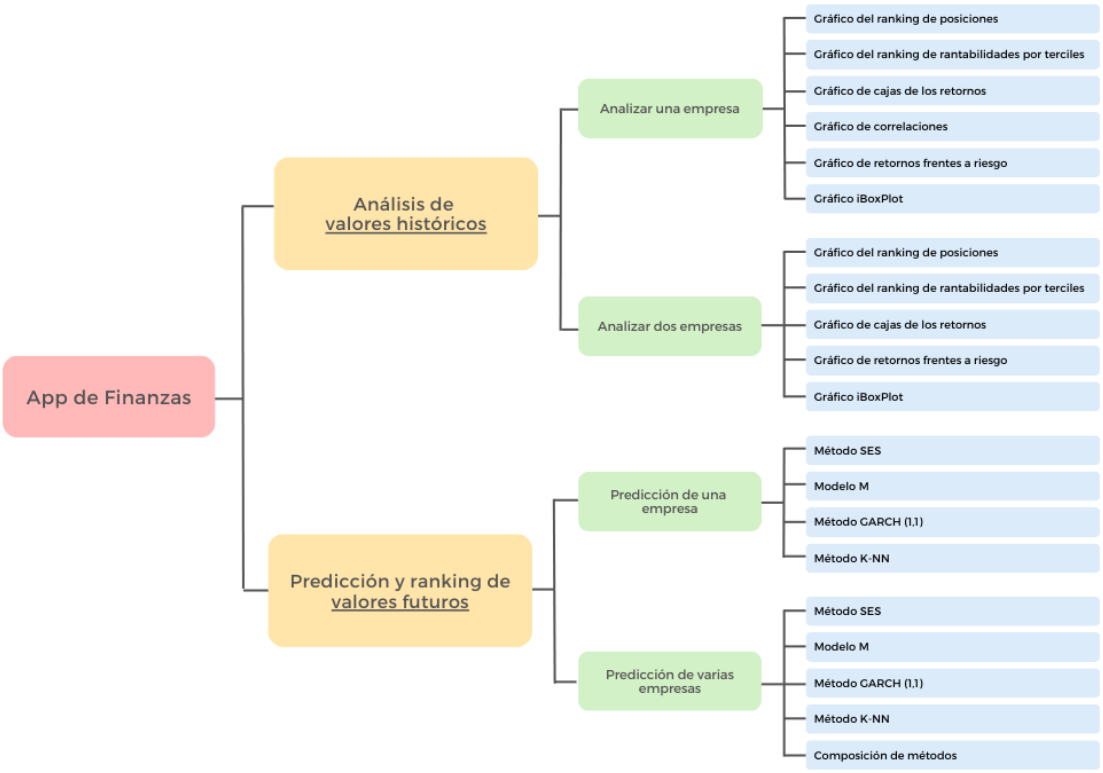


Ilustración 1. Esquema de funcionalidades de la herramienta (Elaboración propia)

Cabe destacar que antes de poder hacer uso de la herramienta es necesario realizar un análisis exhaustivo de los datos para lograr la recolección de los mismos, filtrando aquellos que son de utilidad y desechando los que no lo son. Para ello se hace uso de herramientas como Bloomberg o Yahoo Finance desde donde se pueden descargar datos financieros históricos por días, semanas, meses, trimestres o años. Una vez recopilados, los datos son introducidos en la herramienta para poder comenzar a trabajar. El aspecto de la herramienta al iniciarla será el siguiente.

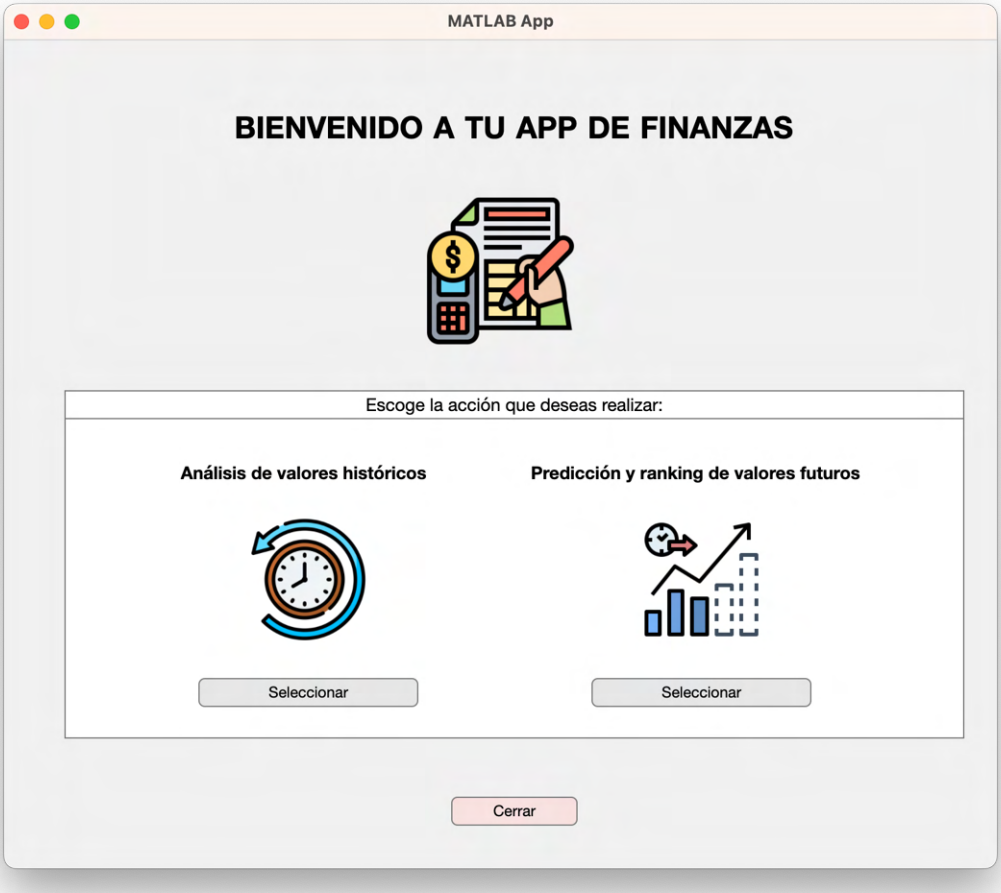


Ilustración 2. Ventana de inicio de la herramienta (Elaboración propia)

Finalmente, se pondrá a prueba la herramienta para probar su funcionamiento, observar los resultados que devuelve y desarrollar las conclusiones pertinentes.

4. Resultados

Se ha analizado la herramienta a partir de una serie de valores pertenecientes al índice bursátil S&P100. Así, con la primera de las funcionalidades de la herramienta es posible analizar la información histórica relativa a una empresa concreta. En este caso, se ha escogido la opción de evaluar los retornos de las empresas del índice frente a su riesgo, destacando aquella que se selecciona como *input* y que en este caso es Amazon. Para ello, el programa calcula la beta, medida que representa el nivel de riesgo sistemático de las acciones de una compañía en un índice, midiendo la volatilidad de la acción en

relación al mercado o índice de referencia. A continuación, se muestra el resultado obtenido.

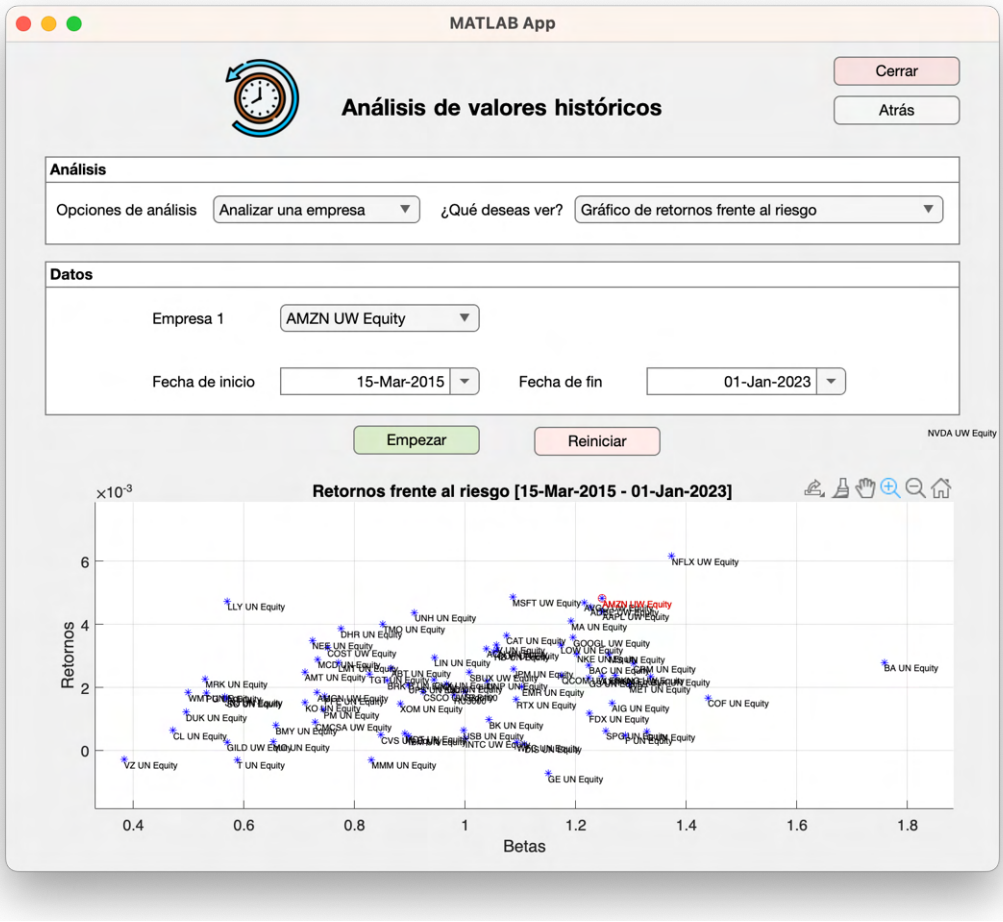
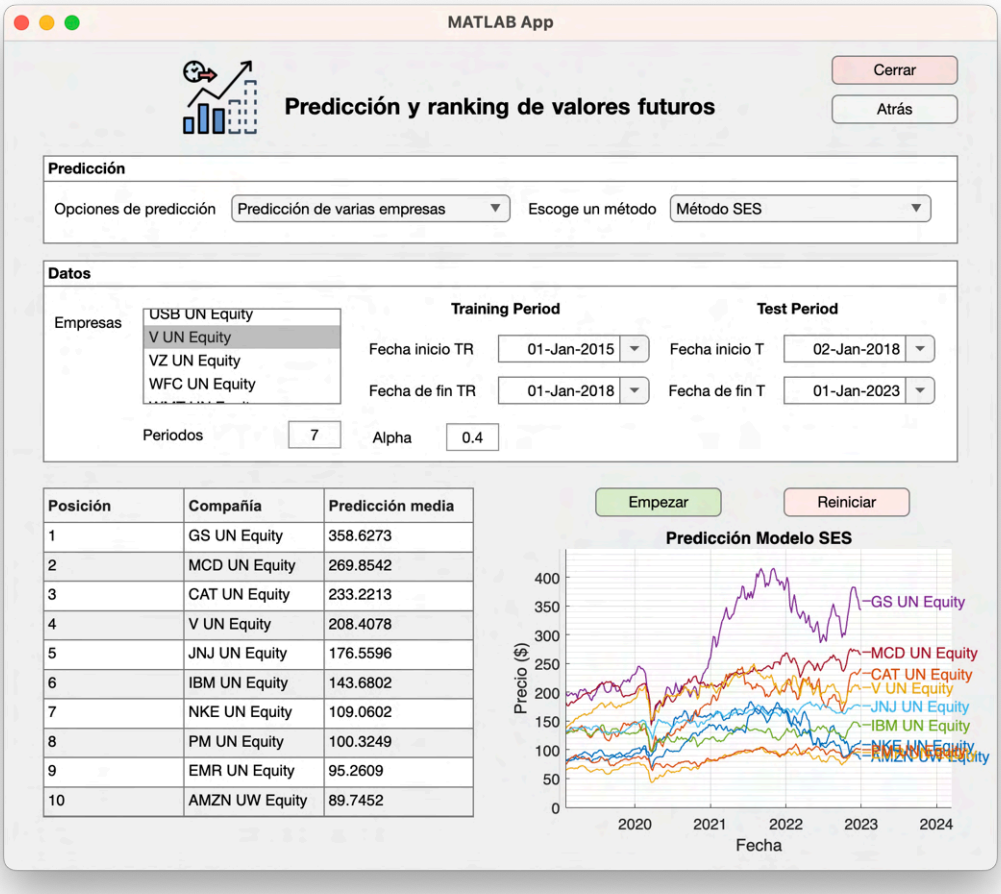


Ilustración 3. Ejemplo de análisis de valores históricos (Elaboración propia)

Por otro lado, la herramienta es capaz de predecir precios y rentabilidades futuros de acciones para una o varias empresas. Se va a mostrar la opción de varias empresas para poder apreciar la capacidad de la herramienta de elaborar un ranking que ordene a las compañías en función de sus retornos esperados. El usuario escogerá las empresas en las que esté interesado en invertir para saber cuáles de ellas le proporcionarán un mayor retorno. Además, deberá seleccionar una serie de *inputs* adicionales en función del método de predicción escogido que, en el caso de este ejemplo, será el método SES (*Single Exponential Smoothing*).



Posición	Compañía	Predicción media
1	GS UN Equity	358.6273
2	MCD UN Equity	269.8542
3	CAT UN Equity	233.2213
4	V UN Equity	208.4078
5	JNJ UN Equity	176.5596
6	IBM UN Equity	143.6802
7	NKE UN Equity	109.0602
8	PM UN Equity	100.3249
9	EMR UN Equity	95.2609
10	AMZN UW Equity	89.7452

Ilustración 4. Ejemplo de predicción y ranking de valores futuros (Elaboración propia)

Una vez conocido el ranking, el usuario está en posición de poder elaborar su cartera de la manera más conveniente, en función del riesgo que esté dispuesto a asumir.

5. Conclusiones

La optimización de carteras es esencial en el ámbito financiero ya que cada inversor tiene sus propias necesidades y objetivos. Así, la construcción de una cartera óptima implica encontrar un equilibrio adecuado entre el rendimiento deseado y el riesgo que está dispuesto a asumir el inversor. Para ello se realizan profundos análisis de datos históricos, tendencias del mercado y características de riesgo de los activos, entre otras cosas, con el objetivo de distribuir de forma óptima el capital invertido y formar así la cartera.

Por otro lado, el uso de datos de intervalo proporciona beneficios significativos en varios sectores, incluido el financiero. Estos datos son especialmente útiles a la hora de representar la incertidumbre, evaluar y gestionar riesgos, optimizar carteras e informar decisiones de inversión. Por ello, en este proyecto se ha desarrollado una herramienta que tiene el propósito de convertirse en un sistema de referencia para los inversores a la hora de tomar decisiones y construir sus carteras, utilizando el análisis de datos de intervalo.

La herramienta desarrollada en este proyecto ofrece dos funciones principales. La primera es analizar datos históricos para comprender el comportamiento de las acciones a lo largo del tiempo, lo que permite al inversor tomar decisiones informadas y ajustar sus estrategias en consecuencia. La segunda función es realizar estimaciones de los valores futuros de las acciones, identificando oportunidades de inversión y optimizando la distribución del capital en la cartera. Esta herramienta resulta muy útil para obtener una visión completa antes de elaborar una estrategia de inversión.

La primera de las ideas extraídas del proyecto tiene que ver con el uso de los datos históricos, donde se ha descubierto la importancia de estudiar tanto los más recientes como los más antiguos. Los datos recientes son más relevantes para comprender el comportamiento a corto plazo, pero los datos más antiguos pueden proporcionar información sobre ciclos, patrones y tendencias a largo plazo de una empresa, industria o mercado.

En cuanto a las predicciones futuras, es relevante prestar atención a los cinco modelos de predicción ofrecidos y expuestos en el Capítulo 6. Cada modelo tiene en cuenta diferentes factores que influyen en el precio de la acción. Se ha observado que el método K-NN presenta el menor error absoluto medio, seguido por el modelo M. A continuación, se encontrarían el método GARCH(1,1) y el de composición de métodos y por último, aquel que presentaría un mayor error sería el método SES. Sin embargo, todos ellos presentan resultados realistas que se recomienda tener en cuenta a la hora de hacer el análisis.

En resumen, se recomienda considerar los valores resultantes de todos los métodos de predicción para obtener una idea realista del resultado pronosticado.

Finalmente, en cuanto a la frecuencia de los datos, se ha concluido que aquellos con los que se ha obtenido un menor error de predicción son los datos semanales, seguidos de los mensuales y posteriormente, los trimestrales. Por ello, a la hora de realizar el estudio se recomienda al usuario hacer uso de datos históricos por semanas.

Además, el ranking obtenido en función de la frecuencia de datos seleccionada será diferente y según el ejemplo estudiado con el método SES, se ha derivado que las frecuencias semanal y mensual devuelven un resultado con menor error que la trimestral frente al caso real, obtenido a partir de datos semanales, destacando que para el primer caso el ranking obtenido es prácticamente igual al caso real, por lo que resulta ser un buen método de estimación de ranking. En la siguiente tabla se puede apreciar el resultado.

Ranking - Real		Ranking - Semana		Ranking - Mes		Ranking - Trimestre	
Posición	Compañía	Posición	Compañía	Posición	Compañía	Posición	Compañía
1	GS UN Equity	1	GS UN Equity	1	GS UN Equity	1	GS UN Equity
2	MCD UN Equity	2	MCD UN Equity	2	MCD UN Equity	2	CAT UN Equity
3	CAT UN Equity	3	CAT UN Equity	3	CAT UN Equity	3	MCD UN Equity
4	V UN Equity	4	V UN Equity	4	V UN Equity	4	V UN Equity
5	JNJ UN Equity	5	JNJ UN Equity	5	JNJ UN Equity	5	JNJ UN Equity
6	IBM UN Equity	6	IBM UN Equity	6	IBM UN Equity	6	IBM UN Equity
7	NKE UN Equity	7	NKE UN Equity	7	PM UN Equity	7	NKE UN Equity
8	PM UN Equity	8	PM UN Equity	8	NKE UN Equity	8	PM UN Equity
9	AMZN UW Equity	9	EMR UN Equity	9	AMZN UW Equity	9	EMR UN Equity
10	EMR UN Equity	10	AMZN UW Equity	10	EMR UN Equity	10	AMZN UW Equity

Tabla 1. Ranking según la frecuencia de los datos (Elaboración propia)

Una vez finalizado el proyecto, se puede concluir que se ha logrado cumplir con los objetivos principales al desarrollar una herramienta efectiva y visual capaz de ayudar a los inversores a realizar sus movimientos en los mercados financieros con una menor incertidumbre. Se dejan las puertas abiertas para posibles desarrollos futuros como la elaboración de un artículo que completen el trabajo realizado durante estos últimos meses.

6. Referencias

Cosín López-Medel, Á. (Julio de 2019). Basic portfolio management based on beta values using interval-valued information. Madrid: Escuela Técnica Superior de Ingeniería (ICAI). Grado en Ingeniería Industrial.

Kolm, P. N., Tütüncü, R., & Fabozzi, F. J. (2014). 60 years of portfolio optimization: Practical challenges and current trends. European Journal of Operational Research, 234(2), 356-371.

Maté, C. G. (2022). Forecasting in FOREX the spot price interval of tomorrow with the same information of today. An analysis of the seven majors using a linear regression model based on interval arithmetic. Knowledge-Based Systems, 258, 109923.

Matlab. (s.f.). *Develop apps using App Designer*. Obtenido de Matlab: https://nl.mathworks.com/help/matlab/app-designer.html#:~:text=App%20Designer%20es%20un%20entorno,componentes%20interactivos%20de%20la%20IU.

NEW APROACHES TO BASIC PORTFOLIO MANAGEMENT BASED ON RANKINGS AND RISK. THE USE OF INTERVAL-VALUED INFORMATION

Author: Domingo Górriz, Teresa.
Supervisor: Maté Jiménez, Carlos.
Collaborating Entity: ICAI – Universidad Pontificia Comillas

ABSTRACT

In recent years, there has been an exponential increase in the amount of data collected, processed, and analyzed thanks to technological developments. While the analysis of crisp data is common in many fields, Interval Analysis (IA) has not been studied as extensively and can be highly valuable in representing uncertainty more accurately. By utilizing interval-valued (IV) data, a more comprehensive understanding of the available information can be obtained, enabling improved decision-making and risk assessment.

The use of IA can be particularly beneficial in financial markets, which is why this project will focus on them. The combination of IA with optimization and forecasting models can offer an intriguing perspective on portfolio management, optimization, and creation. Therefore, one of the objectives of this project is to develop an original, user-friendly, and visually-oriented system that serves as a tool for investors to analyze the past, present, and future behavior of stock prices. This system aims to assist in the formation of optimal portfolios based on desired performance and risk levels. Information from various financial markets such as IBEX35 or S&P100 will be utilized over different time periods to achieve a set of results and draw relevant conclusions.

Keywords: Interval Analysis, Portfolio Optimization, Financial Markets, Risk, Return, Ranking, Forecasting Methods, Matlab.

1. Introduction

Financial markets represent an essential element for the economies of different countries, as well as an aid for the development and enhancement of their economic health.

At a particular level, there are many people for whom these markets represent a form of savings or business, seeking to obtain a certain return on their money from the purchase of shares, bonds, or other securities with which they create a portfolio.

Portfolio optimization consists of the generation of different portfolios of securities with the objective of choosing the one that best satisfies the return and risk objectives of each investor. When creating these portfolios there are different tools that can be used. Since the first Modern Portfolio Theory (MPT) (Kolm, Tütüncü, & Fabozzi, 2014), research on this topic has continued and other tools have been developed (Cosín López-Medel, 2019).

On the other hand, IV data have long been underestimated compared to crisp data. However, while crisp data can present challenges in terms of accuracy, as they do not capture all the information inherent in complex variables, IV data offer a broader and more precise perspective, which is valuable in decision making and accurate understanding of the information.

Throughout this project, various methods of historical data analysis will be studied, as well as predictive models with IV time series (Maté C. , 2022) in order to focus these analyses on portfolio optimization. For this purpose, past records of different financial markets, such as the IBEX35 or the S&P100, will be studied in order to develop a tool to assist the investor in decision making.

2. Definition of the project

The project will consist of two main parts. The first one will include the study of the documentation related to the most important concepts of portfolio management and optimization, performance and risk, forecasting methods and interval data analysis. All this will be collected in the first chapters and will serve as a basis for the understanding and further development of the tool.

The second branch of the project consists of the creation of the tool. It is a known fact that the world of financial markets is an unpredictable environment as they are affected by a large number of factors. Therefore, the aim of this project is to develop a tool that combines the study of some of these factors and statistical methods to try to estimate future stock returns and thus, build an optimized portfolio for each investor.

According to that, the project will seek to make it easier for investors to make investment decisions since it will not only offer them the possibility of analyzing the past performance of different companies operating in a stock market index and comparing them with each other, but will also offer them various methods for creating portfolios based on the use of forecasting models. All this will be offered to the user by means of different graphs and tables that will help him to understand the results in an intuitive and visual way, providing valuable information.

In order to develop the tool, Matlab's App Designer functionality, which is an interactive development environment that allows designing an application and programming its behavior (Matlab, s.f.), will be used. Thus, the developed program will be applied to real data obtained from different stock market indexes to obtain a realistic model. Finally, it will be time to check its accuracy and effectiveness.

3. Description of the tool

The development of the tool is based on the theoretical concepts previously studied and explained throughout the project. The tool will have two main value analysis functions that differ in their situation in time.

The first one deals with the analysis of historical values, i.e., stock prices from the past. Within this option, the user will be able to choose between looking at the information of one company or two, in case he is interested in comparing the behavior of two specific companies. The user must then choose the type of analysis he wishes to obtain, and the program will produce a graph based on a series of input parameters previously selected by the user.

The second of the functionalities offered by the tool is the prediction of future values from the analysis of past data. To do this, the tool chooses a series of historical values of the company or companies of interest to the user and, according to the forecasting method chosen by the user, it will produce an estimate of the future value of the shares. Thus, it will be able to elaborate a ranking that puts companies in order according to their

expected return so that the user can have an idea of the expected results of each company and elaborate its own portfolio according to its level of return and expected risk.

In addition, it should be noted that the user will be free to decide from the tool whether the data to be studied are weekly, monthly, or quarterly.

Below is an outline of the functionalities offered by the tool to facilitate its understanding.

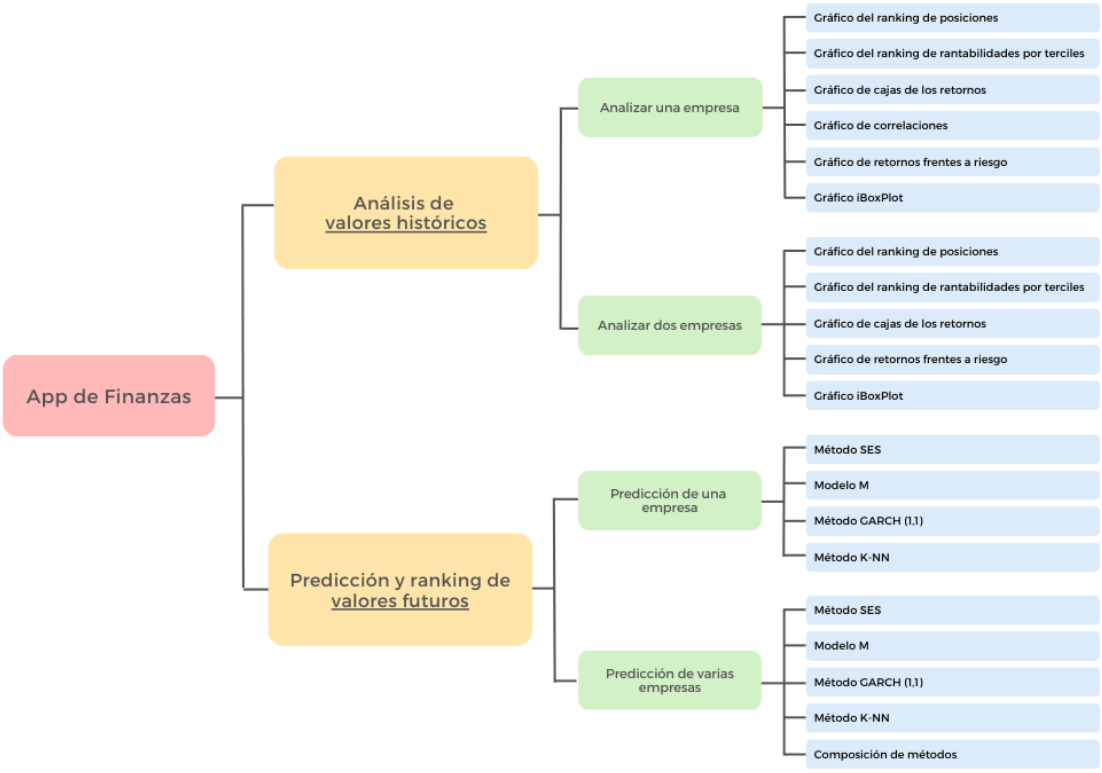


Ilustración 5. Scheme of the functionalities of the tool (Own elaboration)

It should be noted that before being able to use the tool, a thorough analysis of the data is necessary in order to collect them, filtering those that are useful and discarding those that are not. For this purpose, tools such as Bloomberg or Yahoo Finance are needed, from where historical financial data can be downloaded by days, months, or years. Once collected, the data is entered into the tool. The appearance of the tool at startup will be as follows.

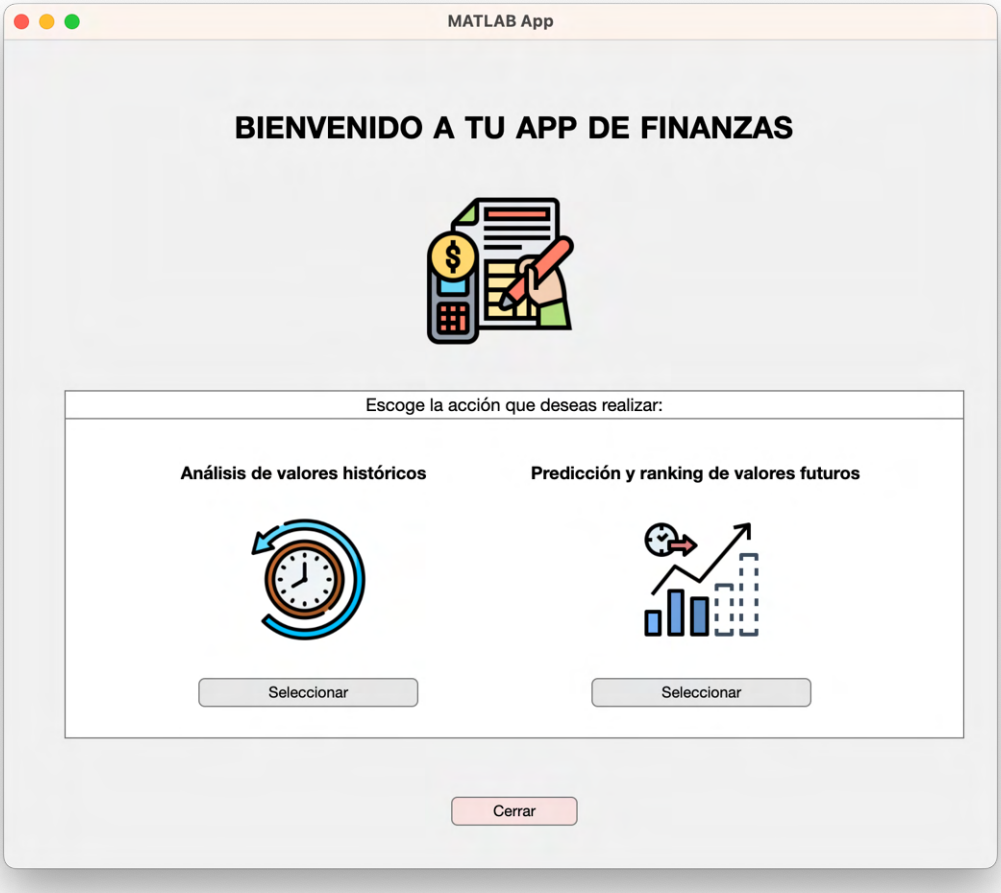


Ilustración 6. Tool startup window (Own elaboration)

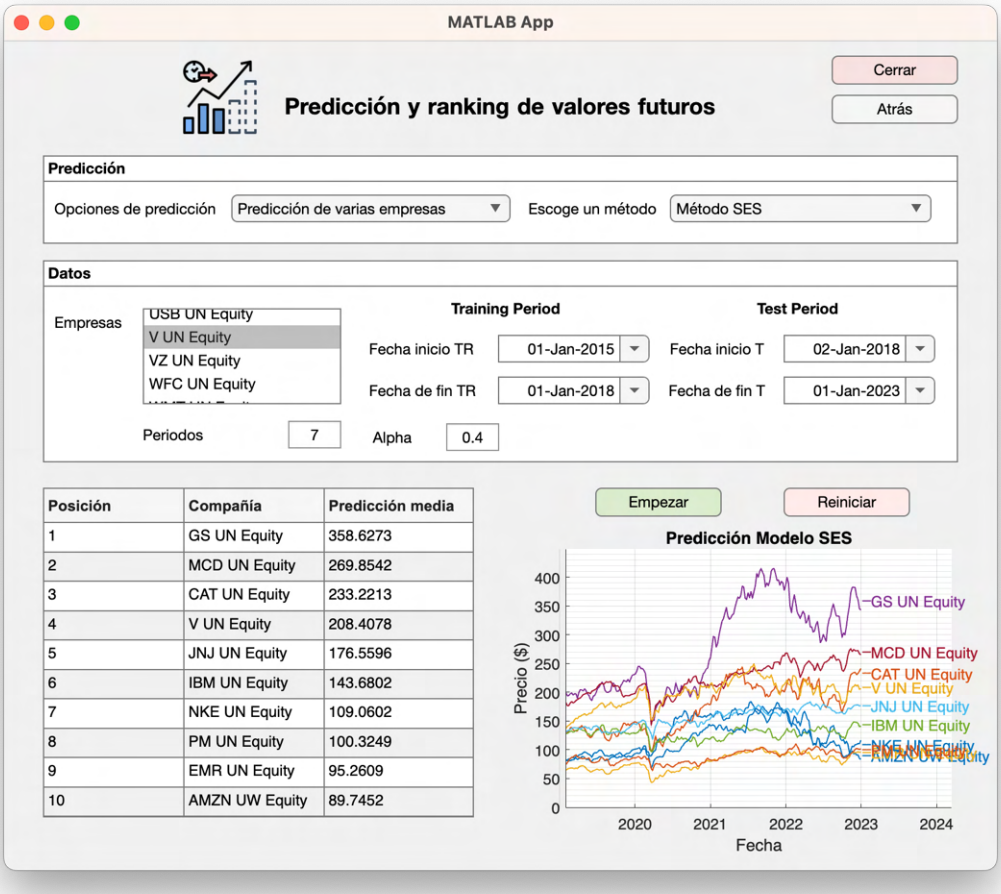
Finally, the tool will be tested to prove its performance, observe the results it returns and develop the pertinent conclusions.

4. Results

The tool has been analyzed based on a series of stocks belonging to the S&P100 stock index. Thus, with the first of the tool's functionalities it is possible to analyze the historical information relating to a specific company. In this case, the option that has been chosen is to evaluate the returns of the companies in the index against their risk, highlighting the company selected as input, which in this case is Amazon. To do this, the program calculates the beta, a measure that represents the level of systematic risk of the shares of a company in an index, measuring the volatility of the share in relation to the market benchmark index. The result obtained is shown below.

Ilustración 7. Example of analysis of historical values (Own elaboration)

On the other hand, the tool is used to forecast future stock values and returns for one or several companies. The option of several companies will be shown in order to appreciate the tool's ability to elaborate a ranking that orders the companies according to their expected returns. The user will choose the companies in which he is interested in investing in order to know which of them will provide him with a higher return. In addition, the user will have to select a series of additional inputs depending on the chosen prediction method, which in the case of this example will be the SES (Single Exponential Smoothing) method.



Posición	Compañía	Predicción media
1	GS UN Equity	358.6273
2	MCD UN Equity	269.8542
3	CAT UN Equity	233.2213
4	V UN Equity	208.4078
5	JNJ UN Equity	176.5596
6	IBM UN Equity	143.6802
7	NKE UN Equity	109.0602
8	PM UN Equity	100.3249
9	EMR UN Equity	95.2609
10	AMZN UW Equity	89.7452

Ilustración 8. Example of forecast and ranking of future values (Own elaboration)

Once the ranking is known, the user is in the position to build his portfolio in the most convenient way, depending on the level of risk he is willing to take.

5. Conclusions

Portfolio optimization is essential in the financial arena as each investor has his own needs and objectives. Thus, the construction of an optimal portfolio involves finding an appropriate balance between the desired return and the risk that the investor is willing to assume. For this purpose, in-depth analyses of historical data, market trends and risk characteristics of the assets, among other things, are carried out in order to optimally allocate the invested capital and build the portfolio.

On the other hand, the use of interval-valued data provides significant benefits in various industries, including finance. These data are particularly useful in representing uncertainty, assessing and managing risk, optimizing portfolios and informing investment decisions. Therefore, the developed tool is intended to become a reference system for investors to make decisions and construct their portfolios using interval data analysis.

The tool developed in this project offers two main functionalities. The first one is to analyze historical data to understand the performance of stocks over time, allowing the investor to make better informed decisions and adjust their strategies accordingly. The

second functionality is to make estimates of future stock values, identifying investment opportunities and optimizing the allocation of capital in the portfolio. This tool is very useful for obtaining a complete picture of the market situation before developing an investment strategy.

The first of the ideas extracted from the project has to do with the use of historical data, where the importance of studying both the most recent and the oldest data has been noticed. Recent data is more relevant for understanding short-term behavior, but older data can provide information on cycles, patterns and long-term trends of a company, industry, or market.

In terms of future predictions, it is relevant to pay attention to the five forecasting models offered and discussed in Capítulo 6. Each model takes into account different factors that influence the share price and return. It has been observed that the K-NN method presents the lowest mean absolute error, followed by the M model. Next, we would find the GARCH(1,1) method and the method composition method and finally, the one that would present a greater error would be the SES method. However, all of them present realistic results that it is recommended to take into account when doing the analysis.

To summarize, it is recommended to consider the values resulting from all the forecasting methods to obtain a realistic idea of the predicted result.

Finally, regarding the frequency of the data, it has been concluded that those with which the lowest prediction error has been obtained are weekly data, followed by monthly data and then quarterly data. Therefore, when carrying out the study, the user is recommended to make use of historical data by weeks.

In addition, the ranking obtained depending on the selected data frequency will be different and, according to the example analyzed with SES method, it has been derived that the weekly and monthly frequencies return a result with a lower error than the quarterly one compared to the real case, obtained from weekly data. The following table shows the result.

Ranking - Real		Ranking - Semana		Ranking - Mes		Ranking - Trimestre	
Posición	Compañía	Posición	Compañía	Posición	Compañía	Posición	Compañía
1	GS UN Equity	1	GS UN Equity	1	GS UN Equity	1	GS UN Equity
2	MCD UN Equity	2	MCD UN Equity	2	MCD UN Equity	2	CAT UN Equity
3	CAT UN Equity	3	CAT UN Equity	3	CAT UN Equity	3	MCD UN Equity
4	V UN Equity	4	V UN Equity	4	V UN Equity	4	V UN Equity
5	JNJ UN Equity	5	JNJ UN Equity	5	JNJ UN Equity	5	JNJ UN Equity
6	IBM UN Equity	6	IBM UN Equity	6	IBM UN Equity	6	IBM UN Equity
7	NKE UN Equity	7	NKE UN Equity	7	PM UN Equity	7	NKE UN Equity
8	PM UN Equity	8	PM UN Equity	8	NKE UN Equity	8	PM UN Equity
9	AMZN UW Equity	9	EMR UN Equity	9	AMZN UW Equity	9	EMR UN Equity
10	EMR UN Equity	10	AMZN UW Equity	10	EMR UN Equity	10	AMZN UW Equity

Tabla 2. Ranking according to data frequency (Own elaboration)

Once the project has been completed, it can be concluded that the main objectives have been achieved by developing an effective and visual tool capable of helping investors to make their movements in the financial markets with less uncertainty. The doors are left open for possible future developments such as the elaboration of an article to complete the work done during these last months.

6. References

Cosín López-Medel, Á. (Julio de 2019). Basic portfolio management based on beta values using interval-valued information. Madrid: Escuela Técnica Superior de Ingeniería (ICAI). Grado en Ingeniería Industrial.

Kolm, P. N., Tütüncü, R., & Fabozzi, F. J. (2014). 60 years of portfolio optimization: Practical challenges and current trends. European Journal of Operational Research, 234(2), 356-371.

Maté, C. G. (2022). Forecasting in FOREX the spot price interval of tomorrow with the same information of today. An analysis of the seven majors using a linear regression model based on interval arithmetic. Knowledge-Based Systems, 258, 109923.

Matlab. (s.f.). *Develop apps using App Designer*. Obtenido de Matlab: https://nl.mathworks.com/help/matlab/app-designer.html#:~:text=App%20Designer%20es%20un%20entorno,componentes%20interactivos%20de%20la%20IU.

Índice de la memoria

Capítulo 1. Introducción 13

1.1 Motivación del proyecto 13

1.2 Objetivos del proyecto 14

1.3 Metodología del proyecto 15

Capítulo 2. Gestión de carteras 17

2.1 Introducción a la gestión de carteras 17

2.1.1 Definición y objetivos 17

2.1.2 Tipos de gestión de carteras 19

2.1.3 Proceso seguido en la gestión de carteras 20

2.2 Retorno y varianza en carteras de inversiones 21

2.2.1 Rendimiento de la inversión, ROI 22

2.2.2 Rendimiento del periodo de tenencia, HPR 22

2.2.3 Rendimiento esperado, ER 23

2.2.4 Varianza y desviación estándar de una cartera de inversiones 24

2.3 Riesgo en carteras de inversión, el uso de beta (β) y la diversificación 24

2.3.1 Riesgo sistemático y riesgo específico 25

2.3.2 Relación entre riesgo y retorno y SML 26

2.3.3 El uso de beta (β) 28

2.3.4 Diversificación 29

2.3.5 Medidas de riesgo 30

2.4 Modelos CAPM 31

2.5 Tolerancia al riesgo 32

2.6 Asignación de activos 33

2.7 Optimización de la cartera 34

2.8 Fondos de inversión 34

2.9 Tecnologías para la gestión de carteras 34

2.9.1 FactSet 35

2.9.2 Bloomberg 38

Capítulo 3. Optimización de carteras 41

3.1 Introducción a la optimización de carteras 41

3.1.1 Activos de inversión *41*
3.1.2 Carteras de inversión *43*
3.1.3 Volatilidad de un activo *44*
3.1.4 Covarianza de una cartera *44*
3.1.5 Diversificación *45*
3.2 Herramientas para la optimización de carteras 48
3.2.1 Optimización con Media-Varianza (Markowitz) y frontera de eficiencia *48*
3.2.2 Optimización de carteras con CVaR *50*
3.2.3 Programación lineal *50*
3.2.4 Carteras Naïve *51*
3.2.5 Ratio de Sharpe *51*
3.2.6 Post-Modern Portfolio Theory (PMP) *52*
3.3 Métodos de asignación de activos 53
3.4 Índices utilizados en la optimización de carteras 54
3.4.1 Índice de Fuerza Relativa *54*
3.4.2 Índice de Mercado *55*
3.4.3 Fondos Cotizados e Índices de Futuros *55*
3.5 El uso de beta (β) 56
3.5.1 Beta (β) *56*
3.5.2 Doble beta *57*
3.5.3 El uso de la beta en el IBEX 35 ..... *59*
3.5.4 El uso de la beta en el S&P 100 ..... *64*

Capítulo 4. Datos de intervalo y regresión* *66

4.1 Introducción a los datos de intervalo 66
4.1.1 Operaciones aritméticas con intervalos *68*
4.1.2 Medidas de distancia para intervalos *69*
4.1.3 Estadísticos descriptivos *70*
4.1.4 Conceptos básicos al analizar predicciones para una ITS *71*
4.2 Análisis de Intervalos (AI) 73
4.2.1 Optimización de carteras haciendo uso de IA *73*
4.3 Regresión lineal con datos de intervalo (ILR) 74
4.3.1 Métodos de ILR *76*
4.3.2 Método CRM *78*

4.3.3 Método de Regresión Robusto para Datos de Intervalo basado en Puntos Medios y Radios Logarítmicos (LN-IRR method) ... *79*
4.3.4 Modelo MG ... *82*

Capítulo 5. CAPM y su uso con métodos de regresión para datos de intervalo ... ***86***
5.1 Introducción ... 86
5.2 CAPM ... 88
5.2.1 CAPM para datos puntuales (CAPM clásico) ... *88*
5.2.2 CAPM para datos de intervalo ... *92*
5.3 Regresión de intervalo y CAPM ... 95
5.3.1 Método de Regresión de Intervalos (Iregr) ... *96*
5.4 Iregr en el marco del CAPM de intervalos ... 100

Capítulo 6. Métodos y modelos de predicción ... ***102***
6.1 Método de Suavizado Exponencial Simple (SES) ... 102
6.2 Modelo M ... 106
6.3 Método de GARCH (1,1) ... 113
6.4 Método K-NN ... 117
6.5 Composición de métodos ... 121

Capítulo 7. Análisis de la herramienta ... ***124***
7.1 Introducción a la herramienta ... 124
7.2 Funcionalidad de la herramienta ... 125
7.2.1 Análisis de valores históricos ... *125*
7.2.2 Predicción y ranking de valores futuros ... *139*
7.3 Próximos pasos ... 151
7.4 Análisis de potenciales mejoras ... 152

Capítulo 8. Conclusiones ... ***155***

Capítulo 9. Bibliografía ... ***158***

ANEXO I: ALINEACIÓN DEL PROYECTO CON LOS ODS ... ***166***

Índice de tablas

Tabla 1. Ranking según la frecuencia de los datos (Elaboración propia)..... 9
Tabla 2. Ranking according to data frequency (Own elaboration)..... 17
Tabla 3. Información de las empresas del IBEX 35 (Elaboración propia)..... 62
Tabla 4. Información de las empresas del S&P 100 (Elaboración propia)..... 64
Tabla 5. Resultados de predicciones a partir del modelo SES para Starbucks (Elaboración propia)..... 104
Tabla 6. Resultados de ranking y predicciones a partir del modelo SES para varias compañías (Elaboración propia)..... 106
Tabla 7. Resultados de predicciones a partir del modelo M para Starbucks (Elaboración propia)..... 111
Tabla 8. Resultados de ranking y predicciones a partir del modelo M para varias compañías (Elaboración propia)..... 112
Tabla 9. Resultados de predicciones a partir del modelo GARCH (1,1) para Starbucks (Elaboración propia)..... 115
Tabla 10. Resultados de ranking y predicciones a partir del modelo GARCH (1,1) para varias compañías (Elaboración propia)..... 116
Tabla 11. Resultados de predicciones a partir del método KNN para Starbucks (Elaboración propia)..... 120
Tabla 12. Resultados de ranking y predicciones a partir del método KNN para varias compañías (Elaboración propia)..... 121
Tabla 13. Resultados de comparación resultados de los métodos estudiados (Elaboración propia)..... 122
Tabla 14. Resultados de ranking y predicciones a partir de composición de métodos (Elaboración propia)..... 123
Tabla 15. Evaluación de resultados de la herramienta 151
Tabla 16. Resultados y error de predicción según la frecuencia de los datos 153
Tabla 17. Ranking según la frecuencia de los datos..... 153

Índice de ilustraciones

Ilustración 1. Esquema de funcionalidades de la herramienta (Elaboración propia) 5
Ilustración 2. Ventana de inicio de la herramienta (Elaboración propia) 6
Ilustración 3. Ejemplo de análisis de valores históricos (Elaboración propia) 7
Ilustración 4. Ejemplo de predicción y ranking de valores futuros (Elaboración propia) 8
Ilustración 5. Scheme of the functionalities of the tool (Own elaboration) 13
Ilustración 6. Tool startup window (Own elaboration) 14
Ilustración 7. Example of analysis of historical values (Own elaboration) 15
Ilustración 8. Example of forecast and ranking of future values (Own elaboration) 16
Ilustración 9. Riesgo sistemático y específico (López, 2020) 26
Ilustración 10. SML (Wikipedia, 2022) 27
Ilustración 11. Desviación estándar versus número de activos en la cartera (Easy Portfolio, 2019) 30
Ilustración 12. FactSet (FactSet, 2023) 36
Ilustración 13. Ejemplo de Apple en FactSet (FactSet, 2023) 37
Ilustración 14. Gestión de carteras en FactSet (FactSet, 2023) 38
Ilustración 15. Terminal Bloomberg (Fortuño, 2022) 39
Ilustración 16. Bloomberg PORT (Ellis, 2019) 40
Ilustración 17. Efectos de la diversificación (Bertani, 2021) 48
Ilustración 18. Diagrama de media-varianza (Cosín López-Medel, 2019) 49
Ilustración 19. Cartera eficiente y frontera de eficiencia (Lasheras, 2020) 52
Ilustración 20. Evolución valor del IBEX 35 (Elaboración propia) 60
Ilustración 21. Retornos frente al riesgo en el IBEX 35 (Elaboración propia) 63
Ilustración 22. Retornos frente al riesgo en el S&P 100 (Elaboración propia) 65
Ilustración 23. Serie de tiempo de intervalos (Arroyo & Maté, 2006) 67
Ilustración 24. Análisis de regresión (Fraschini & Axhausen, 2001) 75
Ilustración 25. Ejemplo de predicciones a partir del modelo SES para Starbucks (Elaboración propia) 104

Ilustración 26. Ejemplo de predicciones a partir del modelo SES para varias compañías (Elaboración propia) 105
Ilustración 27. Ejemplo de predicciones a partir del modelo SES para varias compañías ampliado (Elaboración propia) 105
Ilustración 28. Ejemplo de predicciones a partir del modelo M para Starbucks (Elaboración propia) 111
Ilustración 29. Ejemplo de predicciones a partir del modelo M para varias compañías (Elaboración propia) 112
Ilustración 30. Ejemplo de predicciones a partir del modelo GARCH (1,1) para Starbucks (Elaboración propia) 115
Ilustración 31. Ejemplo de predicciones a partir del modelo GARCH (1,1) para varias compañías (Elaboración propia) 116
Ilustración 32. Elección de los k vecinos más cercanos con K-NN (IBM, 2023) 117
Ilustración 33. Ejemplo de predicciones a partir del método KNN para Starbucks (Elaboración propia) 119
Ilustración 34. Ejemplo de predicciones a partir del método KNN para varias compañías (Elaboración propia) 120
Ilustración 35. Ejemplo de comparación de los métodos estudiados (Elaboración propia) 122
Ilustración 36. Ventana de inicio de la herramienta (Elaboración propia) 125
Ilustración 37. Ventana análisis de valores históricos (Elaboración propia) 126
Ilustración 38. Analizar una empresa (Elaboración propia) 126
Ilustración 39. Analizar una empresa - Gráfico del ranking de posiciones (Elaboración propia) 127
Ilustración 40. Analizar una empresa - Gráfico del ranking de rentabilidades por terciles (Elaboración propia) 128
Ilustración 41. Analizar una empresa - Gráfico de cajas de los retornos (Elaboración propia) 129
Ilustración 42. Analizar una empresa - Gráfico de correlaciones (Elaboración propia) ... 130

Ilustración 43. Analizar una empresa - Gráfico de retornos frente al riesgo (Elaboración propia)..... 132
Ilustración 44. Analizar una empresa - Gráfico iBoxPlot (Elaboración propia) 134
Ilustración 45. Analizar dos empresas - Gráfico del ranking de posiciones (Elaboración propia)..... 135
Ilustración 46. Analizar dos empresas - Gráfico del ranking de rentabilidades por terciles (Elaboración propia)..... 136
Ilustración 47. Analizar dos empresas - Gráfico de cajas de los retornos (Elaboración propia) 137
Ilustración 48. Analizar dos empresas - Gráfico iBoxPlot (Elaboración propia)..... 138
Ilustración 49. Ventana predicción y ranking de valores futuros (Elaboración propia).... 139
Ilustración 50. Predicción de precio de una empresa – Método SES (Elaboración propia) 141
Ilustración 51. Predicción de precio de una empresa – Modelo M (Elaboración propia) . 142
Ilustración 52. Predicción de precio de una empresa – Método GARCH (1,1) (Elaboración propia)..... 143
Ilustración 53. Predicción de precio de una empresa – Método K-NN (Elaboración propia) 144
Ilustración 54. Predicción de rentabilidad y ranking de varias empresas – Método SES (Elaboración propia)..... 146
Ilustración 55. Predicción de rentabilidad y ranking de varias empresas – Modelo M (Elaboración propia)..... 147
Ilustración 56. Predicción de rentabilidad y ranking de varias empresas – Método GARCH (1,1) (Elaboración propia) 148
Ilustración 57. Predicción y ranking de varias empresas – Método K-NN (Elaboración propia)..... 149
Ilustración 58. Predicción y ranking de varias empresas – Composición de métodos (Elaboración propia)..... 150
Ilustración 59. Selección de frecuencia en la herramienta 152
Ilustración 60. Resultados del ranking de todos los métodos (Elaboración propia)..... 157

Índice de ecuaciones

Ecuación 2.1. ROI de un activo 22
Ecuación 2.2. ROI de una cartera 22
Ecuación 2.3. HPR 22
Ecuación 2.4. E(R) de un activo 23
Ecuación 2.5. E(R) de una cartera 23
Ecuación 2.6. E(R) anualizado 23
Ecuación 2.7. Varianza de una cartera 24
Ecuación 2.8. Retorno de una inversión 26
Ecuación 2.9. Pendiente de la SML 27
Ecuación 2.10. Beta de un activo 28
Ecuación 2.11. Beta de una cartera 29
Ecuación 2.12. Volatilidad 30
Ecuación 2.13. Valor en riesgo (VaR) 31
Ecuación 2.14. Ratio de Sharpe 31
Ecuación 2.15. CAPM 32
Ecuación 3.1. Rendimiento de un activo 41
Ecuación 3.2. Tasa de rendimiento de un activo 42
Ecuación 3.3. Relación entre rendimiento y tasa de rendimiento de un activo 42
Ecuación 3.4. Retorno acumulado 42
Ecuación 3.5. Rendimiento total del apalancamiento 43
Ecuación 3.6. Inversión en un activo dentro de una cartera 43
Ecuación 3.7. Retorno de una cartera 44
Ecuación 3.8. Varianza de un activo 44
Ecuación 3.9. Desviación típica de un activo 44
Ecuación 3.10. Covarianza de una cartera 45
Ecuación 3.11. Varianza de la tasa de rendimiento de una cartera 46
Ecuación 3.12. Tasa de rendimiento de una cartera 46

Ecuación 3.13. Varianza de r 46
Ecuación 3.14. Varianza de una cartera 47
Ecuación 3.15. Semivarianza 52
Ecuación 3.16. Volatilidad invertida 53
Ecuación 3.17. U y D en periodo de alza 54
Ecuación 3.18. U y D en periodo a la baja 54
Ecuación 3.19. Fuerza relativa 55
Ecuación 3.20. Índice de fuerza relativa 55
Ecuación 4.1. Límite inferior del intervalo 67
Ecuación 4.2. Límite superior del intervalo 68
Ecuación 4.3. Límite inferior del intervalo en un modelo Naive Forecast 68
Ecuación 4.4. Límite superior del intervalo en un modelo Naive Forecast 68
Ecuación 4.5. Suma de intervalos 68
Ecuación 4.6. Diferencia de intervalos 68
Ecuación 4.7. Producto de un intervalo por un escalar 68
Ecuación 4.8. Producto de intervalos 69
Ecuación 4.9. Inversa de un intervalo 69
Ecuación 4.10. Cociente de un intervalo 69
Ecuación 4.11. Distancia Euclídea 69
Ecuación 4.12. Distancia de Hausdorff 70
Ecuación 4.13. Distancia de Bertoluzza 70
Ecuación 4.14. Media de los centros de una serie de intervalos 70
Ecuación 4.15. Media de los radios de una serie de intervalos 70
Ecuación 4.16. Intervalo de la media de Aumann 70
Ecuación 4.17. Intervalo de la mediana de Hausdorff 71
Ecuación 4.18. Intervalo de la mediana bajo 1-norm 71
Ecuación 4.19. Error de la distancia media (MDE) 71
Ecuación 4.20. Varianza relativa media de intervalo (iARV) 72
Ecuación 4.21. U de Theil para intervalos (iU) 72
Ecuación 4.22. Centro del intervalo 78

Ecuación 4.23. Radio del intervalo 78
Ecuación 4.24. Regresión lineal en CRM 78
Ecuación 4.25. Suma de los cuadrados de las desviaciones en CRM 79
Ecuación 4.26. Predicción en CRM 79
Ecuación 4.27. Modelo de regresión lineal múltiple para datos de intervalo 81
Ecuación 4.28. Simplificación del modelo de regresión lineal múltiple para datos de intervalo 82
Ecuación 4.29. Regresión en modelo MG 83
Ecuación 4.30. Centro del intervalo en modelo MG 83
Ecuación 4.31. Radio del intervalo en modelo MG 83
Ecuación 4.32. Simplificación del intervalo en el modelo MG 84
Ecuación 4.33. Ejemplo intervalo en modelo MG 84
Ecuación 4.34. Estimación del término de intervalo en Mínimos Cuadrados 85
Ecuación 5.1. Matriz de los retornos de los p valores en CAPM 89
Ecuación 5.2. Retorno esperado del valor j en CAPM 89
Ecuación 5.3. Retorno esperado de una cartera en CAPM 89
Ecuación 5.4. Varianza de una cartera en CAPM 90
Ecuación 5.5. Retorno esperado de una cartera A 91
Ecuación 5.6. Condición para el teorema de CAPM 91
Ecuación 5.7. βj en CAPM 91
Ecuación 5.8. Excesos de retorno 92
Ecuación 5.9. Matriz de los retornos en CAPM de intervalos 93
Ecuación 5.10. Exceso de retornos del valor j en CAPM de intervalos 94
Ecuación 5.11. Exceso de retornos de la cartera del mercado en CAPM de intervalo 94
Ecuación 5.12. Simplificación del exceso de retornos del valor j en CAPM de intervalos 94
Ecuación 5.13. Estimación de un intervalo Zjt 95
Ecuación 5.14. Variable dependiente en Iregr 97
Ecuación 5.15. βI en Iregr 97
Ecuación 5.16. αI en Iregr 97
Ecuación 5.17. xm e ym en Iregr 97

Ecuación 5.18. Intervalo cerrado del set de Ecuación 5.15 98
Ecuación 5.19. Intervalo cerrado del set de Ecuación 5.16 98
Ecuación 5.20. Línea de regresión en Iregr 98
Ecuación 5.21. Predicción de un intervalo Y 98
Ecuación 5.22. Variable dependiente en Iregr 99
Ecuación 5.23. Hiperplano Y en Iregr 99
Ecuación 5.24. Intervalo cerrado de βjI 100
Ecuación 6.1. Suavizado Exponencial Simple (SES) 102
Ecuación 6.2. Problema de mínimos cuadrados en el modelo M 107
Ecuación 6.3 Medias de los centros y radios de [X] e [Y] en Modelo M 108
Ecuación 6.4 Varianzas de los centros y radios de [X] en Modelo M 108
Ecuación 6.5 Covarianzas de los centros y radios de [X] e [Y] en Modelo M 108
Ecuación 6.6 Medias de [X] e [Y] en Modelo M 108
Ecuación 6.7 Varianza θ-Fréchet de [X] en Modelo M 109
Ecuación 6.8 Covarianza θ a de [X] e [Y] y de [-X] e [Y] en Modelo M 109
Ecuación 6.9 Estimación de $a0$ en Modelo M 109
Ecuación 6.10 Valor de a si $\sigma x, y < 0$ y $\sigma - x, y < 0$ en Modelo M 109
Ecuación 6.11 Valor de a si $\sigma x, y \geq 0$ y $\sigma x, y \geq \sigma - x, y$ en Modelo M 109
Ecuación 6.12 Valor de a si $\sigma - x, y \geq 0$ y $\sigma - x, y \geq \sigma x, y$ en Modelo M 109
Ecuación 6.13 Valor de $[B]$ en Modelo M 110
Ecuación 6.14 Predicción en Modelo M 110
Ecuación 6.15. Modelo GARCH(p,q) 114
Ecuación 6.16. Distancia Euclídea 118

Capítulo 1. INTRODUCCIÓN

A lo largo de este proyecto se va a realizar un estudio en profundidad sobre la gestión de carteras de valores a partir de la elaboración de rankings y del estudio del comportamiento histórico del precio de las acciones. Para realizar dicho análisis se hará uso de datos de intervalos pasados sobre los precios semanales de apertura y cierre de las acciones de un índice de mercado.

El objetivo del proyecto es la elaboración de una herramienta que sirva de ayuda a los inversores en la toma de decisiones a la hora de formar sus carteras. La herramienta deberá darles la oportunidad de observar datos pasados sobre una compañía concreta o de comparar dos de ellas. Además, los inversores tendrán la oportunidad de seleccionar diversas empresas de su interés y la herramienta les devolverá un ranking que las ordene en función de la rentabilidad esperada en el futuro, según distintos métodos de predicción.

De esta forma, se persigue ofrecer al inversor una visión completa sobre la situación pasada y la esperada de los mercados financieros para poder facilitarle la toma de decisiones, con la finalidad de formar la cartera que mejor se adapte a sus necesidades, objetivos y características.

1.1 MOTIVACIÓN DEL PROYECTO

Es un hecho habitual que los inversores depositen su dinero en un grupo de diferentes activos formando una cartera de la que esperan obtener rentabilidad. Es importante tener en cuenta que el rendimiento que obtendrán de ella dependerá del peso, es decir, el porcentaje del capital total invertido en cada tipo de activo que forma la cartera. Así, los inversores tratan de reducir el riesgo de su inversión depositando su capital en diferentes activos, lo cual recibe el nombre de diversificación.

De cara a tener una cartera bien diversificada, los inversores buscan invertir en activos de diferentes campos y mercados, además de tener un número considerable de activos, pero hay ciertos riesgos que el inversor no puede hacer desaparecer mediante la diversificación. Así, su objetivo será obtener el mayor beneficio posible para un nivel de riesgo que esté dispuesto a asumir (Markowitz, 1952), por lo que la cartera óptima dependerá de la aversión al riesgo de cada inversor.

En el mundo de hoy en día existen una gran cantidad de datos almacenados que gracias a los avances tecnológicos de los últimos años pueden servir de ayuda para predecir hechos futuros y con ello, definir estrategias y planes de actuación. Por ello, el objetivo de la realización de este proyecto es estudiar el comportamiento de datos históricos de intervalo para elaborar una herramienta en Matlab con la que sea posible analizar la evolución de los precios de las acciones junto con el riesgo de cada una de ellas, de cara a poder ayudar a cada inversor a formar su cartera óptima, buscando obtener el mayor rendimiento posible.

1.2 OBJETIVOS DEL PROYECTO

Los objetivos que persigue el proyecto se definen en cada uno de los capítulos que forman el mismo, los cuales se indican a continuación.

- Capítulo 2. Gestión de carteras. Se realiza una introducción a la gestión de carteras, siendo este el proceso que engloba todas las decisiones correspondientes a la formación y administración de una cartera de valores. Se expondrán los conceptos de mayor interés en relación a este tema, que servirán de base para la comprensión y elaboración del resto del proyecto.
- Capítulo 3. Optimización de carteras. En este caso se profundizará sobre el capítulo anterior para definir la idea de la optimización de las carteras para cumplir con los objetivos de retorno y riesgo del inversor. Más adelante, se hará uso de los conceptos teóricos aquí presentados, como el concepto de la beta y doble beta como medidas de riesgo (Nystuen & Nyrnes, 2022), para desarrollar diversas funcionalidades en Matlab.

- Capítulo 4. Datos de intervalo y regresión. El Análisis de Intervalos (AI) es utilizado para representar la incertidumbre en la información y aumentar su precisión, por ello resulta de gran utilidad en los datos financieros (Maté C. , 2012). La regresión con datos de intervalo se utiliza para el modelado, estimación y predicción de datos financieros y se expone en este capítulo con el objetivo de aplicar los conceptos teóricos en el desarrollo práctico de la herramienta.
- Capítulo 5. CAPM y su uso con métodos de regresión para datos de intervalo. En relación con el capítulo anterior, se presenta una adaptación del modelo CAPM en la que los retornos de los activos se representan por variables de intervalos (Gioia, 2009). Se realiza un estudio con diferentes métodos de regresión con intervalos para comprender el concepto.
- Capítulo 6. Métodos y modelos de predicción. Se definen una serie de métodos de predicción de valores que permitirán a los inversores observar la oscilación del precio de las acciones en el futuro (Nowak, 2022). Los métodos propuestos se utilizarán en la herramienta para ofrecer diversos rankings al inversor.
- Capítulo 7. Análisis de la herramienta. Todo lo explicado previamente se lleva a la práctica en la herramienta que se expone en este capítulo y que será ofrecida al usuario final, toda ella elaborada con la funcionalidad de App Designer en Matlab. La herramienta ofrecerá al inversor la información necesaria para escoger su cartera óptima y más tarde se medirá la efectividad de cada uno de los métodos de predicción desarrollados para conocer su comportamiento.
- Capítulo 8. Conclusiones. Se recogerán las principales ideas obtenidas a lo largo del proyecto, así como los resultados más relevantes encontrados en el capítulo anterior.

1.3 METODOLOGÍA DEL PROYECTO

Una vez definida la idea principal del proyecto, se establecen los objetivos que se van a perseguir a lo largo del mismo y se organizan y planifican las tareas a completar para lograrlos. Con ello, es posible decir que el proyecto consta de dos tareas fundamentales: la

primera, el estudio en profundidad de los conceptos teóricos para elaborar toda la documentación relacionada con ellos y que permitirá entender y desarrollar la segunda parte del proyecto, la realización práctica de la herramienta.

Para empezar, se llevará a cabo una profunda investigación de toda la información relacionada a la gestión y optimización de carteras, los datos de intervalo y su uso con métodos de regresión, y los modelos de predicción que más tarde se aplicarán en el estudio de los mencionados datos. Para ello se analizarán libros, artículos o noticias que ofrezcan la información deseada.

Más tarde, se va a hacer uso del lenguaje de programación Matlab para llevar a la práctica los conceptos previos y poder desarrollar un sistema que permita cumplir con los objetivos previamente expuestos, pudiendo facilitar una herramienta de inversión al usuario. Esta herramienta se dividirá en dos secciones: la primera, realizará un análisis de datos pasados de una o dos empresas; la segunda, elaborará una predicción de la rentabilidad de una serie de empresas seleccionadas por el usuario para estimar un ranking que las ordene de mayor a menor rentabilidad partiendo de (González-Rivera, Lee, & Mishra, 2008).

Finalmente, se analizarán los resultados obtenidos para obtener las conclusiones pertinentes que serán recogidas en el último capítulo del proyecto.

En paralelo con todo ello se desarrollará la memoria del proyecto en la que se definen los conceptos más importantes que serán acompañados de diversos ejemplos que faciliten su comprensión, para finalmente exponer la herramienta final. Una vez terminado el trabajo se realizará la presentación y defensa del mismo.

Capítulo 2. GESTIÓN DE CARTERAS

A lo largo de este capítulo se va a introducir lo que se conoce hoy en día como gestión de carteras. Para ello, se definirán conceptos e ideas generales, se describirán los distintos tipos de carteras existentes y se investigarán diferentes herramientas y parámetros que utilizan los inversores y gestores a la hora de evaluar la evolución del comportamiento de una cartera. Además, se describirá el concepto de riesgo en la gestión de carteras y se mostrarán parámetros de medida del nivel de riesgo de un activo o cartera.

2.1 INTRODUCCIÓN A LA GESTIÓN DE CARTERAS

2.1.1 DEFINICIÓN Y OBJETIVOS

La **gestión de carteras** es el proceso por el cual una entidad adopta de forma discrecional e individualizada todas las decisiones relativas a la composición y administración de una cartera de valores de acuerdo con las instrucciones expresadas por el titular de los mismos. Para ello, hacen uso de diferentes herramientas y elaboran la mejor estrategia de acuerdo a las características y objetivos del titular, buscando siempre generar el rendimiento óptimo y minimizar el riesgo, en un periodo de tiempo dado.

La **cartera** queda por tanto representada por la combinación de inversiones de un cliente. Una cartera puede comprender diferentes tipos de activos como pueden ser acciones, deuda, bienes raíces, materias primas o metales, entre otros.

El **gestor de carteras** es la persona que colabora con el cliente para comprender sus necesidades de inversión y se encarga de sugerirle un conjunto de inversiones que encajen con los objetivos que persigue, manteniendo un equilibrio entre el retorno y el riesgo. Normalmente, acude a ellos la gente que no tiene conocimientos sobre el mundo de los mercados de valores y/o las personas que no tienen el tiempo suficiente para dedicarlo a controlar el estado de sus inversiones.

A la hora de gestionar carteras, existen diferentes objetivos que se deben perseguir. El primero de ellos y el principal, tiene que ver con la revalorización del capital, es decir, maximizar la rentabilidad de las inversiones para conseguir aumentar el dinero invertido.

Por otro lado, el gestor de carteras debe tener en cuenta cuál es el objetivo de su cliente a la hora de realizar la inversión o lo que es lo mismo, para qué quiere el cliente utilizar el dinero que gane invirtiendo, lo cual ayudará a saber el tiempo disponible para conseguirlo. Aquí existen un número ilimitado de respuestas que pueden ir desde la compra de una casa hasta el pago de un viaje o los estudios de sus hijos.

Además, se debe buscar la eficiencia a la hora de crear la cartera. Para ello es fundamental entender las condiciones del mercado que constantemente sufre cambios debido a distintos factores como acciones de los gobiernos o cambios en la actividad económica, con tendencias alcistas ("*bull*") o bajistas ("*bear*"). Por ello, cada tipo de activo es más conveniente para unos periodos u otros, por lo que es importante asignar los distintos activos de una manera diversificada y coherente. Durante los **periodos alcistas** se dan condiciones favorables en el mercado, donde los precios aumentan y los inversores son optimistas sobre la continuación de la tendencia creciente de cara al futuro. Por ejemplo, cuando las tasas de empleo son altas en un país y su economía es fuerte. Por otro lado, los **periodos bajistas** representan condiciones desfavorables donde los precios tienden a bajar y, por tanto, los inversores son pesimistas ya que esto implica que la economía se ralentice y que el desempleo aumente, pudiendo llevar a una recesión económica según (Cosín López-Medel, 2019).

Otro de los objetivos clave que se busca a la hora de gestionar inversiones es la mitigación del riesgo manteniendo un nivel de rentabilidad considerable. Para ello hay que tener en cuenta que cuanto mayor sea el porcentaje de retorno esperado mayor será el riesgo existente y cuanto menor sea la rentabilidad que se busca, más pequeño será también el riesgo. Es por ello por lo que es importante buscar un equilibrio en el que ambos sean compatibles.

2.1.2 TIPOS DE GESTIÓN DE CARTERAS

Gestión de carteras activa versus pasiva

- **Gestión de carteras activa:** el gestor participa de forma activa analizando las oportunidades del mercado para buscar la mejor estrategia y conseguir el mayor rendimiento, lo cual conlleva mayores tasas para el cliente. Normalmente, el gestor busca comprar acciones cuando están infravaloradas para venderlas cuando su valor aumenta.
- **Gestión de carteras pasiva:** en este caso el gestor no se encarga de analizar el mercado de forma activa, sino que se crea una estrategia fija que apenas cambia en largos periodos de tiempo. En este caso, prefieren hacer inversiones con retornos bajos y constantes durante periodos más amplios de tiempo, disminuyendo la capacidad de reaccionar a cambios en el mercado.

Gestión de carteras discrecional versus no discrecional

- **Gestión de carteras discrecional**: los gestores de carteras pueden tomar decisiones de inversión de las carteras que llevan y realizar inversiones en nombre de los inversores de acuerdo con lo que consideren que es mejor para el perfil y los objetivos de cada uno.
- **Gestión de carteras no discrecional**: los gestores simplemente pueden aconsejar al inversor sobre lo que creen que es mejor para ellos, pero es el inversor quien tiene la última palabra. En este caso, el gestor no puede llevar a cabo una inversión en nombre de su cliente.

Gestión de carteras tradicional versus moderna

- **Gestión de carteras tradicional:** el objetivo es conseguir un "equilibrio" en la cartera esparciendo las inversiones en diferentes activos y/o bonos. Buscan la diversificación de industrias, invirtiendo en compañías de sectores diferentes. Según explican (Gitman & Joehnk, 2008), este tipo de gestores invierte en acciones de compañías reconocidas que suelen tener éxito y, por tanto, son menos arriesgadas.

Así, consiguen convencer a sus clientes de invertir en ellas en vez de en compañías que no sean tan reconocidas.

- **Gestión de carteras moderna:** utiliza diferentes medidas estadísticas para encontrar una cartera óptima que maximice el retorno esperado dado el grado de riesgo que el cliente está dispuesto a soportar. Las medidas estadísticas utilizadas son el retorno esperado, la desviación estándar de las acciones y de la cartera y la correlación entre las acciones. El primero en introducir este tipo de gestión de carteras fue Harry Markowitz en 1952, quien introdujo el concepto de diversificación de carteras. Esta teoría defiende el uso de una "frontera de eficiencia" en carteras óptimas para maximizar el retorno esperado dependiendo del riesgo, la cual se explica más adelante.

2.1.3 PROCESO SEGUIDO EN LA GESTIÓN DE CARTERAS

A la hora de comenzar con una nueva cartera, el gestor debe seguir una serie de pasos que le lleven a realizar una gestión de la forma más conveniente posible. Los pasos a seguir son los que se detallan a continuación:

1- Objetivo de inversión: deben tener siempre en cuenta cuál es el objetivo que persigue su cliente a la hora de invertir en esa cartera, así como las limitaciones de tiempo y riesgo establecidas para conseguirlo.
2- Perfil del cliente: una vez que el gestor conoce el objetivo del cliente, debe averiguar si este está capacitado para conseguir sus objetivos teniendo en cuenta los requisitos y limitaciones establecidas. Para ello, tendrá que conocer ciertos datos del cliente como sus ingresos propios, sus ingresos familiares o la cantidad que está dispuesto a invertir para conseguir su objetivo. Así, el gestor puede empezar a pensar en la estrategia que quiere seguir.
3- Estimar el mercado: los mercados sufren cambios constantemente y experimentan diferentes estados a lo largo del tiempo. Además, puede haber circunstancias inesperadas que afecten a los mismos. Es por ello por lo que es importante que el

gestor lo analice para ver cuáles son sus puntos fuertes y débiles en cada momento y orientar la inversión hacia el mejor camino.

4- Asignación de activos: el gestor debe decidir qué tipos de activos (metales, materias primas, bienes raíces...) deben ser prioridad para cada cartera dependiendo de los factores estudiados en los pasos anteriores.
5- Selección de la cartera: una vez que se ha decidido el tipo de activos en los que invertir, el gestor debe escoger las compañías o clases concretas en las que hacerlo, así como la cantidad de dinero que va a destinar a cada una.
6- Formulación de la estrategia: aquí se define la estrategia a seguir en la gestión de la cartera (activa, pasiva...) y el parámetro de riesgo que se va a implementar en dicha estrategia.
7- Implementación de la cartera: una vez que las decisiones anteriores han sido tomadas, es hora de pasar a la acción. Es en este punto donde realmente se procede a la ejecución y creación real de la cartera.
8- Evaluación y revisiones: una vez que la inversión se ha realizado, es fundamental que el gestor siga un control de cómo esta está funcionando, teniendo en cuenta los objetivos que definió el cliente en el inicio.
9- Reequilibrio de la cartera: de acuerdo a la evaluación anterior, si algo no está yendo como se esperaba o si se prevé que los objetivos esperados no van a cumplirse, se puede hacer una reestructuración de ciertas inversiones de la cartera y cambiarlas a otro tipo de activos o compañías con las que obtener una mayor rentabilidad.

2.2 RETORNO Y VARIANZA EN CARTERAS DE INVERSIONES

El **retorno** es una medida utilizada por los inversores para medir el comportamiento de un activo o de una cartera. Existen diferentes tipos dependiendo de la situación en la que se encuentren y sirven para ayudar en el proceso de toma de decisiones. Según (Cosín López-Medel, 2019) los más utilizados son el rendimiento de la inversión (ROI, del inglés *Return On Investment*), rendimiento del periodo de tenencia (HPR, del inglés *Holding Period*

Return) y el rendimiento esperado (ER, del inglés *Expected Return*), los cuales se van a explicar a continuación junto con el concepto de varianza.

2.2.1 RENDIMIENTO DE LA INVERSIÓN, ROI

El **ROI** mide la ganancia o pérdida de una inversión y se utiliza para evaluar la eficiencia de la misma o para realizar una comparación de la eficiencia entre dos o más inversiones. Es fácil de calcular y de comprender y se suele expresar en porcentaje. La fórmula utilizada para calcularlo es la siguiente:

$$ROI = \frac{Retorno\ neto\ de\ la\ inversión}{Coste\ de\ la\ inversión} \times 100\%$$

Ecuación 2.1. ROI de un activo

Dentro de una cartera, el ROI depende del peso (w_i) y del retorno (r_i) asociado a cada uno de los n activos que componen la cartera, siendo el peso el porcentaje de inversión en un activo sobre la inversión total de la cartera. Así, el ROI de la cartera en conjunto viene dado por la siguiente expresión:

$$ROI_{cartera} = \sum_{i=1}^{n} w_i \times r_i$$

Ecuación 2.2. ROI de una cartera

2.2.2 RENDIMIENTO DEL PERIODO DE TENENCIA, HPR

El **HPR** se utiliza para obtener el retorno de un activo o de una cartera durante el periodo en el que se tuvieron y viene dato por la ecuación que se muestra a continuación:

$$HPR = \frac{Ingresos + precio\ final - precio\ inicial}{precio\ inicial}$$

Ecuación 2.3. HPR

2.2.3 RENDIMIENTO ESPERADO, ER

El **ER** es el valor del beneficio o de la pérdida que un inversor prevé que va a obtener de una inversión y se calcula haciendo uso rendimientos históricos. Depende de las preferencias del inversor, el nivel de incertidumbre de los activos y del estado de la economía. Existen distintos métodos y modelos para calcular el rendimiento esperado por medio de datos pasados y uno de los más utilizados es mediante el cálculo de la media de los retornos históricos, donde n es el número de periodos que se van a analizar y R_i es el rendimiento obtenido en cada uno de esos periodos.

$$E(R) = \sum_{i=1}^{n} R_i$$

Ecuación 2.4. E(R) de un activo

Así, de acuerdo a esta fórmula es posible calcular el rendimiento esperado de la cartera total, haciendo una media ponderada del peso de cada una de las inversiones en los distintos activos que compongan la misma.

$$E(R)_{cartera} = \sum_{i=1}^{n} w_i \times E(R)_i$$

Ecuación 2.5. E(R) de una cartera

Estas medidas son utilizadas para elegir entre diferentes opciones de inversión, de manera que cuanto mayor sea el rendimiento esperado, mejor será la inversión. De todas formas, es importante tener en cuenta que no todas las inversiones se mantienen con la misma duración, por lo que los inversores hacen uso del rendimiento esperado anualizado, que se calcula mediante la siguiente fórmula.

$$Annualized\ E(R) = \big(1 + E(R)\big)^{\frac{1}{nº\ years}} - 1$$

Ecuación 2.6. E(R) anualizado

De forma que si $E(R)_1 > E(R)_2$, la primera inversión será mejor.

2.2.4 VARIANZA Y DESVIACIÓN ESTÁNDAR DE UNA CARTERA DE INVERSIONES

Otra de las medidas habitualmente utilizadas por los inversores en la gestión de carteras es la varianza o la desviación estándar, siendo esta la raíz cuadrada positiva de la varianza.

La **desviación estándar** muestra la dispersión del retorno medio de la cartera de inversiones y es directamente proporcional al riesgo, es decir, cuanto mayor sea la desviación estándar, más grande será el riesgo de la cartera mientras que cuanto más pequeña sea la desviación, menor riesgo correrá la cartera.

La **varianza** de una cartera de inversiones representa la suma del cuadrado del peso de cada activo (w_i), multiplicado por la varianza del activo (var_i) más el múltiplo de la covarianza entre ese activo y cada uno del resto de otros activos ($covar_{i,j}$) por sus pesos al cuadrado.

$$Var = \sum_{i=1}^{n} w_i^2 \times var_i + \sum_{i=1}^{n}\sum_{j=1}^{n} w_i^2 \times w_j^2 \times covar_{i,j} \quad when\ i \neq j$$

Ecuación 2.7. Varianza de una cartera

2.3 RIESGO EN CARTERAS DE INVERSIÓN, EL USO DE BETA (β) Y LA DIVERSIFICACIÓN

Siempre que alguien está negociando en un mercado existe incertidumbre sobre el valor futuro de las acciones, lo cual irremediablemente implica que haya un riesgo. El **riesgo de inversión** es la oportunidad de que el resultado real de una inversión difiera del resultado esperado. El retorno total de una inversión depende del retorno esperado, pero también del inesperado. Mientras que el primero es calculable gracias al uso de distintas herramientas, el segundo no es posible de conocer, puede ser tanto positivo como negativo y proviene de imprevistos con los que no se cuenta a la hora de calcular el retorno esperado.

El riesgo de la inversión depende de la variabilidad de los posibles resultados, por lo que cuanto mayor sea la variabilidad, mayor será el riesgo existente. Para analizar y estimar el riesgo, se utilizan diferentes medidas como la varianza, la desviación estándar, la volatilidad, el valor en riesgo y el ratio de Sharpe.

2.3.1 RIESGO SISTEMÁTICO Y RIESGO ESPECÍFICO

A la hora de invertir, existen dos tipos de riesgo a tener en cuenta, el riesgo sistemático (en inglés, *systematic risk*) y riesgo específico (*unsystematic risk* en inglés).

El **riesgo sistemático** afecta a un gran número de activos y es un riesgo residual que no puede diversificarse, aunque los valores individuales de la cartera estén óptimamente mezclados. Es el que viene de cambios en el PIB, tasas de interés, inflación, desastres naturales u otros factores que influyen en los precios de los activos y que no se pueden evitar.

El **riesgo específico o no sistemático** influye únicamente en un pequeño número de activos o incluso en uno solo. Este tipo de riesgo es intrínseco al activo en cuestión y es posible reducirlo mediante diversificación. Este tipo de riesgo puede venir de una causa interna de un activo concreto como una mala actuación o comportamiento o una buena inversión de la compañía. También puede venir de una causa externa como, por ejemplo, el cambio en el precio del petróleo que afectaría a las compañías de transporte.

En la siguiente imagen se puede apreciar gráficamente la relación entre el número de activos y el riesgo total, diferenciando los dos tipos de riesgo explicados.

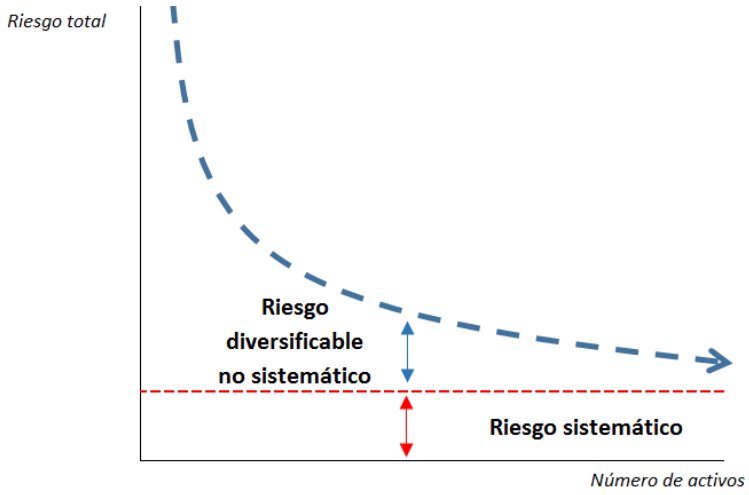


Ilustración 9. Riesgo sistemático y específico (López, 2020)

Para finalizar, el retorno de una inversión viene dado por el retorno esperado más la parte procedente del riesgo esperado y la del específico, según la siguiente ecuación:

$$R = E(R) + riesgo\ sistemático + riesgo\ específico$$

Ecuación 2.8. Retorno de una inversión

2.3.2 RELACIÓN ENTRE RIESGO Y RETORNO Y SML

El riesgo y el retorno están directamente relacionados en el mundo de las inversiones. Por ello, una inversión con un rendimiento esperado alto tiene mucho más riesgo que una con un rendimiento esperado más bajo, lo cual se debe a la incertidumbre del mercado de inversiones y su elevada dispersión.

La **Línea del Mercado de Valores** (SML, del inglés *Security Market Line*) representa la tasa esperada de retorno de un valor individual en función del riesgo sistemático. Es una línea recta con pendiente positiva que muestra la relación entre el retorno esperado y el riesgo sistemático, siendo β el coeficiente que mide el riesgo sistemático. Los inversores hacen uso de esta herramienta para conocer la recompensa de sus inversiones frente al riesgo, y debería ser la misma para todos los activos del mercado.

Normalmente, la SML se utiliza para comparar dos tipos de inversiones diferentes ya que pueden ver cuál de ellas tiene una ratio mejor de recompensa frente al riesgo puesto que unas pueden estar sobrevaloradas mientras que otras estarán infravaloradas. Finalmente, los inversores elegirán aquellos activos que se encuentren por encima de la SML en la siguiente representación debido a que serán las que les permitan obtener una mayor recompensa para un riesgo dado.

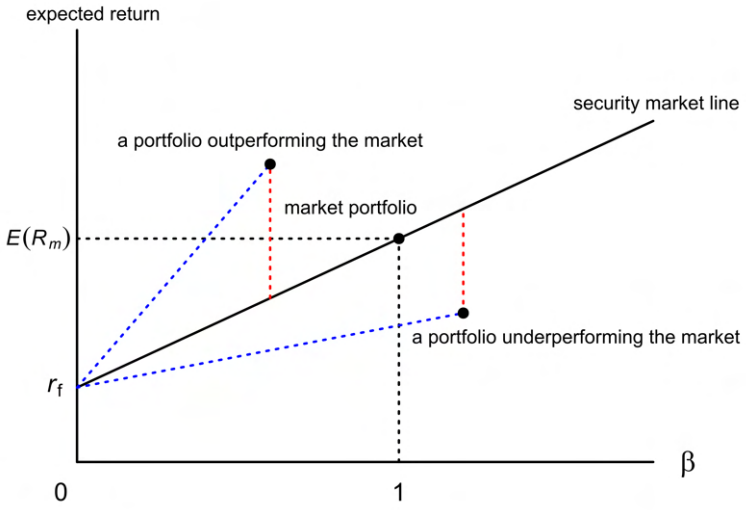


Ilustración 10. SML (Wikipedia, 2022)

La pendiente de la SML se calcula mediante la siguiente ecuación:

$$SML_{pendiente} = \frac{E(R_M) - R_f}{\beta_M} = \frac{E(R_M) - R_f}{1} = E(R_M) - R_f$$

Ecuación 2.9. Pendiente de la SML

Donde $E(R_M)$ es el retorno esperado del mercado, β_M es la beta o riesgo sistemático del mercado y R_f es la tasa de interés libre de riesgo.

2.3.3 EL USO DE BETA (β)

Beta es un factor utilizado para medir el riesgo sistemático de un activo o de una cartera en comparación con una referencia. Beta indica la volatilidad de un valor con respecto a un índice, es decir, indica si un valor sube o baja en mayor o menor medida que el índice del mercado, cuya beta es 1 y representa el retorno medio de todas las acciones o de un número muy grande de ellas como el IBEX 35 o el S&P 500.

En caso de que un activo tenga un valor de beta superior a 1, este activo será más volátil que el mercado, es decir, una variación en el mercado provoca una variación mayor en este activo, lo cual es favorable cuando el mercado está atravesando una etapa de crecimiento, siendo desfavorable si atraviesa una etapa de recesión.

Si la beta del activo es igual a la del mercado, 1, una variación en el mercado provocará la misma variación en ese activo

Por último, si la beta es inferior a 1, una variación en el mercado implica una variación menor en el activo, lo cual será favorable en periodos de recesión, pero desfavorable en periodos de crecimiento.

Además, el retorno esperado depende del valor de las betas, de manera que los activos con mayores betas tendrán mayor riesgo sistemático y, por tanto, sus retornos esperados serán mayores.

La forma más común de estimar la beta de un activo (β_{activo}) es mediante la división de la covarianza de la rentabilidad del activo y la del índice del mercado ($r_{activo}, r_{mercado}$) entre la varianza de la rentabilidad mercado ($r_{mercado}$).

$$\beta_{activo} = \frac{cov(r_{activo}, r_{mercado})}{var(r_{mercado})}$$

Ecuación 2.10. Beta de un activo

Por otro lado, la beta de la cartera completa ($\beta_{cartera}$) viene dada por la beta (β_i) ponderada según su peso (w_i) de cada uno de los n títulos que forman la cartera.

$$\beta_{cartera} = \sum_{i=1}^{n} w_i \times \beta_i$$

Ecuación 2.11. Beta de una cartera

2.3.4 DIVERSIFICACIÓN

De cara a reducir el riesgo los inversores apuestan por tener una cartera compuesta por activos de diferentes tipos en vez de tener activos de una sola clase, buscando un equilibrio entre obtener la máxima recompensa con el mínimo riesgo posible. Esto se conoce con el nombre de **diversificación**. Para reducir el riesgo, según (Gitman & Joehnk, 2008), una de las maneras más efectivas es la combinación de activos que están correlados negativamente ya que la variabilidad se reduce.

Existen diferentes formas de diversificación para reducir el riesgo:

- Diversificación de compañías: se trata de tener acciones de compañías diferentes.
- Diversificación de industrias: consiste en repartir la inversión en acciones de compañías pertenecientes a diferentes industrias y/o mercados.
- Diversificación de tipos de activos: busca invertir en activos de diferentes clases como acciones, bonos o materias primas.
- Diversificación geográfica: trata de invertir en compañías de distintos países y puede ser muy efectiva debido a que la economía de cada país crece de manera diferente.

Teniendo esto en cuenta, es importante saber buscar un equilibrio entre diversificar para reducir el riesgo sistemático y escoger un número razonable de activos en los que invertir ya que cuanto mayor sea el número de activos que se manejan, más complicado será hacer un seguimiento de todos ellos. Por ello, existen estudios que defienden que las carteras mejor diversificadas deben tener entre 20 y 30 activos ya que, según se muestra en la gráfica

siguiente, llega un momento en el que añadir nuevos activos disminuye de manera casi inapreciable la desviación estándar, por lo que deja de merecer la pena ya que añade trabajo.

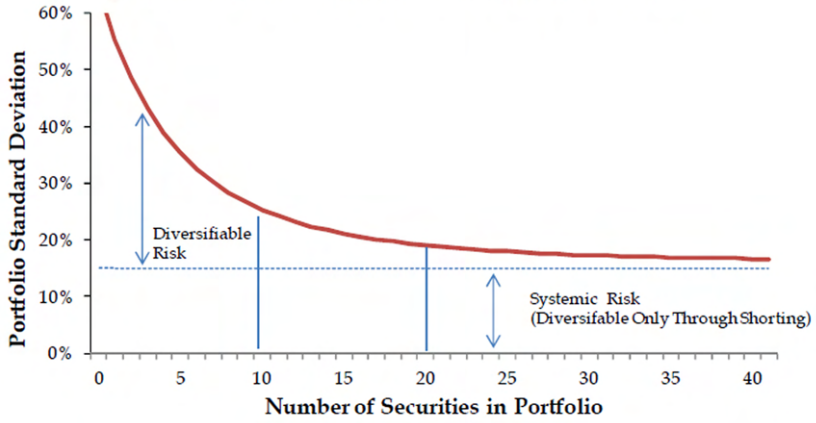


Ilustración 11. Desviación estándar versus número de activos en la cartera (Easy Portfolio, 2019)

2.3.5 MEDIDAS DE RIESGO

Es de vital importancia tener en cuenta ciertas medidas de riesgo a la hora de optimizar una cartera. Por un lado, existen la varianza y la desviación estándar que ya se han explicado en secciones anteriores. Sin embargo, existen otras medidas más sofisticadas que se han desarrollado para proporcionar más información sobre el riesgo a los inversores para ayudarles a optimizar sus carteras. Las más comunes son las siguientes.

- **Volatilidad:** medida estadística de dispersión que muestra el grado de variación de los precios de las acciones a lo largo de un periodo de tiempo determinado. A continuación, se muestra la fórmula para calcularla, donde R_i es el rendimiento de cada uno de los n activos y $\bar{R}$ es el rendimiento medio.

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

Ecuación 2.12. Volatilidad

- **Valor en riesgo:** (VaR, del inglés *Value at Risk*) mide la pérdida (X) que una inversión podría sufrir en condiciones normales de mercado en un intervalo de tiempo y con un cierto nivel de probabilidad o confianza (β). Su fórmula es la siguiente:

$$VaR_\beta = \min\{\alpha \in \mathbb{R}: P(X > x) \leq \alpha\} \quad donde\ \alpha = (1 - \beta)$$

Ecuación 2.13. Valor en riesgo (VaR)

- **Ratio de Sharpe:** fue definido por el Premio Nóbel William F. Sharpe, de la Universidad de Stanford, y mide la relación entre la rentabilidad y la desviación estándar de una inversión. El valor del ratio aumenta cuando el riesgo disminuye debido a la diversificación, por lo que cuantos más activos tenga la cartera, mayor será el valor del ratio de Sharpe.

  Además, cuanto mayor sea el valor del ratio de Sharpe, mejor será la relación de la recompensa frente al riesgo. La fórmula utilizada para calcularlo se muestra a continuación, donde R_p es la rentabilidad de la cartera en el periodo de análisis, R_f es la rentabilidad del activo libre de riesgo en ese mismo periodo y σ_p representa la desviación típica de la rentabilidad de la cartera.

$$Ratio\ de\ Sharpe = S = \frac{R_p - R_f}{\sigma_p}$$

Ecuación 2.14. Ratio de Sharpe

2.4 MODELOS CAPM

El **Modelo de Valoración de Activos Financieros** (CAPM, del inglés *Capital Asset Pricing Model*) muestra la relación entre el rendimiento esperado de un activo y su beta, por lo que permite calcular la rentabilidad que un inversor debe exigir al realizar una inversión en un activo financiero en función del riesgo que está asumiendo.

El modelo CAPM no solo se utiliza para establecer el precio de los activos, sino que también tiene amplia utilidad en las carteras. La ecuación utilizada para calcular el CAPM es la siguiente:

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

Ecuación 2.15. CAPM

donde R_f es la tasa libre de riesgo, $E(R_M)$ es el rendimiento esperado del mercado y β_i es la beta de la inversión.

Gracias a estos modelos, los inversores pueden comparar carteras diferentes añadiendo o eliminando activos de ellas para encontrar aquella que ofrezca un retorno mayor para un riesgo determinado que estén dispuesto a asumir.

En el Capítulo 5. se habla de los modelos CAPM con mayor profundidad, estudiando su aplicación para datos de intervalo.

2.5 TOLERANCIA AL RIESGO

La **tolerancia al riesgo** mide el grado en el cual el inversor acepta cierto riesgo a cambio de obtener un retorno mayor. Como se ha explicado en secciones anteriores, es importante que los inversores definan la tolerancia que tienen frente al riesgo antes de desarrollar sus estrategias de inversión, lo cual depende de los objetivos y de la personalidad del inversor o de su cliente. Existen tres tipos distintos de perfiles según su tolerancia frente al riesgo.

- Inversores conservadores: suelen presentar aversión hacia el riesgo, es decir, buscan una volatilidad baja para conseguir retornos bajos y alto nivel de liquidez, centrándose en la estabilidad de la cartera.
- Inversores moderados: toman un nivel de riesgo justo en sus inversiones, buscando un equilibrio entre la estabilidad de la cartera y su apreciación. Combinan instrumentos de alta volatilidad con otros de volatilidad más baja como los bonos.

- Inversores agresivos: están dispuestos a aceptar unos niveles mayores de riesgo a cambio de obtener una más alta recompensa. Combinan activos de alta volatilidad con la posibilidad de obtener pérdidas importantes.

Es importante que los inversores tengan una idea realista de su tolerancia frente al riesgo a la hora de asignar los activos en los que van a invertir para escoger el riesgo deseado y poder comprar o vender en el momento correcto.

Además, según (Hvaerland, 2019), la tolerancia al riesgo de los inversores también depende del horizonte de tiempo que se esté evaluando, es decir, el periodo de tiempo en el que se va a mantener la inversión antes de ser liquidada. Normalmente, los inversores con horizontes de tiempo más largos están dispuestos a asumir riesgos mayores para más altas recompensas mientras que los inversores con horizontes de tiempo más cortos suelen presentar mayor aversión al riesgo.

2.6 ASIGNACIÓN DE ACTIVOS

La **asignación de activos** es la estrategia financiera que conlleva seleccionar las clases de activos en las que se va a invertir, así como los porcentajes de la inversión total que se van a destinar a cada una, buscando un equilibrio entre el riesgo que se está dispuesto a asumir para alcanzar el rendimiento deseado. Existen dos tipos de comportamientos a la hora de asignar los activos, asignación estratégica y asignación táctica.

Asignación de activos estratégica versus táctica

- **Asignación de activos estratégica:** conlleva seleccionar los tipos de activos y sus proporciones para completar la cartera, destinada a permanecer estable durante un largo periodo de tiempo y haciendo revisiones periódicas para re-equilibrar los pesos en caso de que los valores de mercado cambien el precio de los activos.
- **Asignación de activos táctica:** en este caso, el reparto de activos se hace de forma temporal en función de comportamientos de los mercados o sectores fuertes de ellos.

Se busca tomar las ventajas de las oportunidades económicas que puedan sufrir y los porcentajes asignados a cada activo pueden ir variando, tratando de mantener constantes los objetivos de riesgo y rendimiento.

2.7 OPTIMIZACIÓN DE LA CARTERA

El proceso de **optimización de una cartera** consiste en la generación de diferentes carteras para elegir aquella que mejor cumpla con los objetivos de retorno y riesgo del inversor y se explica en profundidad en el Capítulo 3.

2.8 FONDOS DE INVERSIÓN

Todo lo anterior tiene su aplicación directa en los fondos de inversión. Un **fondo de inversión** es un grupo de inversores que manejan de manera colectiva una cartera diversificada de valores. Para ello, buscan reunir fondos de distintos inversores para invertirlos en diferentes instrumentos financieros. Así, cada inversor debe elegir el fondo que mejor se adapte a sus requisitos de retorno y riesgo a la hora de hacer una inversión.

Una de las principales ventajas que ofrecen los fondos reside en la profesionalidad de sus gestores, que, debido a su experiencia y conocimiento sobre los mercados, normalmente ofrecerán mejores rendimientos para riesgos dados. Además, los inversores se pueden beneficiar del concepto de "economía de escala" y acabar pagando menos tasas de las que pagarían invirtiendo de forma individual. Por último, debido a que existe capacidad para crear una cartera con gran diversificación mediante inversiones en compañías de diferentes industrias, el riesgo de la inversión queda reducido.

2.9 TECNOLOGÍAS PARA LA GESTIÓN DE CARTERAS

Los mercados financieros se mueven a una muy alta velocidad. Para tomar decisiones informadas, los inversores necesitan datos, noticias e investigaciones precisas, rápidas y relevantes, además de tecnología especial para analizar cada riesgo posible, encontrar las

mejores oportunidades y comunicarse con otros actores del mercado en todo el mundo. Hoy en día existen herramientas capaces de poner todo esto al alcance de los profesionales. Para ello, utilizan el poder de la tecnología para llevar transparencia a los mercados de capitales y permitir a los distintos actores ver los mismos datos en tiempo real.

A continuación, se van a presentar una serie de herramientas de software que ofrecen información financiera de las cuales se ha hecho uso para realizar diferentes análisis y obtener datos de cara a la elaboración de este proyecto.

2.9.1 FACTSET

FactSet es una compañía multinacional de datos financieros y software con su centro de operaciones situado en Estados Unidos, fundada en 1978. La compañía provee una plataforma con información financiera y software analítico para profesionales de inversiones y ofrece acceso de análisis de datos a los analistas, gerentes de administración de cartera y banqueros de inversión en instituciones financieras. Además, consolida datos de mercados globales, empresas públicas y privadas y carteras de renta fija y variable (Wikipedia, 2022).

Los negocios de FactSet se centran en tecnología y servicio de atención al cliente y su trabajo consiste en combinar bases dispares de datos en una sola fuente, además de trabajar en el desarrollo de software.

2.9.1.1 Productos y Servicios

FactSet provee información financiera y aplicaciones analíticas a profesionales globales que venden y compran instrumentos financieros, como asesores de cartera de valores, analistas de bolsa o gerentes de riesgo, entre otros. La plataforma de software incluye noticias y cotizaciones en tiempo real, análisis de compañía y de cartera de valores, análisis de industria, optimización y simulación de cartera de valores y herramientas para analizar valores y carteras de valores.

Se puede acceder a FactSet a través de dispositivos inalámbricos como los teléfonos móviles y el software es integrado con aplicaciones de Microsoft Office.

2.9.1.2 Funcionamiento

El objetivo de FactSet es ayudar al desempeño de los profesionales de la inversión en todo el mundo, proporcionando información analítica a partir de datos financieros brutos. Para ello, ofrece a sus clientes las herramientas que necesitan para convertir estos datos en información que pueda usarse para el análisis financiero y la elaboración de informes. En la siguiente imagen se puede observar un ejemplo del aspecto que tiene la herramienta a la hora de navegar en ella (FactSet, 2023).

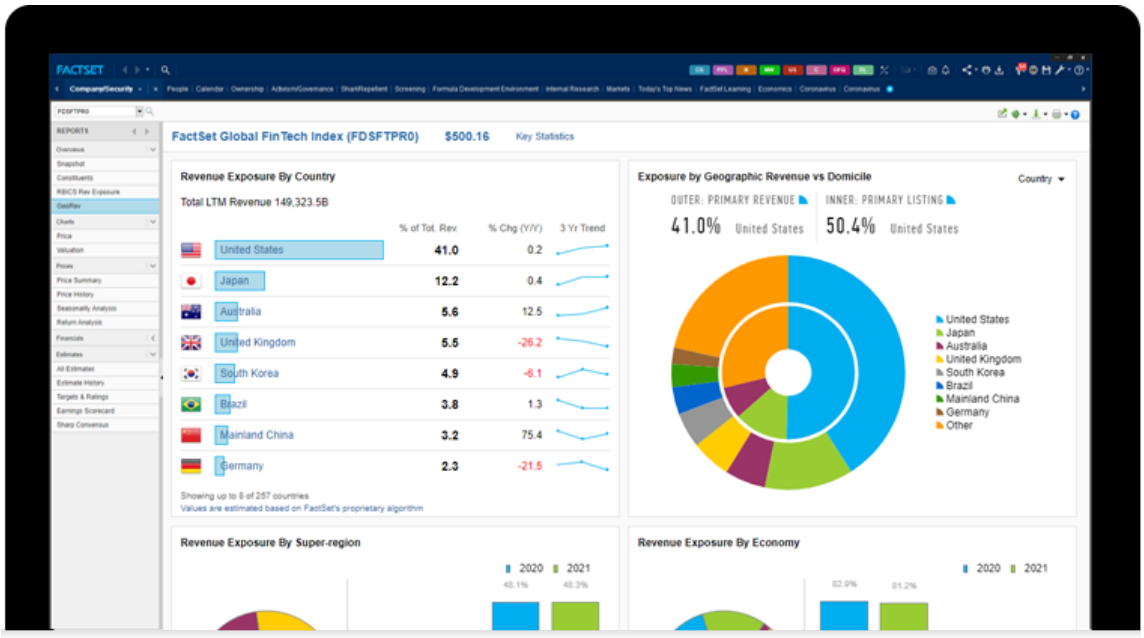


Ilustración 12. FactSet (FactSet, 2023)

A continuación, se puede apreciar un ejemplo de cómo se muestran los datos históricos de mercado para el caso de las acciones de Apple.

Ilustración 13. Ejemplo de Apple en FactSet (FactSet, 2023)

2.9.1.3 Gestión de carteras

Una de las soluciones que FactSet ofrece a sus clientes tiene que ver con la gestión de carteras, proporcionando ayuda para optimizar el proceso de inversión ya sea desde el análisis de empresas, la supervisión de carteras globales o la gestión de ejecución de múltiples activos.

Por un lado, permite gestionar la cartera buscando la maximización de su rendimiento mediante el análisis y la ejecución de ideas de inversión de forma rápida. Para ello, hace uso de *insights* obtenidos a partir de los datos que maneja y con los que opera, además de información como el riesgo o su comportamiento, así como de un sistema de predicción que permite realizar la mejor combinación para obtener el rendimiento más alto posible (FactSet, 2023).

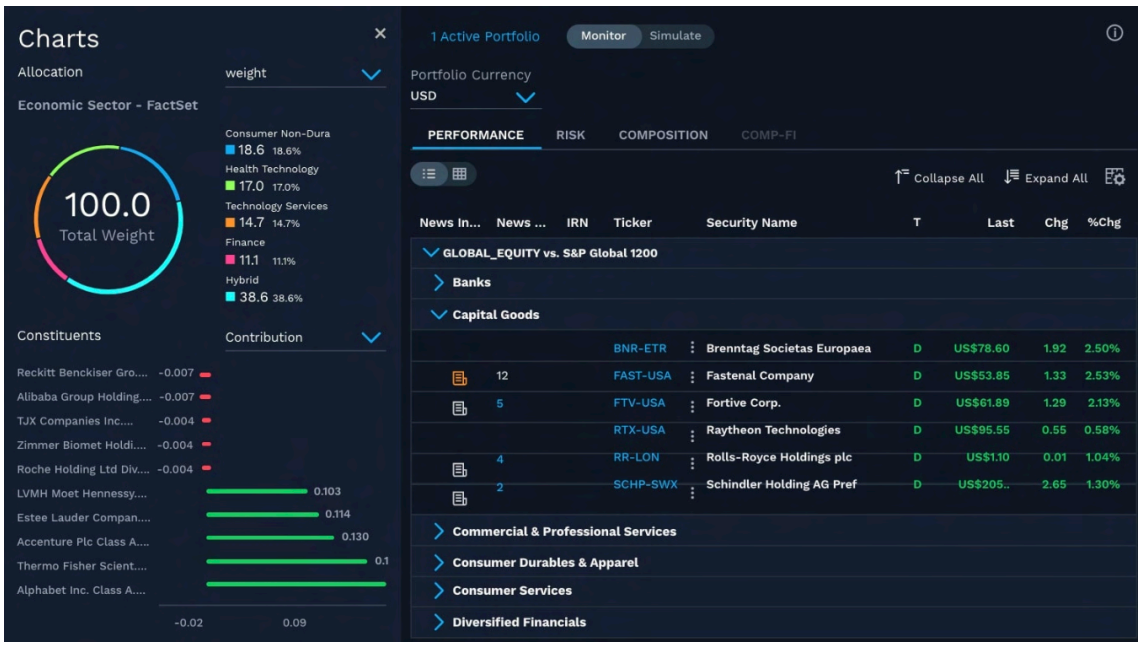


Ilustración 14. Gestión de carteras en FactSet (FactSet, 2023)

Por otro lado, FactSet permite construir carteras óptimas utilizando una amplia gama de datos de mercado integrados y centralizados. Las herramientas de optimización de FactSet ayudan a los clientes a crear y probar fácilmente estrategias que incluyen alfas propias, límites de inversión en activos y sectores, así como restricciones definidas en cualquier aspecto de la gestión de carteras para garantizar el cumplimiento de los requisitos de los fondos.

2.9.2 BLOOMBERG

Bloomberg es una empresa fundada en 1981 en Nueva York dedicada a ofrecer servicios de información financiera. Su producto emblemático es el **Terminal Bloomberg**, que es un sistema informatizado que permite a los profesionales de las finanzas y otros sectores acceder a los servicios que ofrece Bloomberg para monitorizar y analizar en tiempo real los mercados financieros, además de efectuar operaciones de compraventa de activos en su plataforma electrónica. Asimismo, el servicio ofrece noticias, comunicación entre terminales y diferentes aplicaciones que facilitan la labor de los intervinientes en los mercados financieros (Wilipedia, 2022).

Cabe destacar que la mayor parte de las empresas del mundo financiero están suscritas al servicio Bloomberg y sus terminales se han convertido en un parte imprescindible para desarrollar su actividad.

2.9.2.1 Productos y servicios

La compañía proporciona herramientas de software financiero, tales como análisis y plataformas de comercio de capital, servicio de datos y noticias para las empresas financieras y organizaciones en todo el mundo a través de la Terminal Bloomberg, su producto base de ganancias. Además, Bloomberg ha crecido para incluir un servicio mundial de noticias que incluye televisión, radio, internet y otras publicaciones (Bloomberg, 2023). En la siguiente imagen se muestra un ejemplo de cómo se ve la terminal Bloomberg a la hora de navegar por ella.

Ilustración 15. Terminal Bloomberg (Fortuño, 2022)

2.9.2.2 Gestión de carteras

En relación con este proyecto, Bloomberg también puede ser de utilidad para el manejo de las carteras. La complejidad de productos y la volatilidad de los mercados pueden dificultar el equilibrio entre el rendimiento y el riesgo. La solución de Análisis de Riesgos y Cartera de Bloomberg ofrece las herramientas necesarias para implementar con éxito estrategias de gestión de riesgos e inversión de cartera. Esta plataforma ayuda a tomar

decisiones más rápidas y documentadas y permite medir la exposición a riesgos específicos de seguridad y relacionados con el mercado para las carteras.

Más allá de las pérdidas y ganancias de una cartera y su rendimiento contrastado con los parámetros pertinentes, existen otras preocupaciones adicionales de liquidez y flujo de efectivo, cambios regulatorios y acontecimientos del mercado. Desde benchmarking hasta *backtesting* y de procesos de revisión a pruebas de resistencia, las herramientas de análisis de riesgos y cartera que Bloomberg ofrece proporcionan acceso inmediato a un conocimiento profundo para conceder una mayor claridad a las decisiones de los gestores. Para ello, Bloomberg cuenta con una herramienta denominada **PORT** para el análisis de carteras globales de múltiples activos y riesgo de la industria (Bloomberg, 2023).

PORT brinda las herramientas necesarias para implementar con éxito estrategias óptimas de carteras de inversiones, permitiendo medir de manera fácil y precisa el riesgo y rentabilidad de la cartera, ofreciendo una perspectiva integral de la misma (pasado, presente y futuro), optimizando así su flujo de trabajo general de gestión de inversiones. En la siguiente imagen se muestra un ejemplo de la vista de PORT en Bloomberg.

Ilustración 16. Bloomberg PORT (Ellis, 2019)

Capítulo 3. OPTIMIZACIÓN DE CARTERAS

A lo largo de este capítulo se van a cubrir conceptos relacionados con los activos de inversión y con ellos, la creación de carteras de inversión y su optimización, identificando diferentes términos clave a la hora de trabajar con estas ideas.

3.1 INTRODUCCIÓN A LA OPTIMIZACIÓN DE CARTERAS

La. **optimización de una cartera** consiste en la creación de diferentes carteras para elegir aquella que mejor cumpla con los objetivos de retorno y riesgo del inversor.

3.1.1 ACTIVOS DE INVERSIÓN

Antes de comenzar a hablar de carteras es preciso conocer lo que es un activo de inversión. Según (Bertani, 2021), un **activo** es un instrumento de inversión que puede ser comprado y vendido y que normalmente se adquiere con el objetivo de obtener beneficios a partir de él. Algunos ejemplos de activos pueden ser acciones, bonos, ETFs o futuros. Existen varios conceptos que se deben considerar a la hora de hablar de los activos y dos de los más importantes son el precio de apertura y el precio de cierre. El **precio de apertura** es el precio que tiene el activo en el momento en el que se abre el mercado mientras que el **precio de cierre** representa el precio que posee el activo a la hora de cerrar el mercado.

En cuanto a la utilidad de un activo, esta se mide por su rendimiento. El **rendimiento total** de un activo representa la rentabilidad que va a tener el activo, es decir, indica el beneficio que se puede obtener a partir de ese activo y viene dado por la siguiente fórmula.

$$R = \frac{precio\ de\ cierre}{precio\ de\ apertura}$$

Ecuación 3.1. Rendimiento de un activo

Por otro lado, la **tasa de rendimiento** de un activo indica el comportamiento de la inversión a lo largo de un periodo de tiempo, expresado como un porcentaje de su coste inicial y se representa con la siguiente expresión.

$$r = \frac{precio\ de\ cierre - precio\ de\ aprtrura}{precio\ de\ cierre}$$

Ecuación 3.2. Tasa de rendimiento de un activo

Los dos conceptos anteriores se relacionan mediante la siguiente ecuación.

$$R = 1 + r$$

Ecuación 3.3. Relación entre rendimiento y tasa de rendimiento de un activo

Otra idea a tener en cuenta es el **retorno acumulado**, el cual consiste en la cantidad total ganada o perdida a partir de una inversión a lo largo del tiempo, independientemente de la longitud del periodo en cuestión. Así, dada una secuencia de tasas de retorno r_i para $i = 1, \dots, n$, el retorno acumulado se define mediante la siguiente expresión.

$$C_r = \left(\prod_{i=1}^{n} 1 + r_i \right) - 1$$

Ecuación 3.4. Retorno acumulado

Por otra parte, cabe destacar el concepto de **apalancamiento**, el cual consiste en utilizar capital prestado para aumentar el rendimiento potencial de una inversión, con el riesgo de que esto también supondría multiplicar las potenciales pérdidas en el caso de que la inversión no tuviera éxito. Suponiendo la compra de un activo al precio X_0 y teniendo en cuenta el multiplicador m, tomamos un capital prestado igual a $(m-1)X_0$ para adquirir una cantidad mX_0 de un activo. Finalmente, el activo será vendido a un precio mX_1, siendo el gasto inicial únicamente X_0, con un rendimiento total igual al siguiente.

$$rendimiento\ total\ del\ apalancamiento = m\frac{X_1}{X_0}$$

Ecuación 3.5. Rendimiento total del apalancamiento

Por último, el capital prestado debe ser devuelto, pero como resultado el rendimiento obtenido se ha multiplicado por un factor m.

3.1.2 CARTERAS DE INVERSIÓN

Una **cartera de inversión**, como se ha mencionado previamente, es una colección de diferentes inversiones financieras que son los activos anteriormente mencionados. Suponiendo que se tienen n activos disponibles, es posible crear una cartera a partir de la inversión de una cantidad específica de dinero en cada uno de ellos. Así, la cantidad invertida en el activo i para $i = 1, 2, \dots, n$ viene dada por X_{0i} de forma que $\sum_{i=1}^{n} X_{0i} = X_0$, siendo X_0 la inversión total inicial.

Según esto, la cantidad invertida en cada activo se puede expresar como una fracción de la inversión total, donde w_i representa el **peso** del activo i en la cartera, por lo que debe cumplirse $\sum_{i=1}^{n} w_i = 1$. Así, la inversión en cada activo se indica mediante la siguiente expresión.

$$X_{0i} = w_i X_0, \qquad i = 1, 2, \dots, n$$

Ecuación 3.6. Inversión en un activo dentro de una cartera

Por último, siendo R_i el rendimiento total de un activo i, la cantidad de dinero generada a lo largo del periodo por ese activo es $R_i X_{0i} = R_i w_i X_0$ y, por tanto, la cantidad total generada por la cartera es igual a $\sum_{i=1}^{n} R_i w_i X_0$. Como resultado, el retorno total de la cartera es la suma ponderada de los retornos de cada uno de sus componentes, según se muestra a continuación.

$$R = \frac{\sum_{i=1}^{n} R_i w_i X_0}{X_0} = \sum_{i=1}^{n} R_i w_i$$

Ecuación 3.7. Retorno de una cartera

3.1.3 VOLATILIDAD DE UN ACTIVO

A la hora de comprar un activo es imposible saber el dinero por el que lo vamos a vender en el futuro y es por ello por lo que su rendimiento es desconocido en ese momento. Por tanto, se hace uso del **valor esperado** o **valor medio** para indicar el rendimiento que por probabilidad se espera tener de ese activo. Adicionalmente, suele ser de interés conocer también el grado de desviación posible de ese valor medio, lo cual se conoce como **varianza**. Así, para una variable aleatoria y, su varianza se define mediante la siguiente expresión:

$$var(y) = \mathbb{E}[(y - \bar{y})^2]$$

Ecuación 3.8. Varianza de un activo

Además, de manera matemática, la varianza se expresa con el símbolo σ^2, por lo que $\sigma_y^2 = var(y)$ o simplemente, $\sigma^2 = var(y)$.

Por otro lado, se hace uso de la raíz cuadrada de la varianza, denominada **desviación típica** y denotada como σ, como medida de probabilidad de que una variable se desvíe de su valor esperado. Así, su expresión quedaría de la siguiente forma:

$$\sigma_y = \sqrt{\mathbb{E}[(y - \bar{y})^2]}$$

Ecuación 3.9. Desviación típica de un activo

Cabe destacar que el término **volatilidad**, en el mundo de las finanzas, se utiliza tanto para indicar la varianza como la desviación típica de un activo.

3.1.4 COVARIANZA DE UNA CARTERA

A la hora de considerar dos o más variables aleatorias, como los rendimientos de las acciones de una cartera, existe una dependencia entre ellas que se identifica mediante su

covarianza. Así, siendo x_1 y x_2 dos rendimientos de activos con sus rendimientos esperados $\overline{x_1}$ y $\overline{x_2}$, la covarianza entre ellos se puede definir como:

$$cov(x_1, x_2) = \mathbb{E}[(x_1 - \overline{x_1})(x_2 - \overline{x_2})]$$

Ecuación 3.10. Covarianza de una cartera

Además, la covarianza se puede denotar como $ccov(x_1, x_2) = \sigma_{x1,x2}$. Así, en el caso de que dos activos x_1 y x_2 tengan $\sigma_{x1,x2} = 0$ diremos que no están correlacionados, es decir, que el movimiento en el precio de uno de ellos no tiene un efecto notable sobre el precio del otro.

3.1.5 DIVERSIFICACIÓN

A continuación, se va a determinar la varianza de la tasa de rendimiento de una cartera. Se supone la varianza del rendimiento de un activo i denotada como σ_i^2, la varianza de la tasa de rendimiento de una cartera denotada como σ^2 y la covarianza del retorno de un activo i con otro activo j definida por σ_{ij}^2. Se tiene que la varianza del rendimiento de una cartera se puede calcular a partir de las covarianzas del rendimiento de parejas de activos y las proporciones de estos activos en la cartera.

$$\sigma^2 = \mathbb{E}[(r - \bar{r})^2]$$

$$= \mathbb{E}\left[\left(\sum_{i=1}^{n} w_i r_i - \sum_{i=1}^{n} w_i \overline{r_i}\right)^2\right]$$

$$= \mathbb{E}\left[\left(\sum_{i=1}^{n} w_i (r_i - \overline{r_i})\right)\left(\sum_{j=1}^{n} w_j (r_j - \overline{r_j})\right)\right]$$

$$= \mathbb{E}\left[\sum_{i,j=1}^{n} w_i w_j (r_i - \overline{r_i})(r_j - \overline{r_j})\right]$$

$$= \sum_{i,j=1}^{n} w_i w_j \sigma_{ij}$$

Ecuación 3.11. Varianza de la tasa de rendimiento de una cartera

Cuando se trabaja con la gestión de carteras es importante tener en cuenta el concepto de **diversificación**. Una cartera con poca variedad de activos está sujeta a un nivel alto de riesgo, representado por una varianza relativamente alta. Esta varianza puede reducirse al incluir activos adicionales en la cartera, lo cual se conoce como diversificación.

Se tienen n activos de forma que ninguno de ellos está correlacionado con otro activo del grupo y suponiendo que la tasa de rendimiento de cada activo tiene media m y varianza σ^2. Se construye una cartera tomando una proporción igual de cada uno de los activos, de forma que $w_i = \frac{1}{n} \forall i$. La tasa de rendimiento total de la cartera se define por la siguiente expresión.

$$r = \frac{1}{n} \sum_{i=1}^{n} r_i$$

Ecuación 3.12. Tasa de rendimiento de una cartera

Así, la varianza de r se indica a continuación.

$$\sigma(r) = \frac{1}{n^2} \sum_{i=1}^{n} \sigma^2 = \frac{\sigma^2}{n}$$

Ecuación 3.13. Varianza de r

Según lo anterior, la varianza decrece de forma rápida cuando n aumenta. Por otro lado, se puede dar el caso de que los activos estén correlados parcialmente, por lo que el resultado sería distintos y al añadir más activos, la varianza disminuiría más lentamente como se explica a continuación.

Considerando que cada activo tiene una tasa de rendimiento con media m y varianza σ^2 y ahora con una correlación definida por un factor t entre cada pareja de rendimientos, se tiene una covarianza $cov(r_i, r_j) = t\sigma^2$ para $i \neq j$. Si se compone la cartera por proporciones iguales de n de estos activos se tiene una varianza definida por la siguiente expresión.

$$var(r) = \mathbb{E}\left[\sum_{i=1}^{n} \frac{1}{n}(r_i - \bar{r})\right]^2$$

$$= \frac{1}{n^2}\mathbb{E}\left\{\left[\sum_{i=1}^{n} \frac{1}{n}(r_i - \bar{r})\right]\left[\sum_{j=1}^{n} \frac{1}{n}(r_j - \bar{r})\right]\right\}$$

$$= \frac{1}{n^2}\sum_{i,j} \sigma_{ij} = \frac{1}{n^2}\left(\sum_{i=j} \sigma_{ij} + \sum_{i \neq j} \sigma_{ij})\right)$$

$$= \frac{1}{n^2}[n\sigma^2 + t(n^2 - n)\sigma^2]$$

$$= \frac{\sigma^2}{n} + t\sigma^2(1 - \frac{1}{n})$$

$$= \frac{\sigma^2(1-t)}{n} + t\sigma^2$$

Ecuación 3.14. Varianza de una cartera

El resultado de lo anterior se puede apreciar en la siguiente imagen, donde se confirma que en este caso existe un límite inferior para la varianza.

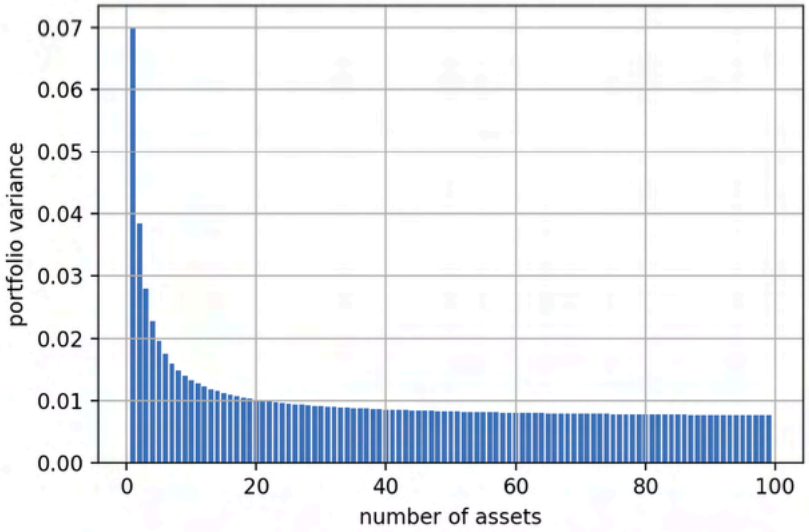


Ilustración 17. Efectos de la diversificación (Bertani, 2021)

3.2 HERRAMIENTAS PARA LA OPTIMIZACIÓN DE CARTERAS

A la hora de generar estas carteras, existen diferentes herramientas que se pueden utilizar. En 1952, Markowitz presentó su Teoría Moderna de Carteras (MPT, del inglés *Modern Portfolio Theory*) y desarrolló el modelo de optimización Media-Varianza. Desde aquí se ha investigado sobre la optimización de las carteras y se han desarrollado nuevas herramientas para ello (Cosín López-Medel, 2019).

3.2.1 OPTIMIZACIÓN CON MEDIA-VARIANZA (MARKOWITZ) Y FRONTERA DE EFICIENCIA

MPT es una teoría que dice que un inversor, con aversión al riesgo, puede construir una cartera de inversión que maximice el retorno esperado para un determinado nivel de riesgo del mercado (Markowitz, 1952). Según Markowitz, el inversor debería diversificar sus inversiones entre todos los activos cuyo retorno esperado sea máximo.

Markowitz desarrolla el **modelo de optimización media-varianza** para generar carteras financieras, buscando maximizar el rendimiento esperado para un nivel de riesgo dado. Para ello, la desviación estándar es utilizada para definir el riesgo de los activos (Kull, 2005). Se debe tener en cuenta que la cartera óptima para cada inversor es diferente ya que cada uno está dispuesto a asumir un riesgo y tiene como objetivo unos niveles de retorno concretos.

La tasa de rendimiento de los activos se puede expresar en un diagrama de dos dimensiones de manera que un activo con rendimiento esperado $\bar{r}$ y desviación típica σ queda representado como un punto en el mencionado diagrama, denominado **diagrama de media-varianza**.

La **frontera de eficiencia** representa todas aquellas combinaciones de puntos que reflejan la menor varianza posible para un nivel dado de rendimiento esperado. Alternativamente, puede definirse como todos aquellos puntos en los cuales se obtiene un rendimiento óptimo dado un nivel de riesgo. A continuación, se representa una imagen de este concepto.

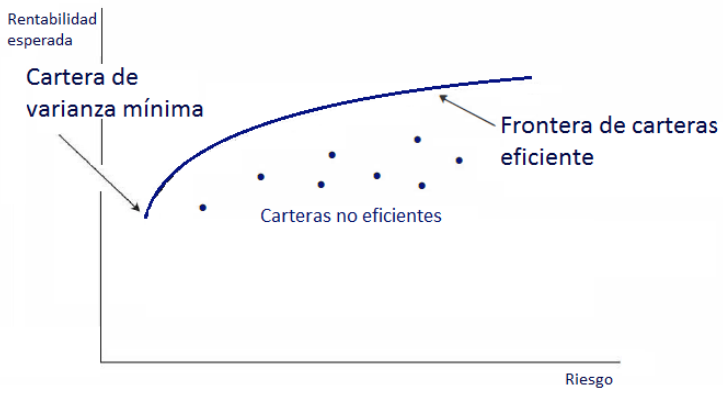


Ilustración 18. Diagrama de media-varianza (Cosín López-Medel, 2019)

A partir de los activos que se muestran en la imagen anterior podría formarse una cartera que tuviera un valor de rendimiento medio esperado, así como una varianza propios, de forma que dicha cartera quedaría representada en el diagrama como un punto más. Este

punto podría encontrarse en diferentes localizaciones del diagrama en función de la covarianza entre los rendimientos de los activos que forman la cartera. La región que cubre todas estas posibles localizaciones recibe el nombre de **región factible**.

(Kull, 2005) defiende que los inversores suelen presentar aversión al riesgo, por lo que elegirán la cartera con el riesgo mínimo cuando el rendimiento esperado sea el mismo, mientras que elegirán la cartera con el rendimiento esperado mayor cuando el riesgo sea igual.

3.2.2 OPTIMIZACIÓN DE CARTERAS CON CVAR

Posteriormente, (Kull, 2005) propone una alternativa al modelo Media-Varianza anterior, haciendo uso del valor en riesgo (VaR) que se ha visto previamente. En este caso, a diferencia del anterior, ya no tiene en cuenta la desviación estándar como medida de riesgo, sino el **VaR condicional (CVaR)**. Para ello, defiende que, a la hora de optimizar carteras, el objetivo principal debe ser minimizar el valor de CVaR o poner límites en esta medida al encontrar el peso óptimo de cada activo.

Este modelo puede ser optimizado mediante el concepto de Programación Lineal, que se explica en el siguiente apartado.

3.2.3 PROGRAMACIÓN LINEAL

El concepto de **Programación Lineal** (LP, del inglés *Linear Programming*) es un método matemático que se utiliza en optimización. Para ello, se define una función objetivo que se busca maximizar o minimizar, además de algunas restricciones que deben cumplirse.

Estos métodos se utilizan en diferentes campos y la gestión de carteras es uno de ellos ya que cuentan con recursos computacionales y herramientas que ayudan a su evaluación.

3.2.4 CARTERAS NAÏVE

La **estrategia Naïve** consiste en tener una cartera en la que cada uno de los N activos que la componen tienen un peso $w_i = 1/N$, es decir, todos tienen el mismo peso. Sin embargo, este método no implica un proceso de optimización como los anteriores, por lo que se suele utilizar por dos razones (DeMiguel, Garlappi, & Uppal, 2007). La primera de ellas, por ser fácil de implementar ya que no necesita realizar la estimación de los retornos de los activos ni optimización. La segunda, porque a pesar de todos los métodos de optimización desarrollados en los últimos años, muchas veces los inversores prefieren hacer uso de reglas más simples de asignación de los activos para asignar su riqueza o la de sus clientes.

3.2.5 RATIO DE SHARPE

El **ratio de Sharpe** es utilizado para ayudar a los inversores a entender el rendimiento de una inversión frente a su riesgo y mide la relación entre la rentabilidad y la desviación estándar de una inversión. El valor del ratio aumenta cuando el riesgo disminuye debido a la diversificación, por lo que cuantos más activos tenga la cartera, mayor será el valor del ratio de Sharpe. Además, cuanto mayor sea el valor del ratio de Sharpe, mejor será la relación de la recompensa frente al riesgo.

Según (Kourtis, 2016), la cartera óptima tiene que ver con la frontera de eficiencia y se corresponde con el punto en el que la línea de la tasa libre de riesgo es tangente a la frontera, como se muestra a continuación.

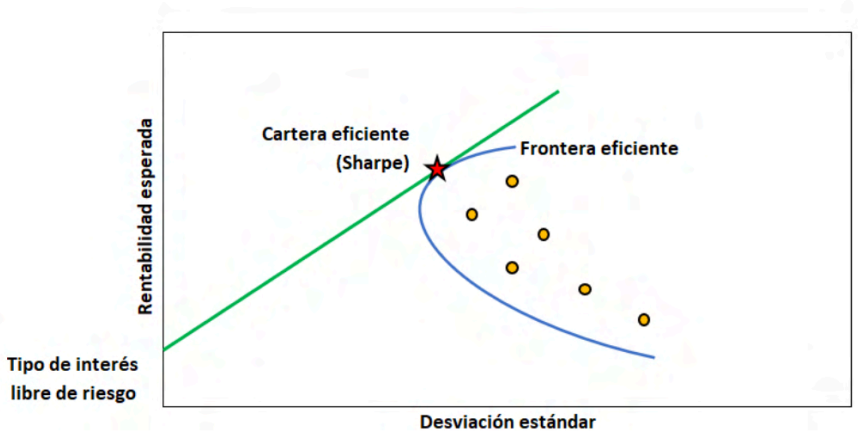


Ilustración 19. Cartera eficiente y frontera de eficiencia (Lasheras, 2020)

Para obtener esta cartera, los inversores deben conocer el retorno esperado y la varianza y covarianza entre los activos que vayan a formar la cartera. Sin embargo, aunque esta cartera puede dar información del máximo retorno por cada unidad de riesgo, el ratio de Sharpe no se podrá obtener debido a que los parámetros reales de la cartera son desconocidos, por lo que este modelo representa una aproximación. El objetivo principal será hacer el ratio de Sharpe lo más grande posible, minimizando la varianza de la cartera.

3.2.6 POST-MODERN PORTFOLIO THEORY (PMP)

La **Teoría Post-Moderna de Carteras** (PMP, del inglés *Post-Modern Portfolio Theory*) surge debido a algunas limitaciones existentes en MPT ya que la varianza que esta utiliza no siempre es la mejor medida de riesgo. Para desarrollar la nueva teoría, Markowitz sugirió un nuevo modelo basado en la **semi-varianza**, la cual comienza por calcular a partir del **Rendimiento Mínimo Aceptable** (MAR, del inglés *Minimum Acceptable Return*), el cual suele ser cero. A continuación, se deben seleccionar aquellos rendimientos iguales o menores al MAR y se calcula la varianza de estos rendimientos según la siguiente expresión:

$$semivarianza = SV(r) = \sum_{i=1}^{n} (\mathbb{E}[(r_i - \bar{r})^2 1_{r_i \leq 0}])$$

Ecuación 3.15. Semivarianza

Donde $1_{r_i \leq 0}$ es una función definida por lo siguiente:

$$1_{r_i \leq 0} = \begin{cases} 1 & si\ r \leq 0 \\ 0 & si\ no \end{cases}$$

3.3 MÉTODOS DE ASIGNACIÓN DE ACTIVOS

La **asignación de activos** consiste en encontrar un balance óptimo entre el riesgo y el rendimiento de una cartera mediante la determinación o el cambio de las proporciones asignadas a cada uno de los componentes de la cartera. Aquí, el **modelo de Markowitz** juega un papel importante siendo este una técnica que busca seleccionar la mejor distribución de activos para una cartera con el objetivo de maximizar los rendimientos para un nivel de riesgo dado.

Existen diversos algoritmos utilizados para la optimización de carteras y la asignación de activos que se pueden considerar como punto de partida a la hora de desarrollar un nuevo método. Uno de los más comunes consiste en asignar a cada activo un peso igual, lo cual se ha observado que puede derivar en buenos resultados.

Otro método de asignación consiste en la volatilidad invertida, donde cada activo recibe un peso proporcional a la inversa de su volatilidad según la siguiente fórmula:

$$w_i = \frac{\frac{1}{\sigma_i}}{\sum_{j=0}^{n} \frac{1}{\sigma_j}}$$

Ecuación 3.16. Volatilidad invertida

El tercer método de asignación es el de paridad de riesgo, cuyo objetivo consiste en crear una cartera donde cada activo contribuye igualmente a la volatilidad total de la misma.

3.4 ÍNDICES UTILIZADOS EN LA OPTIMIZACIÓN DE CARTERAS

A la hora de optimizar una cartera, los inversores hacen uso de distintos indicadores que les sirven de ayuda a la hora de tomar la mejor decisión en función de sus objetivos. Algunos de estos indicadores son los que se presentan a continuación.

3.4.1 ÍNDICE DE FUERZA RELATIVA

El **Índice de Fuerza Relativa** (RSI, del inglés *Relative Strength Index*) es un indicador técnico de tipo oscilador en mercados financieros que mide la velocidad en el movimiento de los precios de las acciones y su magnitud. Se suele medir en periodos de 14 días y puede tener un valor de 0 a 100, siendo 70 y 30 los puntos alto y bajo más comunes.

Así, cada periodo se clasifica como de alza, siendo el precio de cierre mayor que el de apertura, o de baja, siendo el precio de apertura mayor que el precio de cierre. Con ello, para un periodo de alza se calcula U (cambio hacia arriba) y D (cambio hacia abajo) según se muestra a continuación.

$$U = cierre_{actual} - cierre_{anterior}$$

$$D = 0$$

Ecuación 3.17. U y D en periodo de alza

De la misma forma, para un periodo a la baja, los valores de U y D se obtienen a partir de lo siguiente.

$$U = 0$$

$$D = cierre_{anterior} - cierre_{actual}$$

Ecuación 3.18. U y D en periodo a la baja

Si se pretende calcular el valor de U y D promedio de varios periodos se hace uso de una **Media Móvil Suavizada** (SMMA, del inglés *Smoothed Moving Average*) y el cociente de ellas es la **Fuerza Relativa** (RS, del inglés *Relative Strength).*

$$RS = \frac{SMMA(U,n)}{SMMA(D,n)}$$

Ecuación 3.19. Fuerza relativa

Finalmente, se normaliza para obtener el valor del RSI.

$$RSI = 100 - \frac{100}{1 + RS}$$

Ecuación 3.20. Índice de fuerza relativa

Cabe destacar que si el resultado alcanza un valor de 70 se considera que la acción está sobrecomprada, es decir, el especulador debería plantearse venderla. Por otro lado, si alcanza un nivel de 30 se entiende que el valor está sobrevendido, por lo que el especulador debería plantearse comprarla.

3.4.2 ÍNDICE DE MERCADO

Un **índice de mercado** es una cartera hipotética de participaciones de inversión que representa un segmento del mercado financiero. Existen distintos índices de mercado y cada inversor hace uso de varios de ellos para estudiar la situación del mercado en cada momento. Para calcular el valor del índice se hace uso de las participaciones subyacentes. Los índices más conocidos son Dow Jones Industrial Average (DJIA), S&P 500 y Nasdaq Composite Index, entre otros.

3.4.3 FONDOS COTIZADOS E ÍNDICES DE FUTUROS

Un **fondo cotizado** es un tipo de activo cuyo precio presenta una correlación con el valor de un índice, sector, materia prima o activo, que puede ser comprado o vendido en una bolsa de valores de la misma forma que una acción normal.

Un **contrato de futuros** es un acuerdo legal y vinculante para comprar o vender un determinado activo a un precio definido en una fecha futura. Así, el comprador asume la responsabilidad de comprar y recibir el activo cuando el contrato expira y el vendedor del contrato debe entregar dicho activo en la fecha de expiración del contrato.

Los **índices de futuros** son contratos que permiten a una persona adquirir o vender activos en el futuro con el mismo precio que indica el índice hoy.

3.5 EL USO DE BETA (β)

Por lo general, las medidas de riesgo en finanzas se limitan a métodos basados en la varianza y el parámetro más común a la hora de evaluar el riesgo es la beta. El problema con la beta es que pondera igualmente para la varianza al alza y a la baja. La habilidad para descomponer la beta para el alza y para la baja les proporciona a los inversores la posibilidad de definir el riesgo de una cartera de una forma más precisa. Para ello se hace uso de lo que se conoce como doble-beta.

3.5.1 BETA (β)

Como ya se ha mencionado, **beta (β)** se utiliza para medir el riesgo sistemático de un activo o de una cartera en el mercado. La beta indica cómo se mueve un valor en comparación con el resto del mercado o índice de referencia, teniendo en cuenta que un activo se considera de mayor riesgo cuanto mayor sea el valor de su beta. Por ejemplo, si una acción tiene una beta de 1,10, para cada movimiento de un 1% en el índice de referencia, la acción se moverá un 1,10%. Las acciones que poseen betas altas son comúnmente aquellas que históricamente han tenido mayor volatilidad y, por tanto, un mayor riesgo percibido.

La beta es de gran utilidad para los inversores ya que les ayuda a evaluar el riesgo y el retorno esperado de un activo o de una cartera en el mercado. Por ejemplo, si un inversor quiere invertir en una cartera con un nivel de riesgo bajo, escogerá una con un valor de beta bajo mientras que, si el inversor está dispuesto a asumir un riesgo mayor con la posibilidad de obtener un retorno mayor, escogerá una cartera que tenga un valor de beta más elevado.

Existen algunas críticas hacia el uso de la beta como una herramienta de predicción en los mercados financieros por diversas razones. La primera de ellas es que para obtenerla se hace uso de datos pasados, lo cual no tiene por qué ser coherente con lo que vaya a ocurrir en el futuro. Por otro lado, existen diferentes formas de calcular la beta, dependiendo del tiempo que nos remontemos para obtener datos históricos, por lo que el resultado de la beta puede variar en función de ello. El tercer problema que el uso de beta presenta tiene que ver con que la beta solo evalúa el riesgo relacionado con lo que está haciendo el resto del mercado, pero no da información sobre lo que está ocurriendo en un sector específico, por ejemplo. Por último, y como ya hemos comentado antes, en el cálculo de la beta la ponderación es igual para el cálculo al alza (beta plus) y el cálculo a la baja (beta menos) (Nystuen & Nyrnes, 2022). Así, con el objetivo de obtener una mayor precisión a la hora de obtener la información de la beta, surge el concepto de **doble beta**, compuesta por beta plus y beta menos.

3.5.2 DOBLE BETA

La **doble beta** es una medida financiera utilizada para estimar la sensibilidad de los retornos de un activo ante cambios en el mercado general y en el sector específico de la industria en la que opera dicho activo. Es utilizada por los inversores para evaluar el riesgo del activo en particular y es especialmente útil para aquellos inversores que quieren diversificar sus carteras entre distintos sectores de la industria ya que les permite identificar activos que tienen poca correlación con el mercado en general y con una industria en particular. Así, puede ayudar a reducir el riesgo de la cartera y potenciar los retoros.

Beta menos y beta plus son los dos componentes de la doble beta. **Beta plus** analiza el comportamiento de una acción cuando el rendimiento de referencia es positivo. Por ello, este valor de beta se calcula en periodos en los que el mercado tiene una tendencia creciente, es decir, con rendimientos mayores que cero. En conclusión, la beta plus mide la sensibilidad de una acción ante situaciones de crecimiento en el mercado. Los inversores estarán interesados, por tanto, en escoger acciones con valores de beta plus elevados con el objetivo de obtener mayores beneficios.

Por otro lado, la **beta menos** es la beta de una acción medida en los periodos en los que el rendimiento de referencia es negativo. En este caso, los inversores buscarán construir una cartera que minimice este valor de beta, lo cual ayudará a mantener el valor en tiempos en los que el mercado esté en recesión.

Cabe destacar que ambas betas se calculan a partir de la misma fórmula, pero dado que miden diferentes aspectos de la sensibilidad de una acción frente al mercado, sus valores pueden diferir debido a que la acción puede tener un grado de sensibilidad distinto para el caso de periodos de mercado más recesivos y para el caso de periodos en los que el mercado presenta subidas. Esto puede ser debido a factores como el sector en el que trabaja la empresa que se evalúa, una noticia específica de la compañía en cuestión o diferentes condiciones del mercado. Por ejemplo, una acción de una empresa perteneciente a un sector como el del cuidado de la salud puede ser menos sensible a caídas del mercado, es decir, tener un menor valor de beta menos, que para el caso en el que el mercado presenta subidas, es decir, tener un valor mayor de beta plus. Por otro lado, una acción de una empresa que opera en un sector cíclico como el tecnológico puede presentar mayor sensibilidad a caídas del mercado, es decir, tener un mayor valor de beta menos que, a subidas del mercado, teniendo un menor valor de beta plus.

Por todo ello, es importante tener encuentra ambas betas a la hora de analizar la sensibilidad de una acción frente al mercado ya que proporcionan información diferente ante el comportamiento de las acciones frente a distintas condiciones de mercado.

Intervalo de retorno

A la hora de trabajar con datos financieros y analizarlos, seleccionar la frecuencia con la que vamos a escoger la información de los intervalos de los retornos es de vital importancia ya que dependiendo de ella obtendremos unos resultados u otros a la hora de estimar los coeficientes de las betas. Así, teniendo en cuenta un único periodo de estimación, las betas serán diferentes si escogemos datos de retornos diarios, semanales o mensuales. Por ejemplo, hacer uso de los cambios diarios del rendimiento de las acciones a la hora de estimar la beta puede dar lugar a resultados imprecisos si algunas de las acciones suelen

quedarse atrás en comparación con el mercado o, por el contrario, liderar el movimiento del mercado general (Hawawini, 1983). Así, si se expande el intervalo de retorno, se eliminará este efecto de retraso diario. Cuanto mayor sea la expansión temporal, menor será el efecto de retraso percibido de manera que si trabaja con datos semanales, la mayoría de la correlación cruzada intertemporal quedará eliminada mientras que, si se utilizan datos mensuales, este efecto de retraso desaparecerá por completo. Así, hacer uso de datos diarios a la hora de estimar la beta puede causar imprecisión en los resultados mientas que expandir el intervalo de retorno puede acabar con este efecto de manera que cuanto más grande sea esta expansión, menor será el efecto mencionado en el resultado.

Hawawini encontró una manera más directa de predecir el cambio en beta a la hora de modificar el intervalo de retorno. Así, cuando este intervalo se expande, las acciones con un alto Valor de Mercado de Acciones en Circulación (MVSO, del inglés *Market Value of Shares Outstanding*) relativos a la media del mercado sufrirán un cambio decreciente en la beta. Igualmente, los valores que posean un MVSO bajo podrán experimentar un incremento en el valor de beta.

Todo lo anterior deriva en que, al hacer uso de intervalos de retorno diarios, las acciones con un valor de mercado elevado pueden tener un riesgo mayor del que realmente tienen, mientras que las que tengan un valor de mercado menor, aparentemente parecerán menos arriesgadas de lo que son (Hawawini, 1983).

3.5.3 EL USO DE LA BETA EN EL IBEX 35

A la hora de aplicar a la práctica lo explicado anteriormente, se va a hacer uso de información relativa al IBEX 35. El **IBEX 35** es el índice bursátil de referencia en la bolsa española y mide el comportamiento conjunto de las 35 empresas que presentan un mayor y más fuerte interés comprador y vendedor entre los accionistas, entre aquellas que cotizan en el Sistema de Interconexión Bursátil Electrónico (SIBE) en las cuatro bolsas españolas (Madrid, Barcelona, Bilbao y Valencia). Las empresas seleccionadas son aquellas que cumplen mejor los parámetros de capitalización, liquidez y volumen negociado, por lo que no necesariamente son las más grandes del mercado.

El valor del IBEX 35 representa el comportamiento global de las 35 empresas que lo forman y se usa como un punto de referencia para medir cómo se comporta el mercado de valores español. Si el valor del índice aumenta, significa que las acciones de estas compañías han aumentado su valor mientras que, si el índice disminuye, el valor de las acciones de dichas empresas habría sufrido una bajada. Los inversores utilizan el valor del índice como un indicador de la 'salud' de la economía española.

Para trabajar, se hará uso de datos semanales de estas empresas en un periodo que abarca desde el 8 de marzo de 2015 hasta el 17 de marzo de 2023, contando con los precios de apertura, cierre, máximo, mínimo y cierre ajustado. Este último, representa el precio de cierre después de haber realizado los ajustes correspondientes a las distribuciones de *splits* y dividendos correspondientes.

A continuación, se representa la evolución del valor del IBEX 35 a lo largo de los últimos años, aquellos que se han indicado previamente.

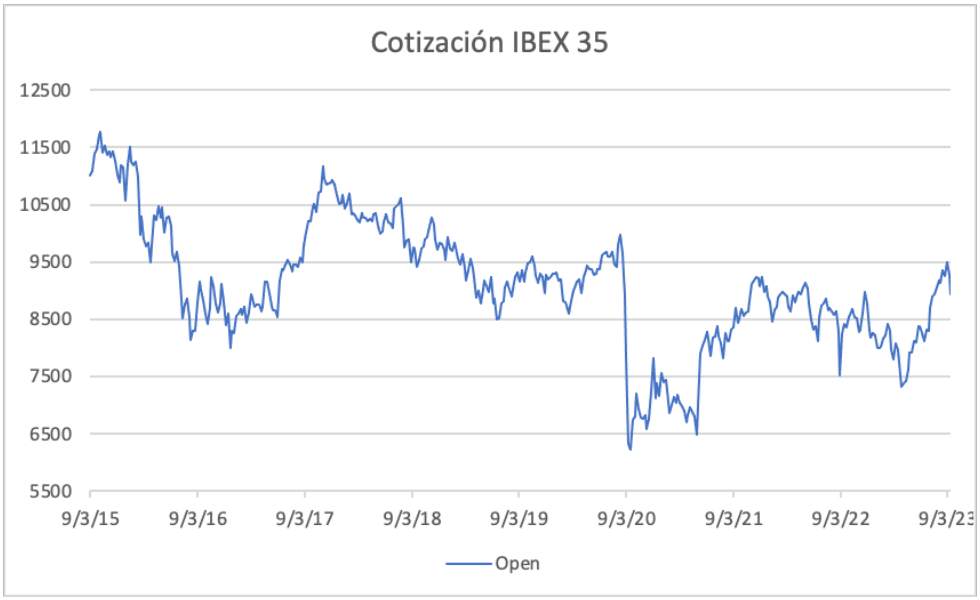


Ilustración 20. Evolución valor del IBEX 35 (Elaboración propia)

En la figura se observan dos principales puntos de recesión en la cotización del IBEX 35. El primero de ellos tiene lugar en el año 2016 y una de las razones a las que se debe es

el triunfo del Brexit en el referéndum realizado en Reino Unido, lo cual hizo que las bolsas europeas, la libra y el euro, entre otras cosas, sufrieran fuertes ventas, ocasionando la mayor caída de la bolsa española hasta el momento (Expansión, 2016). Con ello, el IBEX 35 sufrió una bajada del 12,35%, llegando hasta los 7.787,70 puntos. Por otro lado, el siguiente periodo de recesión que se puede apreciar en la figura tiene lugar en marzo de 2020 debido a la pandemia ocasionada por el Covid-19. En ese momento, la Organización Mundial de la Salud (OMS) hizo oficial la situación de pandemia mundial y desde entonces las bolsas de todo el mundo volvieron a bajar. Así, el índice llegó a perder más de un 15% de su valor, logrando la mayor caída anual desde 2010 y llegando a valores inferiores a los 7.000 puntos. Después de este periodo de incertidumbre, el valor del IBEX 35 comenzó a subir de nuevo progresivamente, aunque volvió a estancarse en febrero de 2022, coincidiendo con el inicio de la guerra entre Rusia y Ucrania.

A continuación, se pueden apreciar los valores medios obtenidos para las actuales empresas que componen el IBEX 35 desde el año 2015 hasta la actualidad. Así, obtenemos la información media de las compañías para el caso de sus precios máximo, mínimo, apertura y cierre. Además, se han calculado las betas en Matlab y se muestran también en la siguiente tabla.

Date	Open	High	Low	Close	Beta
"IBEX35"	"9086.3039"	"9238.5222"	"8908.6891"	"9079.2256"	"1"
"ACS_MC"	"28.9136"	"29.6918"	"28.0517"	"28.8832"	"1.2277"
"ACX_MC"	"10.3438"	"10.6668"	"9.9784"	"10.3135"	"0.93755"
"AENA.MC"	"137.7363"	"141.2584"	"134.3271"	"137.8816"	"0.84262"
"AMS.MC"	"54.4013"	"55.9988"	"52.7829"	"54.4396"	"1.037"
"ANA.MC"	"101.1717"	"104.0872"	"98.347"	"101.4387"	"0.76051"
"ANE.MC"	"34.1952"	"35.3814"	"33.0712"	"34.2548"	"0.13396"
"BBVA.MC"	"5.6609"	"5.8294"	"5.4787"	"5.658"	"1.468"
"BKT.MC"	"6.153"	"6.3281"	"5.9767"	"6.1556"	"1.1007"
"CABK.MC"	"3.0863"	"3.1869"	"2.9807"	"3.084"	"1.2136"
"CLNX.MC"	"29.1011"	"29.9851"	"28.188"	"29.1373"	"0.13174"
"COL.MC"	"7.9203"	"8.1357"	"7.7055"	"7.9244"	"0.94582"
"ELE.MC"	"20.1105"	"20.5419"	"19.6542"	"20.1086"	"0.63171"
"ENG.MC"	"22.5773"	"23.0186"	"22.0639"	"22.5449"	"0.54937"
"FDR.MC"	"12.9594"	"13.4444"	"12.4386"	"12.9809"	"0.57687"
"FER.MC"	"21.4748"	"22.0123"	"20.9444"	"21.5025"	"0.87789"
"GRF.MC"	"21.9487"	"22.5679"	"21.2872"	"21.913"	"0.53229"
"IAG.MC"	"4.8321"	"4.9991"	"4.6397"	"4.8148"	"1.6732"
"IBE.MC"	"8.1737"	"8.3611"	"8.0066"	"8.1931"	"0.6971"
"IDR.MC"	"9.6411"	"9.9513"	"9.3226"	"9.6389"	"0.97981"
"ITX.MC"	"28.0182"	"28.7319"	"27.2841"	"28.0186"	"0.98436"
"LOG.MC"	"19.333"	"19.7296"	"18.8727"	"19.3091"	"0.51109"
"MAP.MC"	"2.3"	"2.3555"	"2.2367"	"2.2953"	"1.072"
"MEL.MC"	"8.7475"	"9.001"	"8.4516"	"8.7242"	"1.3031"
"MRL.MC"	"10.3383"	"10.5741"	"10.0626"	"10.3241"	"0.83316"
"MTS.MC"	"19.8318"	"20.6585"	"18.9946"	"19.8176"	"1.428"
"NTGY.MC"	"21.2862"	"21.8054"	"20.8117"	"21.3192"	"0.78603"
"RED.MC"	"17.867"	"18.2117"	"17.5004"	"17.8579"	"0.49602"
"REP.MC"	"12.7456"	"13.1317"	"12.3712"	"12.7486"	"1.0861"
"ROVI.MC"	"27.7071"	"28.7077"	"26.7058"	"27.7261"	"0.35908"
"SAB.MC"	"1.118"	"1.1596"	"1.072"	"1.1152"	"1.5721"
"SAN.MC"	"3.858"	"3.9743"	"3.7315"	"3.8522"	"1.4757"
"SCYR.MC"	"2.2292"	"2.3088"	"2.1441"	"2.2268"	"1.3773"
"SLR.MC"	"8.0137"	"8.4768"	"7.5756"	"8.0432"	"0.82602"
"TEF.MC"	"7.0552"	"7.2185"	"6.8602"	"7.0291"	"0.93513"
"UNI.MC"	"0.96183"	"0.99613"	"0.92534"	"0.95962"	"0.16109"

Tabla 3. Información de las empresas del IBEX 35 (Elaboración propia)

Finalmente, la imagen siguiente muestra un gráfico que representa el valor de los retornos frente al riesgo para cada una de las compañías del IBEX 35 en el periodo de tiempo seleccionado, de manera puntual y con valores tomados semanalmente. El riesgo de cada una de estas compañías queda representado por el valor de la beta de manera que, cuanto mayor sea el valor de beta, más alto será el riesgo sistemático de los activos de cada compañía estudiada.

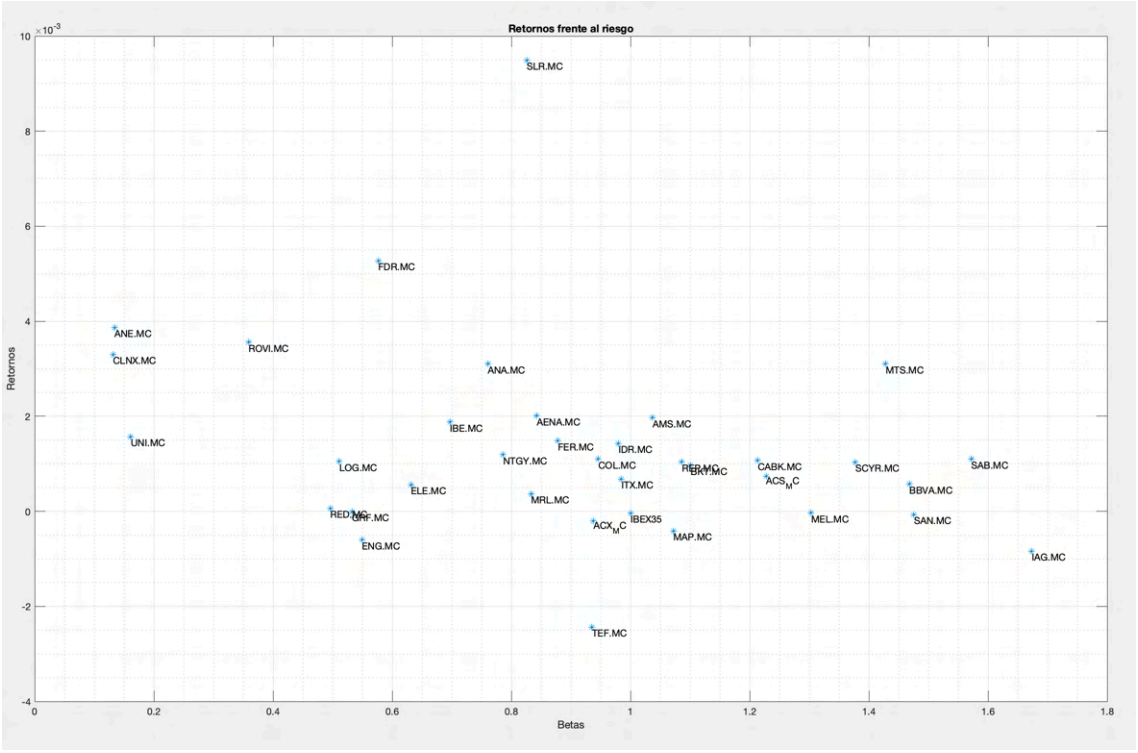


Ilustración 21. Retornos frente al riesgo en el IBEX 35 (Elaboración propia)

Se observa como existe gran variedad en la situación de retorno frente al riesgo de las empresas españolas en el periodo evaluado y esta información es útil a la hora de conformar una cartera. Mientras que algunas empresas muestran un valor de riesgo (beta) bajo, cuando beta está por debajo de 1, como puede ser el caso de Cellnex (CLNX) o Iberdrola (IBE), hay otras empresas que tienen un valor de riesgo mayor, cuando beta está por encima de 1, como por ejemplo IAG o Sabadell (SAB).

Además, se observa que los retornos medios en el periodo evaluado no necesariamente son positivos, como ocurre con Telefónica (TEF) por lo que los inversores deberán tener este factor en cuenta antes de realizar una inversión, buscando invertir en aquellas compañías que les puedan proporcionar un mayor retorno para el nivel de riesgo que estén dispuestos a asumir.

3.5.4 EL USO DE LA BETA EN EL S&P 100

A continuación, se ha realizado el mismo proceso con los datos relativos al índice S&P 100 y las empresas que lo forman. El **S&P 100**, también conocido como OEX u OEX 100, es un índice bursátil de los Estados Unidos, compuesto por las 100 empresas de mayor capitalización que cotizan en la bolsa de Nueva York (NYSE, del inglés *New York Stock Exchange*) y el Nasdaq.

En la siguiente tabla se muestran algunos datos relevantes obtenidos para algunas de las compañías que forman el S&P 100, desde el año 2015 hasta enero de 2023. En concreto, se pueden apreciar las medias de sus precios mínimo, máximo, apertura y cierre, habiendo hecho uso de datos semanales. Además, se han calculado las betas para cada una de ellas para obtener una medida de riesgo de los valores analizados.

Date	Low	High	Open	Close	Beta
"S&P100"	"1329.6488"	"1371.5157"	"1350.7803"	"1353.2579"	"1"
"AAPL UW Equity"	"72.1186"	"76.3514"	"73.9836"	"74.3389"	"1.1941"
"ABT UN Equity"	"75.2951"	"78.8502"	"77.0759"	"77.2989"	"0.80993"
"ACN UN Equity"	"186.3995"	"195.0667"	"190.3503"	"191.112"	"1.0597"
"ADBE UW Equity"	"272.2341"	"289.7538"	"280.9779"	"281.796"	"1.0011"
"AIG UN Equity"	"51.5024"	"54.2336"	"52.8574"	"52.9078"	"1.051"
"AMGN UW Equity"	"193.7134"	"203.2913"	"198.4193"	"198.7005"	"0.70752"
"AMT UN Equity"	"175.4967"	"184.1923"	"179.8855"	"180.2429"	"0.6693"
"AMZN UW Equity"	"87.0416"	"92.6778"	"89.72"	"89.9323"	"1.1218"
"AVGO UW Equity"	"296.6162"	"315.7417"	"305.9669"	"307.1593"	"1.1367"
"AXP UN Equity"	"106.0618"	"111.8824"	"109.0077"	"109.1668"	"1.1431"
"BA UN Equity"	"218.0973"	"233.0454"	"225.2584"	"225.7122"	"1.4725"
"BAC UN Equity"	"26.9935"	"28.5675"	"27.7637"	"27.8134"	"1.1325"
"BK UN Equity"	"45.0406"	"47.3463"	"46.1999"	"46.2276"	"1.0062"
"BKNG UW Equity"	"1750.8955"	"1857.2729"	"1803.1327"	"1806.2453"	"1.1826"
"BLK UN Equity"	"510.8065"	"537.0838"	"523.9234"	"524.9288"	"1.2687"
"BMY UN Equity"	"59.7818"	"62.6687"	"61.2666"	"61.292"	"0.69293"
"BRK B UN Equity"	"204.8466"	"212.2164"	"208.52"	"208.8169"	"0.79368"
"C UN Equity"	"58.095"	"61.4733"	"59.7935"	"59.808"	"1.2089"
"CAT UN Equity"	"135.732"	"143.8003"	"139.6659"	"140.0673"	"1.0337"
"CL UN Equity"	"70.9916"	"73.5417"	"72.2931"	"72.3554"	"0.59054"
"CMCSA UW Equity"	"38.7294"	"40.639"	"39.717"	"39.7559"	"0.83271"
"COF UN Equity"	"92.5598"	"98.4888"	"95.5975"	"95.6517"	"1.1536"
"COP UN Equity"	"57.9598"	"62.2417"	"60.0454"	"60.2283"	"1.2269"
"COST UW Equity"	"272.7855"	"285.1412"	"278.5303"	"279.4665"	"0.76739"
"CRM UN Equity"	"140.6647"	"150.1324"	"145.326"	"145.6546"	"1.0419"
"CSCO UW Equity"	"40.5328"	"42.4674"	"41.4759"	"41.5535"	"0.95775"
"CVS UN Equity"	"78.8385"	"82.8745"	"80.8848"	"80.9018"	"0.82142"
"CVX UN Equity"	"109.7766"	"115.5689"	"112.6109"	"112.8155"	"1.035"
"DHR UN Equity"	"147.8313"	"155.2945"	"151.5499"	"152.0767"	"0.76783"
"DIS UN Equity"	"118.3845"	"124.3097"	"121.4746"	"121.3853"	"1.0754"

Tabla 4. Información de las empresas del S&P 100 (Elaboración propia)

Por último, se representan gráficamente el valor de los retornos frente al riesgo obtenido en función de cada una de las betas calculadas para las diferentes compañías y en el periodo de tiempo seleccionado.

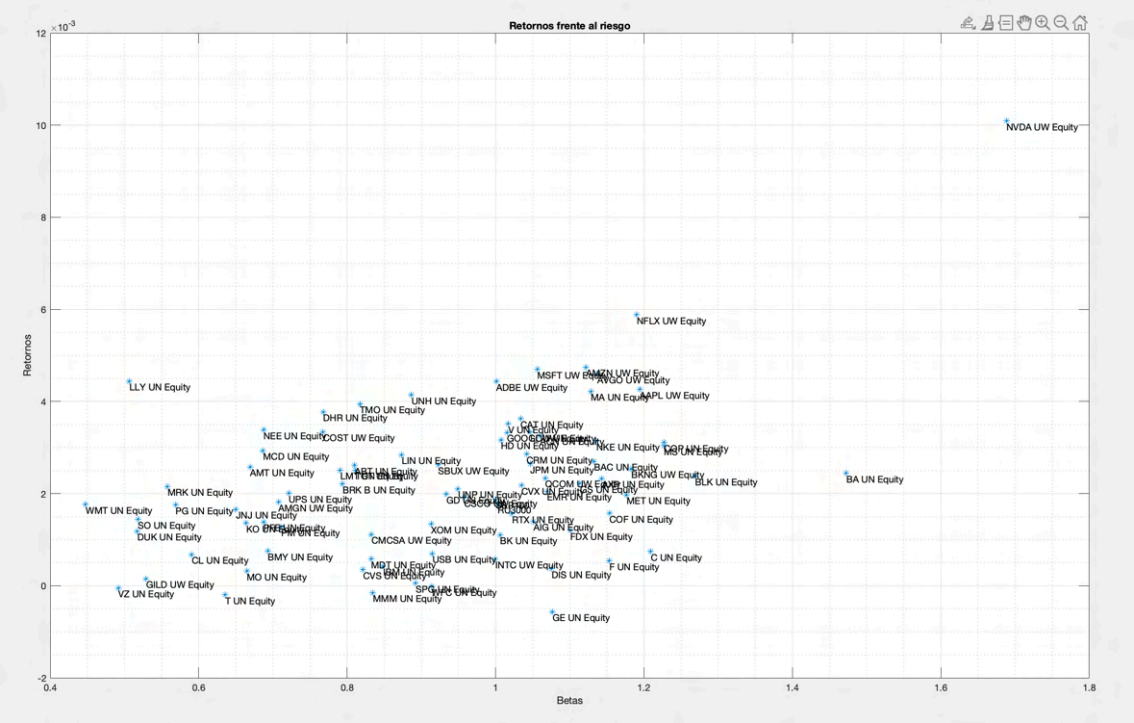


Ilustración 22. Retornos frente al riesgo en el S&P 100 (Elaboración propia)

Se puede apreciar como a medida que aumenta el riesgo sistemático, representado por la beta, el retorno esperado tiende a hacerlo también, lo cual tiene sentido debido a la relación lineal entre ambas variables. Al igual que antes, este gráfico podría servir como referencia para los inversores a la hora de tomar sus decisiones de inversión a partir de información pasada.

Llama la atención la situación de Nvidia Corporation (NVDA), compañía tecnológica americana, ya que presenta el mayor valor de riesgo, con una beta de alrededor de 1,7 pero a la vez es la que ofrece unos retornos mayores. Otras compañías como Netflix (NFLX) o Microsoft (MSFT) a pesar de tener un valor de beta superior a 1 y, por tanto, tener cierto riesgo, presentan retornos elevados, por lo que podrían ser buenas opciones para inversores dispuestos a asumir ciertos riesgos.

Capítulo 4. DATOS DE INTERVALO Y REGRESIÓN

El análisis de datos ha sido un campo de profunda investigación en los últimos años debido a la gran cantidad de información de utilidad que puede ser extraída de ellos, siendo este un paso fundamental para descubrir hechos que ayuden a mejorar el presente y el futuro de nuestra sociedad. La llegada de nuevas tecnologías como el Big Data que se han enfocado hacia diferentes campos ha marcado un antes y un después en la forma de procesar los datos a gran escala. Así, el uso de datos de intervalo en lugar de datos puntuales con estas tecnologías ha ayudado a extraer información y desarrollar distintas herramientas y modelos que permiten obtener soluciones más precisas para los diferentes problemas que se plantean.

El mundo de hoy día es cada vez más incierto y en concreto el campo de las finanzas trabaja con una gran cantidad de datos que presentan un alto grado de incertidumbre. Es por ello por lo que este es uno de los sectores en los que está creciendo el uso del análisis de datos de intervalo ya que estos son de gran utilidad para representar y analizar esta incertidumbre y pueden ayudar a los inversores a la hora de tomar decisiones (Wang, Gurmani, Tao, Liu, & Chen, 2023). Por ejemplo, al estimar los precios de una acción, resulta útil escoger un intervalo en el que es seguro que se encuentra el resultado en lugar de un valor puntual que contenga cierto error.

En este capítulo, se va a realizar una introducción a los datos de intervalo explicando algunos de los conceptos básicos con los que trabajan y además, se introduce la idea de Análisis de Intervalos (IA, del inglés *Interval Analysis*). Finalmente, se presenta la regresión haciendo uso de datos de intervalo y se definirán diferentes modelos de regresión lineal para intervalos.

4.1 INTRODUCCIÓN A LOS DATOS DE INTERVALO

Según (Maté, Arroyo, & González-Rivera, 2016), un **intervalo** $[x]$ en una base $(E, \leq)$ es una pareja $[x] = [x^L, x^U]$ donde $x^L, x^U \epsilon\, E$ son los límites superior e inferior del

intervalo, respectivamente, de forma que $x^L \leq x^U$. Además, es posible representar un intervalo mediante su centro y mediante su radio $\langle x \rangle = \langle x^C, x^R \rangle$, definidos por las siguientes ecuaciones respectivamente, $x^C = \frac{1}{2} * (x^L + x^U)$ y $x^R = \frac{1}{2} * (x^U - x^L)$.

Por otro lado, una **serie de tiempo de intervalos** (ITS, del inglés *Interval Time Series*) es un tipo de proceso estocástico de intervalo, representando este una colección de variables aleatorias de intervalo que se ordenan a lo largo del tiempo, y se denota como $\{[x_t]\} = \{[x^L, x^U]\} = \{\langle x_t^C, x_t^R \rangle\}$ para $t = 1,2, \dots, T$. El análisis ITS es una de las aplicaciones más comunes para investigar, analizar y predecir información en diferentes ámbitos como la estadística, las finanzas o la econometría. Más adelante se definen una serie de métodos utilizados para predecir valores futuros a partir de datos de intervalo que han sido observados.

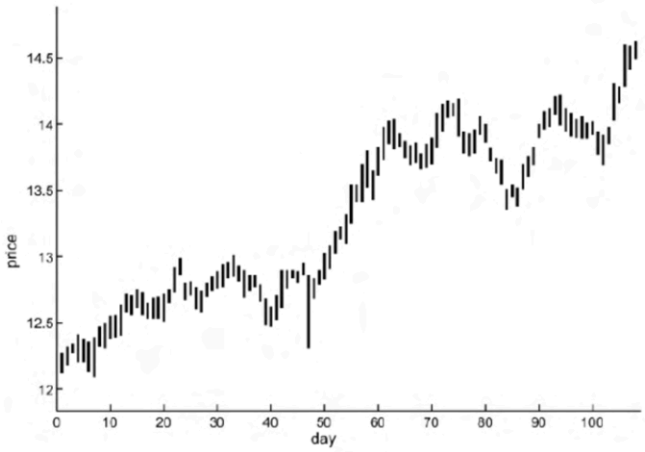


Ilustración 23. Serie de tiempo de intervalos (Arroyo & Maté, 2006)

Además, es interesante conocer el concepto de **Camino Aleatorio para intervalos** (iRW, del inglés *Interval Random Walk*) que se explica como una ITS $[Y_t]$ con una desviación siendo $[Y_t] = [\mu] + [Y_{t-1}] + [\varepsilon_t]$. Así, como consecuencia de ello se obtiene:

$$[Y_t^L] = [\mu^L] + [Y_{t-1}^L] + [\varepsilon_t^L]$$

Ecuación 4.1. Límite inferior del intervalo

$$[Y_t^U] = [\mu^U] + [Y_{t-1}^U] + [\varepsilon_t^U]$$

Ecuación 4.2. Límite superior del intervalo

El caso en el que $[\mu]$= 0, implica iRW sin desviación. El modelo iRW para ITS ofrece la predicción más simple para cada instante en la serie temporal, es decir, un modelo de Naive Forecast. Un modelo **Naive Forecast de intervalo** para una ITS $[Y_t]$ para $t = 2, 3, \ldots, T$, viene dado por $[\hat{Y}_t] = [Y_{t-1}]$, de donde se puede obtener:

$$\hat{Y}_t^L = Y_{t-1}^L$$

Ecuación 4.3. Límite inferior del intervalo en un modelo Naive Forecast

$$\hat{Y}_t^U = Y_{t-1}^U$$

Ecuación 4.4. Límite superior del intervalo en un modelo Naive Forecast

4.1.1 OPERACIONES ARITMÉTICAS CON INTERVALOS

Dados dos intervalos $[x]$ e $[y]$, donde $[x] = [x^L, x^U]$ e $[y] = [y^L, y^U]$ se pueden realizar con ellos las siguientes operaciones aritméticas, entre otras.

- Suma:

$$[x] + [y] = [x^L + y^L, x^U + y^U]$$

Ecuación 4.5. Suma de intervalos

- Diferencia:

$$[x] - [y] = [x^L - y^L, x^U - y^U]$$

Ecuación 4.6. Diferencia de intervalos

- Producto por un escalar:

$$\lambda[x] = [\lambda x^L, \lambda x^U] \; si \; \lambda \geq 0$$

$$\lambda[x] = [\lambda x^U, \lambda x^L] \; si \; \lambda < 0$$

Ecuación 4.7. Producto de un intervalo por un escalar

- Producto:

$$[x] \cdot [y] = [\min S\,, \max S];$$

Ecuación 4.8. Producto de intervalos

donde el conjunto S se define por la siguiente expresión, $S = [x^L y^L, x^U y^L, x^L y^U, x^U y^U]$

- Inversa:

$$\frac{1}{[x]} = \left[\frac{1}{x^U}, \frac{1}{x^L}\right]$$

Ecuación 4.9. Inversa de un intervalo

- Cociente:

$$\frac{[y]}{[x]} = [y] \cdot \left(\frac{1}{[x]}\right)$$

Ecuación 4.10. Cociente de un intervalo

4.1.2 MEDIDAS DE DISTANCIA PARA INTERVALOS

Distancia Euclídea

Con el objetivo de medir la precisión de una estimación es necesario medir la distancia entre la estimación y el valor real. Para el caso de los intervalos, una de las principales formas de medir esta distancia entre dos de ellos $[x]$ e $[y]$ es mediante la **distancia Euclídea**, definida por la siguiente ecuación.

$$d_E([x]\,, [y]) = \frac{1}{\sqrt{2}}\sqrt{(x^L - y^L)^2 + (x^U - y^U)^2}$$

Ecuación 4.11. Distancia Euclídea

Distancia de Hausdorff

La **distancia de Hausdorff** mide cómo de lejos están los intervalos $[x]$ e $[y]$ en el espacio, y se representa mediante la expresión de debajo.

$$d_H([x]\,,[y]) = \max(|x^L - y^L|, |x^U - y^U|) = |x^C - y^C| + |x^R - y^R|$$

Ecuación 4.12. Distancia de Hausdorff

Distancia de Bertoluzza

La **distancia de Bertoluzza** permite poner más peso a las medidas del intervalo que están más cerca del centro, según se indica a continuación.

$$d_B([x]\,,[y]) = \sqrt{(x^C - y^C)^2 + \theta(x^R - y^R)^2}$$

Ecuación 4.13. Distancia de Bertoluzza

4.1.3 ESTADÍSTICOS DESCRIPTIVOS

Siendo $[X]$ una serie de intervalos aleatorios de longitud n compuesto por los intervalos $[x_1], [x_2], \dots, [x_n]$ y siendo las medidas de la media de los centros y de los radios representados por las siguientes ecuaciones:

$$\overline{x^C} = \frac{1}{2n}\sum_{i=1}^{n}\left(x_i^L + x_i^U\right) = \frac{1}{n}\sum_{i=1}^{n} x_i^C$$

Ecuación 4.14. Media de los centros de una serie de intervalos

$$\overline{x^R} = \frac{1}{2n}\sum_{i=1}^{n}\left(x_i^U - x_i^L\right) = \frac{1}{n}\sum_{i=1}^{n} x_i^R$$

Ecuación 4.15. Media de los radios de una serie de intervalos

Se definen una serie de estadísticos que se muestran a continuación.

- Intervalo de la media de Aumann:

$$[\bar{x}] = \left[\overline{x^C} - \overline{x^R}, \overline{x^C} + \overline{x^R}\right]$$

Ecuación 4.16. Intervalo de la media de Aumann

- Intervalo de la mediana de Hausdorff:

$$[Me([X])]_H = [Me\{x_1^C, \dots, x_n^C\} - Me\{x_1^R, \dots, x_n^R\}, Me\{x_1^C, \dots, x_n^C\} + Me\{x_1^R, \dots, x_n^R\}]$$

Ecuación 4.17. Intervalo de la mediana de Hausdorff

- Intervalo de la mediana bajo 1-norm:

$$[Me([X])]_N = [Me\{x_1^L, \dots, x_n^L\}, Me\{x_1^U, \dots, x_n^U\}]$$

Ecuación 4.18. Intervalo de la mediana bajo 1-norm

4.1.4 CONCEPTOS BÁSICOS AL ANALIZAR PREDICCIONES PARA UNA ITS

Para explicar este apartado, se parte de la ITS real $\{[x_t]\}$ y de la ITS estimada $\{[\hat{x}_t]\}$, con $t = 1,2, \dots, T$. A continuación, se van a definir algunas medidas relevantes que tienen que ver con las series de tiempo de intervalos según (Arroyo & Maté, 2006).

Error de la Distancia Media (MDE)

Una manera de evaluar la precisión que tiene un método de predicción de ITS es mediante el uso del **Error de la Distancia Media** (MDE, del inglés *Minimum Distance Error*). Para obtenerlo, se realizar el cálculo de la distancia en cada momento y una vez obtenidos todos los valores, se calcula la media de todas las distancias. La ecuación que define el MDE es la que se muestra a continuación.

$$MDE^2 = \frac{1}{2}\frac{\sum_{t=1}^{T}[(x_t^L - \hat{x}_t^L)^2 + (x_t^U - \hat{x}_t^U)^2]}{T}$$

Ecuación 4.19. Error de la distancia media (MDE)

Cabe destacar que el MDE es equivalente al Error Cuadrático Medio (RMSE, del inglés *Root Mean Squared Error*) que se utiliza para medir la precisión de una estimación, pero con datos puntuales.

Varianza Relativa Media de Intervalo (iARV)

El estadístico de la **Varianza Relativa Media de Intervalo** (iARV, del inglés *Interval Average Relative Variance*) se encarga de realizar una comparación entre las estimaciones del modelo con las estimaciones dadas por los valores medios de una variable de intervalo y se define mediante la siguiente expresión:

$$iARV = \frac{\sum_{t=1}^{T}(x_t^L - \hat{x}_t^L)^2 + \sum_{t=1}^{T}(x_t^U - \hat{x}_t^U)^2}{\sum_{t=1}^{T}(x_t^L - \bar{x}_t^L)^2 + \sum_{t=1}^{T}(x_t^U - \bar{x}_t^U)^2}$$

Ecuación 4.20. Varianza relativa media de intervalo (iARV)

Donde T representa el número de intervalos ajustados, $[\hat{x}_t] = [\hat{x}_t^L, \hat{x}_t^U]$ define el intervalo ajustado número t, $[\bar{x}] = [\bar{x}^L, \bar{x}^U]$ es el intervalo que representa los promedios de las muestras y $\bar{x}^L$ y $\bar{x}^U$ son los promedios del límite inferior y superior, respectivamente.

Cuanto menor sea el valor de iARV, mejor será la predicción ya que implica que hay menos diferencia entre la estimación y los valores reales, siendo el modelo de predicción ideal aquel en el que $iARV = 0$.

Estadístico U de Theil para Intervalos (iU)

El **estadístico U de Theil** representa la comparación entre distintos modelos o métodos de predicción y un modelo Naïve. Para el caso de los datos de intervalo, se hace uso de estadístico U de Theil para intervalos (iU) que compara el modelo utilizado para realizar la estimación con el modelo de iRW, definido previamente. La fórmula para iU viene definida por la siguiente expresión:

$$iU = \sqrt{\frac{\sum_{t=2}^{T}(x_t^L - \hat{x}_t^L)^2 + \sum_{t=2}^{T}(x_t^U - \hat{x}_t^U)^2}{\sum_{t=2}^{T}(x_t^L - x_{t-1}^L)^2 + \sum_{t=2}^{T}(x_t^U - x_{t-1}^U)^2}}$$

Ecuación 4.21. U de Theil para intervalos (iU)

Según el valor obtenido en iU se puede llegar a las siguientes conclusiones en cuanto al modelo de predicción:

- Si $iU = 1$, el modelo observado se comporta igual que el de iRW.
- Si $iU > 1$, el modelo observado es peor que el de iRW.
- Si $iU < 1$, el modelo observado es mejor que el de iRW.

Los tres estadísticos explicados, MDE, iARV e iU, son medidas del tamaño de la desviación de las predicciones que realizan con respecto al valor real, por tanto, cuanto menores sean estos valores, mejor serán las predicciones.

4.2 ANÁLISIS DE INTERVALOS (AI)

Según (Cosín López-Medel, 2019), el origen del **Análisis de Intervalos** (IA, del inglés *Interval Analysis*) se remonta a hace más de 2000 años con el conocido físico y matemático Arquímedes, aunque no fue hasta la década de 1960 cuando Moore publicó el primer libro sobre el IA, donde desarrolló una serie de técnicas para la resolución de problemas que permitían reducir el error y utilizar soluciones delimitadas, con un enfoque inicial hacia la observación de los resultados computacionales surgidos de datos de intervalo.

El IA es por tanto una técnica matemática utilizada para establecer unos límites en los errores de redondeo y medida en el cálculo matemático. El uso del IA puede garantizar resultados realistas y matemáticamente correctos. Así, en vez de representar un valor como un único número, se representa cada valor como un rango de posibilidades. El uso de datos de intervalo es cada vez más común en muchas aplicaciones del día a día como, por ejemplo, para medir la presión de la sangre, los valores máximos y mínimos de temperatura en un periodo de tiempo o los precios de apertura o cierre de una acción en un día, entre otras cosas.

4.2.1 OPTIMIZACIÓN DE CARTERAS HACIENDO USO DE IA

A la hora de tomar decisiones por parte de los inversores, el riesgo y el retorno son dos de los factores más importantes que estos van a tener en cuenta. Sin embargo, según (Serban, Ferrara, & Costea, 2015), existe evidencia de que estos factores no capturan toda la información de interés que es necesaria para optimizar una cartera y esto se debe a que, al

ser datos puntuales, presentan un grado de incertidumbre que puede ser elevado. Esto a la vez, supone un problema a la hora de seleccionar aquellos activos en los que el inversor desea invertir para formar la cartera.

Debido a este inconveniente, han surgido nuevos métodos para modelar una cartera y buscar optimizarla, intentando reducir la incertidumbre debida a la imprecisión de los datos financieros o a su falta de completitud. (Serban, Ferrara, & Costea, 2015) proponen un método de optimización de carteras basado en IA que hace uso de programación lineal. El método en cuestión tiene en cuenta la imprecisión de los datos financieros iniciales y demuestra la efectividad de hacer uso de datos de intervalo para representar la incertidumbre de estos.

Por otro lado, (Li, Wang, & Jin, 2015) desarrollan otro método para la optimización de carteras haciendo uso de datos de intervalo para el riesgo y el retorno de un activo mediante programación lineal.

4.3 REGRESIÓN LINEAL CON DATOS DE INTERVALO (ILR)

Según (Maceda, 2018) se ha hecho habitual uso del análisis de datos puntuales para estudiar y predecir diferentes escenarios de diversos ámbitos en el mundo de la estadística, la economía o el Machine Learning entre otros. Hoy en día se sabe que con la ayuda de los datos de intervalo es posible obtener más información de precisión que haciendo uso únicamente de datos puntuales.

Dentro de este ámbito, el análisis de series de tiempo de intervalos (ITS) ha generado un gran interés, especialmente cuando se aplica al campo de las finanzas, donde se han llevado a cabo diferentes estudios de datos significativos como pueden ser los valores en la bolsa, la fluctuación de las monedas o en relación con este TFM, la gestión de carteras. Además, se han llevado a cabo diferentes análisis de varios métodos de predicción.

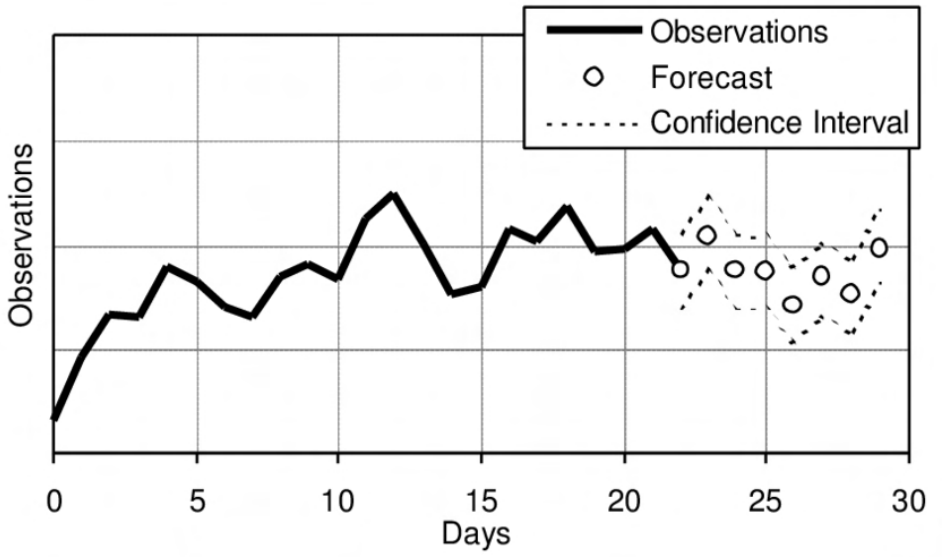


Ilustración 24. Análisis de regresión (Fraschini & Axhausen, 2001)

El **análisis de regresión** consiste en una serie de métodos estadísticos que se utilizan para estimar la relación entre una variable dependiente y una o más variables independientes, también denominadas predictores. Estos métodos se utilizan para el modelado, estimación y predicción, por lo que son de gran interés en el mundo de las finanzas y mientras que empezaron a desarrollarse con datos puntuales, desde hace unos 20 años empezaron a trabajar con datos de intervalo.

A continuación, se van a exponer una serie de ideas relacionadas con la regresión lineal de intervalos. Para empezar, cabe mencionar que, al hablar de la regresión lineal de intervalos (ILR, del inglés *Interval Linear Regression*), se deben destacar 4 paradigmas diferentes:

- Paradigma 1: $\beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n = y$, donde tanto las variables independientes x_n como la variable dependiente y representan datos puntuales. Esta es la regresión que se ha explicado hasta ahora en este proyecto.
- Paradigma 2: $\beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n = [Y]$, donde las variables independientes x_n representan datos puntuales, mientras que la variable dependiente $[Y]$ indica la presencia de un intervalo.

- Paradigma 3: $\beta_0 + \beta_1[X_1] + \cdots + \beta_n[X_n] = y$, donde las variables independientes $[X_n]$ representan datos de intervalo, mientras que la variable dependiente y es una variable puntual.
- Paradigma 4: $\beta_0 + \beta_1[X_1] + \cdots + \beta_n[X_n] = [Y]$, donde tanto las variables independientes $[X_n]$ como la variable dependiente $[Y]$, representan datos de intervalo.

4.3.1 MÉTODOS DE ILR

El análisis de regresión es la técnica estadística más común en el análisis de los datos de intervalo y permite explicar la relación entre una variable dependiente y una o varias variables independientes. Existen diferentes métodos de regresión para variables de intervalo que se han propuesto a lo largo de los años, según explica (Domingo, 2021). (Billard & Diday, 2000) presentaron el **Método del Centro** (CM, del inglés *Centre Method*), para llevar a cabo un análisis de regresión con datos de intervalo estableciendo dos modelos lineales con los mismos coeficientes en los límites inferior y superior de los intervalos, lo que es equivalente a utilizar los puntos medios correspondientes de los intervalos en el modelo de regresión lineal. Más tarde, (Billard & Diday, 2002) desarrollaron el **Método Min-Max** (mMM, del inglés *Min-Max Method*), que permitía construir modelos lineales en los límites superior e inferior de los intervalos, siendo los coeficientes de regresión diferentes. Por otro lado, (Lima Neto & De Carvalho, 2008) propusieron el **Método de Centro y Radio** (CRM, del inglés *Centre and Range Method*) consistente en ajustar dos modelos de regresión lineal según el centro y el radio de los intervalos, y fue propuesto como una alternativa mejorada al modelo CM descrito previamente. Este método se explica con más profundidad más adelante.

Más tarde, (Neto & Carvalho, 2010) presentaron un nuevo método, el **Método del Centro y Radio con Restricciones** (CCRM, del inglés *Constrained Centre and Range Method*), con el objetivo de mantener los radios de los intervalos no negativos, haciendo que los coeficientes de regresión en el modelo para los radios cumplan con esta restricción ya que en el método anterior a veces tenían problemas. Por su parte, (Giordiani, 2015) propone

un modelo de regresión más flexible haciendo uso del **Método Lasso** (del inglés *Least Absolute Shrinkage and Selection Operator*), siendo este un método de análisis de regresión que realiza la selección de variables y regularización para mejorar la exactitud e interpretabilidad del modelo estadístico producido por este, aplicado a datos de intervalo. Además, (Hao & Guo, 2017) desarrollaron el **Método de Centro y Radio Conjuntos con Restricciones** (CCRJM, del inglés *Constrained Centre and Range Joint Model*), el cual mejora la precisión en la predicción del anterior ya que busca un coeficiente para el centro y el radio, pero en este caso, lo hace conjuntamente, es decir, utilizando los valores de los radios para el modelo del centro y viceversa.

Todos estos métodos ya fueron analizados en (Domingo, 2021), por lo que ahora se van a definir nuevos métodos que resultan de interés y que han sido desarrollados más recientemente.

Teniendo en cuenta la información interna de un intervalo, (Xu, 2010) introdujo el **Método de Covarianza Simbólica** (SCM, del inglés *Symbolic Covariance Mehod*), y (Wang, Guan, & Wu, 2012) propusieron el **Método de Información Completa** (CIM, del inglés, *Complete Information Method*).

Según (Zhao, Wang, & Wang, 2022), aunque los métodos anteriores se han defendido de forma satisfactoria, existen algunas preocupaciones que de ellos nacen. Por ejemplo, el límite inferior ajustado puede ser mayor que el límite superior ajustado para los métodos CM, mMM y CRM. Por otro lado, para el CCRM, no tiene sentido que exista una correlación negativa entre la variable dependiente y las independientes en el modelo de regresión. Además, en el caso del CCRJM, la limitación de superposición entre los intervalos observados y ajustados puede tener un impacto negativo en la estimación y producción cuando existen datos atípicos en el conjunto de datos. Por último, SCM y CIM no son adecuados para escenarios donde los radios de los intervalos son relativamente grandes debido a la poca precisión de la predicción y al alto coste del cálculo.

4.3.2 MÉTODO CRM

El **Método del Centro y Radio** o *Centre and Range Method* (CRM), en inglés, fue presentado por (Lima Neto & De Carvalho, 2008) y fue propuesto como una alternativa mejorada al CM de (Billard & Diday, 2000).

Por un lado, el CRM se basa en el CM y a partir de él, añade el concepto de radios para poder llevar a cabo la predicción de forma mejorada.

Por otro lado, el CRM también hace uso de la idea del método minMax para añadir más información que el CM y poder satisfacer las necesidades que supone la complejidad de los datos de intervalo. En este caso el método, como su propio nombre indica, se basa en el centro y en el radio.

A continuación, se muestra cómo se va a representar el intervalo $[W]$ que, en vez de representarse por sus límites superiores e inferiores, $[W_L, W_U]$, se indicará a partir de su centro y radio, es decir, $[W_c, W_r]$, siendo estos:

$$W_c = (W_L + W_U)/2$$

Ecuación 4.22. Centro del intervalo

$$W_r = (W_L - W_U)/2$$

Ecuación 4.23. Radio del intervalo

Para este modelo, la relación de la regresión lineal se muestra de la siguiente forma:

$$y_{ci} = \beta_0^c + \beta_1^c x_{c1} + \cdots + \beta_n^c x_{cn} + \epsilon_{ci}$$

$$y_{ri} = \beta_0^r + \beta_1^r x_{r1} + \cdots + \beta_n^r x_{rn} + \epsilon_{ri}$$

Ecuación 4.24. Regresión lineal en CRM

Además, al igual que en los casos anteriores, es importante saber la suma de los cuadrados de las desviaciones para realizar la predicción posterior. Esta suma se calcula del siguiente modo:

$$S_{CRM} = \sum_{i=1}^{n} ((\epsilon_{ci})^2 + (\epsilon_{ri})^2)$$

Ecuación 4.25. Suma de los cuadrados de las desviaciones en CRM

Por último, al minimizar la ecuación anterior teniendo en cuenta que los centros $[Y_c]$ y $[X_{c1}, X_{c2}, \dots, X_{cn}]$ como en el CM, de forma independiente a los radios $[Y_r]$ y $[X_{r1}, X_{r2}, \dots, X_{rn}]$ con dos regresiones lineales diferentes, se obtienen los valores de $\beta_0^c, \beta_1^c, \dots, \beta_n^c$ y $\beta_0^r, \beta_1^r, \dots, \beta_n^r$ que minimizan S_{CRM}.

Los valores resultantes de $\hat{\beta}^c = (\hat{\beta}_1^c, \hat{\beta}_2^c, \dots, \hat{\beta}_n^c)$ y de $\hat{\beta}^r = (\hat{\beta}_1^r, \hat{\beta}_2^r, \dots, \hat{\beta}_n^r)$ predecirán $\hat{Y}_{CRM}$ como se indica:

$$\hat{y}_{ci} = \hat{\beta}_0^c + \hat{\beta}_1^c x_{c1} + \cdots + \hat{\beta}_n^c x_{cn}$$

$$\hat{y}_{ri} = \hat{\beta}_0^r + \hat{\beta}_1^r x_{r1} + \cdots + \hat{\beta}_n^r x_{rn}$$

Ecuación 4.26. Predicción en CRM

Como se aprecia en el desarrollo de este método, se utiliza el centro y radio, calculados a partir de estos límites superior e inferior para poder llevar a cabo la predicción.

4.3.3 MÉTODO DE REGRESIÓN ROBUSTO PARA DATOS DE INTERVALO BASADO EN PUNTOS MEDIOS Y RADIOS LOGARÍTMICOS (LN-IRR METHOD)

(Zhao, Wang, & Wang, 2022) proponen un nuevo modelo de regresión robusto para datos de intervalo basado en puntos medios o centros (*midpoints*) y el logaritmo de los radios (*log-ranges*) de los intervalos dependientes, **método LN-IRR** (*LN-IRR method*, en inglés), y obtienen los estimadores de parámetros usando la función de pérdida de Huber para tratar con posibles datos atípicos del conjunto de datos.

La **función de pérdida de Huber** se utiliza en estadística para la regresión robusta, que es menos sensible a los datos atípicos, y describe la penalización incurrida por un procedimiento de estimación.

Los autores mencionados comprueban que los resultados obtenidos son competitivos en comparación con otros métodos, especialmente para aquellos conjuntos de datos (*datasets*) donde existen datos atípicos o *outliers* y, además, en los escenarios donde los puntos medios y los radios de variables independientes están relacionados con los de la independiente.

En este caso, construyen un modelo de regresión basado en el logaritmo de los radios de los intervalos, lo cual les permite eliminar la restricción de no negatividad. El logaritmo les permite mantener la relación relativa entre los valores numéricos del conjunto de datos y reducir la escala de los mismos, lo cual disminuye los valores absolutos de los datos originales haciendo los datos más estables. Además, hacen uso de una alternativa al método de Mínimos Cuadrados Ordinarios (MCO) que utilizan los métodos citados anteriormente y que puede dar lugar a problemas en el caso de la presencia de datos atípicos. En su caso, van a utilizar regresión robusta, una técnica que tiene alta eficiencia y permite analizar datos contaminados con *outliers*, presentando resistencia hacia ellos.

El método que proponen se denomina **Método de Regresión Robusto para Datos de Intervalo en Puntos Medios y Radios Logarítmicos** (LN-IRR method, del inglés *Robust Regression Method for Interval-valued data based on Midpoints and Long-Ranges*). Concretamente, proponen dos modelos de regresión para centros y logaritmos de los radios de la variable de intervalo dependiente, respectivamente, con los centros y los radios de las variables de intervalo independientes involucradas. Además, el modelado del logaritmo de los radios permite eliminar la constante de no negatividad para los radios, lo cual es beneficioso para el uso de otras técnicas de regresión. Así, la estimación M basada en la función de pérdida de Huber (Huber, 1973) se puede aplicar al proceso de estimación para poder tolerar la existencia de *outliers* en la variable de intervalo dependiente.

En este enfoque se asegura que los nuevos radios ajustados sean siempre positivos. Así, el modelo es más flexible sin la necesidad de que exista la restricción de no negatividad. Además, el impacto en la presencia de datos anormales en la variable dependiente en el modelo ajustado se puede reducir gracias a la aplicación de la estimación M basada en la función de pérdida de Huber.

El modelo

Para explicar el modelo, se debe considerar $E = \{e_1, e_2, \dots, e_n\}$ como n ejemplos de $p+1$ variables de intervalo $Y, X_1, X_2, \dots, X_p$. Los puntos medios de Y y de X_j $(j = 1, \dots, p)$ se representan por variables de valor único Y^c y X_j^c $(j = 1, \dots, p)$, respectivamente. Por otro lado, los logaritmos de los radios de Y y de X_j $(j = 1, \dots, p)$ se representan por variables de valor único Y^r y X_j^r $(j = 1, \dots, p)$, respectivamente.

Para cada $e_i \in E$, $y_i = [y_i^L, y_i^U]$ y $x_{ij} = [x_{ij}^L, x_{ij}^U]$ son, respectivamente, los intervalos observados número i de Y y de X_j, donde y_i^L e y_i^U representan los límites inferior y superior de y_i y, de forma equivalente, x_{ij}^L y x_{ij}^U representan lo mismo para x_{ij}. Además, se calcula para la observación número i de Y^c, X_j^c, $\ln Y^r$ y X_j^r según las siguientes ecuaciones: $y_i^c = (y_i^L + y_i^U)/2$, $x_{ij}^c = (x_{ij}^L + x_{ij}^U)/2$, $\ln y_i^r = \ln[(y_i^U - y_i^L)/2]$ Y $x_{ij}^r = x_{ij}^U - x_{ij}^L)/2$, respectivamente.

Considerando X_j^c y X_j^r variables independientes, se puede construir el modelo de regresión lineal múltiple para datos de intervalo según se muestra a continuación.

$$\begin{cases} y_i^c = \beta_0^c + \beta_1^c x_{i1}^c + \cdots + \beta_p^c x_{ip}^c + \beta_{p+1}^c x_{i1}^r + \cdots + \beta_{2p}^c x_{ip}^r + \varepsilon_i^c \\ \ln y_i^r = \beta_0^r + \beta_1^r x_{i1}^c + \cdots + \beta_p^r x_{ip}^c + \beta_{p+1}^r x_{i1}^r + \cdots + \beta_{2p}^r x_{ip}^r + \varepsilon_i^r \end{cases}$$

Ecuación 4.27. Modelo de regresión lineal múltiple para datos de intervalo

Donde $\beta^c = (\beta_0^c, \beta_1^c, \dots, \beta_{2p}^c)^T$ y $\beta^r = (\beta_0^r, \beta_1^r, \dots, \beta_{2p}^r)^T$ son los vectores de coeficientes del modelo de regresión del punto medio y del modelo de regresión de logaritmo

del radio, respectivamente. Además, $\varepsilon^c = (\varepsilon_1^c, \dots, \varepsilon_n^c)^T$ y $\varepsilon^r = (\varepsilon_1^r, \dots, \varepsilon_n^r)^T$ son vectores de error aleatorios con media cero y varianza finita.

Así, es posible simplificar el modelo anterior según la siguiente ecuación:

$$\begin{cases} Y^c = X\beta^c + \varepsilon^c \\ \ln Y^r = X\beta^r + \varepsilon^r \end{cases}$$

Ecuación 4.28. Simplificación del modelo de regresión lineal múltiple para datos de intervalo

Donde $X = [1_n \; X^c \; X^r]$ representa la matriz de covarianzas de dimensiones $n \times (2p+1)$ según se indica a continuación lo siguiente,

$$1_n = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}, X^c = \begin{bmatrix} x_1^{cT} \\ \vdots \\ x_n^{cT} \end{bmatrix} = \begin{bmatrix} x_{11}^c & \cdots & x_{1p}^c \\ \vdots & \ddots & \vdots \\ x_{n1}^c & \cdots & x_{np}^c \end{bmatrix}, X^r = \begin{bmatrix} x_1^{rT} \\ \vdots \\ x_n^{rT} \end{bmatrix} = \begin{bmatrix} x_{11}^r & \cdots & x_{1p}^r \\ \vdots & \ddots & \vdots \\ x_{n1}^r & \cdots & x_{np}^r \end{bmatrix},$$

$Y^c = (y_1^c, y_2^c, \dots, y_n^c)^T$ y $\ln Y^r = (\ln y_1^r, \ln y_2^r, \dots, \ln y_n^r)^T$.

Por último, cabe destacar que en este método es posible predecir los valores del centro y del radio del intervalo dependiente utilizando los centros y los radios de las variables de intervalo independientes, lo que hace que el modelo sea aplicable para escenarios más generales.

4.3.4 MODELO MG

El **modelo M_G** (García-Bárzana, Ramos-Guajardo, Colubi, & Kontoghiorghes, 2020) es un modelo de regresión lineal múltiple para intervalos que surge como una extensión del modelo M (Blanco-Fernández, Corral, & González-Rodríguez, 2011) transformando la regresión lineal simple a regresión múltiple en un escenario más flexible. El modelo M_G se basa en aritmética de intervalos y en él, se toma el intervalo en su conjunto para realizar una regresión lineal múltiple. Se propone como una extensión de modelos anteriores y tiene en cuenta las relaciones entre los centros y los radios de los intervalos en una sola ecuación. El modelo ofrece una información mucho más precisa frente a los tradicionales, preservando la coherencia con respecto a la información que proporcionan los datos de intervalo.

Modelo de población

Sea $[Y]$ un intervalo aleatorio de respuesta y sean $[X_1], [X_2], \dots, [X_k]$ los k intervalos aleatorios explicativos, por lo que el parámetro k hace referencia al número de variables independientes. Se asume que las variables aleatorias de valor real centro y radio asociadas a los intervalos aleatorios no están degeneradas, que los intervalos aleatorios considerados tienen varianza finita y estrictamente positiva y que la matriz var-cov de las variables explicativas es invertible. El modelo de regresión lineal flexible múltiple basado en aritmética de conjuntos, denotado como M_G, se formaliza de la siguiente manera:

$$[Y] = \left[\left(\overline{b_1} X^{C^T} + \overline{b_4} X^{R^T}\right) \pm \left(\overline{b_2} X^{R^T} + \overline{b_3} \left|X^{C^T}\right|\right)\right] + \varepsilon$$

Ecuación 4.29. Regresión en modelo MG

Donde $\overline{b_1}$, $\overline{b_4}$ $\epsilon \mathbb{R}^k$, $\overline{b_2}$, $\overline{b_3}$ $\epsilon \mathbb{R}^{k+}$, $X^C = [X_1^C, X_2^C, \dots, X_k^C]^T \epsilon \mathbb{R}^k$, $X^R = [X_1^R, X_2^R, \dots, X_k^R]^T \epsilon \mathbb{R}^k$, y ε es un intervalo aleatorio que representa el error tal que $E(\varepsilon|[X]) = \Delta \epsilon K_c(\mathbb{R})$. $[X]$ y ε no van a estar correlacionadas, es decir, $\sigma_{\varepsilon,[X_i]} = 0$, para todo $i = 1, \dots, k$. Las relaciones lineales por separado para el centro y el radio de los intervalos transferidos se representan según las siguientes ecuaciones:

$$Y^C = X^{C^T} \overline{b_1} + X^{R^T} \overline{b_4} + \varepsilon^C$$

Ecuación 4.30. Centro del intervalo en modelo MG

$$Y^R = X^{R^T} \overline{b_2} + |X^{C^T}| \overline{b_3} + \varepsilon^R$$

Ecuación 4.31. Radio del intervalo en modelo MG

Así, ambas variables Y^C e Y^R se modelan a partir de la información completa proporcionada por los intervalos aleatorios independientes de $[X]$, caracterizados por el vector aleatorio $[X^C, X^R]$. De esta propiedad se puede obtener una conclusión inmediata y es que el modelo M_G permite mayor flexibilidad en la posible relación lineal entre los intervalos aleatorios que otros modelos estudiados previamente basados en aritmética de conjuntos. Sin

embargo, al incluir más coeficientes también aumenta la dificultad del proceso de estimación, al igual que ocurre en los problemas de regresión clásicos.

De cara a establecer una notación simplificada, se definen los siguientes intervalos: $x^M = [X^{C^T}, X^{C^T}]$, $x^S = [-X^{R^T}, X^{R^T}]$, $x^C = [-|X^{C^T}|, |X^{C^T}|]$ y $x^R = [X^{R^T}, X^{R^T}]$. Por lo tanto, el modelo M_G se puede definir igualmente a partir de la siguiente ecuación:

$$[Y] = X^{Bl}B + \varepsilon$$

Ecuación 4.32. Simplificación del intervalo en el modelo MG

Donde $X^{Bl} = (x^M|x^S|x^C|x^R) \ \epsilon \ K_c(\mathbb{R})^{1x4k}$ y $B = (\overline{b_1}|\overline{b_2}|\overline{b_3}|\overline{b_4})^T$.

Así, dada una muestra aleatoria simple $\{(y_j, x_{1,j}, \dots, x_{k,j})\}_{j=1}^n$ obtenida de los intervalos $([Y], [X_1], [X_2], \dots, [X_k])$ obtenemos:

$$[Y] = X^{eBl}B + \varepsilon$$

Ecuación 4.33. Ejemplo intervalo en modelo MG

Donde $[Y] = ([Y]_1, \dots, [Y]_n)^T$, $X^{eBl} = (X^M|X^S|X^C|X^R) \ \epsilon \ K_c(\mathbb{R})^{nx4k}$, $B = (\overline{b_1}|\overline{b_2}|\overline{b_3}|\overline{b_4})^T$ y $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)^T$ tal que $E(\varepsilon|[X]) = 1^n\Delta$. X^M es la matriz de intervalos de dimensión $(n \times k)$ tal que $(X^M)_{j,i} = [X_{i,j}^C, X_{i,j}^C]$ y análogamente X^S, X^C y X^R.

Método de estimación de mínimos cuadrados

El **método de estimación de los mínimos cuadrados** (LS, del inglés *Least Squares*) intenta buscar los parámetros $\hat{B}$ y $\hat{\Delta}$ que minimizan $d_\theta^2(y, X^{eBl}A + 1^nC)$ para $A\epsilon \ \mathbb{R}^{4k\times 1}$, $C\epsilon K_c(\mathbb{R})$ garantizando la existencia de los residuos $\varepsilon = y - H^{X^{eBl}}A$. A partir de aquí, es fácil ver que $(X^{eBl}A)^R = (Xa_2)^R + |X^C|a_3$, de forma que las condiciones siguientes se pueden incluir en el problema de minimización $(Xa_2)^R + |X^C|a_3 \leq y^r$.

Al igual que ocurre en la regresión clásica, la estimación del término de intervalo Δ se puede obtener mediante la siguiente ecuación:

$$\hat{\Delta} = \bar{y} - H^{\overline{X^{eBl}}}\hat{B}$$

Ecuación 4.34. Estimación del término de intervalo en Mínimos Cuadrados

Como resultado de ello, la estimación de mínimos cuadrados del parámetro B se obtiene minimizando la siguiente ecuación:

$$d_{\theta}^{2}\left(y - H^{X^{eBl}}A, \bar{y} - H^{\overline{X^{eBl}}}A\right)$$

Quedando esta definición sujeta a la siguiente condición, $(Xa_2)^R + |X^C|a_3 \leq y^r$.

Siendo $A = (a_1|a_2|a_3|a_4)$ de forma que $a_1, a_4 \epsilon\ \mathbb{R}^k$ y $a_2, a_3 \epsilon\ \mathbb{R}^{k+}$.

Según (Acín, 2022), este problema se puede transformar en un problema de optimización cuadrática para el cálculo de estimadores. Así, el problema presentaría la siguiente estructura:

$$min\frac{1}{2}x^t Hx - f^t x$$

Condicionado a:

$$Ax \geq b$$

En caso de querer profundizar sobre este tema, consultar (García-Bárzana, Ramos-Guajardo, Colubi, & Kontoghiorghes, 2020).

Capítulo 5. CAPM Y SU USO CON MÉTODOS DE REGRESIÓN PARA DATOS DE INTERVALO

El **Modelo de Valoración de Activos Financieros** (CAPM, del inglés *Capital Asset Pricing Model*) proporciona una expresión que nos permite relacionar el rendimiento esperado de un activo con su riesgo sistemático. Sin embargo, a la hora de tomar decisiones que tengan que ver con datos financieros, es importante tener en cuenta que existe cierta incertidumbre debida a que la información de la que se dispone puede ser imprecisa o estar incompleta. Para tratar esta incertidumbre en los datos existe la opción de considerar un intervalo de valores en el que el dato puede encontrarse, en vez de definirlo por un único valor que puede no ser correcto. Teniendo esto en cuenta, el modelo CAPM se puede adaptar y en ese caso, los retornos de los activos que se estén considerando serán variables definidas por intervalos en vez de variables dadas por un valor puntual, de acuerdo a lo que dice (Gioia, 2009). Para estudiar esto, se pueden utilizar distintos métodos de regresión con datos de intervalo que se presentan más adelante.

5.1 INTRODUCCIÓN

En el mundo de hoy en día es cada vez más común el uso de los datos para todo tipo de aplicaciones. Es habitual hacer uso de datos que se conocen, pero también se utilizan datos desconocidos para realizar cálculos o estimaciones. Las posibilidades de los datos y de las tecnologías continúan avanzando a pasos agigantados, impulsadas por personas, empresas, gobiernos o investigadores que buscan adaptarse a las preocupaciones presentes y futuras, obtener nuevos beneficios y mejorar la condición humana.

En el caso de los datos financieros, es habitual la presencia de incertidumbre debido a la imprecisión o a la falta de completitud de los mismos, por lo que al trabajar con ellos es importante estar preparado para hacer uso de información incierta y para ello se utilizan los intervalos.

Además, los datos de intervalo pueden ser considerados cuando existe interés en analizar un fenómeno en un intervalo de tiempo dado (diario, mensual...) de manera que se pueda evaluar no solo su "comportamiento medio" sino también su variabilidad, que es uno de los factores más destacables en el mundo de la economía y de las finanzas.

En relación con la variabilidad, se pueden encontrar diversos trabajos que hablan de métodos para pronosticarla mediante la predicción de intervalos basados en la confianza, sin embargo, estos métodos no son de gran calidad. Debido a esto, con el fin de mejorar la calidad de la predicción por intervalos, (He & Hu, 2007) hacen uso de datos de intervalos en lugar de datos puntuales, lo cual les permite obtener pronósticos de los límites superior e inferior de MCO (Mínimos Cuadrados Ordinarios) para predecir la variabilidad de los mercados de valores. Además, (He & Hu, 2008) presentan evidencia de que es posible alcanzar una mayor precisión en la predicción si los datos de intervalos se combinan con herramientas de álgebra de intervalos.

Como se ha explicado antes, el modelo CAPM permite relacionar el rendimiento esperado de un activo con su grado de riesgo sistemático, de manera que, en función de este riesgo, el valor de retorno será uno u otro. Para el caso de intervalos, se define el concepto de *IntervalCAPM* (**CAPM de Intervalos**) que hace referencia al modelo CAPM en el caso en el que los retornos de un activo determinado y la cartera del mercado sean variables definidas por intervalos.

(Gioia, 2015) describe algunos métodos de regresión para variables de intervalo que están presentes hoy en día en la literatura. Además, realizara una comparación entre ellos en el marco de CAPM de Intervalos.

(Gioia, 2009) hace un uso extensivo de herramientas de álgebra de intervalos, combinadas con técnicas de optimización para considerar el intervalo como una estructura conjunta y obtener un intervalo de soluciones, es decir, el intervalo que contiene todos los posibles valores que se asumen por una función determinada (función *slope* y función *intercept*), cuando los valores observados cambian en su propio intervalo de valores.

5.2 CAPM

Un mercado financiero es un espacio en el que se intercambian instrumentos financieros y se definen sus preferencias. Dentro de este mercado, son los inversores los que guiados por unos objetivos realizarán la compra o venta de los activos financieros, buscando crear una cartera de ellos que les proporcione la mayor rentabilidad posible dado un nivel de riesgo que estén dispuestos a asumir.

Cuanto mayor sea el riesgo asumido en una inversión, mayor será el retorno esperado de la misma, mientras que cuanto menor riesgo se asuma, el rendimiento que el inversor pueda esperar será también más pequeño. Gracias al CAPM, es posible cuantificar el nivel de riesgo y la recompensa esperada por asumirlo.

5.2.1 CAPM PARA DATOS PUNTUALES (CAPM CLÁSICO)

Suponiendo que se tienen p activos denominados como S_j $(j = 1, \dots, p)$ y que t_0 es el instante en el que un inversor decide comprar una cartera con estos activos, el inversor diversificará la cartera según los rendimientos esperados de los activos en un momento de tiempo futuro t_1. Los precios de los activos son determinados en cada instante, de acuerdo a la ley de la oferta y la demanda, y esto ocurre en los dos momentos que se estudian $(t_0\ y\ t_1)$ con la diferencia de que en t_0 los precios son conocidos mientras que en t_1 son estimados de acuerdo al comportamiento de los inversores en t_0.

Así, se tiene que x_j es la parte de la inversión total que se asigna a un activo y se cumple lo siguiente:

$$\sum_{j=1}^{n} x_j = 1$$

donde $x = (x_1, x_2, \dots, x_p)^T$ es la cartera del inversor considerado, siendo X el conjunto de posibles carteras.

En t_1 se asume que los retornos de los p valores son $n \times 1$ variables aleatorias denotadas por R_j $(j = 1, \dots, n)$, siendo n los diferentes estados que se contemplan, y se representan como una matriz de dimensiones $n \times p$ como la siguiente:

$$R^T = \begin{pmatrix} R_{11} & R_{12} & \cdots & R_{1p} \\ R_{21} & R_{22} & \cdots & R_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ R_{n1} & R_{n2} & \cdots & R_{np} \end{pmatrix}$$

Ecuación 5.1. Matriz de los retornos de los p valores en CAPM

Suponiendo una distribución de probabilidad discreta $p = (p_1, p_2, \dots, p_n)$ sobre los resultados de la variable aleatoria R_j $(j = 1, \dots, p)$, el retorno esperado del valor j^{th} es:

$$E(R_j) = \sum_{i=1}^{k} p_i R_{ij}$$

Ecuación 5.2. Retorno esperado del valor j en CAPM

La matriz de covarianza Q asociada a la matriz de retornos R^T, tiene en su fila r^{th} y columna s^{th} la covarianza entre R_r y R_s, que se expresa con $\sigma(R_r, R_s)$.

Así, el retorno esperado de una cartera x y su varianza, se pueden escribir con las siguientes ecuaciones:

$$E(x) = \sum_{j=1}^{n} E(R_j) \cdot x_j$$

Ecuación 5.3. Retorno esperado de una cartera en CAPM

$$\sigma^2(x) = \sum_{j=1}^{n} x_j \cdot \sum_{i=1}^{n} x_i \, \sigma(R_r, R_s)$$

Ecuación 5.4. Varianza de una cartera en CAPM

Según la primera ecuación, el retorno esperado de la cartera es la media ponderada de los retornos de los activos que la componen, mientras que su varianza dependerá de algo más que simplemente las varianzas de estos activos.

De acuerdo con Markowitz, los inversores elegirían una cartera de media-varianza eficiente, es decir, una cartera con el retorno esperado más alto para un determinado nivel de varianza. Dado que la varianza de una cartera x se ve como el riesgo de inversión, la mejor inversión sería por tanto aquella que tenga la mínima varianza, sujeta a un retorno E_0, lo que lleva a la siguiente definición:

$$min_x \sigma^2(x)$$

donde:

$$E(x) = E_0$$

$$\sum_{j=1}^{n} x_j = 1$$

$$x_i \geq 0$$

El CAPM deduce un criterio de decisión compartido por todos los agentes. La hipótesis adicional introducida para derivar el modelo es que el mercado está en equilibrio, es decir, que la suma de las cantidades de todos los activos disponibles en el mercado es igual a la suma de las cantidades de todos los activos requeridos. Así, el modelo proporciona una expresión que pone en relación el rendimiento esperado de un activo con su riesgo.

Se define la cartera del mercado como un vector $A = (A_1, A_2, \dots, A_n)$ donde A_j es la cantidad agregada al activo de riesgo j^{th} disponible en la economía. La cartera del mercado

es una cartera consistente en la suma ponderada de cada activo del mercado, siendo sus pesos las proporciones que indican la presencia de este activo en el mercado, asumiendo que estos activos son infinitamente divisibles. Así, el retorno esperado de la cartera se define como:

$$E(A) = \sum_{j=1}^{n} E(R_j) \cdot A_j$$

Ecuación 5.5. Retorno esperado de una cartera A

A continuación, se define el siguiente teorema:

Teorema 1: *Si el mercado está en equilibro, para cada activo R_j se cumple:*

$$E(R_j) - r_f = (E(A) - r_f) \cdot \beta_j$$

Ecuación 5.6. Condición para el teorema de CAPM

donde r_f es la prima libre de riesgo y β_j es el factor de proporcionalidad, definido por la siguiente expresión:

$$\beta_j = \frac{\sigma(A, j)}{\sigma^2(A)}$$

Ecuación 5.7. β_j en CAPM

siendo $\sigma(A, j)$ la covarianza entre el retorno de la cartera del mercado y el retorno del activo j^{th} y $\sigma^2(A)$, la varianza de la cartera del mercado.

Como se ha definido en la ecuación de arriba, el modelo CAPM indica que, en equilibrio, la diferencia entre la tasa de rendimiento esperada de cada activo con riesgo y la tasa libre de riesgo es proporcional a la diferencia entre la tasa de rendimiento esperada en la cartera del mercado y la tasa libre de riesgo.

Así, la ecuación del Teorema 1 puede tener diferentes usos:

I. Dados los retornos esperados $[E(R_j)]$ y $[E(A)]$ de un activo S_j y de la cartera del mercado, respectivamente, y dada una tasa libre de riesgo r_f, es posible calcular la β_j correspondiente;

II. Dada la β_j de S_j y la tasa libre de riesgo r_f, se puede calcular el retorno esperado $[E(R_j)]$;

III. Dada la tasa libre de riesgo r_f, es posible utilizar un método de estimación para determinar el valor de β_j de un activo de riesgo particular y así calcular el retorno esperado del activo $[E(R_j)]$.

Así, siendo Z_j y Z_m los vectores de los excesos de retorno del activo S_j y de la cartera del mercado respectivamente, en diferentes periodos de tiempo, se pueden definir estos excesos de retorno por medio del modelo de mercado de índice único de exceso de retorno, representado por la siguiente ecuación:

$$Z_j = \alpha + \beta Z_m + e_j$$

Ecuación 5.8. Excesos de retorno

donde e_j es el vector de perturbación y $Z_j = R_j - r_f$ y $Z_m = A - r_f$. Además, los estimadores de β y α son los estimadores de MCO. A partir de la estimación de β es posible calcular la prima de riesgo exigida por el mercado y compararla con la que va implícita en el precio real del activo. Comprar un activo cuya prima de riesgo excede la que predice el CAPM, y vender activos donde la prima de riesgo del CAPM excede la del activo es una regla de decisión para los inversores en los mercados financieros. Por ello, el CAPM es una herramienta de gran utilidad para el análisis de los precios de los activos en los mercados financieros.

5.2.2 CAPM PARA DATOS DE INTERVALO

Como ya se ha mencionado, los intervalos pueden ser útiles para representar la incertidumbre en datos financieros. En caso de definir los retornos con intervalos, un inversor debería tener en cuenta una serie de cosas:

I. Si lo retornos fluctúan con respecto a unos valores fijos, ¿cómo se reflejan estas perturbaciones en el riesgo?

II. ¿En qué intervalo este riesgo cambia?

III. ¿Cuál es el intervalo en el que la estimación del rendimiento esperado del activo que se esté considerando varía?

El método CAPM que se ha definido previamente se puede extender al caso en el que los retornos del activo considerado se definan mediante intervalos (*IntervalCAPM*) y se va a explicar a continuación.

Se supone que el retorno de un valor S_j $(j = 1, \dots, n)$ no solo varía con respecto al tiempo, sino que también va a poder representarse por un intervalo de valores cuando ocurra cada estado, entonces R_j $(j = 1, \dots, n)$ estará definida por variables valoradas en intervalos, denotado por R_j^I $(j = 1, \dots, n)$ y representados por columnas según la siguiente matriz de intervalos:

$$(R^T)^I = \begin{pmatrix} [\underline{R}_{11}, \overline{R}_{11}] & [\underline{R}_{12}, \overline{R}_{12}] & \cdots & [\underline{R}_{1p}, \overline{R}_{1p}] \\ [\underline{R}_{21}, \overline{R}_{21}] & [\underline{R}_{22}, \overline{R}_{22}] & \cdots & [\underline{R}_{2p}, \overline{R}_{2p}] \\ \vdots & \vdots & \ddots & \vdots \\ [\underline{R}_{n1}, \overline{R}_{n1}] & [\underline{R}_{n2}, \overline{R}_{n2}] & \cdots & [\underline{R}_{np}, \overline{R}_{np}] \end{pmatrix}$$

Ecuación 5.9. Matriz de los retornos en CAPM de intervalos

donde $[\underline{R}_{ij}, \overline{R}_{ij}]$ es el intervalo en el que el retorno del valor S_j cae cuando el estado i^{th} ocurre.

El objetivo del CAPM por intervalos es extender el modelo al caso de los retornos por intervalos.

Cuando se da una matriz de retornos de intervalos, los excesos de retornos del valor j^{th} y de la cartera del mercado se definen por las siguientes variables de intervalos:

$$Z_j^I = \left(Z_{jt} = \left[\underline{Z}_{jt}, \overline{Z}_{jt}\right]\right), \;\; t = 1, \dots, n$$

Ecuación 5.10. Exceso de retornos del valor j en CAPM de intervalos

$$Z_m^I = \left(Z_{mt} = \left[\underline{Z}_{mt}, \overline{Z}_{mt}\right]\right), \;\; t = 1, \dots, n$$

Ecuación 5.11. Exceso de retornos de la cartera del mercado en CAPM de intervalo

En este caso concreto, la ecuación de regresión debe tratarse con métodos de regresión por intervalos. Aplicando el método *Iregr,* hace que se trate a la Ecuación 5.8 como una ecuación de regresión por intervalos, que quedaría de la siguiente forma:

$$Z_j^I = \hat{\alpha}_j^{\,I} + \hat{\beta}_j^{\,I} Z_m^I + E^I$$

Ecuación 5.12. Simplificación del exceso de retornos del valor j en CAPM de intervalos

Así, cada vez que un retorno R_{ij} toma un valor particular dentro de su propio intervalo de variación $[\underline{R}_{ij}, \overline{R}_{ij}]$, una serie de puntos en el plano (Z_j, Z_m) se determinan de forma unívoca. Además, la pendiente y el intercepto de la línea de regresión, que da el "mejor" ajuste para esta serie de puntos, son elementos que pertenecen a los intervalos $\hat{\beta}_j^{\,I}$ y $\hat{\alpha}_j^{\,I}$ respectivamente. Por lo tanto, se deduce lo siguiente:

I. El intervalo $\hat{\beta}_j^{\,I}$ es un conjunto de todas las betas del valor S_j cuando cada retorno $R_{ij}, (i = 1, \dots, k)$ varía en su propio intervalo de valores.

II. Recordando que para el activo de valor real S_j el modelo CAPM dice que el intercepto α_j es cero, es posible interpretar el intervalo $\hat{\alpha}_j^{\,I}$ como el conjunto de todos los "errores" que se podrían hacer al hacer uso del CAPM para predecir el rendimiento esperado $E(R_j)$ del valor S_j para cada $R_{ij} \in [\underline{R}_{ij}, \overline{R}_{ij}]$.

Así, la estimación de un intervalo Z_{jt} vendrá dada por:

$$Z_{jt} = \hat{\alpha}_j^{\,I} + \hat{\beta}_j^{\,I} Z_{mt}$$

Ecuación 5.13. Estimación de un intervalo Z_{jt}

donde las operaciones son operaciones de álgebra de intervalos.

Es importante poner esto en relación con métodos de regresión mediante intervalos, como hace (Gioia, 2009), ya que con el método de CAPM por intervalos, los inversores pueden conocer la incertidumbre que existe alrededor de los valores de riesgo e intercepto definidos. Sin embargo, no todos los métodos de regresión por intervalos muestran esta incertidumbre de la misma manera. Según (Gioia, Capital Asset Pricing Model Using Regression Methods for Interval-Valued Variables, 2015), los métodos CM (Método del Centro), mMM (Método MinMax) y CRM (Método del Centro y Radio) calculan la estimación del intervalo del retorno de la cartera dando unos valores de pendiente e intercepto escalares, sin embargo, el método Iregr (del inglés *Optimal Interval of Solution Based Method*) permite dar soluciones por intervalos también para estos datos, siendo $\hat{\alpha}_j^{\,I}$ el intervalo del intercepto y $\hat{\beta}_j^{\,I}$, el intervalo para la pendiente o riesgo.

5.3 REGRESIÓN DE INTERVALO Y CAPM

Existen diferentes métodos para ajustar una regresión lineal de variables valoradas en intervalos. (Gioia, 2015) estudia dos enfoques diferentes que tienen el objetivo común de predecir una variable de intervalo dependiente de una o más variables de intervalo independientes. El primero de los enfoques consiste en estimar el valor de la variable dependiente haciendo uso de funciones de regresión con parámetros escalares. El segundo trata de pronosticar la variable de intervalo dependiente mediante el cálculo de parámetros de intervalo, es decir, definiendo la línea de regresión de intervalo. Esta metodología utiliza técnicas de optimización para calcular un intervalo de soluciones en vez de una solución dada por un intervalo. Así, devuelve un intervalo que contiene todos los valores posibles asumidos por una función determinada cuando los valores observados varían dentro de su propio intervalo de variación.

Considerando una variable de intervalos dependiente y p variables de intervalos independientes, se tiene: $Y^I = \left(Y_i = \left[\underline{y}_i, \overline{y}_i\right]\right)$, $X_j^I = \left(X_{ij} = \left[\underline{x}_{ij}, \overline{x}_{ij}\right]\right)$, con $i = 1, \dots, n$ y $j = 1, \dots, p$. Además, siendo $E = \{e_1, \dots, e_m\}$ un conjunto de ejemplos definidos por las $p + 1$ variables de intervalo Y^I, $X_1^I, \dots, X_p^I$, cada ejemplo $e_i \epsilon E$ se representa por el vector de intervalo $Z_i^I = (X_{i1}, \dots, X_{ip}, Y_i)$, donde X_{ij} e Y_i se corresponden con los valores observados de X_j^I e Y^I, respectivamente.

(Gioia, 2015) en su estudio del modelo CAPM utilizando métodos de regresión para variables dadas en intervalos, añade a los métodos de regresión de CM, mMM y CRM, el método Iregr de regresión de intervalos, que se define a continuación.

5.3.1 MÉTODO DE REGRESIÓN DE INTERVALOS (IREGR)

De acuerdo con los métodos CM, mMM y CRM, la predicción de un intervalo $\hat{Y} = [\hat{\underline{y}}, \hat{\overline{y}}]$ es un intervalo que se calcula a partir de los resultados de métodos de regresión clásicos que minimizan los criterios diferentes al MCO. A continuación, se va a describir el **Método de Regresión de Intervalos** (Iregr, del inglés *Interval Regression*), que hace un uso exhaustivo de álgebra de intervalos combinada con técnicas de optimización y permite considerar el intervalo como una estructura conjunta y a partir de él, calcular un intervalo de soluciones que contenga todos los posibles valores que pueden darse como solución.

Para empezar, se considera el caso de regresión de intervalos simple, teniendo en cuenta las siguientes variables independiente y dependiente, respectivamente.

$$X^I = \left(X_i = \left[\underline{x}_i, \overline{x}_i\right]\right), \qquad i = 1, \dots, n$$

$$Y^I = \left(Y_i = \left[\underline{y}_i, \overline{y}_i\right]\right), \qquad i = 1, \dots, n$$

La variable dependiente Y^I estaría definida por la siguiente expresión:

$$Y^I = \alpha^I + \beta^I X^I + E^I$$

Ecuación 5.14. Variable dependiente en Iregr

donde E^I es la componente que representa el intervalo de error.

Este método tiene en cuenta todos los posibles valores de las componentes x_i e y_i, cada una de las cuales dentro de su intervalo de valores $[\underline{x}_i, \overline{x}_i]$ y $[\underline{y}_i, \overline{y}_i]$, con , $i = 1, \dots, n$. Así, al hacer la regresión entre dos variables de intervalos se obtendrá un conjunto de líneas de regresión, cada una de ellas con el mejor ajuste, según MCO, de una serie de puntos en el plano. Este conjunto de puntos se determina de forma unívoca cada vez que los componentes $x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n$ toman un valor particular dentro de su propio intervalo de variación.

Matemáticamente, calcular la línea de regresión de intervalo entre las dos variables de intervalo definidas X^I e Y^I comienza con obtener los siguientes dos conjuntos:

$$\hat{\beta}^I = \left\{ \hat{\beta}(X,Y) = \frac{\sum_{i=1}^{n}(x_i - x_m)(y_i - y_m)}{\sum_{i=1}^{n}(x_i - x_n)^2}, \qquad /X \in X^I, Y \in Y^I \right\}$$

Ecuación 5.15. $\hat{\beta}^I$ en Iregr

$$\hat{\alpha}^I = \left\{ \hat{\alpha}(X,Y) = y_m - \hat{\beta} x_m, \;\; /X \in X^I, Y \in Y^I \right\}$$

Ecuación 5.16. $\hat{\alpha}^I$ en Iregr

donde x_m e y_m vienen dadas por las siguientes expresiones:

$$x_m = \frac{1}{n}\sum_{i=1}^{n} x_i \;\; ; \;\; y_m = \frac{1}{n}\sum_{i=1}^{n} y_i$$

Ecuación 5.17. x_m e y_m en Iregr

Estos dos conjuntos pueden obtenerse numéricamente solucionando ciertos problemas de optimización, buscando el mínimo y el máximo de las funciones $\hat{\beta}(X,Y)$ y $\hat{\alpha}(X,Y)$ descritas en Ecuación 5.15 y Ecuación 5.16.

Ambas funciones son continuas en un conjunto conectado y compacto, lo que implica que los sets de Ecuación 5.15 y Ecuación 5.16 son los siguientes intervalos cerrados.

$$\hat{\beta}^I = \begin{bmatrix} \min \hat{\beta}(X,Y), & \max \hat{\beta}(X,Y) \\ X \in X^I, Y \in Y^I & X \in X^I, Y \in Y^I \end{bmatrix}$$

Ecuación 5.18. Intervalo cerrado del set de Ecuación 5.15

$$\hat{\alpha}^I = \begin{bmatrix} \min \hat{\alpha}(X,Y), & \max \hat{\alpha}(X,Y) \\ X \in X^I, Y \in Y^I & X \in X^I, Y \in Y^I \end{bmatrix}$$

Ecuación 5.19. Intervalo cerrado del set de Ecuación 5.16

Estos intervalos se pueden interpretar de forma que dado un intercepto $\hat{\alpha}$ en el intervalo $\hat{\alpha}^I$, existe una pendiente $\hat{\beta}$ en el intervalo $\hat{\beta}^I$, que hacen que la línea de regresión Y sea la única línea que realiza el mejor ajuste, en el contexto de MCO, de una serie dada de puntos $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ en el plano $(x_i \in X_i, y_i \in Y_i,\ i = 1, \dots, n)$ siendo la línea de regresión Y la que se muestra debajo.

$$Y = \hat{\alpha} + \hat{\beta} X$$

Ecuación 5.20. Línea de regresión en Iregr

Así, la predicción de un intervalo $Y = \left[\underline{y}, \overline{y}\right]$ de Y^I se obtendría a partir de la siguiente expresión:

$$Y = \hat{\alpha}^I + \hat{\beta}^I X$$

Ecuación 5.21. Predicción de un intervalo Y

Una vez visto el caso de regresión simple, es interesante analizar la opción desde el punto de vista de la regresión múltiple, es decir, considerando una variable dependiente y p variables independientes, todas ellas evaluadas en intervalos, siendo estas las que se muestran debajo.

$$Y^I = \left(Y_i = \left[\underline{y}_i, \overline{y}_i\right]\right), X_j^I = \left(X_{ij} = \left[\underline{x}_{ij}, \overline{x}_{ij}\right]\right),\ \ i = 1, \dots, n,\ \ j = 1, \dots, p$$

Según esto, la variable dependiente se representaría de la siguiente forma:

$$Y^I = \beta_0^I + \beta_1^I X_1^I + \beta_2^I X_2^I + \cdots + \beta_p^I X_p^I + E^I$$

Ecuación 5.22. Variable dependiente en Iregr

donde E^I es la componente que representa el intervalo de error. Además, se debe tener en cuenta todos los posibles valores de las componentes x_{ij} e y_i, cada una de las cuales dentro de su intervalo de valores $[\underline{x}_{ij}, \overline{x}_{ij}]$ y $[\underline{y}_i, \overline{y}_i]$, con , $i = 1, \dots, n$.

En este caso, el método calcula una serie de hiperplanos, ya no líneas, donde cada uno de ellos realizará el mejor ajuste posible, en el contexto de MCO, de una serie de puntos en el espacio. El hiperplano se determina de forma unívoca cada vez que los componentes tomen un valor particular dentro de su propio intervalo de variación.

A continuación, se toma $X^I = (1, X_1^I, X_2^I, \dots, X_p^I)$ y una matriz de intervalos de dimensiones $n \times (p+1)$, donde la primera columna es es un vector de unos de dimensión $n \times 1$, y el resto de las columnas serán p variables independientes de intervalo. El método consiste en definir el siguiente vector de intervalos $\hat{\beta}^I$.

$$\hat{\beta}^I = (\hat{\beta}_0^I, \hat{\beta}_1^I, \dots, \hat{\beta}_p^I) = \{\hat{\beta} = \left(\hat{\beta}_0(X,Y), \hat{\beta}_1(X,Y), \dots, \hat{\beta}_p(X,Y)\right) = (X^T X)^{-1} X^T Y,$$

$$X \in X^I, Y \in Y^I\}$$

$\hat{\beta}^I$ representa un conjunto de todos los posibles vectores $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p)$ coeficientes de un hiperplano Y de la forma que se muestra debajo.

$$Y = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \cdots + \hat{\beta}_p X_p$$

Ecuación 5.23. Hiperplano Y en Iregr

Generando así, al igual que para el caso de la regresión simple, el mejor ajuste, según MCO, de una serie de puntos en el espacio dado por x_{ij} e y_i $(i = 1, \dots, n,\ j = 1, \dots, p)$, cada uno dentro de su propio intervalo de variación.

De forma matemática, se debe resolver el siguiente problema de optimización para cada $j = 1, \dots, p$ para obtener $\hat{\beta}_j^I$.

$$\hat{\beta}_j^I = \begin{bmatrix} \min \hat{\beta}_j(X,Y), & \max \hat{\beta}_j(X,Y) \\ X \in X^I, Y \in Y^I & X \in X^I, Y \in Y^I \end{bmatrix}$$

Ecuación 5.24. Intervalo cerrado de $\hat{\beta}_j^I$

Finalmente, según (Gioia, 2015) el método Iregr es sensible a ciertas características de los datos de entrada y particularmente, funciona bien cuando no hay datos atípicos o *outliers*, cuando los radios de los intervalos reflejan fluctuaciones naturales, excluyendo eventos excepcionales, y cuando los intervalos están separados de forma correcta.

5.4 IREGR EN EL MARCO DEL CAPM DE INTERVALOS

Haciendo uso de los métodos de regresión de intervalos citados previamente (CM, mMM, CRM) para CAPM con datos de intervalo, se puede estimar un valor de intervalo para el rendimiento esperado de un activo S_j. Sin embargo, estos métodos no ofrecen información sobre la pendiente y el intercepto de la línea de regresión, es decir, no permiten saber qué pasa con el riesgo y el intercepto cuando los retornos varían dentro del intervalo de variación, lo cual supone un inconveniente si se desea saber ciertos datos como si los retornos fluctúan con respecto a unos valores fijos, cómo se reflejan estas perturbaciones en el riesgo o en qué intervalo puede variar el riesgo.

Para dar respuesta a estas preguntas, (Gioia, 2015) defiende que el método Iregr puede ser utilizado ya que devuelve el intervalo de la pendiente $\hat{\beta}_j^I$, así como el intervalo del intercepto $\hat{\alpha}_j^I$, de la línea de regresión de intervalo, que se definen en la Ecuación 5.12. $\hat{\beta}_j^I$ y $\hat{\alpha}_j^I$ representan un conjunto de estimaciones de todas las pendientes e interceptos de una determinada SML, cuando el retorno varía dentro de su propio intervalo de valores. Así, un inversor que utilice el método de CAPM de intervalos y calcule $\hat{\beta}_j^I$ y $\hat{\alpha}_j^I$ para un activo

determinado, tiene la oportunidad de conocer no solo unos valores escalares del riesgo y del intercepto, sino también la incertidumbre que existe alrededor de estas cantidades.

Capítulo 6. MÉTODOS Y MODELOS DE PREDICCIÓN

A la hora de invertir no solo es importante tener en cuenta el análisis del comportamiento histórico de una acción, sino que para los inversores también es de gran utilidad poder hacerse una idea de cómo va a oscilar el precio de esa acción en el futuro. Por ello, existen un conjunto de herramientas y técnicas estadísticas de las cuales hacen uso para pronosticar y predecir el comportamiento del precio de las acciones ante ciertos eventos. Esto es lo que se conoce como modelos de predicción (Nowak, 2022).

A lo largo de este capítulo se van a presentar una serie de métodos de predicción de valores que más tarde serán utilizados en la herramienta en la que se ha trabajado a lo largo del proyecto.

6.1 MÉTODO DE SUAVIZADO EXPONENCIAL SIMPLE (SES)

El método de **Suavizado Exponencial Simple** (SES, del inglés *Simple Exponential Smooth*) es una técnica de predicción de series temporales que asume que los valores futuros de una serie pueden ser estimados asignando pesos decrecientes a las observaciones pasadas, ya que considera más importantes los datos más recientes (López Mínguez, 2007).

A la hora de implementar el método de Suavizado Exponencial Simple (SES) para predecir precios futuros de acciones, el primer paso es separar el conjunto de datos en dos subconjuntos, uno de test y otro de entrenamiento. El primero de ellos será utilizado para evaluar el comportamiento del modelo mientras que el segundo será necesario para ajustar el modelo. A continuación, se realizará el cálculo de la predicción para los datos de test a partir de la ecuación siguiente.

$$F_{t+1} = \alpha Y_t + (1 - \alpha) F_t$$

Ecuación 6.1. Suavizado Exponencial Simple (SES)

Donde F_{t+1} es el valor estimado para el momento $t+1$, Y_t es el valor real en t, F_t es el valor pronosticado en el momento t y α es el parámetro de suavizado, que puede tomar cualquier valor entre 0 y 1. El valor inicial de predicción para F_1 se puede establecer como la primera observación de la serie o como la media de las primeras observaciones.

Conviene destacar que α es un parámetro que permite controlar el nivel de suavizado del modelo y determina el peso asignado a la observación más reciente. Cuanto mayor sea el valor de α, mayor peso tendrán las observaciones más recientes del modelo a la hora de realizar la estimación, lo que dará como resultado pronósticos más sensibles, y potencialmente con mayor volatilidad.

El siguiente paso será comparar las predicciones del modelo con los datos reales del subconjunto de test de cara a poder evaluar el comportamiento del modelo. Para ello, se ha calculado el RMSE (*Root Mean Square Error*).

A continuación, se muestra un ejemplo del resultado obtenido para una sola compañía haciendo uso del modelo SES. En este caso, se ha evaluado el comportamiento de Starbucks desde el 12 de Febrero de 2014 hasta el 13 de Enero de 2023, haciendo uso de un valor de $\alpha = 0{,}4$.

Ilustración 25. Ejemplo de predicciones a partir del modelo SES para Starbucks (Elaboración propia)

Se observa que la predicción estimada devuelve unos valores más suavizados que los datos reales. Además, se puede apreciar en verde el valor medio estimado para fechas posteriores al periodo de estudio, las cuales se indican en la siguiente tabla.

Fecha	Mínimos	Máximos
20-Jan-2023	98.298	103.53
27-Jan-2023	98.298	103.53
03-Feb-2023	98.298	103.53
10-Feb-2023	98.298	103.53

Tabla 5. Resultados de predicciones a partir del modelo SES para Starbucks (Elaboración propia)

Por otro lado, se ha extendido la función anterior para que pueda comparar el comportamiento de los precios de distintas empresas del índice S&P 100 para elaborar una estimación de los valores futuros y así poder realizar un ranking en función de ello, ordenándolas de mayor a menor valor esperado. El resultado obtenido, para una serie de empresas del sector tecnológico, es el siguiente.

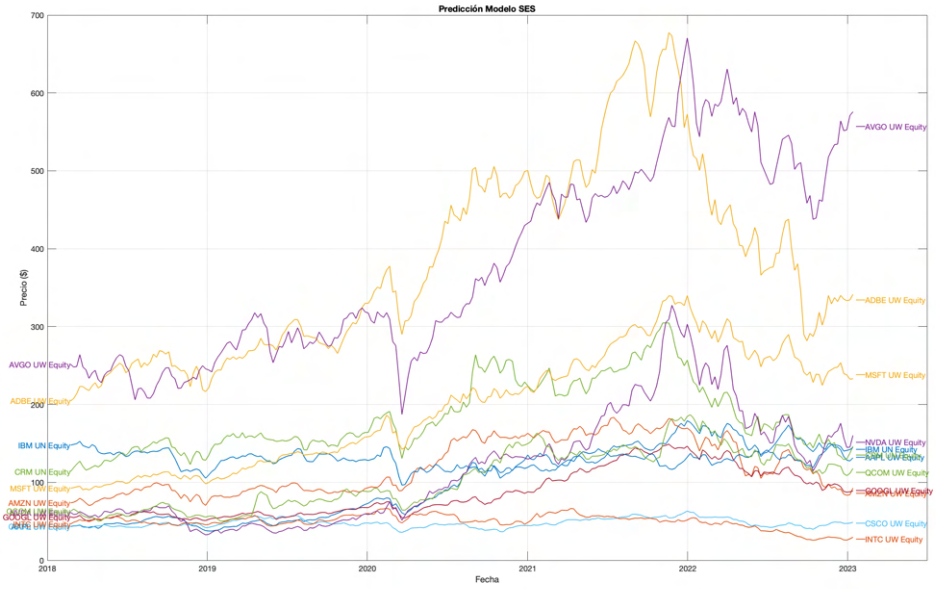


Ilustración 26. Ejemplo de predicciones a partir del modelo SES para varias compañías (Elaboración propia)

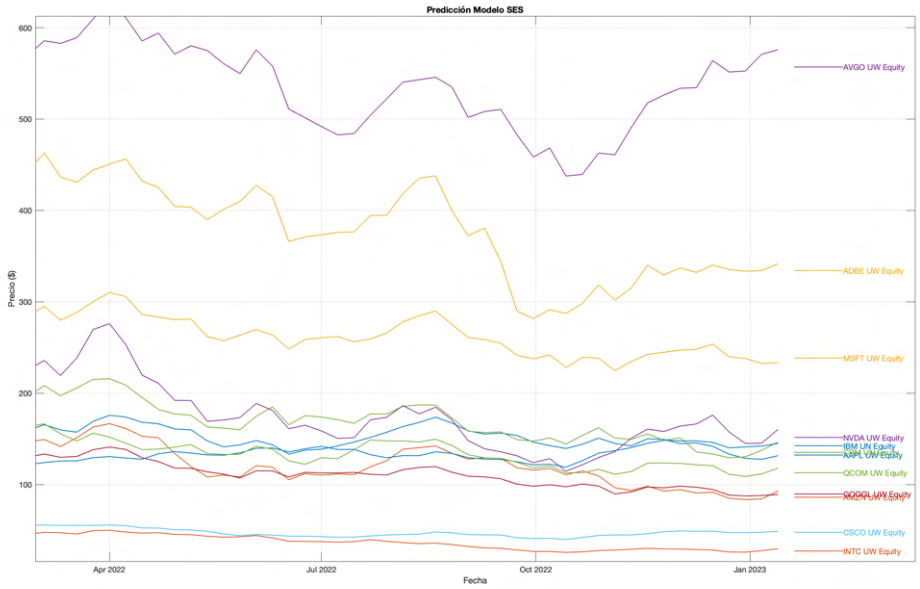


Ilustración 27. Ejemplo de predicciones a partir del modelo SES para varias compañías ampliado (Elaboración propia)

En la siguiente tabla es posible apreciar los resultados obtenidos de forma numérica.

Posición	Compañía	Predicción media
"1"	"AVGO UW Equity"	"556.973"
"2"	"ADBE UW Equity"	"334.2094"
"3"	"MSFT UW Equity"	"238.3455"
"4"	"NVDA UW Equity"	"151.5491"
"5"	"IBM UN Equity"	"142.4461"
"6"	"CRM UN Equity"	"135.1715"
"7"	"AAPL UW Equity"	"132.2621"
"8"	"QCOM UW Equity"	"112.9079"
"9"	"GOOGL UW Equity"	"89.6218"
"10"	"AMZN UW Equity"	"86.0377"
"11"	"CSCO UW Equity"	"47.6953"
"12"	"INTC UW Equity"	"27.2581"

Tabla 6. Resultados de ranking y predicciones a partir del modelo SES para varias compañías (Elaboración propia)

De acuerdo con esto, se observa que los valores obtenidos siguen tendencias similares a los valores reales de la serie ya que el método SES es un método conservador que proporciona una referencia sobre la futura rentabilidad de la empresa, sin correr grandes riesgos.

El método SES es relativamente simple y sencillo de implementar. Sin embargo, asume que la serie no sigue ninguna tendencia o estacionalidad por lo que, si los datos mostraran estos patrones, será necesario hacer uso de métodos de predicción más avanzados.

Cabe destacar que la herramienta comparará las rentabilidades de diferentes acciones en lugar de los precios ya que de esta manera es posible obtener una información más completa con el análisis.

6.2 MODELO M

A la hora de investigar la relación entre diferentes elementos aleatorios, los análisis de regresión permiten buscar un efecto de causalidad de una variable (o varias) aleatoria(s) sobre otra. Así, las técnicas de regresión han sido siempre relevantes en diferentes campos. La mayoría de los métodos de regresión asumen que los elementos aleatorios que aparecen pueden ser formalizados como variables aleatorias reales. Sin embargo, existe un número relevante de situaciones prácticas para las cuales los atributos no pueden tomar un valor real,

sino que toman un valor de intervalo. Así, el análisis estadístico de estos datos es de gran interés y con ello surge el **Modelo de Regresión Lineal Simple Basado en Aritmética de Intervalos** o, en inglés, *Interval Arithmetic-based Simple Linear Regression Model*, también conocido como Modelo M, (Sinova, Colubi, & González-Rodríguez, 2012).

El **Modelo M** es un modelo de regresión lineal simple que para predecir hace uso de aritmética de intervalos. En él, haciendo uso de una serie de operaciones aritméticas basadas en el centro y el radio del intervalo, se pueden calcular los coeficientes $\hat{a}$ y $[\hat{B}]$. Finalmente, a partir de ellos se obtendrá la predicción.

Un modelo lineal basado en aritmética de intervalos habitual se centraría simultáneamente en el centro y en el radio del intervalo de datos estudiado, asegurando que el modelo de intervalo está bien definido. Sin embargo, también sería posible analizar estos atributos por separado, o incluso tratarlos como variables aleatorias independientes, pero de esta manera no sería posible garantizar que un modelo lineal para el radio tuviera sentido, ya que no nos aseguraría que el radio fuera mayor o igual a cero (Domingo, 2021).

Con el objetivo de poder realizar una primera aproximación, asumiendo un modelo lineal, se va a tratar de encontrar un valor real para $\hat{a}$, así como unos valores de intervalo para $[\hat{B}]$, de forma que quede la expresión $\hat{a}[X] + [\hat{B}]$ lo más similar posible al intervalo de la muestra $[Y]$.

Para empezar, cabe tener en cuenta que el problema de mínimos cuadrados utilizado para encontrar $\hat{a}$ y $\widehat{[B]}$ es el que se muestra a continuación.

$$\frac{1}{n}\sum_{i=1}^{n}(d_\theta([Y], \hat{a}[X] + [\hat{B}]))$$

Ecuación 6.2. Problema de mínimos cuadrados en el modelo M

Llegados a este punto, hay que seguir los siguientes pasos para poder calcular el valor de $\hat{a}$ (Maté, 2022).

Paso 1. Calcular las estimaciones de la muestra de los momentos siguientes para las variables aleatorias reales del centro y del radio.

- Las medias de los centros y radios de $[X]$ e $[Y]$ son:

$$\overline{centx} = \frac{1}{n}\sum_{i=1}^{n} centx_i\,, \qquad \overline{radx} = \frac{1}{n}\sum_{i=1}^{n} radx_i$$

$$\overline{centy} = \frac{1}{n}\sum_{i=1}^{n} centy_i\,, \qquad \overline{rady} = \frac{1}{n}\sum_{i=1}^{n} rady_i$$

Ecuación 6.3 Medias de los centros y radios de [X] e [Y] en Modelo M

- Las varianzas de los centros y radios de $[X]$ son:

$$s^2_{centx} = \frac{1}{n}\sum_{i=1}^{n} (centx_i - \overline{centx})^2\,, \qquad s^2_{radx} = \frac{1}{n}\sum_{i=1}^{n} (radx_i - \overline{radx})^2,$$

Ecuación 6.4 Varianzas de los centros y radios de [X] en Modelo M

- La covarianza de los centros de $[X]$ e $[Y]$ y la covarianza de los radios de $[X]$ e $[Y]$:

$$s_{centx,centy} = \frac{1}{n}\sum_{i=1}^{n} (centx_i - \overline{centx})(centy_i - \overline{centy})$$

$$s_{radx,rady} = \frac{1}{n}\sum_{i=1}^{n} (radx_i - \overline{radx})(rady_i - \overline{rady})$$

Ecuación 6.5 Covarianzas de los centros y radios de [X] e [Y] en Modelo M

Paso 2. Calcular las estimaciones de la muestra de los momentos siguientes del conjunto aleatorio en intervalos.

- Las medias de los intervalos $[X]$ e $[Y]$:

$$[\overline{X}] = \left[\overline{centx} - \overline{radx}, \overline{centx} + \overline{radx}\right], \qquad [\overline{Y}] = \left[\overline{centy} - \overline{rady}, \overline{centy} + \overline{rady}\right]$$

Ecuación 6.6 Medias de [X] e [Y] en Modelo M

- La varianza θ-Fréchet del intervalo $[X]$:

$$\hat{\sigma}_x^2 = s_{centx}^2 + \theta s_{radx}^2$$

Ecuación 6.7 Varianza θ-Fréchet de [X] en Modelo M

- La covarianzaθ-Fréchet entre $[X]$ e $[Y]$ y entre $[-X]$ e $[Y]$:

$$\widehat{\sigma_{x,y}} = s_{centx,centy} + \theta s_{radx,rady}, \qquad \widehat{\sigma_{-x,y}} = -s_{centx,centy} + \theta s_{radx,rady}$$

Ecuación 6.8 Covarianza θ a de [X] e [Y] y de [-X] e [Y] en Modelo M

Paso 3. Calcular las estimaciones de la muestra $\widehat{a_0}$ y $\hat{a}$ para los datos disponibles:

- Si todas las x_i son valores reales, entonces $\widehat{a_0} = \infty$;
- si no,

$$\widehat{a_0} = min_{i:radx_i \neq 0} \frac{rady_i}{radx_i}$$

Ecuación 6.9 Estimación de $\widehat{a_0}$ en Modelo M

donde la estimación de $\hat{a}$ será según se muestra a continuación,

- si $\widehat{\sigma_{x,y}} < 0$ y $\widehat{\sigma_{-x,y}} < 0$, entonces,

$$\hat{a} = 0$$

Ecuación 6.10 Valor de $\hat{a}$ si $\widehat{\sigma_{x,y}} < 0$ y $\widehat{\sigma_{-x,y}} < 0$ en Modelo M

- si $\widehat{\sigma_{x,y}} \geq 0$ y $\widehat{\sigma_{x,y}} \geq \widehat{\sigma_{-x,y}}$, entonces,

$$\hat{a} = min\left\{\widehat{a_0}, \frac{\widehat{\sigma_{x,y}}}{\hat{\sigma}_x^2}\right\}$$

Ecuación 6.11 Valor de $\hat{a}$ si $\widehat{\sigma_{x,y}} \geq 0$ y $\widehat{\sigma_{x,y}} \geq \widehat{\sigma_{-x,y}}$ en Modelo M

- si $\widehat{\sigma_{-x,y}} \geq 0$ y $\widehat{\sigma_{-x,y}} \geq \widehat{\sigma_{x,y}}$, entonces,

$$\hat{a} = -min\left\{\widehat{a_0}, \frac{\widehat{\sigma_{-x,y}}}{\hat{\sigma}_x^2}\right\}$$

Ecuación 6.12 Valor de $\hat{a}$ si $\widehat{\sigma_{-x,y}} \geq 0$ y $\widehat{\sigma_{-x,y}} \geq \widehat{\sigma_{x,y}}$ en Modelo M

Finalmente, una vez que $\hat{a}$ ha sido calculado, es el momento de obtener el valor $[\hat{B}]$, dado por la siguiente expresión.

$$[\hat{B}] = [Y] - \hat{a}[X]$$

Ecuación 6.13 Valor de $[\hat{B}]$ en Modelo M

A partir de estos valores es posible predecir $[\hat{Y}_{Modelo\ M}]$ como se muestra a continuación:

$$[\hat{Y}] = \hat{a}[X] + [\hat{B}]$$

Ecuación 6.14 Predicción en Modelo M

Según esta idea y en relación con el resto de los métodos de predicción que se están estudiando, se ha hecho lo correspondiente haciendo uso de este modelo. Su desarrollo ha partido del procedimiento explicado anteriormente con aritmética de intervalos para obtener el valor de los coeficientes $\hat{a}$ y $[\hat{B}]$, y poder calcular la predicción a partir de ellos.

Las variables $[Y]$ y $[X]$ vienen dadas por $[Y(t)]$ e $[Y(t-1)]$, respectivamente, por lo que se tiene que el modelo es autorregresivo (M-AR) ya que la variable de interés dependerá de observaciones pasadas. Además, la función irá variando el conjunto de estimación en el que se ubican las estimaciones de $\hat{a}$ y $[\hat{B}]$, por lo que el modelo también será dinámico (M-ARD).

Así, la función podrá calcular el valor de la predicción $[\hat{Y}]$ para el número de periodos indicados por el usuario, expresados en este caso en semanas.

Para empezar, se ha estudiado una sola compañía para comprobar de forma sencilla el funcionamiento del código elaborado. La compañía es Starbucks y los datos estudiados abarcan desde Febrero de 2014 hasta Enero de 2023. Así, la predicción obtenida para las 4 semanas posteriores a la última semana de la que se tiene registro de datos son las siguientes.

Fecha	Mínimos	Máximos
20-Jan-2023	103.07	106.17
27-Jan-2023	101.98	105.06
03-Feb-2023	100.92	103.98
10-Feb-2023	99.873	102.92

Tabla 7. Resultados de predicciones a partir del modelo M para Starbucks (Elaboración propia)

El gráfico siguiente muestra los resultados obtenidos de la predicción (verde) junto con los datos reales que se tenían hasta la fecha (azul), a partir de los cuales se ha desarrollado la predicción.

Ilustración 28. Ejemplo de predicciones a partir del modelo M para Starbucks (Elaboración propia)

Así, es posible observar que se espera que el valor del precio de las acciones de Starbucks comience a decrecer a partir de Enero de 2023.

Por otro lado, una extensión de la función es capaz de recoger la información de varias compañías para realizar la predicción del precio de las acciones futuras de todas ellas de cara a ordenarlas en un ranking, en función de cuál se espera que tenga un mayor precio de la acción. Se escogen un total de 12 compañías tecnológicas que operan en el S&P100

para poder apreciar un ejemplo, durante las mismas fechas que el caso anterior. El resultado es el siguiente.

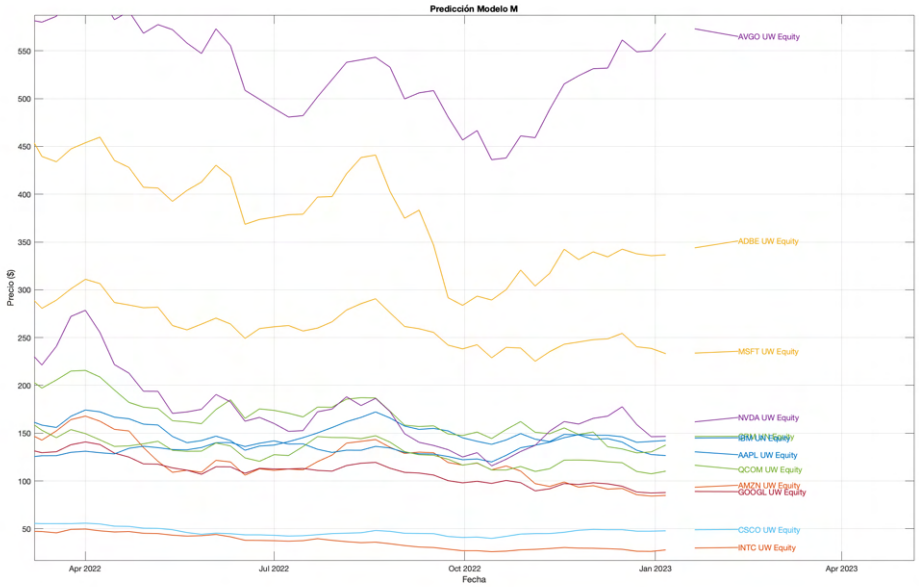


Ilustración 29. Ejemplo de predicciones a partir del modelo M para varias compañías (Elaboración propia)

A continuación, se muestra una tabla con la media de las predicciones de cada compañía y el ranking en el que según esa media se posicionaría cada una.

Posición	Compañía	Predicción media
"1"	"AVGO UW Equity"	"569.0295"
"2"	"ADBE UW Equity"	"347.5084"
"3"	"MSFT UW Equity"	"234.7183"
"4"	"NVDA UW Equity"	"164.124"
"5"	"CRM UN Equity"	"146.6234"
"6"	"IBM UN Equity"	"145.0599"
"7"	"AAPL UW Equity"	"128.9344"
"8"	"QCOM UW Equity"	"114.2768"
"9"	"AMZN UW Equity"	"94.4655"
"10"	"GOOGL UW Equity"	"88.8772"
"11"	"CSCO UW Equity"	"49.1046"
"12"	"INTC UW Equity"	"30.092"

Tabla 8. Resultados de ranking y predicciones a partir del modelo M para varias compañías (Elaboración propia)

Según se apreciaba en las gráficas previas, la compañía de la que se espera un mayor valor es Broadcom Inc. (*AVGO UW Equity*), la cual presenta una diferencia significativa con respecto a las compañías que le siguen.

6.3 MÉTODO DE GARCH (1,1)

Los **Modelos Autorregresivos de Heterocedasticidad Condicional** (ARCH, del inglés *Autoregressive Conditional Heteroskedasticity*) son utilizados para el modelado de series de tiempo de datos financieros que presentan agrupaciones de volatilidad variables en el tiempo, lo que significa que tienen periodos de oscilaciones alternados con otros periodos de relativa estabilidad. Su principal objetivo es modelar la volatilidad del rendimiento de un activo para mejorar la estimación de los parámetros, así como la exactitud de la predicción (Domingo, 2021).

Estos modelos tratan la forma en la que la volatilidad evoluciona con el paso del tiempo. En el caso de los modelos GARCH, se utiliza una función generalizada para describir la evolución de la volatilidad. El **modelo GARCH** (del inglés, *Generalized ARCH*) representa una extensión del modelo ARCH y puede ofrecer buenas estimaciones, por lo que se utiliza como un método de predicción para hallar la volatilidad en el futuro (López Mínguez, 2007).

El modelo GARCH asume que la varianza condicional de una serie de tiempo es una función de términos de error y de varianzas condicionales pasados. Por ello, defiende la idea de que la volatilidad de la serie varía con el tiempo y puede ser influenciada por observaciones pasadas. Así, la ecuación más general para el método GARCH (p,q) es la siguiente.

$$a_t = \sigma_t \varepsilon_t$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \, a_{t-i}^2 + \sum_{j=1}^{q} \beta_j \, \sigma_{t-j}^2$$

$$donde\ \varepsilon_{t\ iid}\sim N(0,1), \alpha_0 > 0, \alpha_i \geq 0, \beta_j \geq 0, \sum_{i=1}^{\max(m,s)} (\alpha_i + \beta_i) < 1$$

Ecuación 6.15. Modelo GARCH(p,q)

Esta última restricción implica que la varianza condicional de a_t es finita. Se van a considerar α_i como parámetros ARCH, capturando el impacto de los errores cuadrados pasados en la varianza condicional, y β_j como parámetros GARCH, capturando el impacto de las varianzas condicionales pasadas en la varianza condicional actual. Además, σ_t^2 representa la varianza condicional en el momento t y a_t^2 será el residuo estandarizado en t, calculado como la diferencia entre el valor observado y el valor de predicción, dividida entre la desviación estándar condicional. Finalmente, p y q indican el número de periodos de retraso a considerar para los términos ARCH y GARCH, respectivamente.

En este capítulo se va a trabajar con el modelo GARCH en Matlab para poder obtener la predicción de rentabilidad para cada empresa, conociendo los rangos de valores en los que la compañía oscilará en periodos futuros. Para empezar, se ha estudiado el ejemplo de Starbucks con datos desde Febrero de 2014 y Enero de 2023. El modelo trabaja con los datos prediciendo la volatilidad que se espera para cada momento de tiempo y así realizar una estimación del valor de la acción, en este caso, durante el periodo de test y en los 4 periodos posteriores al final de este periodo. Se ha trabajado con un modelo GARCH (1,1) y el resultado ha sido el siguiente.

Ilustración 30. Ejemplo de predicciones a partir del modelo GARCH (1,1) para Starbucks (Elaboración propia)

En la imagen se aprecia el conjunto de test real (azul) junto con los valores estimados para ese periodo (rojo), que son muy similares a los datos de test reales. Además, la función realiza la predicción del precio de la acción una vez que el periodo del conjunto de datos estudiado ha terminado (verde), para 4 semanas, lo cual se puede apreciar numéricamente a continuación.

Fecha	Mínimos	Máximos
20-Jan-2023	103.01	108.47
27-Jan-2023	102.9	108.57
03-Feb-2023	102.8	108.68
10-Feb-2023	102.69	108.79

Tabla 9. Resultados de predicciones a partir del modelo GARCH (1,1) para Starbucks (Elaboración propia)

Por otro lado, se ha continuado desarrollando el modelo para que sea capaz de trabajar con la información de varias compañías de manera simultánea y poder elaborar el ranking de las mismas en función de los precios de sus valores, ordenándolos de mayor a menor, durante 4 semanas a partir de la última fecha de estudio de los datos.

Siguiendo el ejemplo anterior y haciendo uso de una serie de compañías que operan en la industria tecnológica, para esas mismas fechas, el resultado obtenido de forma gráfica es el siguiente.

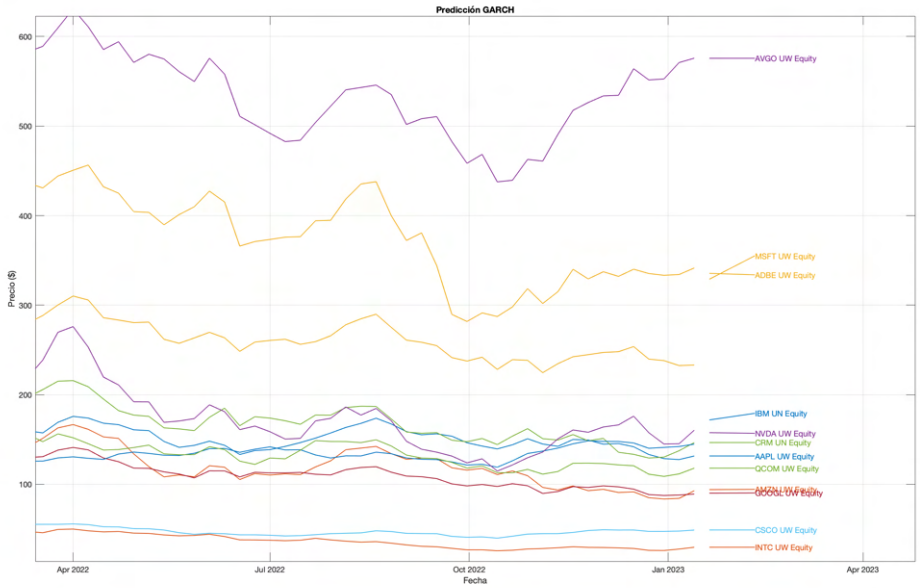


Ilustración 31. Ejemplo de predicciones a partir del modelo GARCH (1,1) para varias compañías (Elaboración propia)

A continuación, se presenta la tabla de resultados obtenida, la cual es similar a la del ejemplo anterior, aunque presenta algunos cambios.

Posición	Compañía	Predicción media
"1"	"AVGO UW Equity"	"575.6922"
"2"	"MSFT UW Equity"	"342.047"
"3"	"ADBE UW Equity"	"334.6904"
"4"	"IBM UN Equity"	"175.5066"
"5"	"NVDA UW Equity"	"157.2296"
"6"	"CRM UN Equity"	"146.5769"
"7"	"AAPL UW Equity"	"131.5752"
"8"	"QCOM UW Equity"	"117.9549"
"9"	"AMZN UW Equity"	"94.1362"
"10"	"GOOGL UW Equity"	"90.1344"
"11"	"CSCO UW Equity"	"48.9335"
"12"	"INTC UW Equity"	"29.6563"

Tabla 10. Resultados de ranking y predicciones a partir del modelo GARCH (1,1) para varias compañías (Elaboración propia)

Llama la atención cómo algunas compañías presentan un salto entre último valor recopilado de sus acciones y la estimación esperada con el modelo GARCH (1,1), como es el caso de Microsoft (*MSFT UW Equity*). Aunque este salto es positivo e implicaría un mayor valor de la acción, es importante prestar atención a la tendencia que ha seguido el precio de la acción durante los últimos periodos antes de tomar una decisión de inversión.

6.4 MÉTODO K-NN

El **método K-NN** (del inglés, *K-Nearest Neighbors*) representa un algoritmo simple y efectivo utilizado en clasificación y regresión. Es en el campo de la regresión cuando se utiliza para predecir y, por tanto, en el que se va a trabajar en este capítulo. Mediante este método, es posible estimar el valor de una variable objetivo a partir de observaciones conocidas, basándose en los k puntos más cercanos. En la siguiente imagen se puede ver un ejemplo del funcionamiento de este modelo a la hora de escoger los vecinos más cercanos.

Ilustración 32. Elección de los k vecinos más cercanos con K-NN (IBM, 2023)

Para poder trabajar con este método es necesario escoger una medida de distancia que permita medir la similitud entre los vectores de características. En este caso se toman distancias temporales y la distancia a utilizar será la distancia Euclídea (López Mínguez, 2007). Así, la distancia Euclídea entre dos series de datos d en los instantes r y s, será:

$$D(X_r^d, X_s^d) = \sqrt{\frac{\sum_{i=1}^{d} \alpha^{d-t+1}(x_{r-i+1} - x_{s-i+1})^2}{\sum_{i=1}^{d} \alpha^{d-i+1}}}$$

Ecuación 6.16. Distancia Euclídea

Con esta fórmula se dará más peso a los datos más cercanos en el tiempo, en función del valor de α. A continuación, es preciso escoger el valor de k, que representa el número de vecinos a considerar a la hora de hacer las predicciones. Es necesario elegir este valor de manera cautelosa ya que tendrá gran influencia en la precisión de la predicción. Así, un valor de k pequeño puede implicar una mayor varianza y predicciones más sensibles mientras que un valor de k más grande puede significar un mayor sesgo.

El siguiente paso será encontrar a los k vecinos más cercanos. Para ello, dado un momento de tiempo para el que se desea conocer la predicción, habrá que calcular la distancia entre este punto y el resto de los puntos de datos conocidos, seleccionando aquellos que tengan menores distancias a él. Así, hay que tener en cuenta que los k vecinos más cercanos serán una serie de d datos históricos, X_t^d, representando observaciones consecutivas, donde d debe ser un número a optimizar.

$$X_t^d = (x_t, x_{t-1}, \dots, x_{r-d+1})$$

Después de haber obtenido los k vecinos más cercanos, representados por $X_{T_K}^d$ con $k = 1,2, \dots, K$, habrá que estimar la predicción, para lo cual se calculará la media de los valores pertenecientes a los k vecinos más cercanos.

La función va a tomar de entrada un parámetro $kRange$ que representa el vector de k valores que se van a testear para obtener el valor óptimo. Para ello, va a iterar los intervalos de test y los diferentes valores de k dentro del rango para encontrar el mejor valor de predicción a partir del RMSE (*Root Mean Squared Error*). Al probar distintos valores de k, es posible identificar aquel que proporciona una mejor sensibilidad ante cambios abruptos en los datos.

A continuación, se muestra un ejemplo del funcionamiento del método. Para empezar, se ha estudiado como método de predicción para los valores de las acciones de una sola empresa. Tomando el caso de Starbucks, con un periodo de training desde el 12 de Febrero de 2020 hasta el 13 de Febrero de 2021 y un periodo de test desde el 20 de Febrero de 2021 hasta el 13 de Enero del 2023. Además, se ha considerado un rango de *k* igual a 15 para que la función escoja el número de vecinos óptimo entre 1 y 15, y un número de periodos a predecir (semanales) igual a 4. El resultado obtenido es el siguiente.

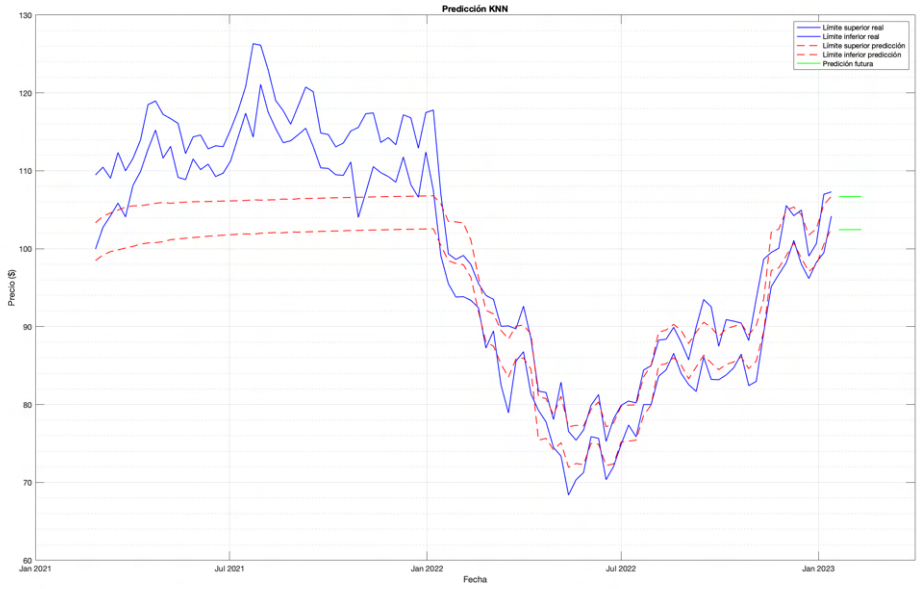


Ilustración 33. Ejemplo de predicciones a partir del método KNN para Starbucks (Elaboración propia)

En la imagen se puede apreciar como los primeros valores de la predicción se alejan de la realidad, pero a partir de cierto momento (Junio de 2022), la predicción empieza a seguir una tendencia muy parecida a la de los valores reales. Finalmente, los valores obtenidos para las 4 semanas posteriores a los datos estudiados se muestran en verde, y se puede apreciar con más claridad en la tabla de debajo.

Fecha	Mínimos	Máximos
20-Jan-2023	102.44	106.68
27-Jan-2023	102.45	106.69
03-Feb-2023	102.46	106.69
10-Feb-2023	102.47	106.7

Tabla 11. Resultados de predicciones a partir del método KNN para Starbucks (Elaboración propia)

El siguiente paso ha sido continuar con el desarrollo de la función para que sea capaz de comparar el comportamiento de diferentes compañías en el S&P100, realizar la predicción del precio de sus valores durante una serie de periodos semanales futuros y con ello poder elaborar un ranking ordenando las compañías de mayor a menor precio de sus acciones.

En el siguiente ejemplo se han evaluado una serie de compañías tecnológicas que operan en el índice. El resultado obtenido se muestra en la siguiente imagen.

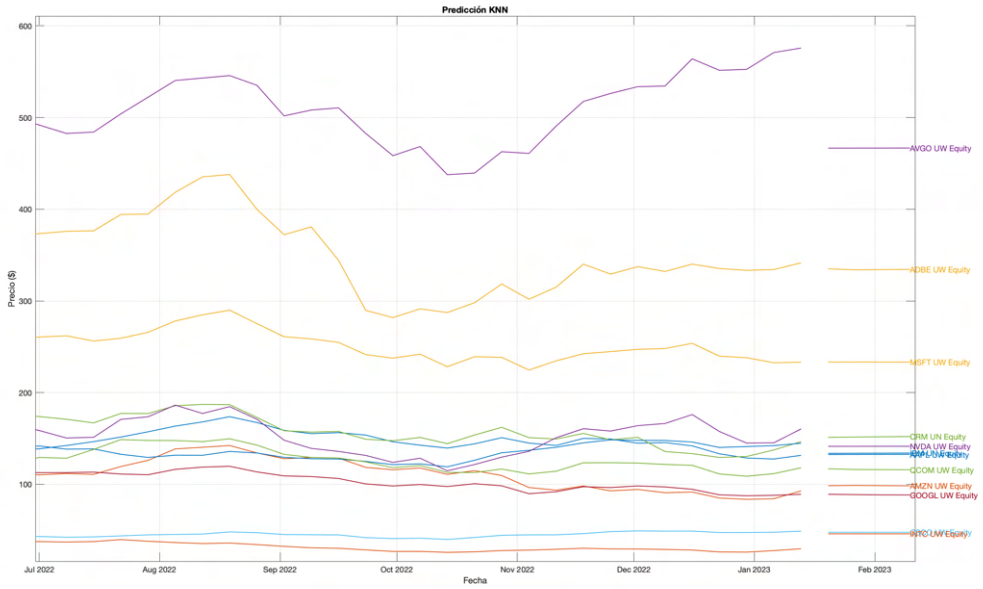


Ilustración 34. Ejemplo de predicciones a partir del método KNN para varias compañías (Elaboración propia)

De acuerdo con la gráfica, es posible apreciar cómo existen algunas acciones, como las de Broadcom Inc. (*AVGO UW Equity*), cuya predicción va a dar un gran salto con respecto a los últimos datos que se tienen hasta la fecha mientras que hay otras que mantendrán un precio similar al último registrado, como Microsoft (*MSFT UW Equity*), por lo que es posible que según este método, haya menos riesgo a la hora de invertir en acciones que sigan un comportamiento más estable aunque otras puedan ofrecer mayor rentabilidad.

Para una mayor aclaración, el ranking obtenido como resultado es el que se muestra debajo.

Posición	Compañía	Predicción media
"1"	"AVGO UW Equity"	"466.696"
"2"	"ADBE UW Equity"	"334.3504"
"3"	"MSFT UW Equity"	"233.2904"
"4"	"CRM UN Equity"	"151.6237"
"5"	"NVDA UW Equity"	"141.3938"
"6"	"IBM UN Equity"	"133.8655"
"7"	"AAPL UW Equity"	"132.4972"
"8"	"QCOM UW Equity"	"116.184"
"9"	"AMZN UW Equity"	"98.4865"
"10"	"GOOGL UW Equity"	"88.6237"
"11"	"CSCO UW Equity"	"47.6506"
"12"	"INTC UW Equity"	"45.9092"

Tabla 12. Resultados de ranking y predicciones a partir del método KNN para varias compañías (Elaboración propia)

6.5 COMPOSICIÓN DE MÉTODOS

Para llevar a cabo este último método, se va a hacer uso de los resultados obtenidos en los métodos anteriores. Para ello, se indicarán las empresas a evaluar y a partir de ellas cada uno de los métodos elaborará su propio ranking haciendo uso de los procedimientos explicados anteriormente. Finalmente, este nuevo método se encargará de elaborar un nuevo ranking una vez realizada la media de las valoraciones futuras de acciones obtenidas de cada uno de los rankings.

En la siguiente imagen se puede apreciar una comparación de los resultados obtenidos en cada uno de los métodos.

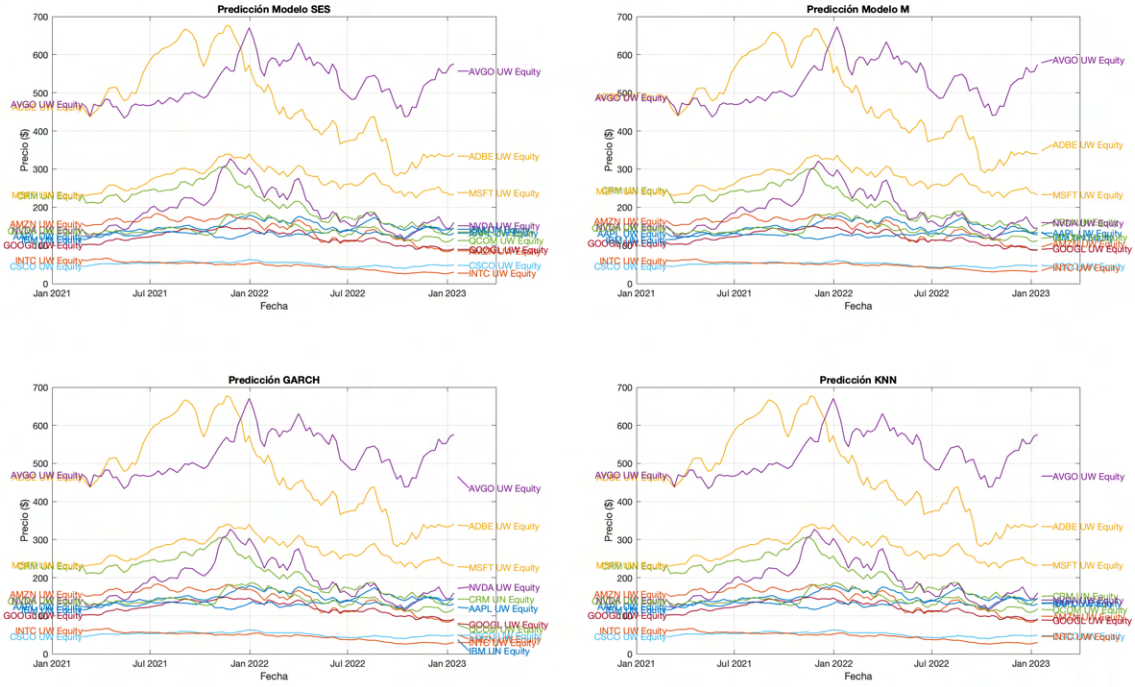


Ilustración 35. Ejemplo de comparación de los métodos estudiados (Elaboración propia)

A continuación, se observa el resultado obtenido según cada método para cada compañía. De acuerdo con la tabla, cabe concluir que es interesante comparar los resultados de todos los métodos antes de tomar una decisión ya que hay algunos que presentan mayor desviación frente al resto, especialmente en el caso del método GARCH(1,1).

Compañía	Predicción SES	Predicción Modelo M	Predicción GARCH(1,1)	Predicción K-NN
"AAPL UW Equity"	"132.2621"	"132.2921"	"121.259"	"132.4972"
"AMZN UW Equity"	"86.0377"	"100.7161"	"44.6282"	"98.4865"
"ADBE UW Equity"	"334.2094"	"356.0752"	"337.076"	"334.3504"
"AVGO UW Equity"	"556.973"	"582.8029"	"450.4896"	"466.696"
"CRM UN Equity"	"135.1715"	"156.4828"	"143.6985"	"151.6237"
"CSCO UW Equity"	"47.6953"	"47.0323"	"48.8279"	"47.6506"
"GOOGL UW Equity"	"89.6218"	"90.8956"	"78.2782"	"88.6237"
"IBM UN Equity"	"142.4461"	"127.6739"	"22.6011"	"133.8655"
"INTC UW Equity"	"27.2581"	"38.5922"	"29.6378"	"45.9092"
"MSFT UW Equity"	"238.3455"	"233.1701"	"228.272"	"233.2904"
"NVDA UW Equity"	"151.5491"	"159.408"	"173.6281"	"141.3938"
"QCOM UW Equity"	"112.9079"	"120.1095"	"71.0753"	"116.184"

Tabla 13. Resultados de comparación resultados de los métodos estudiados (Elaboración propia)

Finalmente, la función calculará la media entre estos resultados para poder realizar su propio ranking a partir del valor obtenido. En la siguiente tabla se puede apreciar el resultado final de esta función.

Posición	Compañía	Predicción media
"1"	"AVGO UW Equity"	"514.2404"
"2"	"ADBE UW Equity"	"340.4278"
"3"	"MSFT UW Equity"	"233.2695"
"4"	"NVDA UW Equity"	"156.4948"
"5"	"CRM UN Equity"	"146.7441"
"6"	"AAPL UW Equity"	"129.5776"
"7"	"IBM UN Equity"	"106.6466"
"8"	"QCOM UW Equity"	"105.0692"
"9"	"GOOGL UW Equity"	"86.8548"
"10"	"AMZN UW Equity"	"82.4671"
"11"	"CSCO UW Equity"	"47.8015"
"12"	"INTC UW Equity"	"35.3493"

Tabla 14. Resultados de ranking y predicciones a partir de composición de métodos (Elaboración propia)

Capítulo 7. ANÁLISIS DE LA HERRAMIENTA

Con el objetivo de recopilar todo lo explicado hasta ahora en relación al análisis de precios de las acciones y su aplicación a la gestión de carteras optimizadas se ha diseñado una herramienta orientada funcionar como un sistema de análisis para los inversores que les permita conocer con mayor profundidad tanto datos pasados del precio de acciones como datos estimados para el futuro.

7.1 INTRODUCCIÓN A LA HERRAMIENTA

La herramienta desarrollada integra todas las funcionalidades estudiadas hasta ahora y permite realizar dos funciones principales. La primera de ellas, analizar datos pasados de una empresa, o comparar los de dos compañías entre sí, con el objetivo de que el inversor obtenga una idea de la forma en la que se ha comportado el precio de las acciones de la compañía a lo largo de los años. La segunda, realizará una estimación de precios futuros de una acción o de las rentabilidades en el caso de escoger varias acciones en base a datos históricos. Para realizar el estudio se ha hecho uso de datos semanales pertenecientes al índice bursátil S&P100, entre 2010 y 2023.

Cabe destacar que se ha trabajado con App Designer, un entorno de desarrollo interactivo de Matlab que permite diseñar una aplicación y programar su comportamiento.

Una vez desarrollada la herramienta, el primer paso para poder hacer uso de ella es su inicialización desde Matlab.

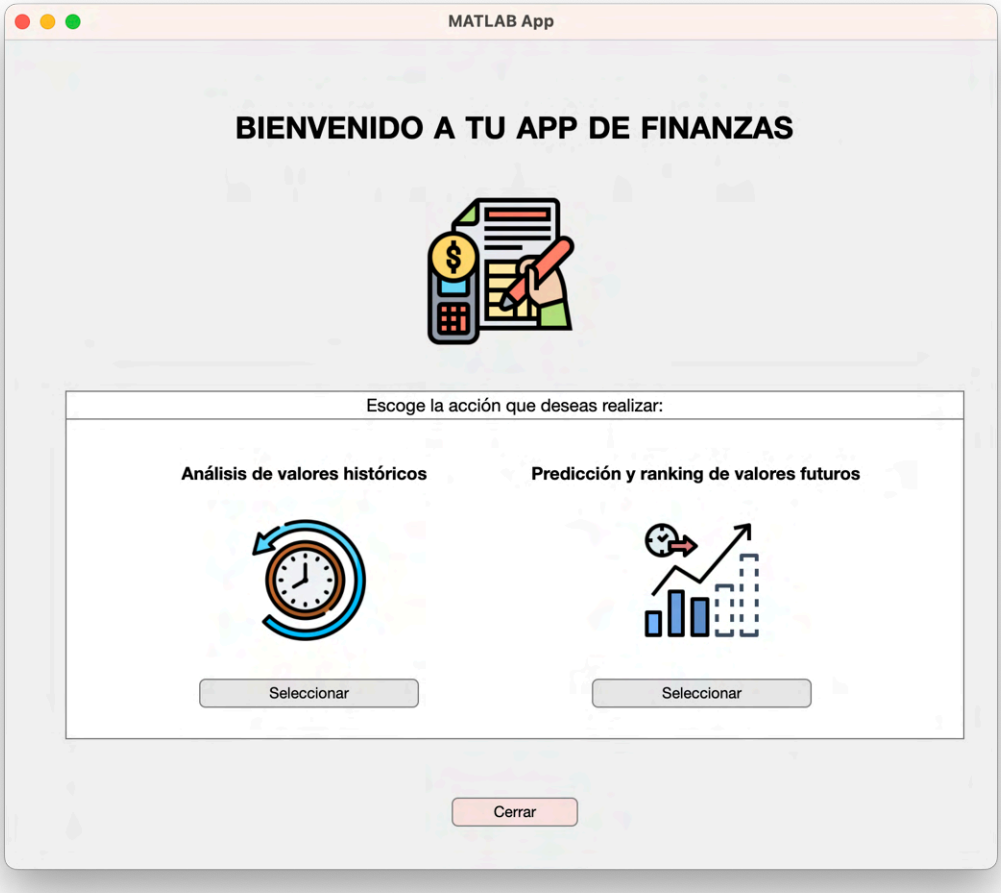


Ilustración 36. Ventana de inicio de la herramienta (Elaboración propia)

En la pantalla de inicio se aprecian las dos opciones que tiene el usuario para trabajar con la herramienta y que se exponen con mayor profundidad a continuación.

7.2 FUNCIONALIDAD DE LA HERRAMIENTA

7.2.1 ANÁLISIS DE VALORES HISTÓRICOS

Al seleccionar esta opción, el usuario va a ser dirigido a una nueva ventana en la que tendrá que elegir si el número de empresas a analizar, que va a ser 1 o 2, así como el tipo de gráfico que desea visualizar.

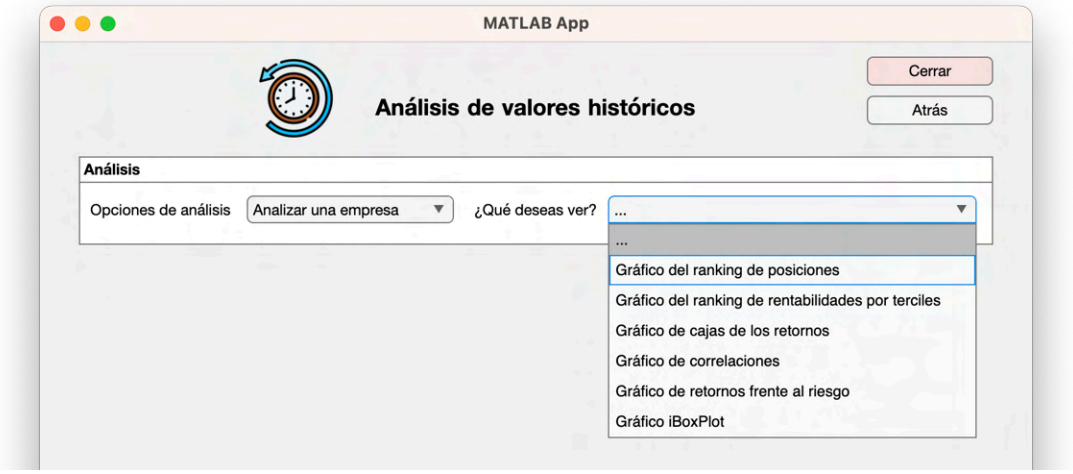


Ilustración 37. Ventana análisis de valores históricos (Elaboración propia)

7.2.1.1 Analizar una empresa

Una vez seleccionada esta opción, el usuario tendrá la oportunidad de escoger el tipo de análisis que quiere realizar sobre los datos pasados de una empresa determinada. Para ver el funcionamiento de la aplicación se va a estudiar el caso de Amazon (*AMZN UW Equity*), con datos entre el 1 de Enero de 2019 y el 1 de Enero de 2023. Una vez seleccionada la opción deseada, el usuario introduce estos datos de entrada escogiendo la empresa desde una lista desplegable y las fechas desde un calendario mensual. Finalmente, la aplicación mostrará una gráfica con el resultado obtenido.

Ilustración 38. Analizar una empresa (Elaboración propia)

Gráfico del ranking de posiciones

En la gráfica se puede observar la posición de la empresa en el ranking de todas las empresas en función de su rentabilidad en cada semana, siendo la primera aquella con mayor rentabilidad. La función muestra la evolución de la posición de Amazon en el ranking (azul) y además, se puede apreciar la posición media de la compañía (verde).

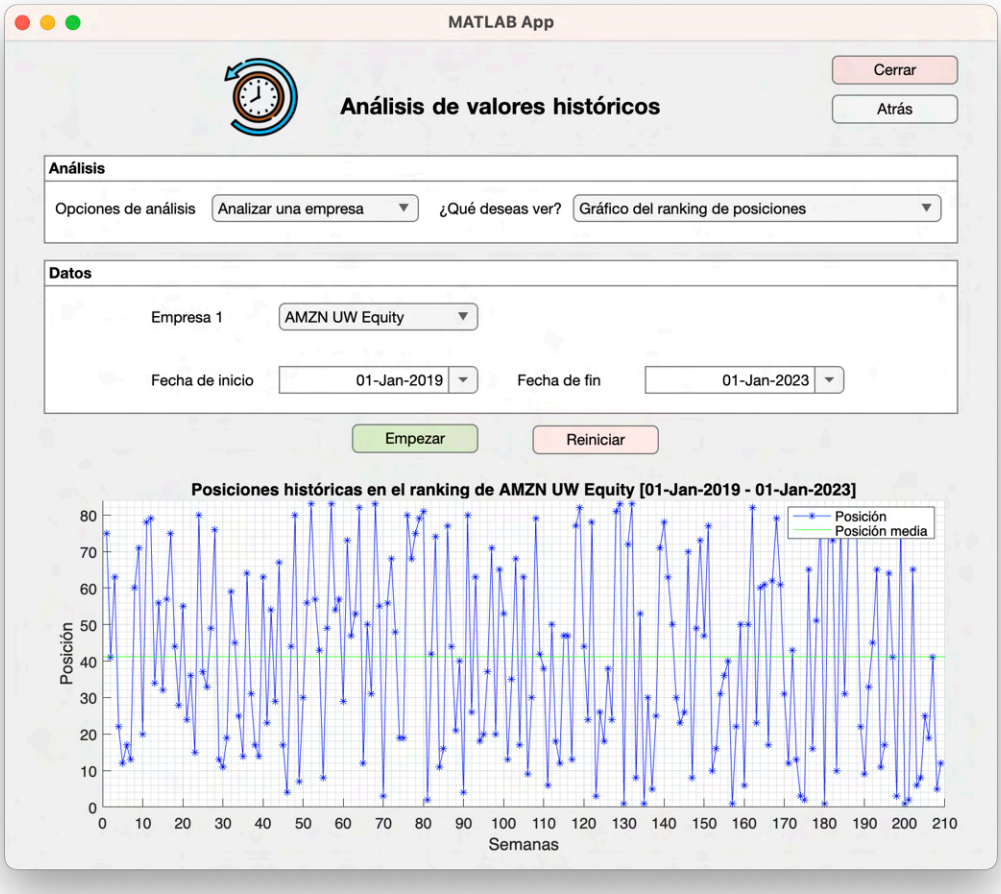


Ilustración 39. Analizar una empresa - Gráfico del ranking de posiciones (Elaboración propia)

En el caso de Amazon y para el periodo estudiado, se observa que la compañía varía constantemente y de forma rápida en cuanto a su posición en el ranking de rentabilidades. Además, presenta una posición media de casi 42, por lo que en general, se encontraría en una posición central en el ranking de rentabilidades. Para pasar al siguiente análisis, el usuario debe pulsar el botón 'Reiniciar'.

Gráfico del ranking de rentabilidades por terciles

En este apartado, la gráfica va a analizar las rentabilidades semanales de cada empresa para mostrarlas en el ranking, representando también el rango de rentabilidades del resto de compañías en esa misma semana y dividiendo ese rango en 3 secciones en función del 30% mejores (verde), el 30% peores (rojo) y el 40% intermedias (amarillo). El resultado obtenido es el siguiente.

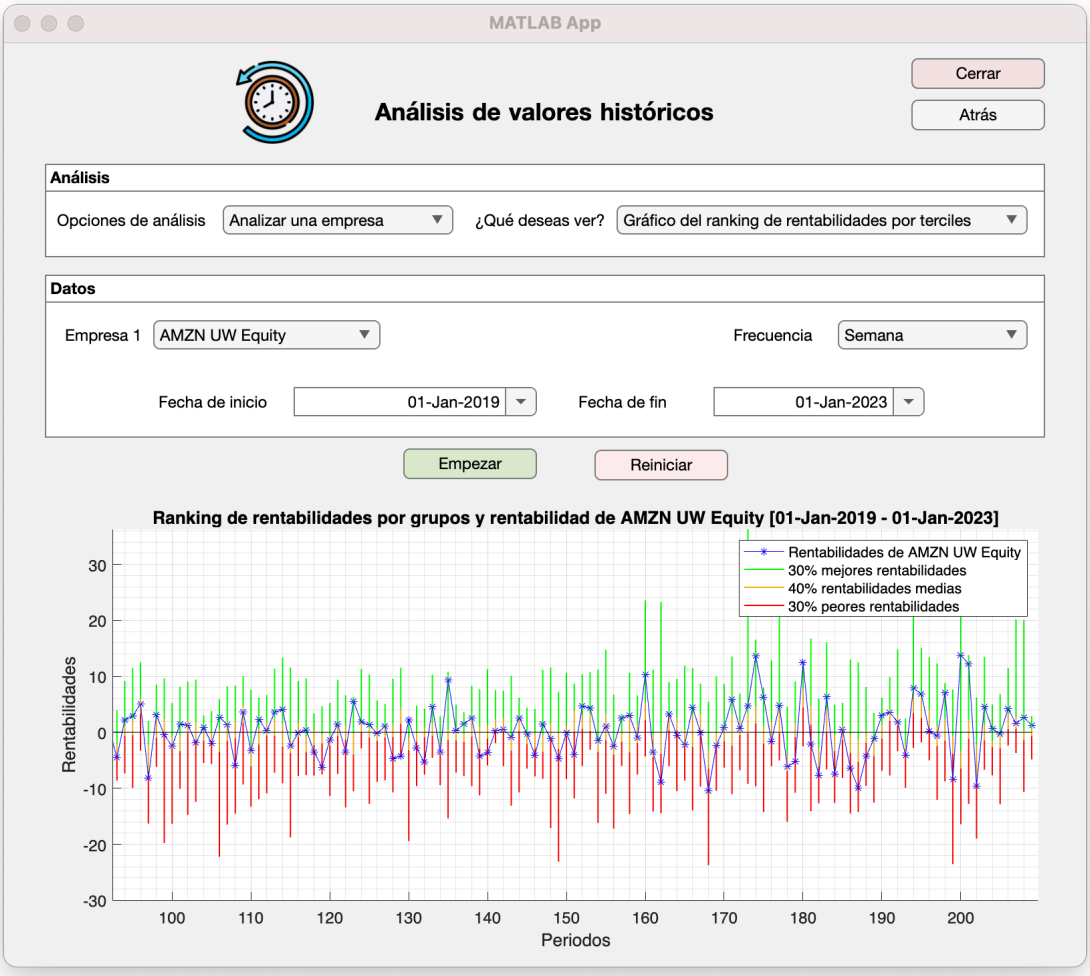


Ilustración 40. Analizar una empresa - Gráfico del ranking de rentabilidades por terciles (Elaboración propia)

Como se indicaba anteriormente, Amazon se mantiene en posiciones generalmente intermedias en el ranking de rentabilidades semanales del S&P100, predominando su presencia en la sección amarilla (rentabilidades medias). Se observa que en las últimas semanas se ha mantenido en rentabilidades positivas y algo más estables que hasta entonces.

Gráfico de cajas de los retornos

El siguiente paso es evaluar los retornos de Amazon para representar un gráfico de cajas. Para ello se evalúan los retornos de la empresa a lo largo del periodo seleccionado por el usuario ver cómo están distribuidos los datos y descubrir si existe algún *outlier*. El resultado obtenido es el siguiente.

Ilustración 41. Analizar una empresa - Gráfico de cajas de los retornos (Elaboración propia)

De esta gráfica destaca que las rentabilidades entre las que se encuentra la empresa en la mayoría del periodo estudiado se encuentran entre 2% y algo menos de -2%, siendo su mediana ligeramente inferior a 0%. A pesar de ello, se encuentran *outliers* que implican rentabilidades superiores a 10% e inferiores a -10% que, aunque sean casos excepcionales, se ha de tener en cuenta que es posible que ocurran.

Gráfico de correlaciones

A continuación, se estudian las correlaciones de Amazon con el resto de las compañías en el periodo indicado previamente para ver la relación de su comportamiento con el del resto de empresas.

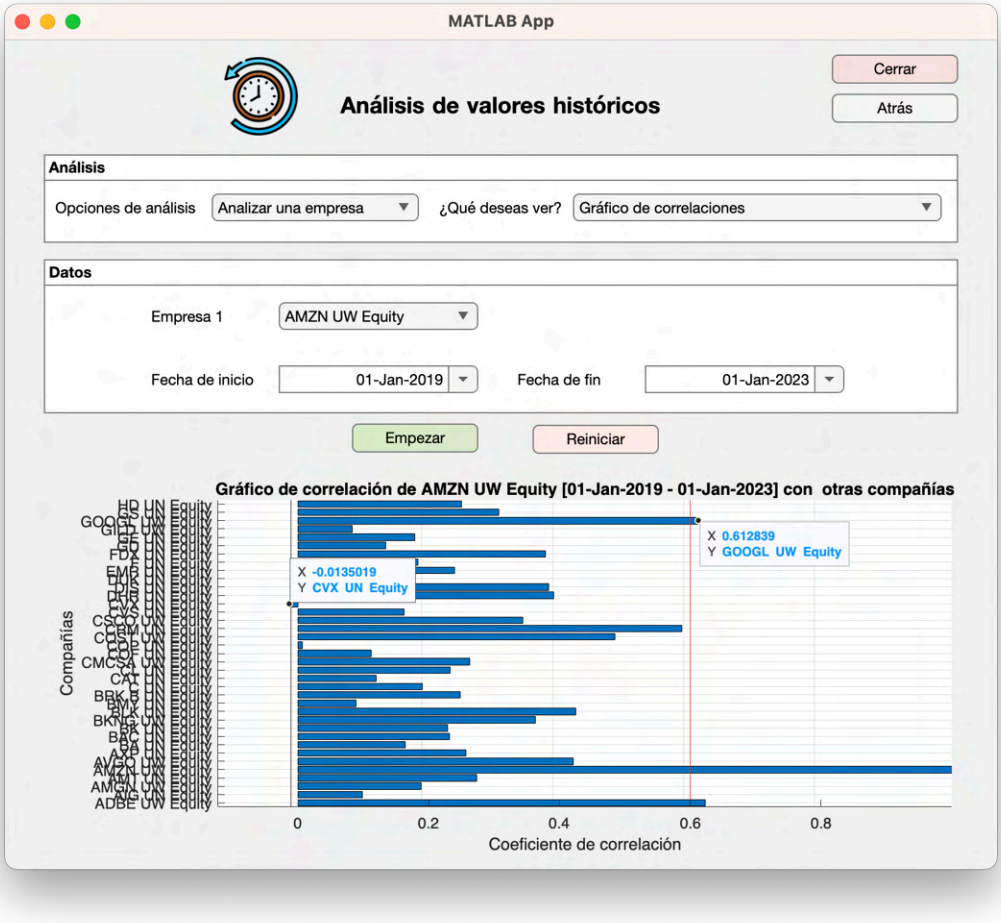


Ilustración 42. Analizar una empresa - Gráfico de correlaciones (Elaboración propia)

Para empezar, la correlación consigo misma será de 1, como era de esperar. A continuación, se ha marcado un límite de 0,6 a partir del cual la correlación se considera significativa y esto ocurre, por ejemplo, en el caso de Google (*GOOGL UW Equity*), lo que indica que cuando Amazon atraviesa un buen momento, es probable que a Google le ocurra lo mismo. Esto tiene sentido ya que ambas compañías operan en el sector tecnológico y por ello, hay factores que afectan de igual forma al rendimiento de ambas como pueden ser tendencias en el mercado.

Por otro lado, existen compañías como la petrolera Chevron Corporation (*CVX UN Equity*) con las que Amazon tiene correlación negativa, lo cual implica que cuando una esté atravesando una buena época, será probable que la otra atraviese un mal periodo y viceversa.

Gráfico de retornos frente al riesgo

En este caso se ha calcula la beta, representante del riesgo de una compañía en el mercado, en relación a su retorno. Así, se obtiene un gráfico que permite comparar a la compañía de análisis con el resto de las compañías del mercado y que resulta una herramienta de gran utilidad a la hora de tomar una decisión de inversión.

Cabe recordar que la beta representa el riesgo de una acción ya que mide la volatilidad de la acción en relación con el resto del mercado. Si es igual a 1, indica que el precio de la acción se mueve igual que el del mercado, si es mayor que 1, la acción tiene mayor volatilidad que el mercado y si beta es menor que 1, la volatilidad de la acción será menor que la del mercado.

En este caso se ha ampliado el periodo de estudio hasta el año 2015 para tener un resultado con más información. En el caso de Amazon, es posible concluir que presenta una rentabilidad media cercana al 5%, por encima de la de muchas otras empresas analizadas. De acuerdo con ella, su beta es algo superior al 1,2, lo que indica que hay mayor riesgo de invertir en la compañía ya que su volatilidad será mayor que la del mercado.

Ilustración 43. Analizar una empresa - Gráfico de retornos frente al riesgo (Elaboración propia)

Gráfico iBoxPlot

Por último, se lleva a cabo un análisis de los precios históricos de la acción para representar en forma de iBoxPlot. La función iBoxPlot (Maté, 2021) se utiliza para representar de forma visual la distribución de un conjunto de datos, en este caso de intervalo. Así, esta es capaz de generar un gráfico a partir de los datos históricos de intervalo de una compañía en concreto y permite ver si existen *outliers* en el caso de que se den cambios abruptos al principio o al final de la pendiente que une los valores máximos y/o la que hace lo mismo con los valores mínimos.

Así, el gráfico mostraría lo siguiente:

- Intervalo mínimo: engloba desde el valor mínimo de los mínimos del intervalo hasta el valor mínimo de los máximos del intervalo.
- Intervalo máximo: incluye desde el valor máximo de los mínimos del intervalo hasta el valor máximo de los máximos del intervalo.
- Intervalo de la mediana: rango desde la mediana de los límites inferiores hasta la mediana de los límites superiores del intervalo.
- Intervalo del primer cuartil (Q1): engloba desde el primer cuartil de los límites inferiores hasta el primer cuartil de los límites superiores.
- Intervalo del tercer cuartil (Q3): rango desde el tercer cuartil de los límites inferiores hasta el tercer cuartil de los límites superiores.

El resultado obtenido para el caso de estudio de Amazon es el siguiente.

Ilustración 44. Analizar una empresa - Gráfico iBoxPlot (Elaboración propia)

Debido a que la pendiente del iBoxPlot no presenta cambios muy bruscos, podría concluirse que no existen cambios muy llamativos en los valores de la acción en el periodo estudiado. En caso de ocurrir, estos predominarían en los valores más bajos de la acción, donde parece que hay un mayor salto entre el periodo anterior y posterior a Q1.

En conclusión, Amazon sería una compañía en la que, a partir de los datos recopilados entre 2019 y 2023, se observaría que no tiene mucho riesgo de inversión, teniendo en consecuencia unas rentabilidades no muy elevadas y relativamente estables y manteniéndose en posiciones intermedias en comparación con el resto de las compañías en

función de sus rentabilidades semanales. Por tanto, podría ser una apuesta de inversión de poco riesgo y segura, pero sin gran potencial de obtención de grandes ganancias.

7.2.1.2 Analizar dos empresas

De manera análoga, es posible realizar el procedimiento anterior, pero estudiando dos empresas simultáneamente para poder compararlas entre ellas. En este caso, se van a analizar Amazon y Google en un periodo más corto para poder ver el resultado con mayor claridad. Este periodo comprende desde Enero de 2021 hasta Enero de 2023.

Gráfico del ranking de posiciones

Ilustración 45. Analizar dos empresas - Gráfico del ranking de posiciones (Elaboración propia)

Se observa que ambas compañías tienen cambios bruscos en cuanto a su posición en el ranking de rentabilidades y que, aunque ambas tienen una posición media parecida, Amazon, situada por delante del puesto 50, se encuentra en un lugar ligeramente mejor al de Google, localizada detrás del puesto 50.

Gráfico del ranking de rentabilidades por terciles

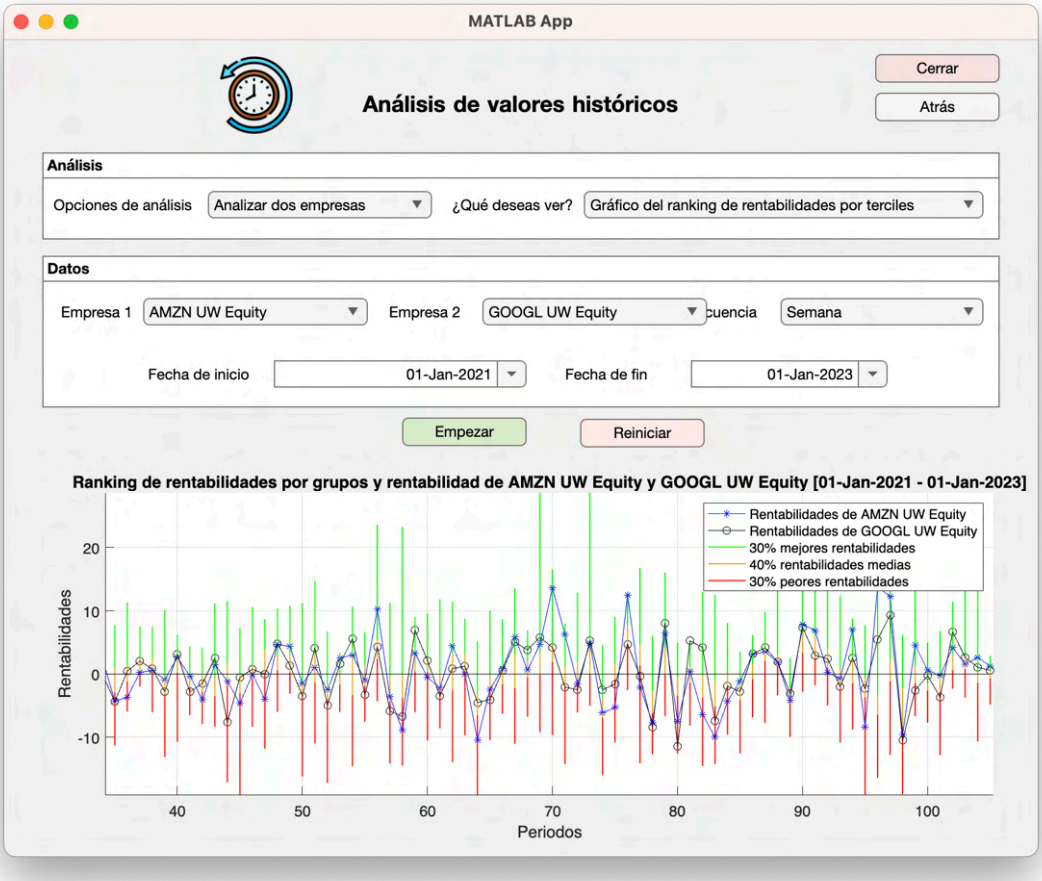


Ilustración 46. Analizar dos empresas - Gráfico del ranking de rentabilidades por terciles (Elaboración propia)

Según este gráfico es posible ver cómo las dos compañías tienen tendencias similares en cuanto a sus rentabilidades semanales y en comparación con el resto. Como se ha mencionado antes, al ser dos compañías que operan en la misma industria existen factores que afectan a ambas por igual, lo cual se puede ver reflejado en los resultados que alcanzan.

Gráfico de cajas de los retornos

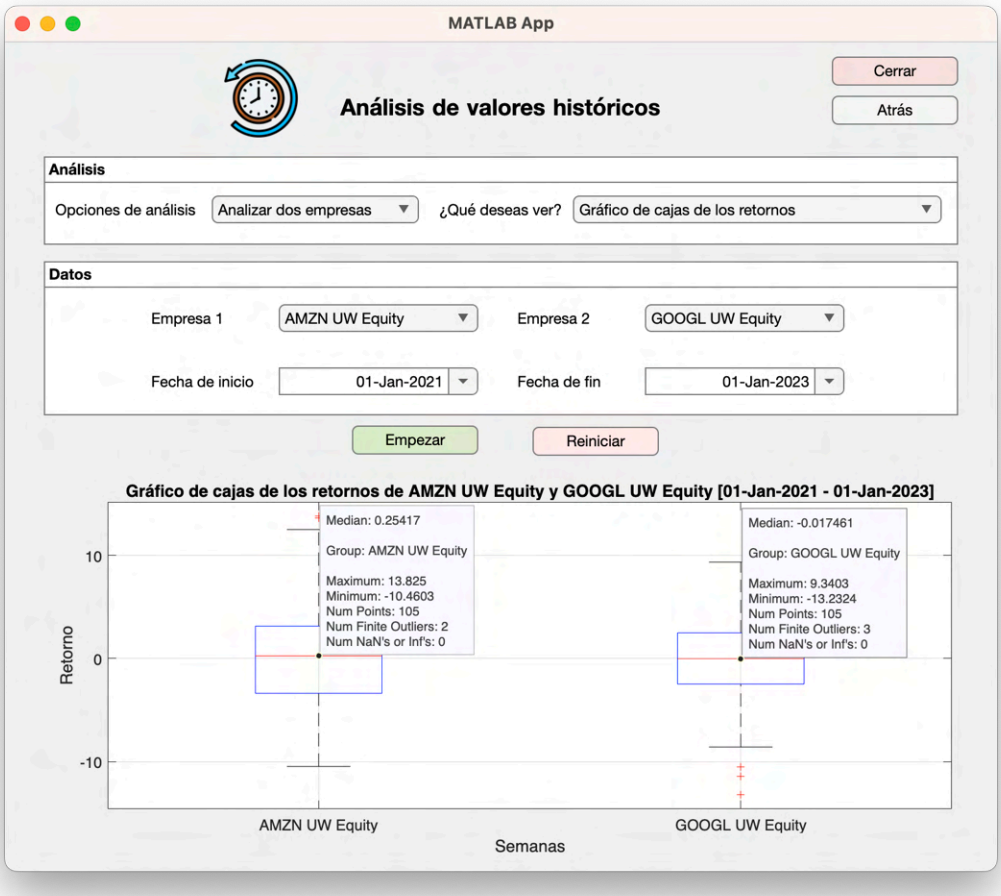


Ilustración 47. Analizar dos empresas - Gráfico de cajas de los retornos (Elaboración propia)

Comparando los resultados procedentes del gráfico de cajas, se observa que Amazon presenta un rango mayor en el que se encuentran la mayoría de sus rentabilidades mientras que Google presenta un rango algo más limitado. Sin embargo, el retorno que representa la mediana es muy similar en ambos casos, con la diferencia de que en el caso de Amazon es positivo y en el de Google es negativo. Además, los *outliers* de Amazon representan valores positivos mientras que los de Google están en la zona negativa, por lo que parece algo mejor escoger Amazon en caso de que haya que tomar una decisión de inversión entre las dos.

Gráfico de los retornos frente al riesgo

Según el gráfico representado en la Ilustración 43, el retorno de Google en función de su beta sería inferior al de Amazon, situándose también con una beta más baja, lo cual implica menor riesgo.

Gráfico iBoxPlot

Como ya se ha analizado el caso de Amazon en la Ilustración 44, en este caso se representa el resultado para Google.

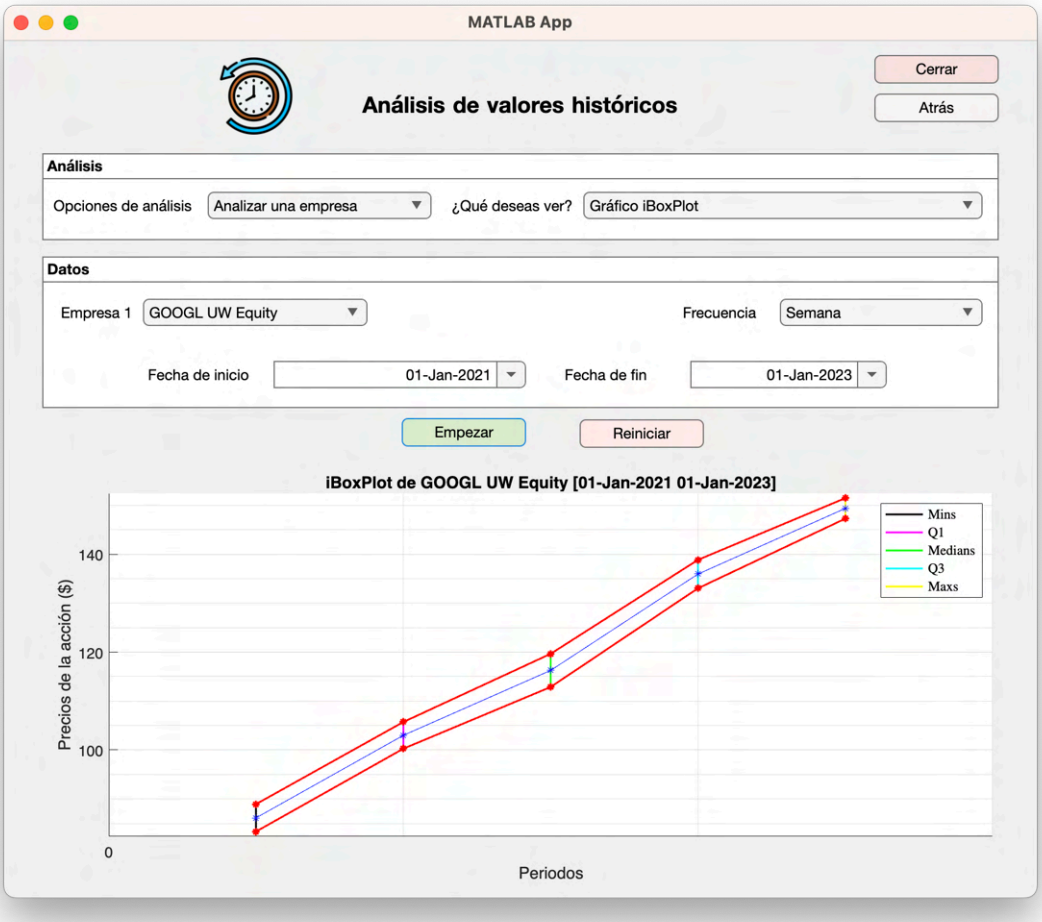


Ilustración 48. Analizar dos empresas - Gráfico iBoxPlot (Elaboración propia)

A partir de este resultado se puede concluir que debido a que la tendencia de ambas pendientes parece constante, en el periodo evaluado Google no será una compañía que presente cambios bruscos en el precio de sus acciones, por lo que no tendrá mucho riesgo de inversión.

En conclusión, podría decirse que tanto Google como Amazon, analizando datos de los últimos años, son compañías que no presentan un gran riesgo de inversión ya que normalmente sus retornos son estables y no sufren cambios significativos. Por tanto, serían una buena opción para inversores con aversión al riesgo, aunque podría ser complicado obtener gran rentabilidad a partir de ellas.

7.2.2 PREDICCIÓN Y RANKING DE VALORES FUTUROS

En el caso de escoger esta opción, se lleva al usuario a la siguiente ventana. En ella deberá escoger si quiere realizar la predicción del precio de las acciones de una o varias empresas, así como el método de predicción a utilizar.

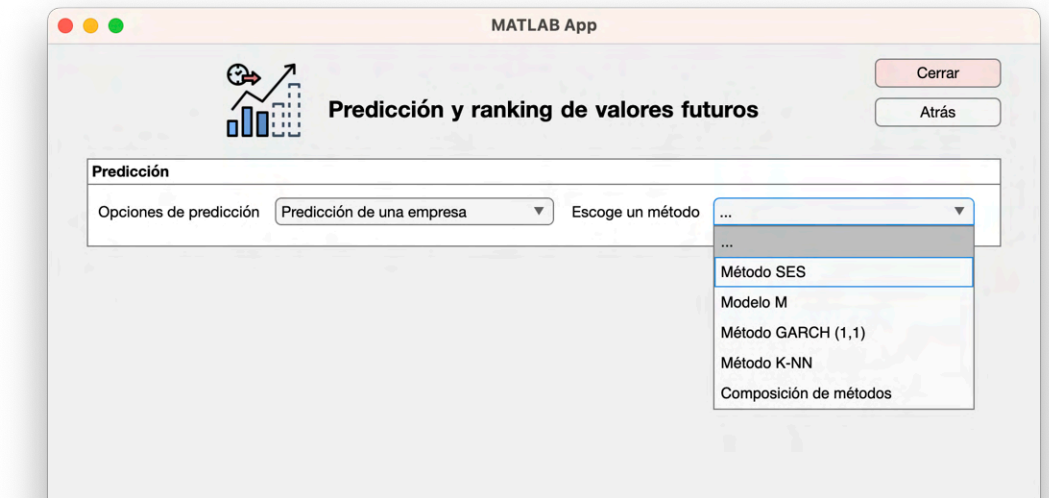


Ilustración 49. Ventana predicción y ranking de valores futuros (Elaboración propia)

7.2.2.1 Predicción de precio de una empresa

Para elaborar la predicción para una sola empresa, la aplicación comenzará por preguntar al usuario la empresa que desea estudiar junto con los datos de entrada necesarios para el estudio del método solicitado.

En este caso, se va a seguir trabajando con Amazon para poder tener una mayor proyección del estudio de la compañía. Los datos utilizados serán desde el 1 de Enero de 2019 hasta el 1 de Enero de 2022 para el conjunto de entrenamiento o *training* y del 2 de Enero de 2022 al 30 de Septiembre de 2023 para el conjunto de prueba o *test*. Además, es va a realizar la predicción para las 4 semanas posteriores a la última semana de la que se tienen datos, lo cual representa el número de periodos de predicción.

Una vez seleccionados los datos de entrada, la aplicación mostrará una gráfica con el resultado de predicción, así como una tabla con los datos numéricos en forma de intervalo.

Método SES

En este caso se utiliza el método de suavizado exponencial simple en el que el usuario introduce el valor de α explicado previamente y que se ha elegido como 0,4 para obtener la respuesta, trabajando con los centros y los radios de los intervalos.

El método SES devolverá una serie de valores constantes para todos los periodos seleccionados, representados por el intervalo [132,2025; 141,7707].

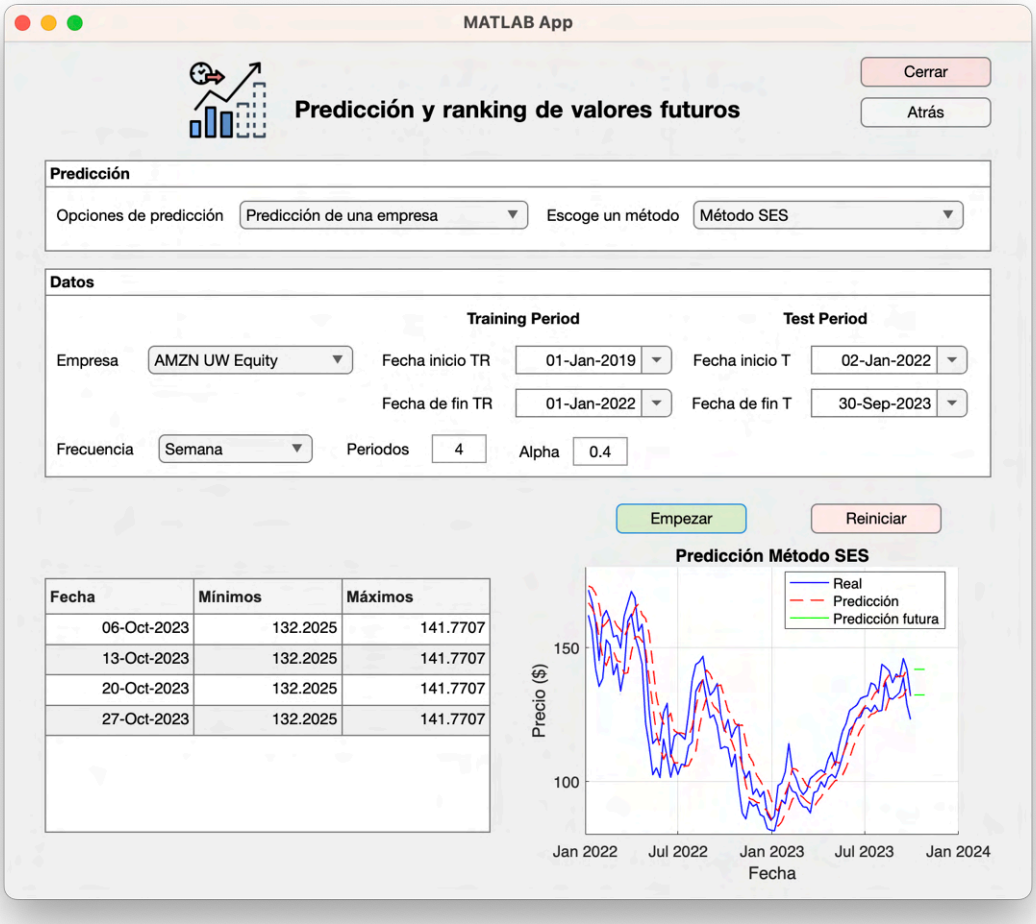


Ilustración 50. Predicción de precio de una empresa – Método SES (Elaboración propia)

Modelo M

El Modelo M hará uso de técnicas de regresión para realizar la predicción de los valores de Amazon en los 7 periodos posteriores a la última fecha del conjunto de test. El resultado obtenido es el que se muestra a continuación. Cabe destacar que en este caso se aprecia un resultado algo más conservador al del método anterior.

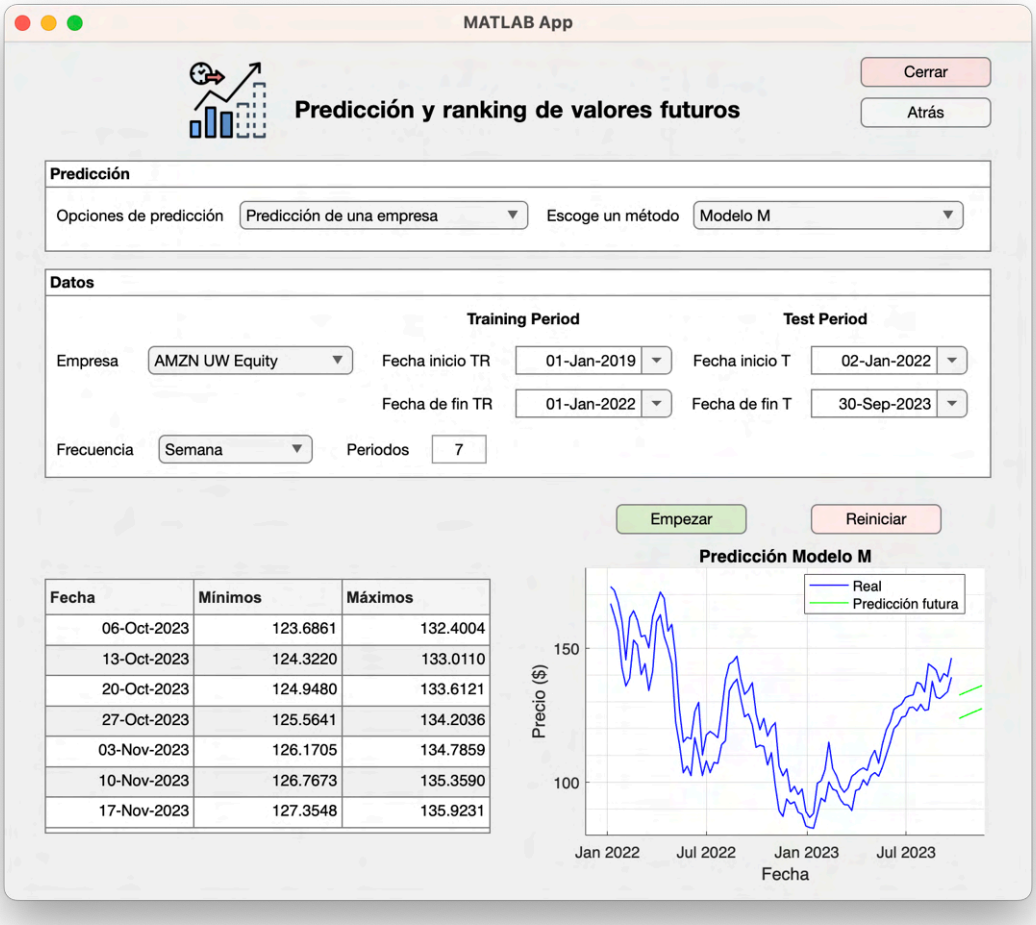


Ilustración 51. Predicción de precio de una empresa – Modelo M (Elaboración propia)

Método GARCH (1,1)

El método GARCH (1,1) estudia la volatilidad del rendimiento de un activo para realizar la predicción de precios futuros del mismo. Así, seleccionado los parámetros p y q iguales a 1, el resultado obtenido es el siguiente.

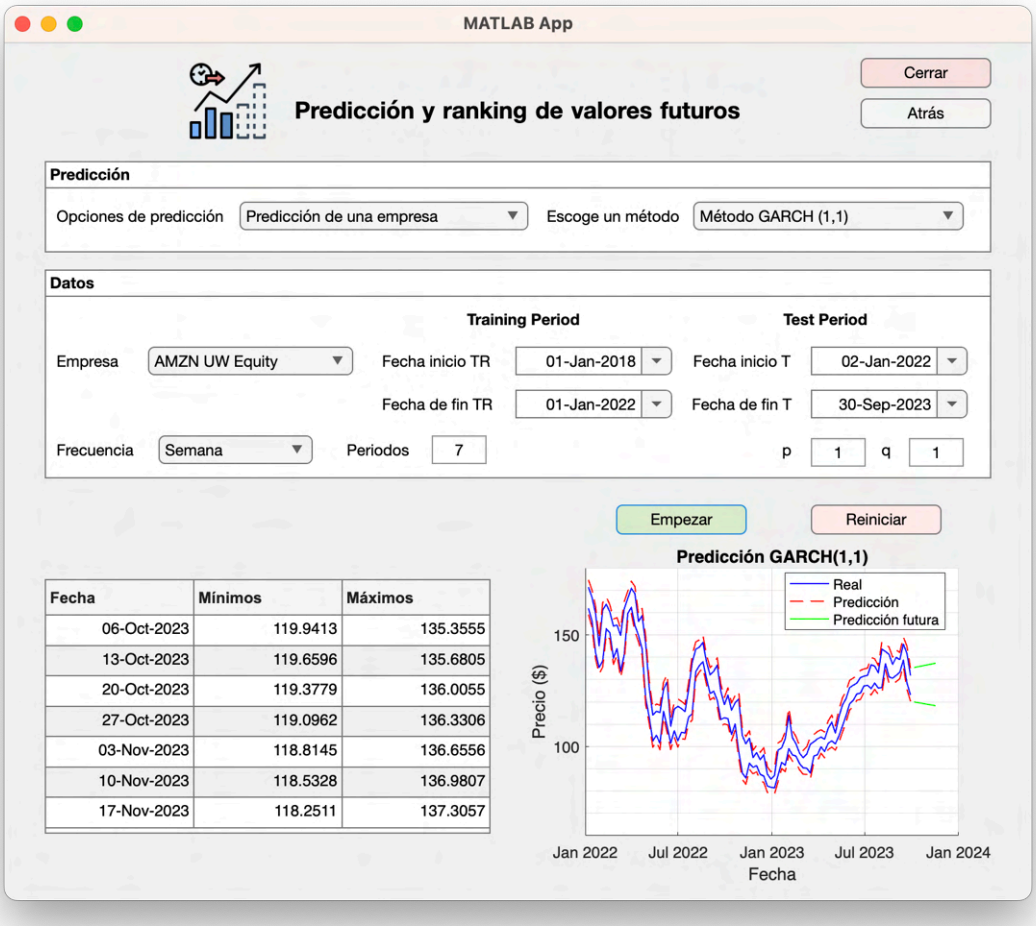
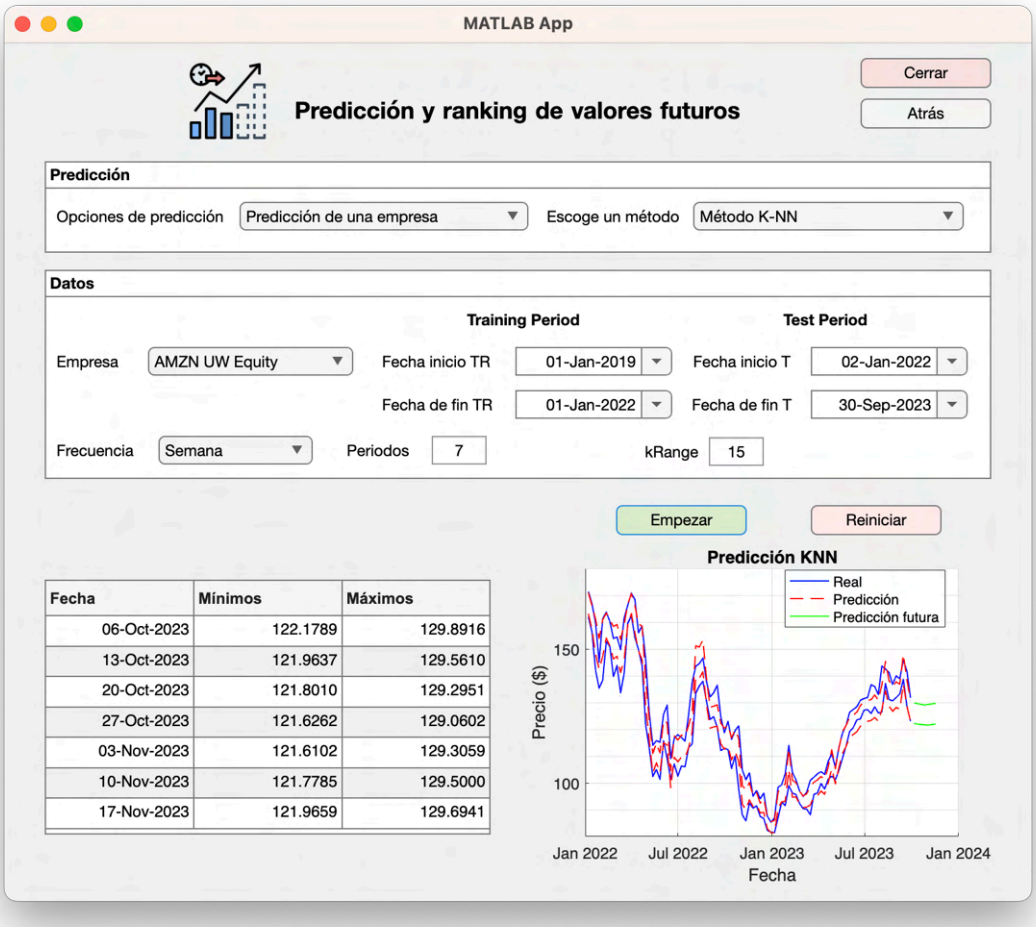


Ilustración 52. Predicción de precio de una empresa – Método GARCH (1,1) (Elaboración propia)

Según esto, el intervalo de precios de la acción irá aumentando su rango a medida que avance el tiempo. Esta opción presenta un resultado con valores menores al del método SES y con un rango de valores más amplio que el del modelo M.

Método K-NN

En este caso, a partir de los *k* vecinos más cercanos, escogidos por la acción a partir de optimización y habiendo establecido un límite de 15, que escoge el usuario, el programa realiza una media entre los valores de los mismos para poder estimar la predicción para Amazon como se muestra a continuación.



Fecha	Mínimos	Máximos
06-Oct-2023	122.1789	129.8916
13-Oct-2023	121.9637	129.5610
20-Oct-2023	121.8010	129.2951
27-Oct-2023	121.6262	129.0602
03-Nov-2023	121.6102	129.3059
10-Nov-2023	121.7785	129.5000
17-Nov-2023	121.9659	129.6941

Ilustración 53. Predicción de precio de una empresa – Método K-NN (Elaboración propia)

Los intervalos obtenidos como predicción para periodos futuros son similares a los casos anteriores, lo cual hace pensar que todos ellos podrían devolver resultados realistas.

7.2.2.2 Predicción de rentabilidad y ranking de varias empresas

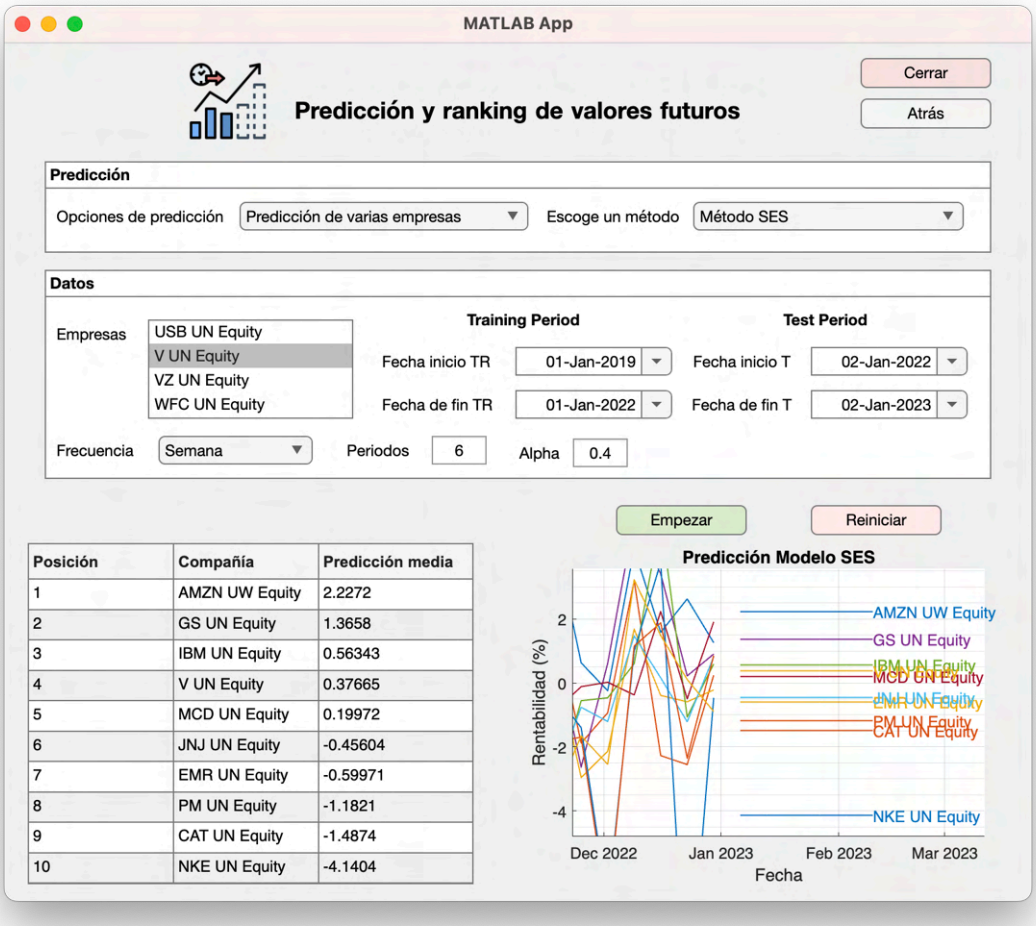
Por otro lado, la aplicación es capaz de estudiar el comportamiento de diferentes empresas de forma simultánea y realizar la predicción de sus rentabilidades futuras de cara a elaborar un ranking a partir de los resultados medios de estos. En este casos e estudian las rentabilidades ya que esta información permite comparar de forma más realista a las empresas entre sí.

Para observar el funcionamiento de la misma, se van a escoger una serie de compañías operando en distintas industrias que van a ser las siguientes: Amazon (*AMZN UW Equity*), Caterpillar Inc. (*CAT UN Equity*), Emerson Electric Co. (*EMR UN Equity*), Goldman Sachs (*GS UN Equity*), IBM Corp. (*IBM UN Equity*), Johnson & Johnson (*JNJ UN Equity*), McDonald's Corp. (*MCD UN Equity*), Nike Inc. (*NKE UN Equity*), Philip Morris Inc. (*PM UN Equity*) y Visa Inc. (*V UN Equity*).

Además, los periodos de estudio tanto para entrenamiento como para prueba serán los mismos que antes y se realizará la predicción para las semanas posteriores al último dato coleccionado.

Método SES

Para un valor de $\alpha = 0{,}4$ se obtiene el siguiente resultado donde destaca Nike, que se encuentra en última posición, siendo Amazon la primera. En este caso no se recomendaría invertir en Nike mientras que sí que se esperaría obtener rentabilidad positiva con Amazon, entre otras.

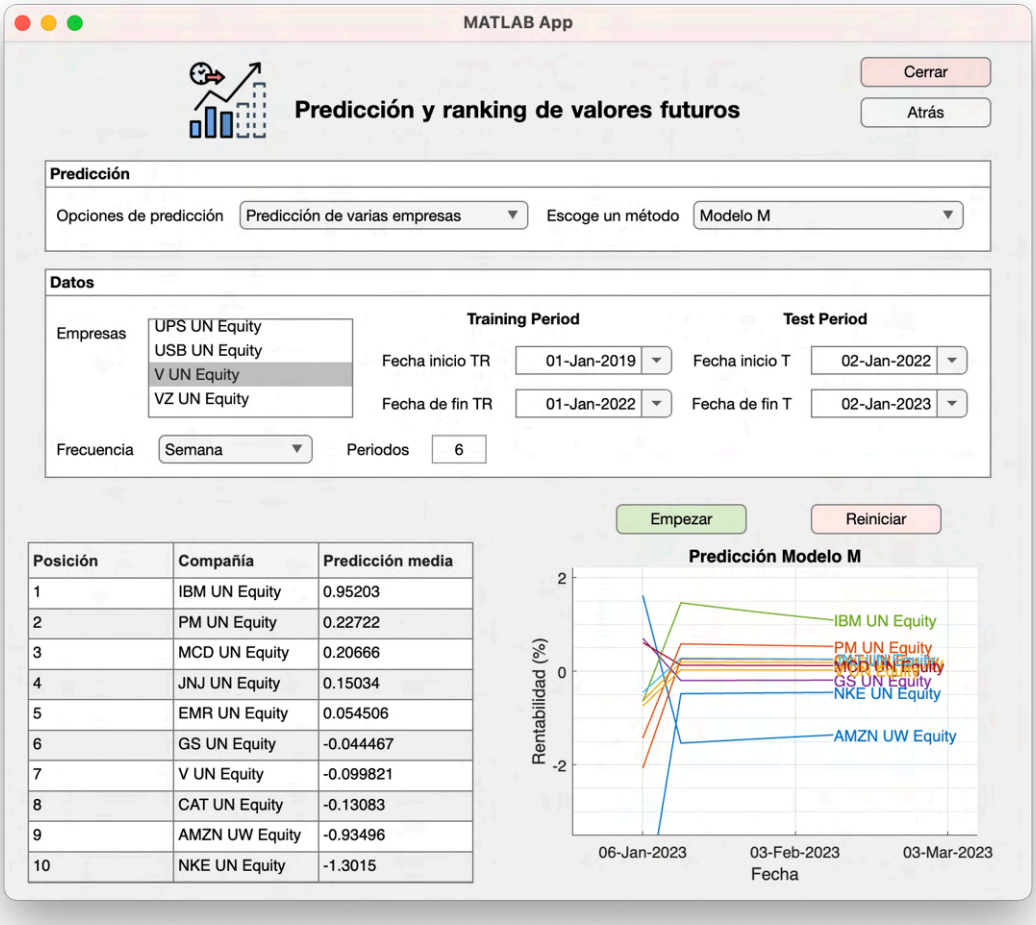


Posición	Compañía	Predicción media
1	AMZN UW Equity	2.2272
2	GS UN Equity	1.3658
3	IBM UN Equity	0.56343
4	V UN Equity	0.37665
5	MCD UN Equity	0.19972
6	JNJ UN Equity	-0.45604
7	EMR UN Equity	-0.59971
8	PM UN Equity	-1.1821
9	CAT UN Equity	-1.4874
10	NKE UN Equity	-4.1404

Ilustración 54. Predicción de rentabilidad y ranking de varias empresas – Método SES (Elaboración propia)

Modelo M

A continuación, se presenta el resultado obtenido en el caso del modelo M, que muestra resultados de predicción distintos a los del caso anterior, destacando que en este caso IBM será el que espera una mayor rentabilidad. Llama la atención que el valor de las rentabilidades es generalmente bajo en comparación con el análisis anterior.

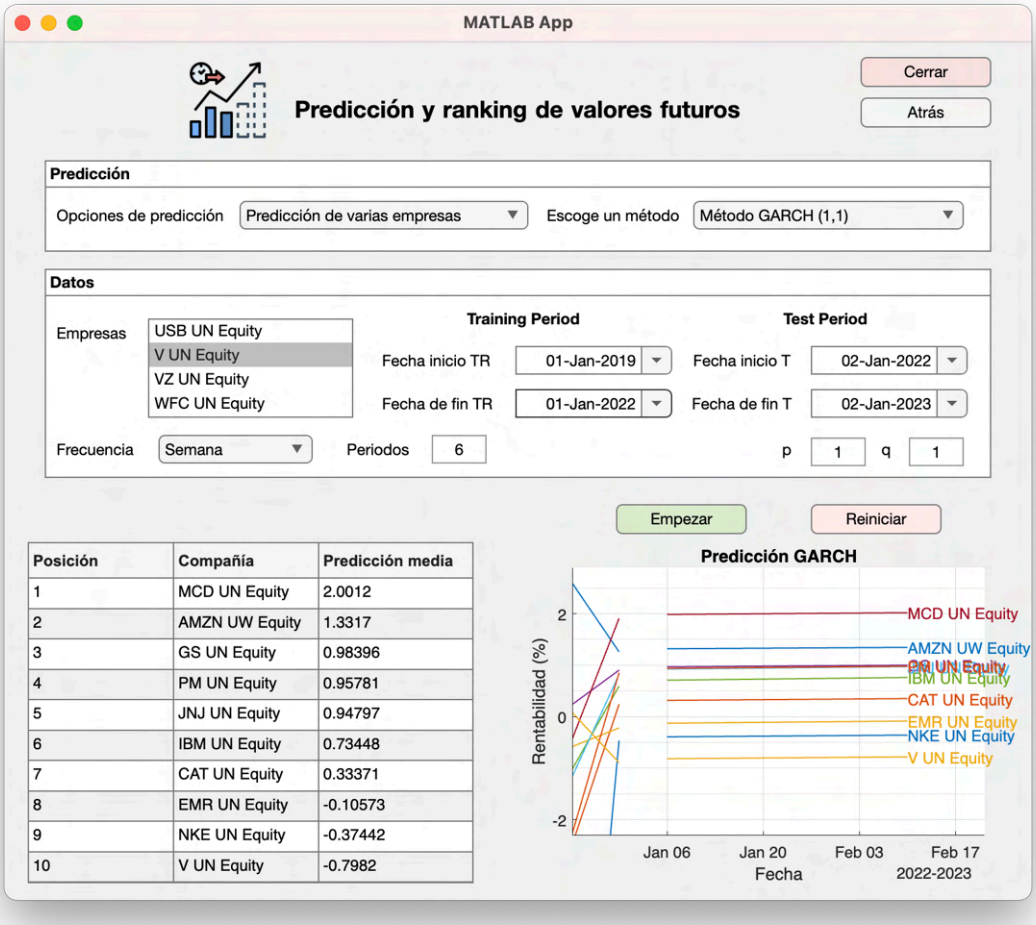


Posición	Compañía	Predicción media
1	IBM UN Equity	0.95203
2	PM UN Equity	0.22722
3	MCD UN Equity	0.20666
4	JNJ UN Equity	0.15034
5	EMR UN Equity	0.054506
6	GS UN Equity	-0.044467
7	V UN Equity	-0.099821
8	CAT UN Equity	-0.13083
9	AMZN UW Equity	-0.93496
10	NKE UN Equity	-1.3015

Ilustración 55. Predicción de rentabilidad y ranking de varias empresas – Modelo M (Elaboración propia)

Método GARCH (1,1)

En este análisis se estudian las compañías desde el método GARCH (1,1) y se observan diferencias en los resultados con respecto a los casos anteriores. Ahora, al igual que en el caso anterior, no hay ningún valor que resulte alarmante y sería McDonald's la compañía en la que más se recomendaría invertir, seguida de Amazon.

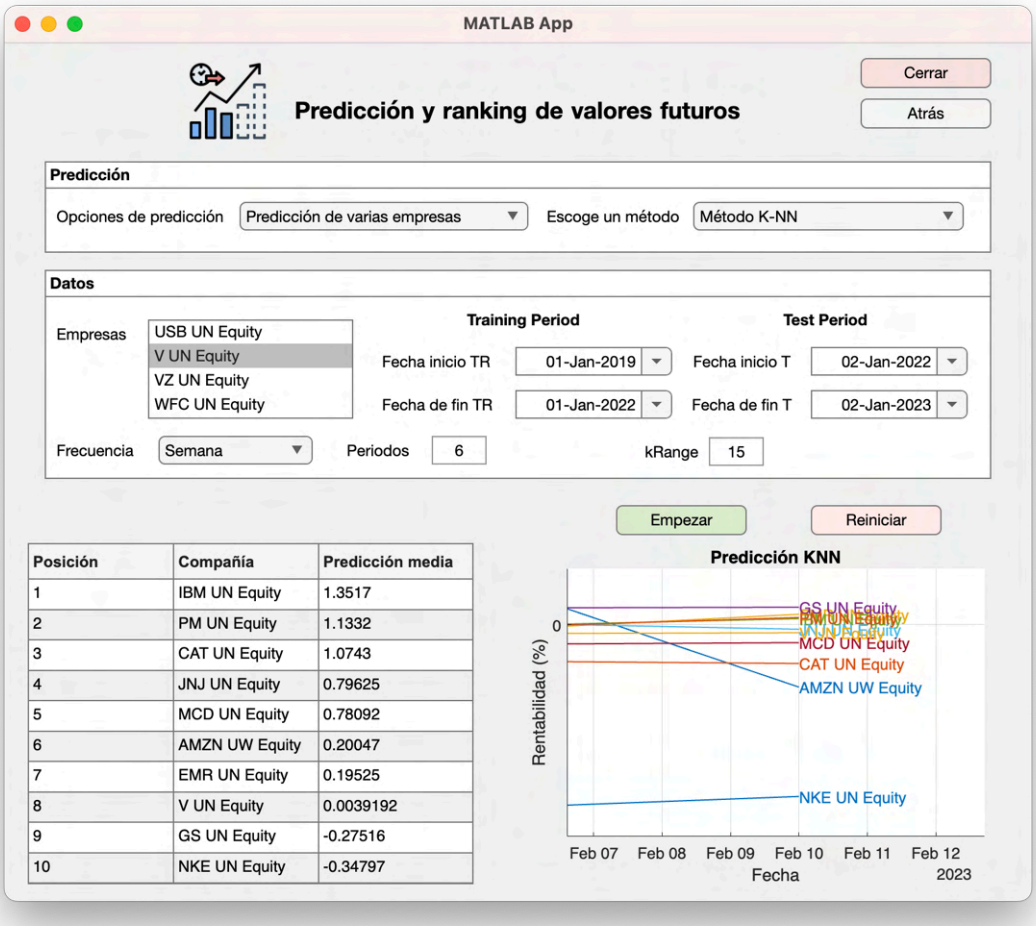


Posición	Compañía	Predicción media
1	MCD UN Equity	2.0012
2	AMZN UW Equity	1.3317
3	GS UN Equity	0.98396
4	PM UN Equity	0.95781
5	JNJ UN Equity	0.94797
6	IBM UN Equity	0.73448
7	CAT UN Equity	0.33371
8	EMR UN Equity	-0.10573
9	NKE UN Equity	-0.37442
10	V UN Equity	-0.7982

Ilustración 56. Predicción de rentabilidad y ranking de varias empresas – Método GARCH (1,1) (Elaboración propia)

Método K-NN

Volviendo a escoger un valor de $kRange = 15$, el resultado obtenido para este método es el siguiente. Al igual que en los anteriores, Nike sigue teniendo el valor más bajo mientras que otras como IBM, se mantienen en las primeras posiciones. Destaca que Amazon tiene una tendencia decreciente en el tiempo.

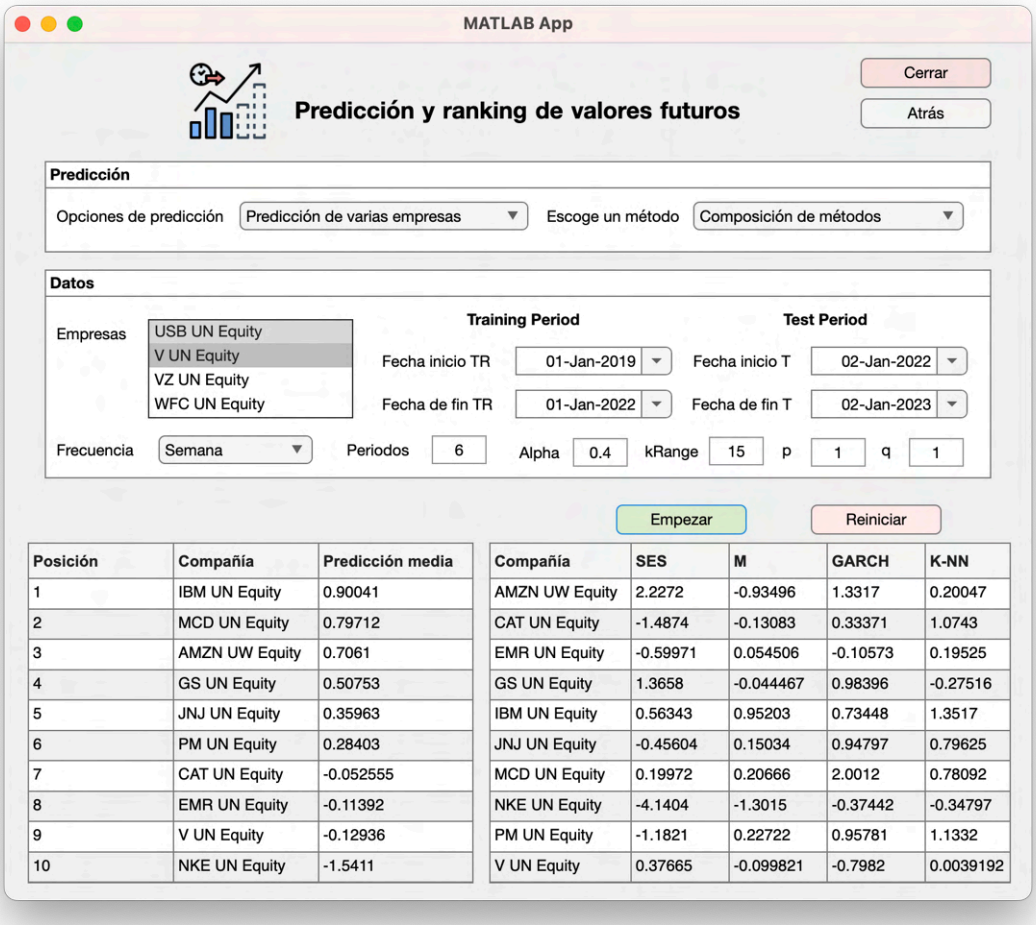


Posición	Compañía	Predicción media
1	IBM UN Equity	1.3517
2	PM UN Equity	1.1332
3	CAT UN Equity	1.0743
4	JNJ UN Equity	0.79625
5	MCD UN Equity	0.78092
6	AMZN UW Equity	0.20047
7	EMR UN Equity	0.19525
8	V UN Equity	0.0039192
9	GS UN Equity	-0.27516
10	NKE UN Equity	-0.34797

Ilustración 57. Predicción y ranking de varias empresas – Método K-NN (Elaboración propia)

Composición de métodos

Finalmente, se va a utilizar este método para estudiar simultáneamente los 4 anteriores y poder realizar una predicción media a partir de ellos. Así, se obtiene el siguiente resultado.



Posición	Compañía	Predicción media
1	IBM UN Equity	0.90041
2	MCD UN Equity	0.79712
3	AMZN UW Equity	0.7061
4	GS UN Equity	0.50753
5	JNJ UN Equity	0.35963
6	PM UN Equity	0.28403
7	CAT UN Equity	-0.052555
8	EMR UN Equity	-0.11392
9	V UN Equity	-0.12936
10	NKE UN Equity	-1.5411

Compañía	SES	M	GARCH	K-NN
AMZN UW Equity	2.2272	-0.93496	1.3317	0.20047
CAT UN Equity	-1.4874	-0.13083	0.33371	1.0743
EMR UN Equity	-0.59971	0.054506	-0.10573	0.19525
GS UN Equity	1.3658	-0.044467	0.98396	-0.27516
IBM UN Equity	0.56343	0.95203	0.73448	1.3517
JNJ UN Equity	-0.45604	0.15034	0.94797	0.79625
MCD UN Equity	0.19972	0.20666	2.0012	0.78092
NKE UN Equity	-4.1404	-1.3015	-0.37442	-0.34797
PM UN Equity	-1.1821	0.22722	0.95781	1.1332
V UN Equity	0.37665	-0.099821	-0.7982	0.0039192

Ilustración 58. Predicción y ranking de varias empresas – Composición de métodos (Elaboración propia)

Además de observar la tabla con el ranking y la predicción media como hasta ahora, se puede apreciar para cada compañía el resultado obtenido según cada método. Según esto, cabe concluir que es importante prestar atención a todos los métodos ya que cada uno tiene en cuenta una serie de factores a la hora de realizar la predicción y por ello, los resultados de las estimaciones pueden ser diferentes entre ellos.

Finalmente, se van a mostrar los datos reales que obtuvieron los valores de las empresas seleccionadas durante los 6 periodos (semanas) estimados para poder realizar una comparación con los resultados obtenidos y ver la efectividad de los distintos modelos.

Semana	Resultado real (%)	Método SES		Modelo M		Método GARCH (1,1)		Método K-NN		Composición de métodos	
		Predicción (%)	Error (VA)	Predicción (%)	Error (VA)	Predicción (%)	Error (VA)	Predicción (%)	Error (VA)	Predicción (%)	Error (VA)
AMZN	-1,44	2,23	3,66	-0,93	0,50	1,33	2,77	0,20	1,64	0,71	2,14
CAT	0,28	-1,49	1,77	-0,13	0,41	0,33	0,05	1,07	0,79	-0,05	0,34
EMR	1,90	-0,60	2,50	0,05	1,84	-0,11	2,00	0,20	1,70	-0,11	2,01
GS	-1,13	1,37	2,50	-0,04	1,08	0,98	2,11	-0,28	0,85	0,51	1,64
IBM	1,25	0,56	0,69	0,95	0,30	0,73	0,52	1,35	0,10	0,90	0,35
JNJ	1,88	-0,46	2,33	0,12	1,76	0,95	0,93	0,80	1,08	0,36	1,52
MCD	0,43	0,20	0,23	0,21	0,22	2,00	1,58	0,78	0,36	0,80	0,37
NKE	0,26	-4,14	4,40	-1,30	1,57	-0,37	0,64	-0,35	0,61	-1,54	1,81
PM	0,78	-1,18	1,97	0,23	0,56	0,96	0,17	1,13	0,35	0,28	0,50
V	-0,47	0,38	0,85	-0,10	0,37	-0,80	0,33	0,00	0,48	-0,13	0,34
Error Medio		2,09		0,86		1,11		0,80		1,10	

Tabla 15. Evaluación de resultados de la herramienta

Según lo anterior, es posible concluir que, de acuerdo al error absoluto medio calculado para cada método, el más fiable será el método K-NN, seguido del modelo M. A continuación, el método GARCH(1,1) y el de composición de métodos, tendrían resultados similares y por último, el que mayor error daría y por tanto, el método menos recomendado sería el método SES.

7.3 PRÓXIMOS PASOS

Una vez obtenidos los resultados que al inversor le interesen de la aplicación, es hora de asignar a cada compañía el porcentaje de acciones que se considere conveniente para poder formar la cartera. Para ello, el inversor podrá tomar distintos enfoques en función del riesgo que esté dispuesto a asumir de acuerdo al objetivo de rentabilidad que desee obtener. Un ejemplo de formación de una cartera podría ser asignando un porcentaje del total de la inversión a las compañías que hayan quedado en las primeras posiciones del ranking que más interese al inversor. Por ejemplo, asignar un 40% a la compañía que quede en primera posición, un 30% a la segunda, un 20% a la tercera y un 10% a la cuarta.

El sistema de asignación será distinto para cada inversor en función de diferentes factores como los ya expuestos anteriormente o, por ejemplo, el grado de diversificación que desee alcanzar ya que inversores que tiendan a diversificar preferirán distribuir sus

inversiones en más compañías que otros que tiendan menos a la diversificación, los cuales optarán por concentrar sus recursos hacia un número menor de compañías.

7.4 ANÁLISIS DE POTENCIALES MEJORAS

Como último paso antes de la finalización del proyecto, se ha añadido la posibilidad de que el usuario seleccione la frecuencia de los datos con los que va a realizar su análisis. Así, el usuario tiene un desplegable en la sección de elección de parámetros en el que podrá escoger si la frecuencia de los datos con los que va a realizar el estudio es semanal, mensual o trimestral, tal y como se muestra a continuación.

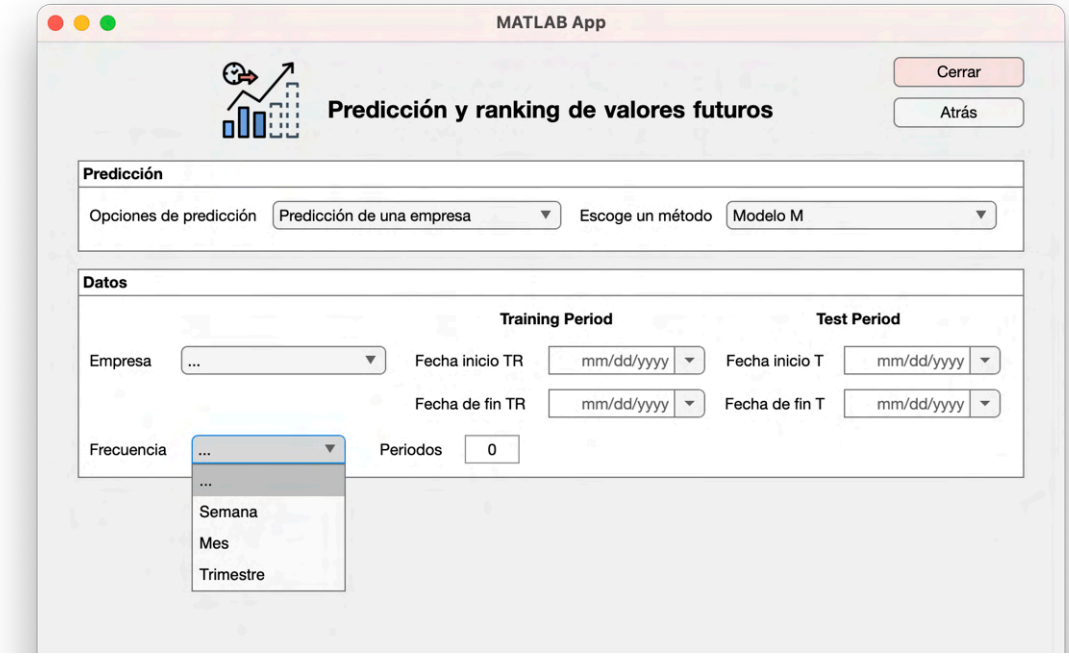


Ilustración 59. Selección de frecuencia en la herramienta

A partir de aquí, se ha realizado la predicción y ranking de precios futuros para varias empresas, las mismas que se analizaron en 7.2.2 y con los mismos parámetros para poder llevar a cabo una comparación realista. Los resultados obtenidos para cada una de las opciones, así como el error de predicción medio, se muestran a continuación.

Semana	Resultado real	Método SES		Modelo M		Método GARCH (1,1)		Método K-NN		Composición de métodos	
		Predicción	% Error	Predicción	% Error	Predicción	% Error	Predicción	% Error	Predicción	% Error
AMZN	$ 98,00	$ 89,75	8,41%	$ 83,22	15,08%	$ 83,11	15,20%	$ 59,01	39,78%	$ 78,77	19,62%
CAT	$ 250,13	$ 229,71	8,16%	$ 256,34	2,48%	$ 236,26	5,54%	$ 134,31	46,30%	$ 250,55	0,17%
EMR	$ 89,36	$ 93,57	4,71%	$ 94,14	5,35%	$ 94,45	5,70%	$ 59,70	33,19%	$ 88,61	0,84%
GS	$ 361,57	$ 350,08	3,18%	$ 332,58	8,02%	$ 332,94	7,92%	$ 226,60	37,33%	$ 323,96	10,40%
IBM	$ 137,63	$ 138,38	0,54%	$ 132,67	3,60%	$ 135,96	1,21%	$ 125,60	8,74%	$ 148,20	7,68%
JNJ	$ 165,83	$ 172,65	4,11%	$ 173,16	4,42%	$ 260,17	56,89%	$ 121,55	26,70%	$ 188,68	13,78%
MCD	$ 267,04	$ 265,16	0,71%	$ 263,15	1,46%	$ 257,69	3,50%	$ 151,47	43,28%	$ 231,86	13,18%
NKE	$ 125,07	$ 107,98	13,66%	$ 104,02	16,83%	$ 115,49	7,66%	$ 61,01	51,22%	$ 99,32	20,58%
PM	$ 101,75	$ 95,43	6,21%	$ 96,38	5,28%	$ 97,64	4,04%	$ 85,56	15,91%	$ 101,24	0,50%
V	$ 224,83	$ 207,19	7,85%	$ 207,31	7,79%	$ 312,54	39,01%	$ 109,14	51,46%	$ 210,70	6,28%
% Error Medio		5,76%		7,03%		14,67%		35,39%		9,30%	

Mes	Resultado real	Método SES		Modelo M		Método GARCH (1,1)		Método K-NN		Composición de métodos	
		Predicción	% Error	Predicción	% Error	Predicción	% Error	Predicción	% Error	Predicción	% Error
AMZN	$ 98,00	$ 110,55	12,81%	$ 87,98	10,22%	$ 88,22	9,98%	$ 54,83	44,05%	$ 85,39	12,87%
CAT	$ 250,13	$ 200,27	19,93%	$ 306,23	22,43%	$ 180,49	27,84%	$ 116,49	53,43%	$ 200,87	19,69%
EMR	$ 89,36	$ 83,74	6,29%	$ 87,93	1,60%	$ 94,82	6,11%	$ 54,75	38,73%	$ 80,31	10,13%
GS	$ 361,57	$ 328,07	9,27%	$ 332,16	8,14%	$ 349,69	3,29%	$ 211,87	41,40%	$ 305,45	15,52%
IBM	$ 137,63	$ 127,96	7,03%	$ 125,28	8,97%	$ 139,05	1,03%	$ 117,23	14,82%	$ 127,38	7,45%
JNJ	$ 165,83	$ 165,23	0,36%	$ 176,96	6,71%	$ 163,61	1,34%	$ 117,30	29,26%	$ 155,77	6,07%
MCD	$ 267,04	$ 254,46	4,71%	$ 281,17	5,29%	$ 213,79	19,94%	$ 143,84	46,14%	$ 223,31	16,38%
NKE	$ 125,07	$ 98,38	21,34%	$ 76,46	38,87%	$ 109,45	12,49%	$ 58,66	53,10%	$ 85,74	31,45%
PM	$ 101,75	$ 87,83	13,68%	$ 92,08	9,50%	$ 96,77	4,90%	$ 81,32	20,08%	$ 89,50	12,04%
V	$ 224,83	$ 199,13	11,43%	$ 219,08	2,56%	$ 194,68	13,41%	$ 103,64	53,90%	$ 179,13	20,33%
% Error Medio		10,68%		11,43%		10,03%		39,49%		15,19%	

Trimestre	Resultado real	Método SES		Modelo M		Método GARCH (1,1)		Método K-NN		Composición de métodos	
		Predicción	% Error	Predicción	% Error	Predicción	% Error	Predicción	% Error	Predicción	% Error
AMZN	$ 98,00	$ 140,83	43,71%	$ 94,32	3,75%	$ 6,89	92,97%	$ 41,75	57,40%	$ 70,95	27,60%
CAT	$ 250,13	$ 188,35	24,70%	$ 500,90	100,26%	$ 161,06	35,61%	$ 84,43	66,25%	$ 233,69	6,57%
EMR	$ 89,36	$ 84,02	5,98%	$ 68,21	23,67%	$ 88,68	0,76%	$ 48,40	45,84%	$ 72,33	19,06%
GS	$ 361,57	$ 317,61	12,16%	$ 253,05	30,01%	$ 317,16	12,28%	$ 184,21	49,05%	$ 268,01	25,88%
IBM	$ 137,63	$ 121,67	11,60%	$ 109,13	20,71%	$ 108,71	21,01%	$ 109,18	20,67%	$ 112,17	18,50%
JNJ	$ 165,83	$ 163,33	1,51%	$ 164,44	0,84%	$ 146,54	11,63%	$ 101,98	38,50%	$ 144,07	13,12%
MCD	$ 267,04	$ 236,45	11,46%	$ 280,14	4,90%	$ 67,22	74,83%	$ 113,64	57,45%	$ 174,36	34,71%
NKE	$ 125,07	$ 119,58	4,39%	$ 58,31	53,38%	$ 100,84	19,37%	$ 53,28	57,40%	$ 83,00	33,64%
PM	$ 101,75	$ 88,17	13,35%	$ 83,59	17,85%	$ 87,53	13,98%	$ 69,19	32,00%	$ 82,12	19,29%
V	$ 224,83	$ 203,32	9,57%	$ 226,85	0,90%	$ 169,30	24,70%	$ 83,76	62,74%	$ 170,81	24,03%
% Error Medio		13,84%		25,63%		30,71%		48,73%		22,24%	

Tabla 16. Resultados y error de predicción según la frecuencia de los datos

Ranking - Real		Ranking - Semana		Ranking - Mes		Ranking - Trimestre	
Posición	Compañía	Posición	Compañía	Posición	Compañía	Posición	Compañía
1	GS UN Equity	1	GS UN Equity	1	GS UN Equity	1	GS UN Equity
2	MCD UN Equity	2	MCD UN Equity	2	MCD UN Equity	2	CAT UN Equity
3	CAT UN Equity	3	CAT UN Equity	3	CAT UN Equity	3	MCD UN Equity
4	V UN Equity	4	V UN Equity	4	V UN Equity	4	V UN Equity
5	JNJ UN Equity	5	JNJ UN Equity	5	JNJ UN Equity	5	JNJ UN Equity
6	IBM UN Equity	6	IBM UN Equity	6	IBM UN Equity	6	IBM UN Equity
7	NKE UN Equity	7	NKE UN Equity	7	PM UN Equity	7	NKE UN Equity
8	PM UN Equity	8	PM UN Equity	8	NKE UN Equity	8	PM UN Equity
9	AMZN UW Equity	9	EMR UN Equity	9	AMZN UW Equity	9	EMR UN Equity
10	EMR UN Equity	10	AMZN UW Equity	10	EMR UN Equity	10	AMZN UW Equity

Tabla 17. Ranking según la frecuencia de los datos

De las tablas anteriores se ha podido concluir que, para estimar el valor del precio de la acción en el futuro, la mejor opción de análisis desde el punto de vista de la composición

de métodos es aquella que utiliza una frecuencia de datos semanales, siendo su error de predicción medio del 9,30%, seguida de la opción de datos mensuales, con un error de predicción medio del 15,19% y por último quedaría la opción de datos trimestrales, la cual ha dado como resultado medio de error de predicción un 22,24%.

Por otro lado, al prestar atención al ranking obtenido en cada una de las casuísticas para el método SES y compararlo con el ranking real, obtenido a partir de datos semanales, se puede observar que las opciones que ofrecen un resultado más parecido a la realidad son las dos primeras, con dos fallos frente a diez valores estudiados en el ranking, mientras que la opción de datos con frecuencia trimestral ofrece cuatro errores frente a los diez estudiados. Cabe destacar que para el primer caso el ranking obtenido es prácticamente igual al caso real, coincidiendo las 8 primeras posiciones del ranking, por lo que resulta ser un buen método de estimación de ranking.

Finalmente, se ha demostrado que el método que presenta menor porcentaje error de predicción es el Método SES, destacando su fiabilidad en este contexto. Sin embargo, es el método K-NN aquel que presenta un error mayor, lo que sugiere que no va a ser tan sólido como única fuente de análisis. Igualmente, como ya se había subrayado en los apartados anteriores, la estrategia más segura para el inversor antes de tomar una decisión será combinar información y resultados procedentes de distintos métodos para poder elaborar la cartera de valores que mejor se ajuste a sus necesidades.

Capítulo 8. CONCLUSIONES

La optimización de carteras representa un campo de vital importancia en el mundo financiero. Cada inversor tiene unas características, preferencias y objetivos particulares que deben ser tenidos en cuenta a la hora de construir su cartera óptima de manera que encuentre el balance correcto entre el rendimiento que desea obtener y el riesgo que está dispuesto a asumir. El objetivo de la optimización de carteras es realizar una distribución óptima del capital a partir del análisis de datos históricos, tendencias de mercado y características de riesgo de los distintos activos

Por otro lado, el uso de datos de intervalo proporciona ventajas significativas en la representación y transmisión de la información en distintos ámbitos, siendo uno de ellos el mundo de los mercados financieros. Así, el uso de estos datos es de gran utilidad a la hora de representar la incertidumbre, medir y gestionar riesgos, optimizar carteras y, con todo ello, tomar decisiones de inversión.

Debido a ello, a lo largo de este proyecto se ha elaborado una herramienta que tiene como objetivo convertirse en un sistema de referencia para inversores a la hora de tomar sus decisiones de inversión y construir sus carteras, a partir del análisis de datos de intervalo. Así, la herramienta consta de dos funcionalidades principales. La primera de ellas, analizar datos históricos que ayuden al inversor a entender el comportamiento de la acción a lo largo del tiempo para construir juicios informados sobre su potencial desarrollo futuro y poder adaptar sus estrategias en función de ello. La segunda, realizar una estimación de los valores futuros de las acciones para identificar oportunidades de inversión y optimizar la distribución del capital que forma la cartera, identificando oportunidades y momentos óptimos de entrada o salida para cumplir con los objetivos deseados.

Una vez probado el funcionamiento de la herramienta se puede concluir que resulta de gran utilidad hacer uso de todas sus funcionalidades de cara a obtener una visión completa antes de elaborar una estrategia de inversión. En cuanto a los datos pasados, es interesante

estudiar los más recientes porque es más probable que tengan un comportamiento similar a los que se vayan a obtener a corto plazo, pero no se deben de olvidar los más antiguos ya que pueden ofrecer información sobre ciclos de vida, patrones o tendencias seguidos por una empresa, industria o mercado.

Por otro lado, desde el punto de vista de las predicciones futuras resulta relevante prestar interés a los cinco modelos de predicción ofrecidos ya que cada uno de ellos realiza sus estimaciones en función de distintos factores que influyen en la determinación del precio y por tanto, rentabilidad de la acción. Se ha obtenido que aquel que devuelve un menor error de predicción es el método K-NN, seguido del modelo M, con unos resultados medios de 0,80 y 0,86 de error absoluto medio respectivamente. A continuación, el método de composición de métodos y el método GARCH (1,1) serían los próximos modelos a elegir, con unos valores de error medio de 1,10 y 1,11. Finalmente, el método menos recomendable a la hora de estimar la rentabilidad sería el método SES, con un resultado de error medio absoluto de 2,09.

Finalmente, cabe concluir que las posiciones en los rankings obtenidas de los distintos métodos son similares, de manera que hay algunas compañías que se sitúan en la primera mitad del ranking mientras que hay otras que fluctúan en la segunda mitad, por lo que se deriva que el análisis es coherente en todos los métodos.

Ranking - SES		Ranking - Modelo M		Ranking - GARCH (1,1)	
Posición	**Compañía**	**Posición**	**Compañía**	**Posición**	**Compañía**
1	AMZN UW Equity	1	IBM UN Equity	1	MCD UN Equity
2	GS UN Equity	2	PM UN Equity	2	AMZN UW Equity
3	IBM UN Equity	3	MCD UN Equity	3	GS UN Equity
4	V UN Equity	4	JNJ UN Equity	4	PM UN Equity
5	MCD UN Equity	5	EMR UN Equity	5	JNJ UN Equity
6	JNJ UN Equity	6	GS UN Equity	6	IBM UN Equity
7	EMR UN Equity	7	V UN Equity	7	CAT UN Equity
8	PM UN Equity	8	CAT UN Equity	8	EMR UN Equity
9	CAT UN Equity	9	AMZN UW Equity	9	NKE UN Equity
10	NKE UN Equity	10	NKE UN Equity	10	V UN Equity

Ranking - K-NN		Ranking - Composición		Ranking - Real	
Posición	**Compañía**	**Posición**	**Compañía**	**Posición**	**Compañía**
1	IBM UN Equity	1	IBM UN Equity	1	EMR UN Equity
2	PM UN Equity	2	MCD UN Equity	2	JNJ UN Equity
3	CAT UN Equity	3	AMZN UW Equity	3	IBM UN Equity
4	JNJ UN Equity	4	GS UN Equity	4	PM UN Equity
5	MCD UN Equity	5	JNJ UN Equity	5	MCD UN Equity
6	AMZN UW Equity	6	PM UN Equity	6	CAT UN Equity
7	EMR UN Equity	7	CAT UN Equity	7	NKE UN Equity
8	V UN Equity	8	EMR UN Equity	8	V UN Equity
9	GS UN Equity	9	V UN Equity	9	GS UN Equity
10	NKE UN Equity	10	NKE UN Equity	10	AMZN UW Equity

Ilustración 60. Resultados del ranking de todos los métodos (Elaboración propia)

En conclusión, se recomienda tener en cuenta los valores resultantes de todos los métodos posibles para poder obtener una idea realista del resultado pronosticado.

Adicionalmente, en el caso de poder elegir la frecuencia de los datos históricos, se ha demostrado que la mejor estrategia para el inversor vendría de los resultados obtenidos de escoger datos semanales, donde se ha destacado que coincidían las 8 primeras posiciones del ranking, lo cual puede deberse a que presentarán menor desviación entre sí y, por tanto, devolverán unos resultados más precisos.

Finalmente, una vez desarrollado el proyecto, es posible concluir que se han cumplido los objetivos principales del mismo gracias al desarrollo de la herramienta explicada, de la cual han podido derivarse sugerentes conclusiones. Se dejan las puertas abiertas para posibles desarrollos futuros en la línea del artículo de (Alzaman, 2024) en relación con las redes neuronales que perfeccionen el trabajo realizado durante estos últimos meses y que con estas palabras llega a su fin.

Capítulo 9. BIBLIOGRAFÍA

Acín, C. (2022). *La causalidad de Granger en el análisis y la previsión de series temporales clásicas, de intervalo y de histograma. Aplicación en mercados financieros.* Escuela Técnica Superior de Ingeniería (ICAI). Trabajo de Fin de Grado.

Alzaman, C. (2024). Deep learning in stock portfolio selection and predictions. *Expert Systems with Applications*, 237, 121404.

Arroyo, J., & Maté, C. (2006). *Introducing interval time series: Accuracy measures.* Proc in Computational Statistics.

Bankinter. (22 de Julio de 2021). *Riesgo financiero: ¿qué es el VaR, la Beta y el índice VIX?* Obtenido de Bankinter: https://www.bankinter.com/blog/finanzas-personales/inversion-riesgo-financiero-que-es-var-vix-beta#:~:text=La%20Beta%20(%CE%B2)%20es%20la,mercado%20en%20el%20que%20cotiza.

Bertani, F. (2021). *Deep Learning methods for Portfolio Optimization.* Scuola di Scienze. Sede di Bologna Dipartimento di Informatica, Scienza e Ingegneria. Tesi di Laurea Magistrale in Finanza Computazionale.

Billard, L., & Diday, E. (2000). Regression analysis of interval-valued data. *Data Analysis, Classification and Related Methods: Proceedings of the Seventh Conference of the International Federation of Classifications SOcietes*, (págs. 369-374).

Billard, L., & Diday, E. (2002). *Symbolic regression analysis.* Data Analysis, Classification, Clustering and Data Analysis, Proceedings of the Eighteenth Conference of the International Federation of Classifications Societies, 281-288.

Blanco-Fernández, A., Corral, N., & González-Rodríguez, G. (2011). *Estimation of a flexible simple linear model for interval data based on set arithmetic.* Computational Statistics & Data Analysis 55(9):2568–2578.

Bloomberg. (2023). *Portfolio Analytics*. Obtenido de Bloomberg: https://www.bloomberg.com/professional/product/portfolio-risk-analytics/

Bloomberg. (2023). *Servicio Bloomberg Professional*. Obtenido de Bloomberg: https://www.bloomberg.com/latam/#

Cosín López-Medel, Á. (Julio de 2019). Basic portfolio management based on beta values using interval-valued information. Madrid: Escuela Técnica Superior de Ingeniería (ICAI). Grado en Ingeniería Industrial.

DeMiguel, V., Garlappi, L., & Uppal, R. (2007). *Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy?* The Review of Financial Studies 22(5), 1915-1953.

Domingo, T. (2021). *Cambios de régimen en modelos de regresión con bases de datos de intervalos. Aplicación a mercados financieros.* Madrid: Escuela Técnica Superior de Ingenieria (ICAI). Trabajo de Fin de Grado.

Easy Portfolio. (10 de Junio de 2019). *What is a Diversified Portfolio?* Obtenido de Easy Portfolio: https://easyportfol.io/blog/diversified-portfolio/

Ellis, S. (26 de Julio de 2019). *Portfolio Risk Analysis in PORT*. Obtenido de YouTube: https://www.youtube.com/watch?v=dzPMzLb3ONQ

Expansión. (24 de Junio de 2016). *El Ibex 35 pierde un 12,3%, la mayor caída de su historia, tras el Brexit*. Obtenido de Expansión: https://www.expansion.com/mercados/cronica-bolsa/2016/06/24/576cc78e468aeb11758b4578.html

FactSet. (2023). *FactSet*. Obtenido de https://www.factset.com/

FactSet. (2023). *Portfolio Management*. Obtenido de FactSet: https://www.factset.com/solutions/portfolio-management-and-trading/portfolio-management

Fortuño, M. (22 de Abril de 2022). *Estas son las razones por las que a los profesionales de la inversión les fascina tanto el terminal Bloomberg*. Obtenido de El blog salmónn: https://www.elblogsalmon.com/productos-financieros/estas-razones-que-a-profesionales-inversion-les-fascina-terminal-bloomberg

Fraschini, E. M., & Axhausen, K. (Febrero de 2001). *Day on day dependencies in travel: First results using ARIMA modelling.* Obtenido de ResearchGate: https://www.researchgate.net/figure/Time-series-data-forecast-series-and-confidence-interval_fig1_228375739

García-Bárzana, M., Ramos-Guajardo, A. B., Colubi, A., & Kontoghiorghes, E. J. (2020). *Multiple linear regression models for random intervals: a set arithmetic approach.* Computational Statistics 35, 755–773.

Gioia, F. (2009). *Portfolio selection models with interval data.* Statistica Applicata, 21, 43-75.

Gioia, F. (2015). *Capital Asset Pricing Model Using Regression Methods for Interval-Valued Variables.* Reliable Computing 21:53-73.

Gioia, F., & Lauro, C. (2005). *Basic statistical methods for interval data.* Statistica Applicata, 17:75–104.

Giordiani, P. (2015). *Linear regression analysis for interval-valued data based on the Lasso technique.* Sapienza University of Rome.

Gitman, L. J., & Joehnk, M. D. (2008). *Fundamentals of investing, 10th edition.*

González-Rivera, G., Lee, T., & Mishra, S. (2008). Jumps in cross-sectional rank and expected returns: a mixture model. *Journal of Applied Econometrics*, 23(5), 585-606.

Hao, P., & Guo, J. (2017). *Constrained Center and Range Joint Model for Interval-valued symbolic data regression.* Computational Statistics and Data Analysis, 116-138.

Hawawini, G. (1983). *Why beta shifts as the return interval changes.* Financial analysts journal, 39(3), 73–77.

He, L., & Hu, C. (2007). *Impactsofintervalmeasurementonstudiesofeconomicvariability - evidence from stock market variability forecasting.* Journal of Risk Finance, 12, 489–507.

He, L., & Hu, C. (2008). *Impacts of interval computing on stock market variability forecasting.* Journal of Computational Economics, 33(3):263–276.

Huber, P. J. (1973). *Robust Regression: Asymptotics, Conjectures and Monte Carlo.* Ann. Statist. 1(5): 799-821.

Hvaerland, C. (2019). *Time Horizon Often Determines Risk Tolerance.* Wells Fargo Investment Institute.

Iberdrola. (s.f.). *ODS 9: Industria, Innovación e Infraestructura*. Obtenido de Iberdrola: https://www.iberdrola.com/sostenibilidad/comprometidos-objetivos-desarrollo-sostenible/ods-9-industria-innovacion-e-infraestructura#:~:text=El%20ODS%209%20se%20refiere,sostenible%20y%20fomentar%20la%20innovaci%C3%B3n.

IBM. (2023). *What is the k-nearest neighbors algorithm?* Obtenido de IBM: https://www.ibm.com/topics/knn

Kolm, P., Tütüncü, R., & Fabozzi, F. (2014). 60 years of portfolio optimization: Practical challenges and current trends. *European Journal of Operational Research*, 234(2), 356-371.

Kourtis, A. (2016). *The Sharpe ratio of estimated efficient portfolios.* Elsevier, 72-78.

Kull, A. (13 de May de 2005). *Sharing risk - An economic perspective.* Obtenido de https://www.actuaries.org/ASTIN/Colloquia/Zurich/Kull.pdf

Lasheras, E. (2020). *Trabajo Final de Máster. Optimización de carteras estocásticas en el marco de la solvencia II.* Obtenido de Barcelona: Facultad de Economía y Empresa Universidad de Barcelona, Máster en Ciencias Actuariales y Financieras: https://diposit.ub.edu/dspace/bitstream/2445/148957/1/TFM-CAF_Lasheras_2020.pdf

Li, C., Wang, D., & Jin, J. (2015). *Portfolio selection models based on interval numbers.* International Conference on Natural Computation.

Lima Neto, E., & De Carvalho, F. (2008). Centre and Range method for fitting a linear regression model to symbolic interval data. *Computational Statistics and Data Analysis*, 1500-1515.

López Mínguez, I. (2007). *Sistema de análisis de cambios en las posiciones de las empresas bursátiles en un ranking de rendimiento. Aplicaciones al IBEX-35.* Madrid: Escuela Técnica Superior de Ingeniería (ICAI). Proyecto de Fin de Carrera.

López, J. F. (1 de Marzo de 2020). *Riesgo específico*. Obtenido de Economipedia: https://economipedia.com/definiciones/riesgo-especifico.html

Maceda, P. (2018). *Basic analytic system for interval valued data.* Madrid: Escuela Técnica Superior de Ingeniería (ICAI). Trabajo Final de Máster.

Markowitz, H. M. (1952). *Portfolio Selection.* Journal of Finance 7:77–91.

Maté, C. (2012). El análisis de intervalos. Aplicaciones en ingeniería. *Anales de mecánica y electricidad*, Vol. 89, pp. 20-27.

Maté, C. (2021). *Combining Interval Time Series Forecasts. A First Step in a Long Way (Research Agenda).* Revista Colombiana de Estadística, 123-157.

Maté, C. (2021). *The iBoxplot: a visual tool for interval-valued data based on order statistics of lows and highs.*

Maté, C. (2022). Forecasting in FOREX the spot price interval of tomorrow with the same information of today. An analysis of the seven majors using a linear regression model based on interval arithmetic. *Knowledge-Based Systems*, 258, 109923.

Maté, C. (2022). *Forecasting in FOREX the spot price interval of tomorrow with the same information of today. An analysis of the seven majors using a linear regression model based on interval arithmetic.* Knowledge-Based Systems, 258, 109923.

Maté, C., Arroyo, J., & González-Rivera, G. (2016). *Forecasting with Interval and Histogram Data: Some Financial Applications.* Handbook of Empirical Economics and Finance, 247-279.

Matlab. (s.f.). *Develop apps using App Designer*. Obtenido de Matlab: https://nl.mathworks.com/help/matlab/app-designer.html#:~:text=App%20Designer%20es%20un%20entorno,componentes%20interactivos%20de%20la%20IU.

Neto, E. L., & Carvalho, F. D. (2010). *Constrained linear regression models for symbolic interval-valued variables.* Computational Statistics and Data Analysis, 54, 333-347.

Nowak, S. (7 de Diciembre de 2022). *Modelos predictivos ¿Qué son y para qué se usan?* Obtenido de Nuclio Digital School: https://nuclio.school/modelos-predictivos-que-son-y-usos/#:~:text=Los%20modelos%20predictivos%2C%20tambi%C3%A9n%20conocidos,el%20comportamiento%20ante%20un%20evento.&text=Y%20tienen%20el%20objetivo%20de%20predecir%20y%20pronosticar%20resultados%20probable

Nystuen, A., & Nyrnes, A. Q. (2022). *Trading based on the dual-beta model: Evidence from the Norwegian Stock Market.* Norwegian University of Science and Technology (NTNU). Master's thesis in Economics and Business administration.

Pacto Mundial. (s.f.). *ODS 17: Alianzas para lograr los objetivos*. Obtenido de Pacto Mundial: https://www.pactomundial.org/ods/17-alianzas-para-lograr-los-objetivos/#:~:text=El%20ODS%2017%20pretende%20fortalecer,en%20las%20esferas%20p%C3%BAblica%2C%20p%C3%BAblico%2D

Serban, F., Ferrara, M., & Costea, A. (2015). *Portfolio Optimization Using Interval Analysis.* Economic computation and economic cybernetics studies and research / Academy of Economic Studies 49(1), 139-154.

Sinova, B., Colubi, A., & González-Rodríguez, G. (2012). *Interval arithmetic-based simple linear regression between interval data: Discussion ans sensitivity analysis on the choice of the metric.* Information Sciences, 199, 109-124.

UNDP. (s.f.). *Los ODS en acción*. Obtenido de UNDP: https://www.undp.org/es/sustainable-development-goals

United Nations. (s.f.). *Objetivo 10: Reducir la desigualdad en y entre los países*. Obtenido de United Nations: https://www.un.org/sustainabledevelopment/es/inequality/

United Nations. (s.f.). *Objetivo 4: Garantizar una educación inclusiva, equitativa y de calidad y promover oportunidades de aprendizaje durante toda la vida para todos*. Obtenido de United Nations: https://www.un.org/sustainabledevelopment/es/education/

Wang, H., Guan, R., & Wu, J. (2012). *Linear regression of interval-valued data based on complete information in hypercubes.* Journal of Systems Science and Systems Engineering volume 21, pages422–442.

Wang, P., Gurmani, S., Tao, Z., Liu, J., & Chen, H. (2023). Interval time series forecasting: A systematic literature review. *Journal of Forecasting.*, 1-37.

Wikipedia. (31 de Marzo de 2022). *FactSet Research Systems* . Obtenido de Wikipedia: https://es.wikipedia.org/wiki/FactSet_Research_Systems

Wikipedia. (3 de Abril de 2022). *Línea del mercado de valores*. Obtenido de Wikipedia: https://es.wikipedia.org/wiki/L%C3%ADnea_del_mercado_de_valores

Wilipedia. (Noviembre de 2022). *Bloomberg L.P.* Obtenido de Wilipedia: https://es.wikipedia.org/wiki/Bloomberg_L.P.

Xu, W. (2010). *Symbolic data analysis: interval-valued data regression.* Atenas: Graduate Faculty of The University of Georgia.

Zhao, Q., Wang, H., & Wang, S. (2022). *Robust regression for interval-valued data based on midpoints and log-ranges.* Advances in Data Analysis and Classification.

ANEXO I: ALINEACIÓN DEL PROYECTO CON LOS ODS

Los Objetivos de Desarrollo Sostenible (ODS), también llamados Objetivos Globales, fueron fijados por las Naciones Unidas en 2015 como una iniciativa universal para erradicar la pobreza, proteger el medio ambiente y asegurar que todas las personas gocen de paz y prosperidad en el año 2030.

Los ODS son un conjunto de 17 objetivos interconectados entre sí, debido a que las acciones tomadas en un área afectan a los resultados en otras áreas, buscando un equilibrio entre la sostenibilidad social, económica y ambiental (UNDP, s.f.). Así, la creatividad, el conocimiento, la tecnología y los recursos financieros de toda la sociedad son necesarios para lograr los ODS en todos los contextos.

Con el fin de cumplir con estos objetivos, cada uno de los ciudadanos de los países miembros debería involucrarse de acuerdo con sus capacidades. Por ello, en el contexto de este proyecto se destaca su alineación con algunos de ellos.

El primero de ellos es el **ODS 4, Educación de calidad**, el cual busca garantizar una educación inclusiva, equitativa y de calidad, además de promover oportunidades de aprendizaje para las personas (United Nations, s.f.). Cada vez son más las personas que tratan de obtener rentabilidad de su dinero a partir de la inversión en mercados financieros. Sin embargo, muchas de ellas carecen del conocimiento necesario para efectuar esas inversiones y en muchos casos acaban teniendo el efecto contrario al deseado, llegando a perder grandes cantidades de dinero. Debido a que los primeros capítulos de este proyecto presentan una serie de conceptos teóricos de gran utilidad a la hora de invertir, estos podrían servir de base a nuevos inversores ser más conscientes de lo que están haciendo y entender cómo invertir de forma correcta. Por ello, el proyecto colabora con el ODS 4 al educar a nuevos inversores sobre lo que el mundo de la inversión y los mercados financieros requiere y cómo es su funcionamiento.

Otro de los objetivos con los que el proyecto colabora es el **ODS 9, Industria, Innovación e Infraestructura**, referido a la necesidad de construir infraestructuras resilientes, promover una industrialización sostenible y fomentar la innovación (Iberdrola, s.f.). Por medio del uso de la tecnología, la innovación y una serie de recursos y conceptos financieros, se han desarrollado un conjunto de funciones para formar una herramienta que ayuda a los inversores a tomar decisiones de manera sostenible en los mercados financieros a la hora de invertir en ellos.

Por otro lado, el **ODS 10, Reducción de las desigualdades**, es otro de los objetivos en los que este proyecto estaría colaborando, el cual busca reducir las desigualdades entre países y dentro de ellos (United Nations, s.f.). El proyecto podría ayudar a personas a tomar decisiones de inversión de manera informada, de forma que redujeran el riesgo de su inversión y les facilitara la obtención de una mayor rentabilidad de su capital. Con ello, los usuarios de la herramienta podrían tener una participación más segura a la hora de invertir, pudiendo obtener mayores ingresos y potencialmente reducir las desigualdades económicas entre personas y países, creando un sistema financiero más inclusivo y equilibrado.

Por último, el proyecto puede realizar una aportación al **ODS 17, Alianzas para lograr los objetivos**, el cual pretende movilizar e intercambiar conocimientos, capacidad técnica, tecnológica y recursos financieros para alcanzar la Agenda propuesta en todos los países, contribuyendo al desarrollo sostenible de forma conjunta (Pacto Mundial, s.f.). Así, todos los conocimientos obtenidos a lo largo del proyecto, así como la herramienta desarrollada y los resultados y conclusiones derivados de ella, son expuestos en esta memoria de manera que pueda ser ofrecida a cualquiera que tenga interés en ella para poder continuar con los conocimientos adquiridos. El proyecto podría ser utilizado por una serie de inversores para colaborar de forma conjunta con sus inversiones en la búsqueda de un desarrollo sostenible.

En conclusión, este proyecto se compromete con la consecución de los 17 ODS mediante su aportación a algunos de ellos. Así, la utilización del conocimiento expuesto en él y la herramienta desarrollada pueden servir de ayuda a los inversores en el aprendizaje de conceptos financieros, su aplicación a la hora de invertir para obtener rentabilidad y reducir desigualdades mediante las innovaciones que ofrece el sistema y las potenciales alianzas

entre inversores con objetivos sostenibles representan los principales medios mediante los cuales este proyecto demuestra su preocupación y compromiso con lograr unos objetivos que permitan desarrollar un mundo mejor.