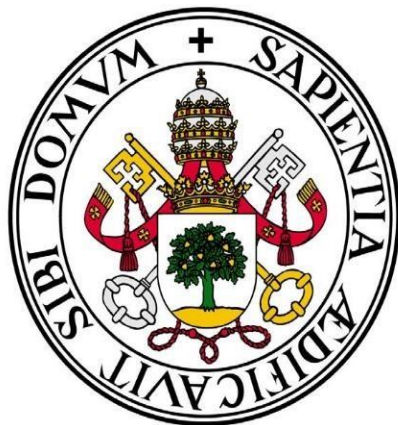




Universidad de Valladolid

Facultad de Ciencias Económicas y Empresariales

Trabajo de Fin de Grado

Grado en ADE

**Análisis del desempeño de Modelos de
Lenguaje Grande (LLM) en la predicción de
la rentabilidad y el valor liquidativo de los
fondos de inversión**

Felipe Meneses Soto

Tutelado por:

Gabriel de la Fuente Herrero

Valladolid a 10 de febrero de 2025

Resumen

Este Trabajo de Fin de Grado evalúa la eficacia de un Modelo de Lenguaje Grande para predecir la rentabilidad y el valor liquidativo de distintos fondos de inversión. Se plantean tres situaciones en las que el modelo recibe información de los fondos de manera incremental: primero con información básica de los fondos, segundo incluyendo información descriptiva de los fondos y finalmente con acceso a Internet. Para validar la precisión de las predicciones, se ha realizado un análisis estadístico de las respuestas del modelo. Los resultados muestran que, aunque se encuentran limitaciones en la predicción de la rentabilidad, la estimación del valor liquidativo produce resultados más cercanos al valor real.

Palabras clave: Fondos de inversión, Modelos de Lenguaje Grande, Predicción financiera, Rentabilidad, Valor liquidativo

Abstract

This Bachelor's Thesis evaluates the effectiveness of Large Language Models in predicting the return and net asset value of various investment funds. Three situations are considered, where the model receives fund information incrementally: first with basic fund data, then including descriptive information, and finally with Internet access. To validate the accuracy of the predictions, statistical tests have been conducted on the model's responses. The results show that, although there are limitations in the prediction of profitability, the estimation of net asset valuation returns values closer to the real value.

Key words: Investment funds, Large language models, Financial prediction, return, net asset value

ÍNDICE

1. INTRODUCCIÓN	4
2. FONDOS DE INVERSIÓN	5
2.1. Tipos de fondos de inversión	5
2.2. Evaluación de un fondo de inversión	7
2.2.1. <i>Rentabilidad</i>	7
2.2.2. <i>Riesgo</i>	8
3. INTELIGENCIA ARTIFICIAL	10
3.1. Modelos de Lenguaje de Gran Tamaño (LLMs)	11
3.2. El Mercado de la Inteligencia Artificial	12
4. METODOLOGÍA	13
4.1. Primer caso: información limitada de los fondos	14
4.2. Segundo caso: información ampliada de los fondos	16
4.3. Tercer caso: acceso a Internet	17
5. ANÁLISIS Y RESULTADOS	17
5.1. Primer caso: información limitada de los fondos	17
5.2. Segundo caso: información ampliada de los fondos	29
5.3. Tercer caso: acceso a Internet	33
6. CONCLUSIONES Y LÍNEAS FUTURAS DE INVESTIGACIÓN	37
REFERENCIAS	39
ANEXO A: TABLAS	41
ANEXO B: GRÁFICOS	43

1. INTRODUCCIÓN

La Inteligencia Artificial está presente cada vez más en las decisiones de la población. Este crecimiento del uso de la herramienta también se observa en el sector financiero. En el ámbito de los fondos de inversión la predicción de rentabilidades es un elemento clave en la toma de decisiones financieras, tanto para gestoras como para inversores individuales. En un contexto de riesgo, la capacidad de realizar mejores predicciones permite maximizar los beneficios del inversor. Es por ello por lo que los modelos de lenguaje de gran tamaño son una oportunidad fundamental en esta toma de decisiones. Se pueden observar diferentes aplicaciones de la inteligencia artificial en los servicios financieros (Fernández, 2019). No obstante, este trabajo se centra en la posibilidad de predecir los cambios de valor en los fondos de inversión tomando como base el trabajo de Kim et al. (2024), *Financial Statement Analysis with large Language Models*, que hacen una comparativa entre las predicciones de cambios de beneficios en acciones realizadas por expertos financieros y por modelos de lenguaje de gran tamaño.

Este Trabajo de Fin de Grado analiza los beneficios de incorporar nuevas tecnologías como instrumento de predicción financiera para la selección de fondos de inversión. El objetivo concreto es evaluar la capacidad de los modelos de lenguaje avanzados para predecir la rentabilidad y el valor liquidativo de fondos de inversión en el corto plazo. En particular, buscamos determinar el desempeño del modelo en la predicción de la rentabilidad y el valor liquidativo a un mes en comparación con los valores reales. También investigamos la capacidad del modelo de ordenar correctamente las predicciones. Por último, analizamos la sensibilidad de los resultados respecto a la información disponible por el modelo, comparando el éxito de las predicciones con información básica de cada fondo, con información detallada de cada fondo y con libre acceso a la información disponible en Internet.

El resto de este trabajo se estructura de la siguiente forma. En el siguiente apartado, se presenta una introducción a los fondos de inversión, qué son, qué tipos de fondos de inversión hay y cómo pueden ser evaluados. En el apartado tercero, explicamos qué es la Inteligencia Artificial y, especialmente, los modelos de lenguaje grande, también

analizando la importancia del uso de la IA en el sector financiero actual. En el apartado cuarto, se describe la metodología de nuestro análisis, dónde explicamos las diferentes situaciones a las que sometemos a nuestro modelo y qué herramientas hemos utilizado. En el apartado quinto, se presenta el análisis estadístico de los resultados. Por último, se enumeran las conclusiones de nuestro trabajo y posibles líneas futuras de investigación.

2. FONDOS DE INVERSIÓN

Las **instituciones de inversión colectiva** (IIC) son entidades que reúnen capital, ya sea a través de fondos, bienes o derechos del público, para gestionarlo e invertirlo de forma conjunta en bienes, valores y otros instrumentos, financieros o no. El objetivo de las instituciones de inversión colectiva es generar rendimientos a los inversores. Se pueden diferenciar entre Fondos y Sociedades de Inversión.

Los fondos de inversión (FI) carecen de personalidad jurídica y pertenecen al conjunto de inversores. Una Sociedad Gestora es la encargada de tomar las decisiones de inversión y representar a la entidad, careciendo de propiedad, y un Depositario es el encargado de custodiar los activos del fondo. Por otra parte, las Sociedades de Inversión (SI) son sociedades anónimas. En un fondo de inversión se realiza una aportación al patrimonio común y el inversor recibe una parte alícuota del patrimonio del fondo, llamado participación. En el caso de una sociedad de inversión, la aportación es realizada al capital social, adquiriendo una parte alícuota del capital social, acción (INVERCO, 2024).

2.1. Tipos de fondos de inversión

Existen diferentes formas de clasificar los fondos de inversión. No obstante, nos centraremos en la clasificación según la vocación inversora del fondo, es decir, del tipo de activos en los que el gestor decide invertir las aportaciones de los partícipes.

- Fondos monetarios: el gestor invierte en renta fija y la duración media de la

cartera es igual o inferior a 6 meses. Los monetarios a corto plazo tienen duración media de la cartera igual o inferior a 60 días. Existen tres tipos de mercados monetarios:

- Mercados monetarios interbancarios: formados por operaciones de crédito y préstamos como derivados a corto plazo, swaps de tipos de interés a corto plazo, depósitos interbancarios y otros activos con periodo de vencimiento inferior a una semana.
 - Mercados de activos empresariales: compuesto por pagarés de empresa.
 - Mercado monetario de deuda pública: deuda pública que emite en España el Tesoro Público. (Bankinter, 2024)
- Fondos de renta fija: solo invierte en renta fija, emisiones de deuda como bonos, obligaciones, Letras del Tesoro o pagarés, realizadas por gobiernos o empresas. El término “fija” hace referencia al plazo de vencimiento, su precio puede variar porque cotizan en los mercados (Santander, s.f.-a).
 - Fondos de renta variable: más del 75% de las aportaciones son invertidas en renta variable. Tienen un riesgo superior a los de renta fija.
 - Fondos mixtos. Podemos diferenciar los fondos de renta fija mixta, que invierten como máximo 30% en renta variable, y los fondos de renta variable mixta, donde el variable se sitúa entre el 30 y el 75% (Santander, s.f.-b).
 - Fondos total o parcialmente garantizados: aseguran total o parcialmente la inversión inicial de los partícipes. En alguna circunstancia aseguran también una determinada rentabilidad (Comisión Nacional del Mercado de Valores, s.f.-a).
 - Fondos de gestión pasiva: estos replican un índice bursátil o financiero. Incluyen a los fondos con un objetivo concreto de rentabilidad no garantizado y los fondos cotizados (ETF, *Exchange Traded Funds*).
 - Fondos de retorno absoluto: su objetivo es conseguir una rentabilidad positiva a medio plazo. Estos fondos invierten en renta variable, renta fija, divisas o multi activo. Suelen incluir los fondos de inversión libre (*Hedge Funds*) (BBVA, s.f.).
 - Fondos globales: el resto de los fondos que no encajan con las anteriores descripciones.

En función del tipo de fondo, el inversor puede asumir más o menos riesgo, en función de su mayor o menor aversión al riesgo. Los fondos de inversión libre tienen altos niveles de riesgo, pueden invertir en cualquier tipo de activo siguiendo la estrategia que se considere más adecuada. Estos fondos se dirigen a clientes profesionales o no profesionales con una inversión mínima de 100.000 euros (Comisión Nacional del Mercado de Valores, s.f.-b).

2.2. Evaluación de un fondo de inversión

Para analizar los fondos de inversión debemos tener en cuenta sus dos características principales: la rentabilidad y el riesgo.

2.2.1. Rentabilidad

La rentabilidad de un fondo se basa en el valor liquidativo de las participaciones del fondo. El valor liquidativo es el precio unitario de la participación en un momento determinado. Se calcula dividiendo el patrimonio del fondo de inversión entre el número de participaciones que hay circulando en ese instante. Generalmente, es la entidad gestora quien calcula el valor liquidativo diariamente¹. El valor liquidativo se puede calcular a partir de la siguiente ecuación:

$$VL_t = \frac{VP_t}{N_t} = \frac{VP_{t-1} + RN_t}{N_t}$$

Donde:

- VL_t : es el valor liquidativo de una participación en el momento t.
- VP_t : es el patrimonio del fondo de inversión en el momento t.
- N_t : es el número de participaciones en circulación en el momento t.
- VP_{t-1} : es el patrimonio del fondo de inversión en el momento t - 1.
- RN_t : rendimiento neto del fondo en el momento t.

Para calcular el valor del patrimonio se debe tener en cuenta el precio de mercado de los activos y valores mobiliarios que integran la cartera del fondo en el momento t, sustrayendo gastos y comisiones asociadas (Comisión Nacional del Mercado de Valores,

¹ No obstante, existen algunas excepciones al cálculo diario: los fondos de inversión libre, los fondos de inversión inmobiliaria y los fondos de carácter financiero.

s.f.-c).

Los inversores buscan maximizar la rentabilidad de sus inversiones. Para obtener mayor rentabilidad es necesario asumir mayor riesgo. La rentabilidad se puede calcular de diferentes formas. La rentabilidad bruta muestra la variación desde la suscripción de las participaciones hasta la fecha, midiendo el rendimiento real para el inversor (Esteban, 2025).

$$R = \frac{VL_t - VL_0}{VL_0}$$

Al no considerar ningún gasto ni comisión, la rentabilidad bruta no es de especial interés para los inversores. Por ello, se emplea la rentabilidad anual efectiva que incluye gastos y comisiones. Esta rentabilidad toma la forma de equivalencia financiera, donde si se considera una inversión con una duración de n años, G_0 (gastos asociados a las comisiones de suscripción) y G_n (gastos y comisiones asociados al reembolso), con una rentabilidad efectiva anual de i_e .

$$(VL_0 + G_0)(1 + i_e)^n = (VL_n - G_n)$$

Si el fondo de inversión distribuye parte de los beneficios entre los partícipes, debemos modificar la ecuación de la rentabilidad efectiva. Lo que nos lleva a la siguiente ecuación:

$$(VL_0 + G_0)(1 + i_e)^n = \sum_{k=1}^n (Q_k - G_k)(1 + i_e)^{n-k} + VL_n$$

En esta ecuación, Q_k representa el beneficio del año k -ésimo con unos gastos G_k (Aguado Peña, 2018).

2.2.2. Riesgo

Otro factor a tener en cuenta a la hora de evaluar un fondo de inversión es el riesgo inherente a dicha inversión. El riesgo es la suma de factores que pueden hacer que la rentabilidad de una inversión sea diferente a la esperada (ya sea superior o inferior). A mayor rentabilidad, mayor será el nivel riesgo que se ha de asumir.

Se puede diferenciar el riesgo sistemático y no sistemático. El sistemático, también conocido como riesgo de mercado, es aquel que no es específico a un activo

determinado, depende de factores genéricos que afectan a la evolución de precios en los mercados de valores. Este riesgo no puede ser eliminado mediante la simple diversificación de la cartera. Para aproximarlos se suele utilizar el coeficiente beta de su serie histórica. Por otra parte, el riesgo no sistemático es aquel específico al activo y que puede ser reducido mediante la diversificación de la cartera (Comisión Nacional del Mercado de Valores, s.f-c). Existen diferentes medidas para determinar el riesgo de un activo financiero: la desviación estándar, la beta del mercado y el *Value at Risk* (VaR).

Desviación estándar muestral: estadístico que mide la dispersión de las observaciones respecto al promedio de una serie. Nos permite medir la volatilidad de la rentabilidad de un fondo de inversión ya que determina cuánto varían los precios respecto a su media.

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x - \mu)^2}{n}}$$

Donde:

- n : número de observaciones de la muestra.
- x : valores observados.
- μ : media de los valores observados.

La desviación estándar permite calcular el rango probable. En una variable aleatoria que sigue una distribución normal, el 68% de los valores se encuentra dentro de una desviación estándar de la media, el 95% se encuentra dentro de dos desviaciones estándares de la media y el 99,7% dentro de tres. Esto quiere decir que el precio estará con un 68% de probabilidad dentro de una desviación estándar de la media, suponiendo que los precios siguen una distribución normal y gaussiana, aunque no siempre sea así. En el análisis financiero, se usa principalmente el 95% al abarcar la mayor parte de los valores posibles (Lock et al., 2021).

Por otra parte, la ratio de Sharpe indica la rentabilidad de una inversión teniendo en cuenta su riesgo. Este indicador nos muestra el exceso de rentabilidad que podemos obtener por cada unidad de volatilidad.

$$\text{Ratio de Sharpe} = \frac{R_i - R_f}{\sigma_i}$$

Donde:

- R_i : es la rentabilidad media esperada del activo financiero.
- R_f : es la rentabilidad del activo libre de riesgo.
- σ_i : es la desviación estándar del activo financiero.

Con este indicador podremos determinar qué fondo nos retorna una relación superior entre la rentabilidad obtenida y el riesgo asumido.

Otra medida del riesgo es el coeficiente beta, que mide la cantidad de riesgo sistemático y como variará el rendimiento de un activo con relación al mercado.

$$\beta = \rho_{im} * \frac{\sigma_i}{\sigma_m}$$

Donde:

- ρ_{im} : es el coeficiente de correlación entre la rentabilidad del activo y la rentabilidad del mercado.
- σ_i : es la desviación estándar de la rentabilidad del activo.
- σ_m : es la desviación estándar de la rentabilidad del mercado.

Con este indicador podemos determinar si el precio variará a la par que el mercado ($\beta = 1$), tendrá mayor volatilidad y variará más ($\beta > 1$) o tendrá menor volatilidad y menos riesgo sistemático ($\beta < 1$).

Por último, el VaR (*Value at Risk*) se emplea para medir la pérdida potencial máxima que se espera observar con un grado de confianza en un periodo de tiempo. Por ejemplo, si tenemos una cartera de inversiones con un VaR anual de un 3% de confianza de 2 millones. Hay una probabilidad del 3% de perder más de 2 millones y una probabilidad del 97% de que la pérdida sea inferior o igual (Chicu, 2020).

3. INTELIGENCIA ARTIFICIAL

En 1927 nació en Boston John McCarthy, matemático por el Instituto Tecnológico de California y doctorado en matemáticas por la Universidad de Princeton (The Editors of

Encyclopaedia Britannica, 2024). En 1955, McCarthy acuñó el término de Inteligencia Artificial (IA) definiéndolo como:

“La ciencia e ingeniería de la creación de máquinas inteligentes, especialmente programas informáticos inteligentes. Se relaciona con la tarea similar de usar ordenadores para comprender la inteligencia humana, pero la IA no tiene por qué limitarse a métodos biológicamente observables.” (McCarthy, 2004)

Desde entonces, la IA ha evolucionado y se ha desarrollado en todos los sectores, convirtiéndose en una herramienta clave para los negocios que quieren prosperar y mantener una ventaja competitiva frente a sus rivales. En el sector financiero, la inteligencia artificial se ha abierto paso con múltiples aplicaciones, entre ellas:

- Gestión de riesgos: uso de modelos para evaluar y gestionar los riesgos asociados a las inversiones.
- Asesoramiento financiero: desarrollo de herramientas de apoyo a la toma de decisiones para inversores individuales e institucionales.
- Control de blanqueo de capitales: detectando anomalías o patrones al analizar un mayor volumen de datos.
- Calificación crediticia: mejorando las evaluaciones y acelerando los plazos de concesión de créditos.
- Predicciones de beneficios: uso de modelos para analizar balances financieros y predecir cambios en los beneficios de empresas (Kim et al., 2024).

3.1. Modelos de Lenguaje Grandes (LLMs)

Los **modelos de lenguaje grande** (LLM) son modelos de aprendizaje profundo que se entrenan con grandes cantidades de datos (AWS, 2023). Los LLMs están basados en *Machine Learning* (ML). También pueden ser descritos como modelos matemáticos generativos de una distribución estadística de *tokens*. Son modelos generativos porque se pueden tomar muestras de ellos, es decir, podemos realizarles preguntas. Cuando se realiza una pregunta al modelo, el modelo busca cuál es la respuesta más probable dado una distribución estadística de palabras (Shanahan, 2024).

La principal ventaja de los modelos de lenguaje de gran tamaño es su versatilidad y flexibilidad para realizar tareas completamente diferentes. Mientras otros modelos como el *Machine Learning* usaban una tabla numérica para representar cada palabra, los LLMs usan vectores multidimensionales que permiten relacionar palabras con significados similares. De esta forma, los modelos pueden comprender el contexto de las palabras y frases con significados similares (AWS, 2023).

La manera en la que funciona estos modelos es mediante la ingesta de cantidades elevadas de datos. Gracias a esta gran cantidad de datos, los modelos son capaces de detectar patrones ocultos entre los datos y clasificar o predecir en función de ellos. La mayoría de los autores diferencia entre aprendizaje supervisado y no supervisado.

El aprendizaje supervisado muestra la habilidad de los algoritmos para generalizar conocimiento procedente de datos etiquetados, de esta manera el modelo puede emplearse para predecir nuevos casos sin etiquetar. Por otra parte, el aprendizaje no supervisado se basa en la agrupación de datos en *clusters* sobre datos que no han sido previamente clasificados. En este caso, los algoritmos deben reconocer las relaciones ocultas o características que les permitan agrupar los casos coherentemente. En el caso que una pequeña cantidad de datos esté categorizada nos encontramos ante un modelo “semi-supervisado” (Alloghani et al., 2020).

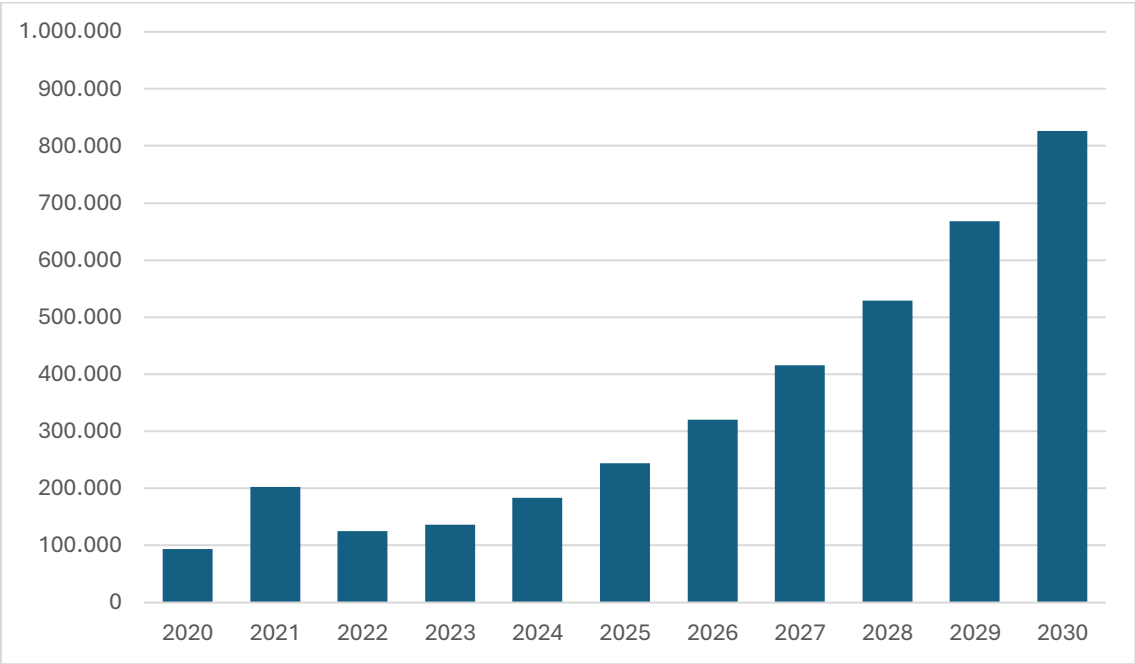
3.2. El Mercado de la Inteligencia Artificial

La cuarta revolución industrial ya está aquí. Los avances tecnológicos actuales y la velocidad y alcance que está teniendo esta revolución no tiene precedentes históricos. El ritmo de la evolución tiene carácter exponencial, no lineal. Los avances tecnológicos emergentes son cada vez mayores: la IA, la robótica, el internet de las cosas (IoT)... Esta revolución permite elevar los niveles de ingresos globales y mejorar la calidad de vida de la población.

El mercado de la Inteligencia Artificial no para de crecer. Como se muestra en el gráfico

1, el tamaño del mercado mundial de la IA es de 184.000 miles de millones de USD en 2024 y se espera que crezca hasta los 826.700 miles de millones de USD para 2030 (Statista, 2024). Por ello, el futuro de los mercados financieros pasa por la integración de la IA en la toma de decisiones de inversión.

Gráfico 1. Tamaño del mercado de la inteligencia artificial a nivel mundial de 2020 a 2030 (en miles de millones de dólares)



Fuente: Statista (2024)

El impacto de la Inteligencia Artificial es muy elevado. Los beneficios procedentes de esta tecnología han seducido tanto a consumidores como a empresas. Además, la disrupción generada también afecta a gobiernos, encargados de regular el uso de estas tecnologías. La revolución que supone la IA está redefiniendo el significado del ser humano, obligando a revisar los límites morales y éticos (Schwab, 2020).

4. METODOLOGÍA

Con el fin de analizar la capacidad de un modelo de lenguaje grande en la predicción de rentabilidades y valores liquidativos de los fondos de inversión, empleamos la versión

GPT-4o de ChatGPT. Esta versión permite realizar tareas más complejas con un mayor número de archivos. Los fondos de inversión utilizados el conjunto de los comercializados en España y cuya información se obtiene de INVERCO². La información solicitada al LLM consiste la predicción de las rentabilidades y valores liquidativos para el mes posterior a la información proporcionada, siendo la información proporcionada la disponible en octubre de 2024 y la predicción solicitada la correspondiente a noviembre de 2024. Nuestro análisis compara los resultados en tres casos diferentes según el volumen y tipo de la información suministrada. A continuación, se describen estos tres casos.

4.1. Primer caso: información limitada de los fondos

En este primer escenario se solicita al LLM las predicciones de rentabilidad y valor liquidativo de los fondos de inversión que INVERCO incluye en las seis categorías diferentes. La tabla 1 muestra la división de las categorías de fondos utilizadas para este análisis además de la cantidad de fondos de cada categoría y la proporción que supone frente al total de fondos. En esta tabla se han sido suprimido los fondos de inversión que presentan valores de liquidación iguales a cero.

Tabla 1. Fondos de inversión incluidos en el análisis

Tipo de fondo	Número	Porcentaje
Fondos de Inversión Libre	132	16.28
Fondos Índice	48	5.92
Fondos Monetarios	13	1.60
Fondos de Renta Fija a corto plazo	204	25.15
Fondos de Renta Fija a largo plazo	349	43.03
Fondos de Renta Variable	65	8.01
Total	811	100.00

Fuente: INVERCO (2024). Esta tabla muestra un desglose de los fondos de inversión incluidos en el análisis de este trabajo y el porcentaje que representa cada uno sobre el total.

La información proporcionada al LLM comprende, en primer lugar, el informe mensual

² INVERCO es la Asociación de Instituciones de Inversión Colectiva y Fondos de Pensiones, y agrupa a la práctica totalidad de las Instituciones de Inversión Colectiva españolas (Fondos y Sociedades de Inversión), a los Fondos de Pensiones españoles, y a las Instituciones de Inversión Colectiva extranjeras registradas en la CNMV a efectos de su comercialización en España, así como a diversos Miembros Asociados.

de INVERCO, en el que se incluyen las siguientes variables:

- | | |
|--------------------|---------------------------|
| - V. Liquidación | - Partícipes |
| - Rentabilidad mes | - Suscripciones |
| - Rentabilidad año | - Suscripciones año |
| - Rent. 1 año | - Reembolsos |
| - Rent. 3 años | - Reembolsos año |
| - Rent. 5 años | - Suscripciones netas |
| - Rent. 10 años | - Suscripciones netas año |
| - Rent. 15 años | - Patrimonio |
| - Rent. 20 años | - Var. patrimonio mes |
| - Rent. 25 años | - Var. patrimonio año |

Además de esta información, proporcionamos información respecto a las siguientes variables macroeconómicas.

- Tipo de interés
- Inflación
- Tipo de cambio EUR-USD
- Tasa de paro
- Precio del oro

También, proporcionamos una serie de informes económicos al LLM con análisis y proyecciones de entidades financieras internacionales como el Banco Central Europeo, la Reserva Federal Estadounidense, Kiel Institute, Cambridge Associates, la Comisión Europea, la OCDE, el World Bank, el Fondo Monetario Internacional. Además, se proporciona informes de bancos internacionales: BBVA, JPMorgan, Deutsche Bank, CaixaBank, BlackRock, Goldman Sachs o Lombard Odier.

Para el desarrollo de este primer caso, hemos anonimizado los datos para que el modelo LLM no pudiera adivinar cuál es cada uno de los fondos de inversión. Rechazamos la hipótesis de que sea capaz de reconocer el fondo de inversión o el año fiscal en base a los resultados de Kim et al. (2024). Este caso usa información general, utilizando información limitada de cada fondo y sin acceso a internet. El *prompt*³ utilizado le

³ *prompt* es una entrada o estímulo proporcionado a un modelo de lenguaje para generar una

proporciona al *GPT* las siguientes instrucciones específicas:

- Actuar como experto en fondos de inversión, analizando la información proporcionada en los archivos específicos contenidos en el proyecto.
- Realización de predicciones individuales para cada fondo.
- Evitar generalizaciones y basar el análisis en toda la información proporcionada, evitando basar el análisis exclusivamente en la rentabilidad y valor liquidativo del fondo del mes.
- Posibilidad de cambio de tendencias, si un fondo tiene rentabilidades positivas un mes, el mes siguiente puede disminuir su rentabilidad y viceversa.

Estas instrucciones han sido perfeccionadas para obtener los datos necesarios para nuestro análisis.

4.2. Segundo caso: información ampliada de los fondos

En este segundo caso, toda la información ya proporcionada en la situación anterior se amplía con la información del folleto del fondo disponible en la CNMV⁴. Nuestro análisis se centra en una de las categorías de fondos, los fondos de inversión libre, para la que se realiza un muestreo aleatorio estratificado según la rentabilidad de octubre, y en caso de coincidencia, el tamaño del fondo, resultando en una muestra de 36 fondos.

El folleto CNMV tiene información detallada sobre el nombre del fondo y su gestora, la política de inversión, el perfil de riesgo, información sobre el procedimiento de suscripción y reembolso, información comercial, comisiones y gastos, información sobre rentabilidad y otra información como el régimen de información periódica o la fiscalidad del fondo.

Este escenario busca determinar si un aumento de la información proporcionada sobre cada fondo conlleva una mejor predicción.

respuesta o acción específica.

⁴ CNMV: Comisión Nacional del Mercado de Valores, organismo encargado de la supervisión e inspección de los mercados de valores españoles y de la actividad de los que intervienen en estos.

4.3. Tercer caso: acceso a Internet

Partiendo de los fondos de inversión analizados en el segundo caso, en esta tercera etapa se vuelven a solicitar predicciones permitiendo al LLM utilizar Internet para completar su análisis con la información que considere necesaria. El objetivo de este escenario es determinar si, con acceso a toda la información disponible, el LLM puede realizar predicciones más acertadas. Buscamos determinar si el modelo es capaz de encontrar la información necesaria para mejorar su rendimiento, incorporando valores en tiempo real.

5. ANÁLISIS Y RESULTADOS

En esta sección vamos a analizar las predicciones realizadas por el modelo GPT-4o de las rentabilidades y valores liquidativos de los diferentes fondos de inversión para cada uno de los tres niveles de información.

5.1. Primer caso: información limitada de los fondos

Para este primer caso, se ha realizado el análisis estadístico tanto para la muestra total como segregando las categorías de los fondos. En la tabla 2 podemos observar los estadísticos descriptivos globales de nuestra muestra. Se ha tenido en cuenta la rentabilidad en octubre (ROCT), la rentabilidad real en noviembre (RNOV), la predicción de la rentabilidad (PREDICRENT), la diferencia absoluta entre la rentabilidad de noviembre y la predicha (DIFRENTAB), la diferencia relativa de la muestra (DIFRENTREL), el valor liquidativo del fondo en octubre (VOC), el valor liquidativo del fondo en noviembre (VNOV), la predicción del valor liquidativo (PREDICV), la diferencia absoluta del valor liquidativo de la muestra (DIFVAB) y la diferencia relativa del valor liquidativo (DIFVREL). En esta tabla podemos observar la media de la muestra, la desviación estándar, y los valores mínimo y máximo para cada variable analizada.

Tabla 2. Estadísticos descriptivos para el conjunto de los fondos de inversión comercializados en España a 31 de octubre de 2024

Variable	Obs	Media	Desviación estándar	Min	Max
ROCT	811	-0,313	1,059	-3,81	6,44
RNOV	811	0,884	2,71	-6,41	47,73
PREDICRENT	811	-0,407	1,688	-5	8,116
DIFRENTAB	811	1,873	2,576	0,001	42,367
DIFRENTREL	769	445,177	1126,856	0,512	16766,667
VOCT	811	491,1	5186,386	0	99753,06
VNOV	811	498,248	5256,542	0	101230,46
PREDICV	811	490,387	5175,467	0	99367,091
DIFVAB	811	9,991	106,967	0	2919,301
DIFVREL	810	5,115	17,752	0	100

Esta tabla muestra el número de observaciones, la media, la desviación estándar, el mínimo y máximo de cada variable para la muestra global de fondos.

Para la rentabilidad, la media del error absoluto es de 1,87 puntos porcentuales, mientras que la media del error relativo es del 445%. En el caso del valor liquidativo, el error absoluto es 9,99 puntos porcentuales frente a un error relativo del 5%. En las tablas 3 y 4 podemos ver la misma información estadística desglosada para los fondos de inversión libre y los fondos de renta fija a largo plazo, el resto de las categorías están en el Anexo A. Podemos destacar que la categoría de fondos con menor error relativo en la predicción de la rentabilidad son los fondos de renta fija a largo plazo con un error relativo igual a 214%.

Tabla 3. Estadísticos descriptivos para los fondos de inversión libre comercializados en España a 31 de octubre de 2024

Fondos Inversión libre	Obs	Media	Desviación estándar	Min	Max
ROCT	132	-0,447	1,614	-3,7	6,44
RNOV	132	1,577	6,035	-6,41	47,73
PREDICRENT	132	-0,521	2,085	-4,851	8,116
DIFRENTAB	132	2,987	5,136	0,013	42,367
DIFRENTREL	108	328,677	523,174	10,164	3199,928
VOCT	132	1940,567	12651,34	0	99753,06
VNOV	132	1975,082	12823,815	1,182	101230,46
PREDICV	132	1939,032	12623,967	0	99367,091
DIFVAB	132	42,188	261,741	0,002	2919,301
DIFVREL	132	17,743	35,076	0,072	100

Esta tabla muestra el número de observaciones, la media, la desviación estándar, el mínimo y máximo de cada variable para los fondos de inversión libre.

Tabla 4. Estadísticos descriptivos para los fondos de renta fija a largo plazo comercializados en España a 31 de octubre de 2024

Fondos renta fija l-p	Obs	Media	Desviación estándar	Min	Max
ROCT	349	-0,162	0,409	-1,48	0,58
RNOV	349	0,866	0,569	-1,38	3,23
PREDICRENT	349	-0,227	1,402	-4,093	3,778
DIFRENTAB	349	1,544	1,215	0,023	6,491
DIFRENTREL	338	213,776	204,737	1,603	1881,548
VOCT	349	166,911	706,17	0	8181,47
VNOV	349	169,258	712,894	0	8261,733
PREDICV	349	166,143	701,365	0	8115,7
DIFVAB	349	3,841	15,767	0,001	146,033
DIFVREL	348	2,952	11,794	0,023	100

Esta tabla muestra el número de observaciones, la media, la desviación estándar, el mínimo y máximo de cada variable para los fondos de renta fija a largo plazo.

A continuación, se ha realizado una serie de contrastes de igualdad de medias. Primero, comprobamos si las variables se distribuyen según una distribución normal utilizando el test de Shapiro-Wilk. El test de Shapiro-Wilk devuelve un estadístico W que mide la desviación entre los datos observados y una distribución normal esperada. Si el estadístico W está cerca de 1, los valores siguen una distribución similar a una normal (Universidad de Barcelona, 2025). En la tabla 5 podemos ver los resultados de esta prueba para toda la muestra, podemos descartar que las muestras sigan una distribución normal. Aunque la rentabilidad predicha se acerca más a 1, sigue estando considerablemente alejada.

Tabla 5. Test de Shapiro-Wilk para toda la muestra de fondos

Variable	Obs	W	V	z	Prob>z
ROCT	811	0,781	114,117	11,630	0,000
RNOV	811	0,299	365,289	14,486	0,000
PREDICRENT	811	0,986	7,223	4,854	0,000
VOCT	811	0,059	490,335	15,209	0,000
VNOV	811	0,059	490,605	15,210	0,000
PREDICV	811	0,059	490,288	15,209	0,000

La tabla muestra el valor del estadístico W, que mide el grado de ajuste de los datos a una distribución normal. La hipótesis nula a contrastar es que los datos siguen una distribución normal.

Aunque hemos concluido que los datos de forma general no siguen una distribución normal. Al realizar este mismo análisis de forma segregada para cada fondo podemos determinar que la predicción de la rentabilidad para alguno de ellos sí sigue una distribución normal. En la tabla 6 podemos observar los diferentes valores que resulta de hacer el test de Shapiro-Wilk para cada categoría de fondo. Para los fondos de inversión libre rechazamos la hipótesis de que las predicciones de rentabilidad siguen una distribución normal. No obstante, para el resto de las categorías de fondos no descartamos esta hipótesis y podemos considerar que la rentabilidad predicha sigue una distribución normal. El resto de las variables que han sido analizadas no siguen una distribución normal.

Tabla 6. Test de Shapiro-Wilk para la variable PREDICRENT en función de la categoría del fondo

PREDICRENT	Obs	W	V	z	Prob>z
FIL	132	0,953	4,956	3,604	0,000
IND	48	0,975	1,123	0,246	0,403
MON	13	0,916	1,474	0,760	0,224
RFCP	204	0,997	0,428	-1,955	0,975
RFLP	349	0,998	0,487	-1,700	0,955
RV	65	0,980	1,155	0,313	0,377

La tabla muestra el valor del estadístico W, que mide el grado de ajuste de los datos a una distribución normal. La hipótesis nula a contrastar es que los datos siguen una distribución normal.

Los gráficos 2 a 4 representan el histograma de nuestra muestra y una distribución normal. Podemos observar como parece que las estimaciones para la rentabilidad hechas por el GPT siguen una distribución normal, lo que nos podría llevar a creer que está haciendo predicciones aleatoriamente. Por otra parte, los análisis no nos hacen pensar que ocurra lo mismo para la predicción del valor liquidativo.

Gráfico 2. Distribución de la rentabilidad de octubre frente a una distribución normal

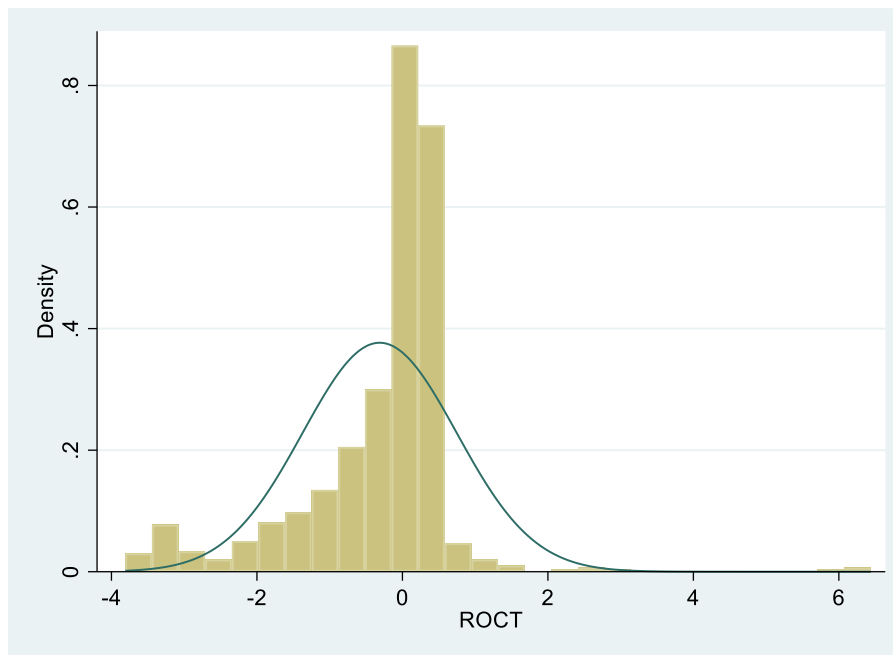


Gráfico 3 Distribución de la rentabilidad de noviembre frente a una distribución normal

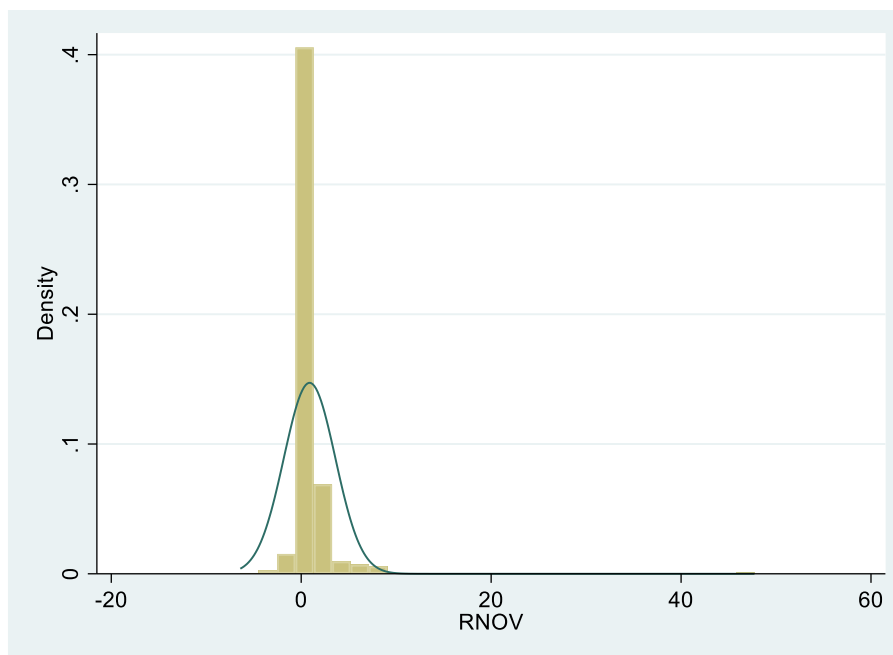
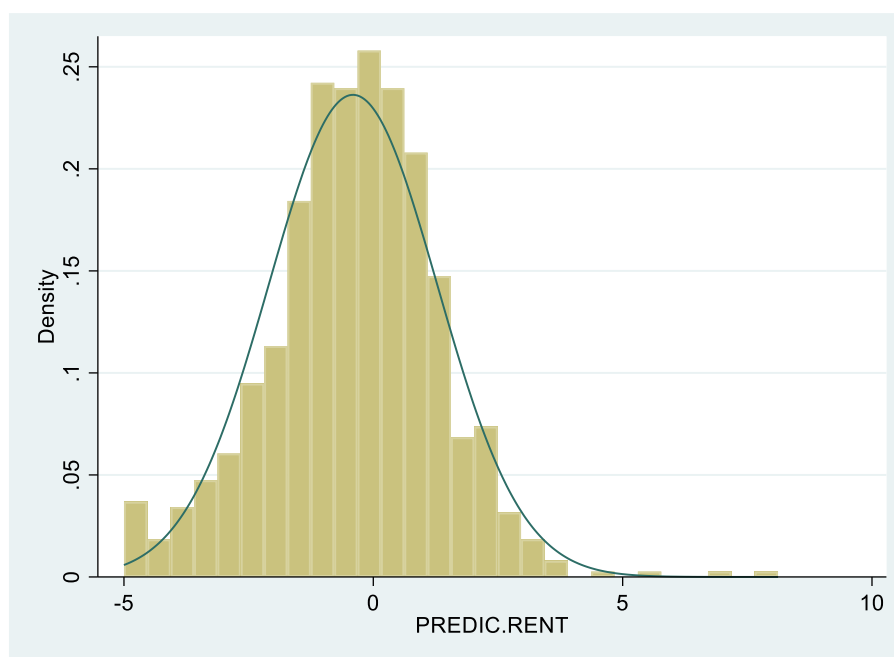


Gráfico 4 Distribución de la predicción de rentabilidad frente a una distribución normal



El siguiente test que hemos realizado es el de Wilcoxon Signed-Rank, que nos permite comparar la rentabilidad actual (RNOV) con la rentabilidad predicha (PREDICRENT). El resultado de estos contrastes, realizados tanto para toda la muestra como para cada una de las categorías, es que existen diferencias estadísticamente significativas entre ambas variables. Solo podemos considerar que no hay diferencias significativas para los fondos monetarios. No obstante, esta categoría de fondos cuenta tan solo con 13 observaciones, lo cual limita la fiabilidad del contraste estadístico. Los resultados de este test se encuentran en la tabla 7.

Tabla 7. Test de Wilcoxon Signed-Rank para la variable rentabilidad en el primer caso

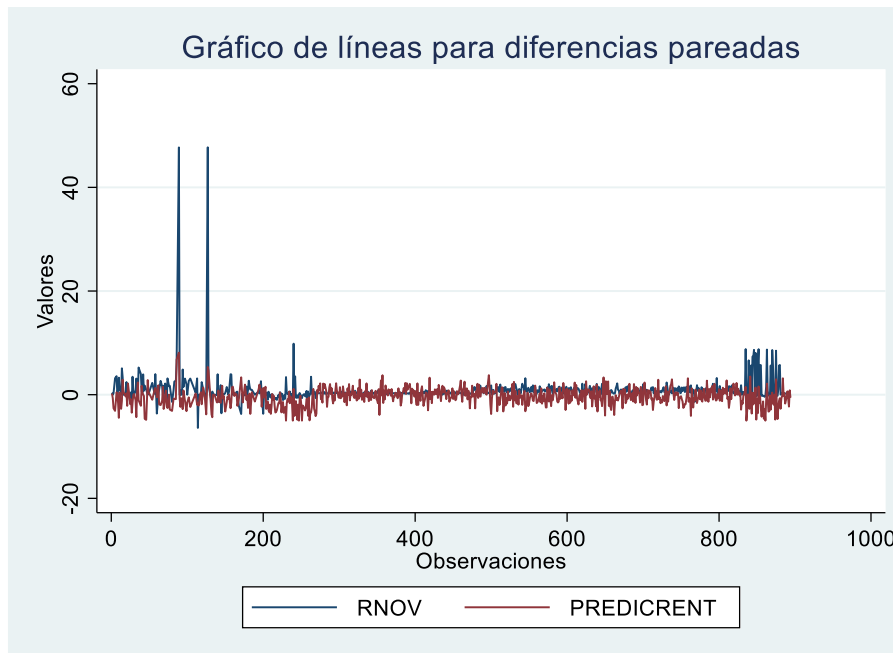
Tipo	Signo	Obs	Suma rangos	Esperado	Z	p-valor
General	Positivo	104	7215	4389	13,404	0
	Negativo	28	1563	4389		
	Cero	0	0	0		
	Todos	132	8778	8778		
FIL	Positivo	93	6932	4389	5,776	0
	Negativo	39	1846	4389		
	Cero	0	0	0		
	Todos	132	8778	8778		
IND	Positivo	45	1146	588	5,723	0
	Negativo	3	30	588		

	Cero	0	0	0		
	Todos	48	1176	1176		
MON	Positivo	9	69	45,500		
	Negativo	4	22	45,500	1,642	0,1005
	Cero	0	0	0		
	Todos	13	91	91		
RFCP	Positivo	113	11970	10455		
	Negativo	91	8940	10455	1,795	0,0727
	Cero	0	0	0		
	Todos	204	20910	20910		
RFLP	Positivo	252	50539	30537,500		
	Negativo	97	10536	30537,500	10,604	0
	Cero	0	0	0		
	Todos	349	61075	61075		
RV	Positivo	58	2086	1072,500		
	Negativo	7	59	1072,500	6,623	0
	Cero	0	0	0		
	Todos	65	2145	2145		

La tabla muestra los resultados del test de Wilcoxon Signed-Rank aplicado a las diferencias entre las rentabilidades predichas por el modelo y las rentabilidades reales observadas. El objetivo es evaluar si estas diferencias son estadísticamente significativas, ayudando a determinar si el modelo ofrece estimaciones coherentes o si las predicciones se aproximan a un comportamiento aleatorio.

Para visualizar estas diferencias hemos representado en el gráfico 5 el gráfico de líneas para diferencias pareadas de la rentabilidad de noviembre y la predicha. En este gráfico podemos observar las 811 muestras de nuestro análisis y cómo hay valores tanto positivos como negativos. La línea azul, que representa la rentabilidad real, muestra varios picos significativos entre 0 y 200, lo que muestra valores atípicos que se alejan mucho de los valores predichos. La línea roja, sobre las predicciones, es mucho más consistente y cercana a 0, lo que nos indica que el modelo tiende a hacer predicciones más conservadoras y no es capaz de capturar los valores atípicos.

Gráfico 5 Gráfico de líneas pareadas para toda la muestra



Si tomamos las categorías, los gráficos 6 y 7 muestran el gráfico de líneas para diferencias pareadas para los fondos de inversión libre y los fondos de renta fija a largo plazo, respectivamente. Estas categorías son las que muestran valores más diferenciados. Es en los fondos de inversión libre donde observamos valores reales más atípicos y donde la rentabilidad predicha para los fondos de renta fija a largo plazo parece estar más alejada de la rentabilidad real. En el Anexo B podemos observar el gráfico de líneas pareadas para el resto de las categorías.

Gráfico 6 Gráfico de líneas para diferencias pareadas para los FIL

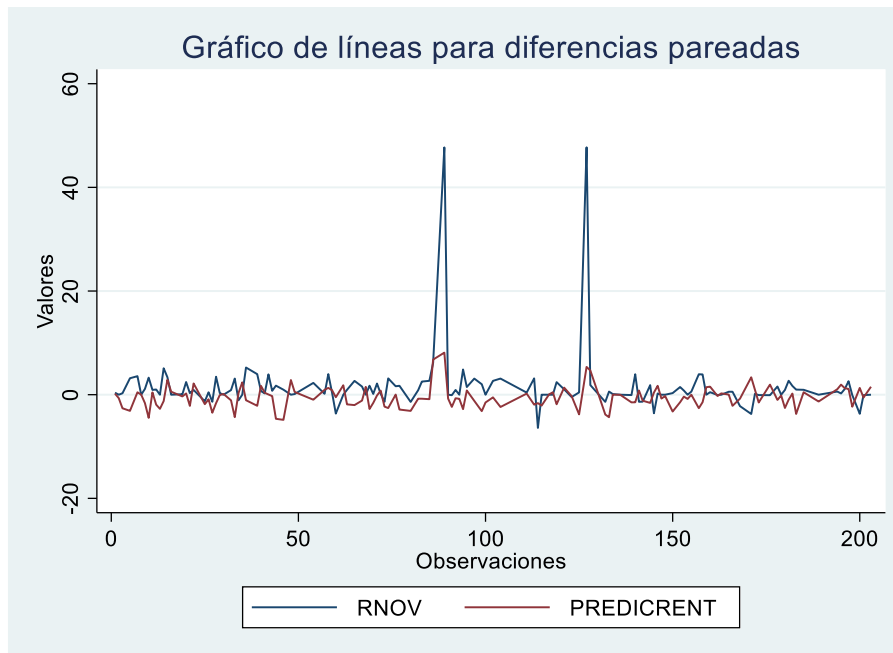
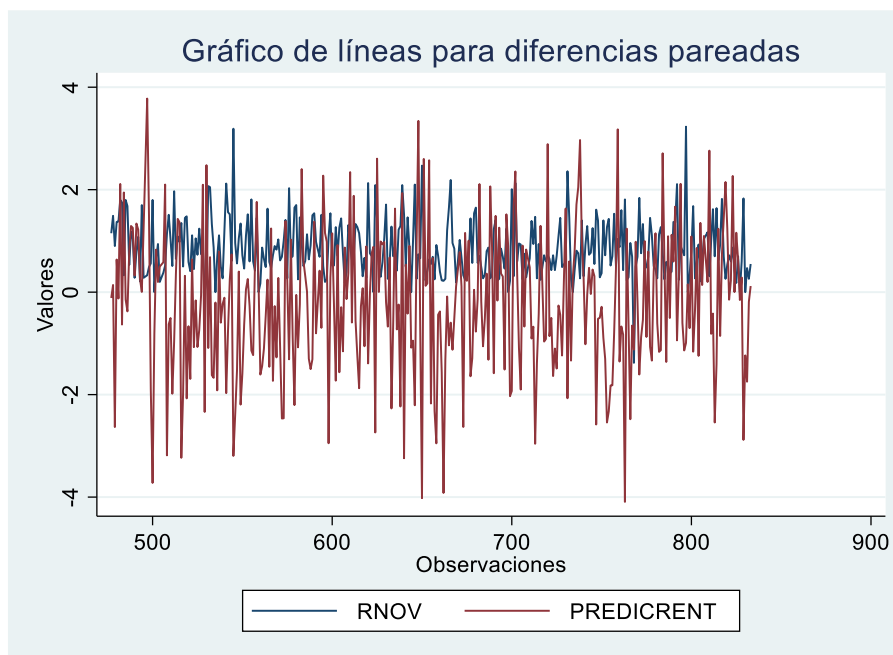
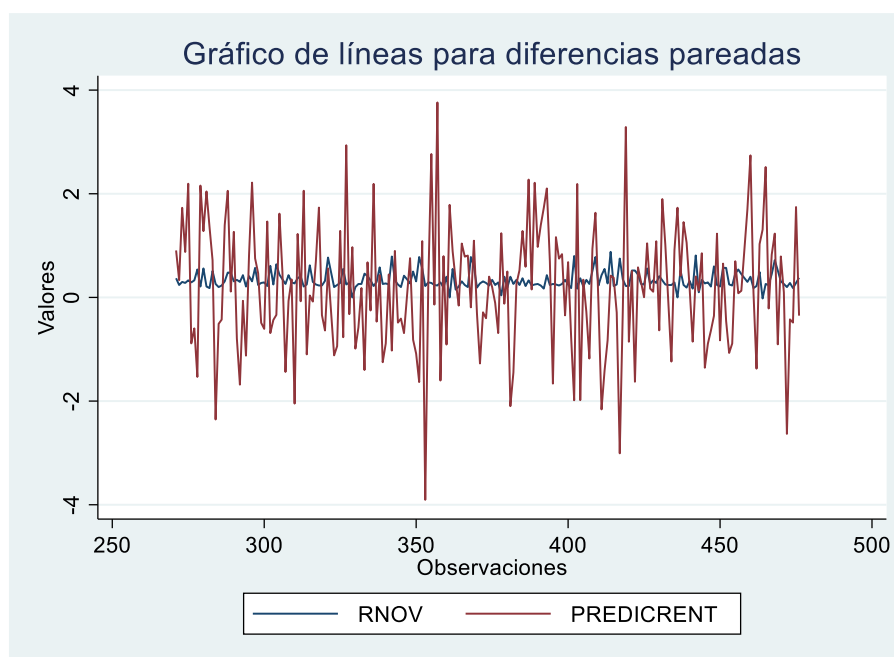


Gráfico 7 Gráfico de líneas para diferencias pareadas para los fondos RF I-p



En el caso de los fondos de renta fija a corto plazo podemos observar cómo los valores predichos son mucho más variantes que los valores actuales, visible en el gráfico 8.

Gráfico 8 Gráfico de líneas para diferencias pareadas para fondos RF c-p



Los resultados del test de Wilcoxon Signed-Rank para el valor liquidativo se encuentran en la tabla 8.

Tabla 8. Test de Wilcoxon Signed-Rank para la variable valor liquidativo en el primer caso

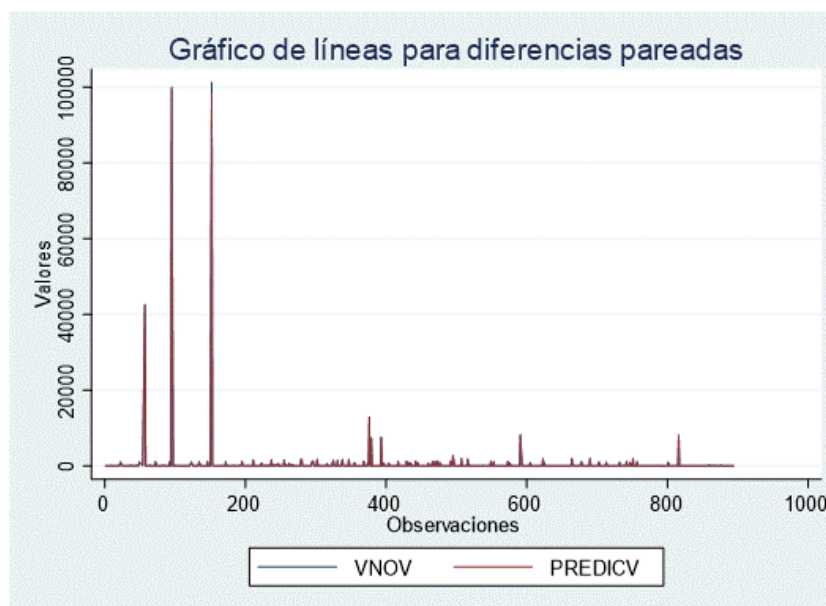
Tipo	Signo	Obs	Suma rangos	Esperado	Z	p-valor
General	Positivo	582	254081	164633	13,404	0
	Negativo	229	75185	164633		
	Cero	0	0	0		
	Todos	811	329266	329266		
FIL	Positivo	104	7215	4389	6,419	0
	Negativo	28	1563	4389		
	Cero	0	0	0		
	Todos	132	8778	8778		
IND	Positivo	45	1155	588	5,815	0
	Negativo	3	21	588		
	Cero	0	0	0		
	Todos	48	1176	1176		
MON	Positivo	9	68	45,500	1,572	0,1159
	Negativo	4	23	45,500		
	Cero	0	0	0		
	Todos	13	91	91		
RFCP	Positivo	113	11709	10455	1,485	0,1374
	Negativo	91	9201	10455		
	Cero	0	0	0		

	Todos	204	20910	20910		
RFLP	Positivo	252	47582	30537,500		
	Negativo	97	13493	30537,500	9,037	0
	Cero	0	0	0		
	Todos	349	61075	61075		
RV	Positivo	59	2096	1072,500		
	Negativo	6	49	1072,500	6,689	0
	Cero	0	0	0		
	Todos	65	2145	2145		

La tabla muestra los resultados del test de Wilcoxon Signed-Rank aplicado a las diferencias entre los valores liquidativos predichos por el modelo y los valores liquidativos reales observados. El objetivo es evaluar si estas diferencias son estadísticamente significativas, ayudando a determinar si el modelo ofrece estimaciones coherentes o si las predicciones se aproximan a un comportamiento aleatorio.

Si observamos el gráfico de líneas para diferencias pareadas para el valor liquidativo en el gráfico 9 podemos observar una similitud mucho mayor, con la ausencia de valores evidentemente diferenciados como ocurría en el caso de la predicción de la rentabilidad. Se incluyen en el Anexo B estos análisis gráficos de las categorías segregadas.

Gráfico 9 Gráfico de líneas pareadas para la muestra total del valor liquidativo



Por último, hemos realizado un test sobre la muestra para comprobar si el GPT ordena correctamente las muestras, aunque falle en la estimación de la rentabilidad concreta. Para eso realizamos un test de Kendall, ya que es el adecuado para variables que no están distribuidas según una normal.

En la tabla 9 podemos observar cómo las categorías de fondos índice y fondos monetarios están correctamente ordenados. También podemos observar que los fondos de renta fija a largo plazo han sido ordenados inversamente. De forma general y para el resto de las categorías no podemos deducir que el orden de las predicciones esté relacionado con el orden de las rentabilidades reales.

Tabla 9. Test de Kendall para la muestra de fondos de la variable rentabilidad

Categoría	Obs	Tau-a	Tau-b	Puntuación de Kendall	Prob>z
Global	811	-0,0018	-0,0018	-588	0,9393
FIL	132	-0,0766	-0,0779	-662	0,1921
IND	48	0,4592	0,4609	518	0
MON	13	0,4615	0,4738	36	0,0314
RFCP	204	-0,0494	-0,0501	-1022	0,2945
RFLP	349	-0,1414	-0,1419	-8585	0,0001
RV	65	0,0635	0,0637	132	0,4582

La tabla muestra los resultados de aplicar el test de Kendall para determinar si las predicciones de rentabilidad han sido ordenadas correctamente. Los valores que se presentan son el número de observaciones de cada categoría, la puntuación de Kendall y la probabilidad para determinar si es significativo.

En el caso de la variable de valor liquidativo podemos observar en la tabla 10 que es más efectiva la predicción de esta variable.

Tabla 10. Test de Kendall para la muestra de fondos de la variable valor liquidativo

Categoría	Obs	Tau-a	Tau-b	Puntuación de Kendall	Prob>z
Global	811	0,8814	0,8819	289497	0
FIL	132	0,6388	0,6462	5523	0
IND	48	0,9716	0,9716	1096	0
MON	13	0,9744	0,9744	76	0
RFCP	204	0,9519	0,9520	19710	0
RFLP	349	0,9306	0,9307	56513	0
RV	65	0,9760	0,9760	2030	0

La tabla muestra los resultados de aplicar el test de Kendall para determinar si las predicciones de valor liquidativo han sido ordenadas correctamente. Los valores que se presentan son el número de observaciones de cada categoría, la significancia de la correlación.

5.2. Segundo caso: información ampliada de los fondos

En esta sección vamos a analizar cómo ha variado la efectividad de las predicciones al introducir información detallada de cada fondo. En la tabla 11 podemos observar la muestra a analizar. Se observan dos variables: la rentabilidad real (rreal) y la rentabilidad predicha sin Internet (rpredsin) y el valor liquidativo real (vreal) y el valor liquidativo predicho sin Internet (vpredsin).

Tabla 11. Estadísticos de la muestra para el segundo caso

Variable	Obs	Media	Desviación estándar	Min	Max
rreal	36	0,061	1,178	-3,08	1,94
rpredsin	36	0,519	0,936	-1,5	2,1
vreal	36	3920,787	17205,983	1,255	96643,554
vpredsin	36	3947,363	17344,136	1,279	97510,341

Esta tabla muestra el número de observaciones, la media, la desviación estándar, el mínimo y máximo de cada variable para los fondos de inversión libre elegidos para el segundo caso.

De forma análoga al análisis para la parte anterior, hemos hecho un contraste de medias para determinar si nuestras muestras siguen una distribución normal o no. Para eso, hemos vuelto a hacer un test de Shapiro-Wilk sobre nuestra nueva población. En la tabla 12 podemos observar los resultados. Podemos concluir que la hipótesis nula no se rechaza en el caso de la rentabilidad predicha sin acceso a Internet, es decir, esta variable sí sigue una distribución normal. Como en el primer caso, podemos pensar que el GPT está haciendo predicciones aleatorias en este caso.

Tabla 12. Test de Shapiro-Wilk para distribución normal de la segunda parte

Variable	Obs	W	V	z	Prob>z
rreal	36	0,939	2,214	1,662	0,048
rpredsin	36	0,952	1,752	1,172	0,121
vreal	36	0,243	27,588	6,937	0,000
vpredsin	36	0,243	27,602	6,938	0,000

La tabla muestra el valor del estadístico W, que mide el grado de ajuste de los datos a una distribución normal. La hipótesis nula a contrastar es que los datos siguen una distribución normal.

Si representamos gráficamente la distribución de la rentabilidad real y la predicha sin uso de Internet frente a una distribución normal obtenemos los siguientes gráficos. En

los gráficos 10 y 11 la línea azul es nuestra variable y la línea roja representa la distribución normal.

Gráfico 10 Distribución de la rentabilidad real y una distribución normal

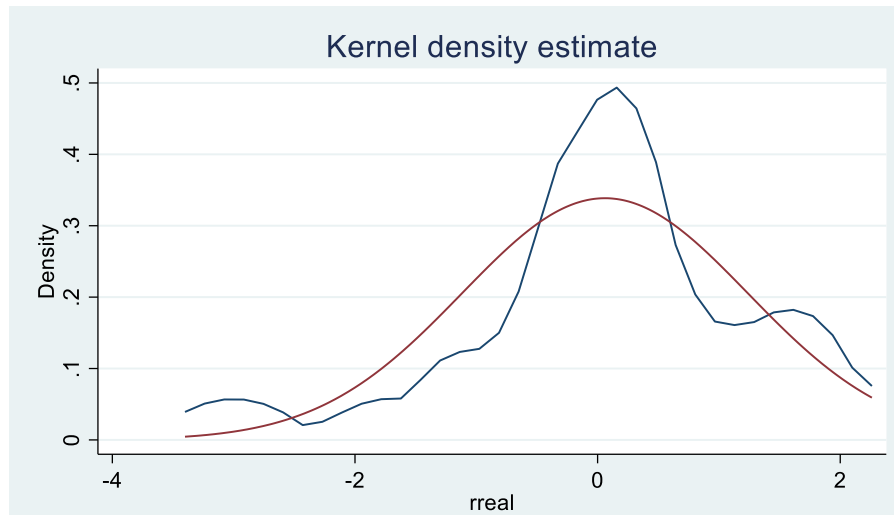
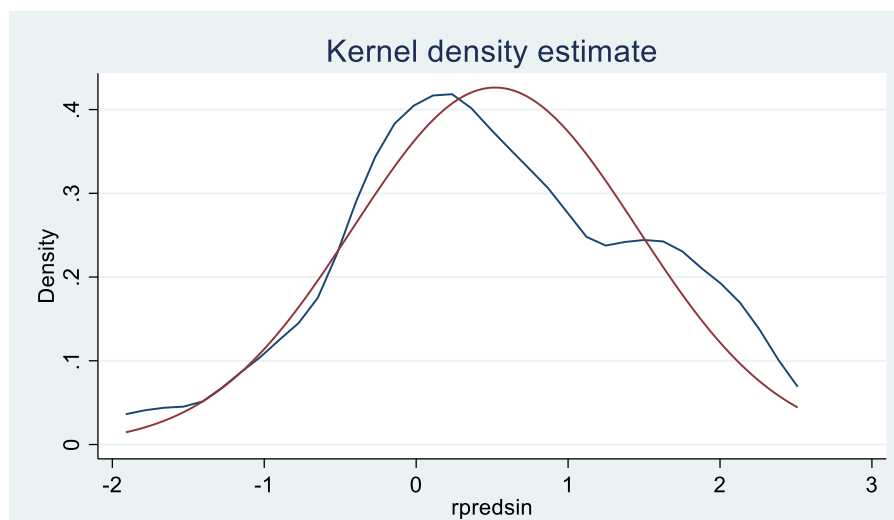


Gráfico 11 Distribución de la rentabilidad predicha sin Internet y una distribución normal



Gráficamente podemos observar las mismas conclusiones que hemos obtenido del test de Shapiro-Wilk, la situación que más se parece a una distribución normal es en la predicción de la rentabilidad sin acceso a Internet. En el caso del valor liquidativo podemos observar en los gráficos 12 y 13 que no se asemejan a la distribución normal. Para esta variable hemos representado el histograma de la muestra.

Gráfico 12 Distribución del valor liquidativo real y una distribución normal

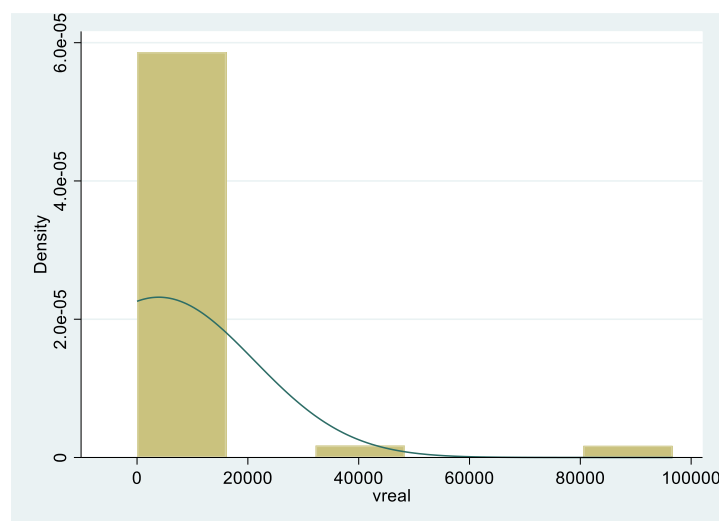
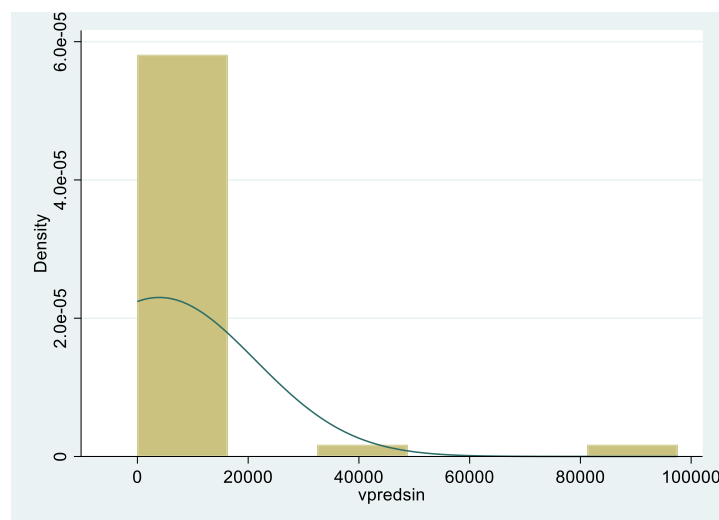


Gráfico 13 Distribución del valor liquidativo predicho sin Internet y una distribución normal



Consecutivamente hemos realizado el test de Wilcoxon Signed-Rank para nuestra población, cuyos resultados se encuentran en las tablas 13 y 14. Este análisis estadístico nos indica que las muestras de rentabilidad real y la rentabilidad predicha sin uso de Internet son muy similares, esto se puede visualizar en el gráfico 14.

Tabla 13. Test de Wilcoxon Signed-Rank para la rentabilidad predicha en el segundo caso

Tipo	Signo	Obs	Suma rangos	Esperado	Z	p-valor
General	Positivo	0	0	333	-5,232	0
	Negativo	36	666	333		

Cero	0	0	0
Todos	36	666	666

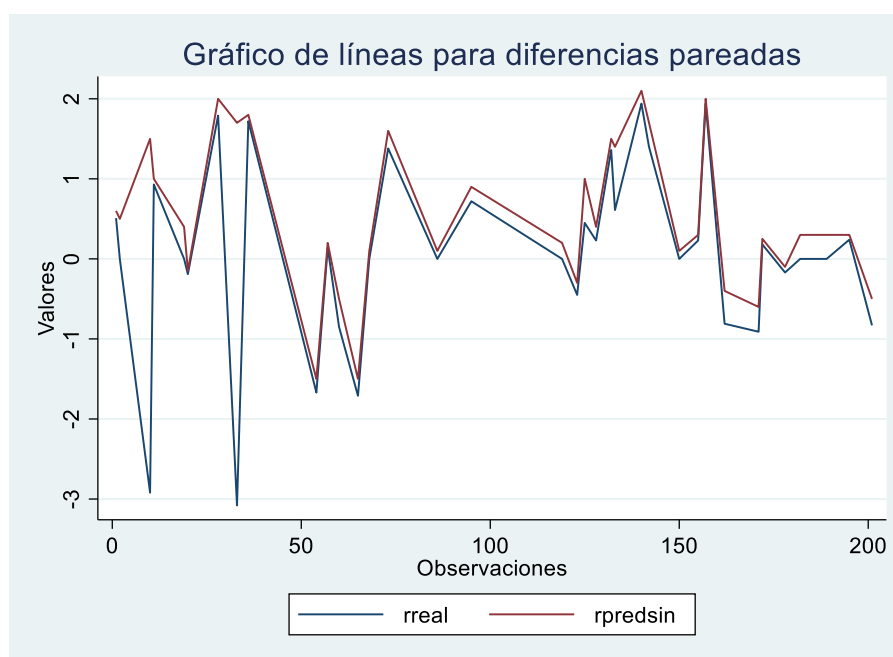
La tabla muestra los resultados del test de Wilcoxon Signed-Rank aplicado a las diferencias entre las rentabilidades de los fondos de inversión libre elegidos para el segundo caso predichas por el modelo y las rentabilidades reales observadas. El objetivo es evaluar si estas diferencias son estadísticamente significativas, ayudando a determinar si el modelo ofrece estimaciones coherentes o si las predicciones se aproximan a un comportamiento aleatorio.

Tabla 14. Test de Wilcoxon Signed-Rank para el valor liquidativo predicho en el segundo caso

Tipo	Signo	Obs	Suma rangos	Esperado	Z	p-valor
General	Positivo	9	166	333	-2,624	0,0087
	Negativo	27	500	333		
	Cero	0	0	0		
	Todos	36	666	666		

La tabla muestra los resultados del test de Wilcoxon Signed-Rank aplicado a las diferencias entre los valores liquidativos de los fondos de inversión libre elegidos para el segundo caso predichas por el modelo y los valores liquidativos reales observadas. El objetivo es evaluar si estas diferencias son estadísticamente significativas, ayudando a determinar si el modelo ofrece estimaciones coherentes o si las predicciones se aproximan a un comportamiento aleatorio.

Gráfico 14 Gráfico de líneas para diferencias pareadas entre la rentabilidad real y la predicha sin Internet



Nuestro análisis estadístico termina realizando el test de Kendall para determinar si nuestras muestras se han ordenado correctamente. En la tabla 15 podemos observar como la rentabilidad real y la predicha sin utilizar Internet se ordenan igual, el valor liquidativo no se ordena de la misma forma. También, podemos observar cómo el valor liquidativo se ordena igual para el valor real como el predicho en ambos casos.

Tabla 15. Test de Kendall entre las variables del segundo del análisis

Categoría	Obs	Tau-a	Tau-b	Puntuación de Kendall	Prob>z
rreal – rpredsin	36	0,6937	0,7206	437	0
vreal – vpredsin	36	0,9905	0,9921	624	0

La tabla muestra los resultados de aplicar el test de Kendall para determinar si las predicciones han sido ordenadas correctamente. Los valores que se presentan son el número de observaciones de cada categoría, la puntuación de Kendall y la significancia de la correlación.

5.3. Tercer caso: acceso a Internet

En esta sección vamos a analizar cómo ha variado la efectividad de las predicciones al introducir al permitir que nuestro LLM tenga acceso a Internet. En la tabla 16 podemos observar la muestra a analizar en este caso. Se observan dos variables: la rentabilidad real (rreal) y la rentabilidad predicha con Internet (rpredcon) y el valor liquidativo real (vreal) y el valor liquidativo predicho con Internet (vpredcon).

Tabla 16. Estadísticos de la muestra para el tercer caso

Variable	Obs	Media	Desviación estándar	Min	Max
rreal	36	0,061	1,178	-3,08	1,94
rpredcon	36	0,167	0,197	-0,6	0,55
vreal	36	3920,787	17205,983	1,255	96643,554
vpredcon	36	3922,762	17209,452	1,258	96643,554

Esta tabla muestra el número de observaciones, la media, la desviación estándar, el mínimo y máximo de cada variable para los fondos de inversión libre elegidos para el tercer caso.

Siguiendo el mismo procedimiento que en la sección anterior, realizamos un test de Shapiro-Wilk para evaluar si las muestras de rentabilidad presentan una distribución normal. En la tabla 17 se pueden observar los resultados obtenidos. Podemos concluir que la hipótesis nula se puede rechazar en el caso de la rentabilidad predicha con acceso a Internet, es decir, en este caso la variable no sigue una distribución normal. Como en los anteriores casos, podemos pensar que el GPT está haciendo predicciones aleatorias en este caso.

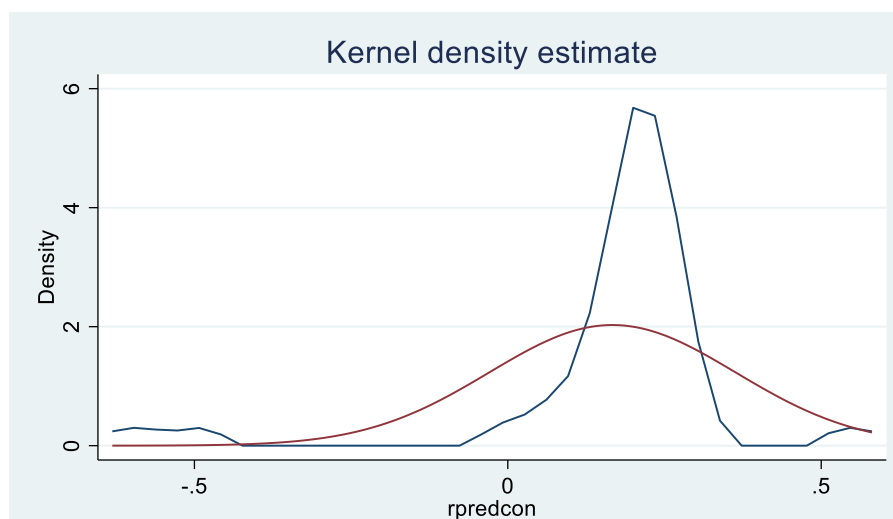
Tabla 17. Test de Shapiro-Wilk para distribución normal del tercer caso

Variable	Obs	W	V	z	Prob>z
rreal	36	0,939	2,214	1,662	0,048
rpredcon	36	0,628	13,560	5,452	0,000
vreal	36	0,243	27,588	6,937	0,000
vpredcon	36	0,243	27,587	6,937	0,000

La tabla muestra el valor del estadístico W, que mide el grado de ajuste de los datos a una distribución normal. La hipótesis nula a contrastar es que los datos siguen una distribución normal.

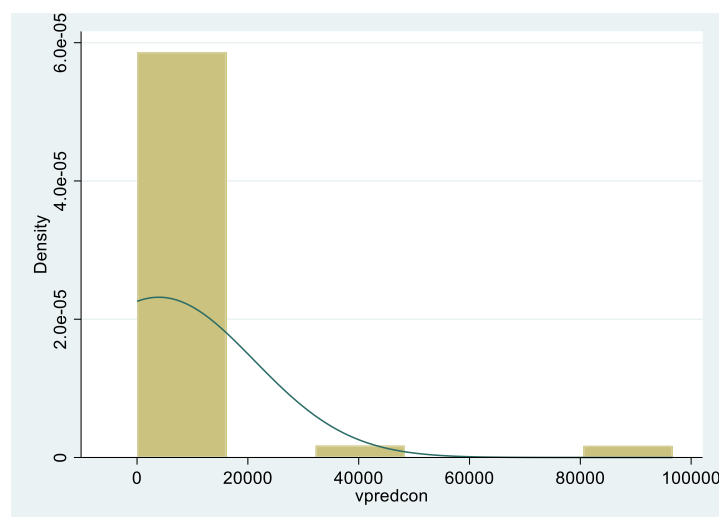
En el anterior apartado hemos observado la distribución de la rentabilidad real frente a una distribución normal. Si representamos gráficamente la distribución de la rentabilidad predicha con uso de Internet frente a una distribución normal obtenemos el gráfico 15, donde la línea azul es nuestra variable y la línea roja representa la distribución normal.

Gráfico 15 Distribución de la rentabilidad predicha con Internet y una distribución normal



Gráficamente podemos observar las mismas conclusiones que hemos obtenido del test de Shapiro-Wilk, la distribución de la rentabilidad predicha con Internet no se asemeja a una distribución normal. En el caso del valor liquidativo podemos observar en el gráfico 16 que no se asemeja a la distribución normal.

Gráfico 16 Distribución del valor liquidativo predicho con Internet y una distribución normal



Consecutivamente hemos realizado el test de Wilcoxon Signed-Rank para nuestra población, los resultados se encuentran en las tablas 18 y 19. Este análisis estadístico nos indica que si comparamos la rentabilidad real con la predicha utilizando Internet obtenemos pocas similitudes, visible en el gráfico 17.

Tabla 18. Test de Wilcoxon Signed-Rank para la rentabilidad predicha en el tercer caso

Tipo	Signo	Obs	Suma rangos	Esperado	Z	p-valor
General	Positivo	11	274,500	333	-0,919	0,3580
	Negativo	25	391,500	333		
	Cero	0	0	0		
	Todos	36	666	666		

La tabla muestra los resultados del test de Wilcoxon Signed-Rank aplicado a las diferencias entre las rentabilidades de los fondos de inversión libre elegidos para el tercer caso predichas por el modelo y rentabilidades reales observadas. El objetivo es evaluar si estas diferencias son estadísticamente significativas, ayudando a determinar si el modelo ofrece estimaciones coherentes o si las predicciones se aproximan a un comportamiento aleatorio.

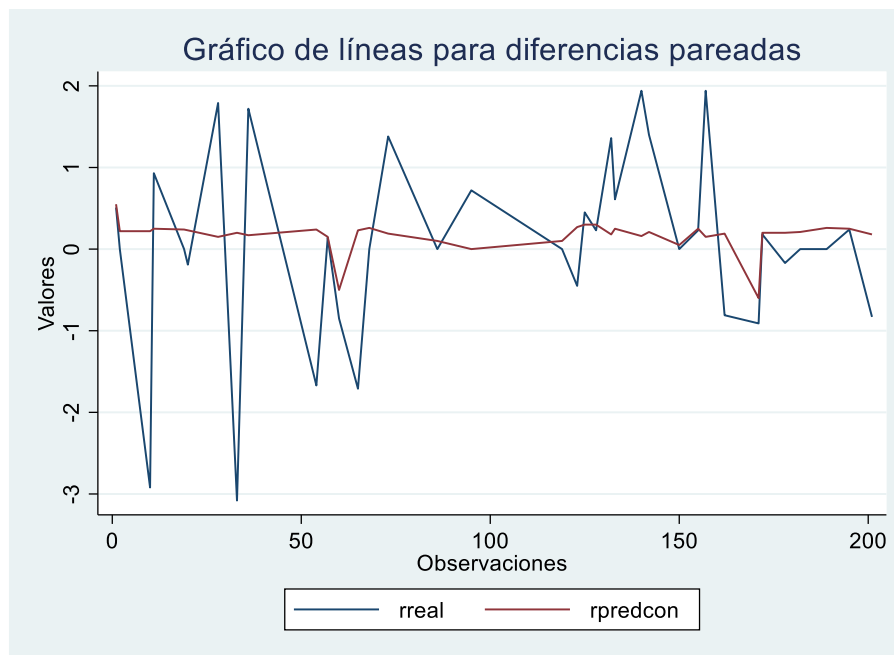
Tabla 19. Test de Wilcoxon Signed-Rank para el valor liquidativo predicho en el tercer caso

Tipo	Signo	Obs	Suma rangos	Esperado	Z	p-valor
General	Positivo	2	40	332,500	-4,595	0
	Negativo	33	625	332,500		
	Cero	1	1	1		
	Todos	36	666	666		

La tabla muestra los resultados del test de Wilcoxon Signed-Rank aplicado a las diferencias entre los valores liquidativos de los fondos de inversión libre elegidos para el tercer caso predichas por el modelo y los valores liquidativos reales observadas. El objetivo es evaluar si estas diferencias son estadísticamente

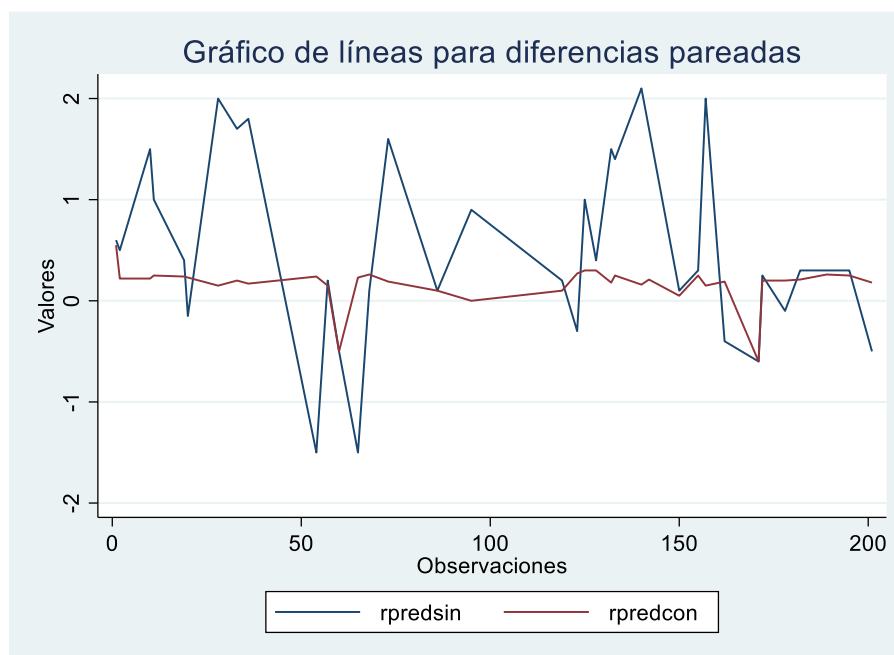
significativas, ayudando a determinar si el modelo ofrece estimaciones coherentes o si las predicciones se aproximan a un comportamiento aleatorio.

Gráfico 17 Gráfico de líneas para diferencias pareadas entre la rentabilidad real y la predicha con Internet



Por último, comparamos las rentabilidades predichas, del segundo caso y este caso, entre sí para determinar si hay similitudes entre ambas estimaciones. No obstante, como podemos ver en el gráfico 18, no hay similitud entre ellas.

Gráfico 18 Gráfico de líneas para diferencias pareadas entre la rentabilidad predicha sin Internet y la predicha con Internet



Nuestro análisis estadístico termina realizando el test de Kendall para determinar si nuestra muestra se ha ordenado correctamente. En la tabla 20 podemos observar como la rentabilidad real y la predicha utilizando Internet no se ordenan igual. También, podemos observar cómo el valor liquidativo se ordena igual. Además, hemos realizado este test entre las predicciones en el apartado anterior y este para comprobar si existe correlación entre las predicciones. Podemos determinar que

Tabla 20. Test de Kendall entre las variables de la segunda parte del análisis

Categoría	Obs	Tau-a	Tau-b	Puntuación de Kendall	Prob>z
rreal – rpredcon	36	0,0016	0,0017	1	1
rpredsint – rpredcon	36	0,0032	0,0033	2	0,9891
vreal – vpredcon	36	0,9889	0,9897	623	0
vpredsint – vpredcon	36	0,9825	0,9833	619	0

La tabla muestra los resultados de aplicar el test de Kendall para determinar si las predicciones han sido ordenadas correctamente. Los valores que se presentan son el número de observaciones de cada categoría, la puntuación de Kendall y la significancia de la correlación.

6. CONCLUSIONES Y LÍNEAS FUTURAS DE INVESTIGACIÓN

Este trabajo ha analizado la eficiencia del modelo de lenguaje de gran tamaño GPT-4o en la predicción de rentabilidad y valor liquidativo de fondos de inversión en diferentes situaciones. Los resultados nos permiten extraer varias conclusiones relevantes.

En primer lugar, el modelo demuestra mayor precisión en la estimación del valor liquidativo frente a la rentabilidad, lo que sugiere que el modelo trabaja mejor con variables menos volátiles. La rentabilidad presenta más dispersión y errores más significativos, lo que nos apunta a dificultades del LLM para capturar fluctuaciones más pronunciadas del mercado, especialmente en casos como veíamos en los valores atípicos de los fondos de inversión libre donde el modelo no ha sido capaz de identificarlos.

Además, encontramos que la introducción de los folletos de los fondos de inversión mejora considerablemente la calidad de las estimaciones, proporcionando un contexto más completo al modelo. No obstante, el uso de Internet no conlleva una mejora de la calidad de las predicciones como podríamos esperar. Esto puede deberse a que el LLM no es capaz de identificar la información necesaria y relevante para mejorar estas predicciones.

A pesar de las limitaciones observadas en la estimación exacta de las variables de rentabilidad y valor liquidativo. El modelo muestra cierta capacidad para ordenar los

fondos según su desempeño. No obstante, en unos casos podemos observar como las predicciones pueden asemejarse a una distribución normal, lo que se puede deber a un comportamiento aleatorio a la hora de realizar predicciones por parte del GPT.

Este trabajo puede ampliarse mediante la inclusión de una mayor cantidad de datos históricos, permitiendo al modelo aprender y mejorar sus predicciones. También se podría hacer uso de nuevas funcionalidades del modelo GPT-4o, automatizando tareas de recolección de datos lo que facilitaría la actualización constante del conjunto de información, mejorando así la capacidad del modelo para generar predicciones más ajustadas y en tiempo real.

Otra línea de investigación sería hacer un análisis de los modelos de lenguaje especializados en finanzas, aquellos diseñados y entrenados específicamente para llevar a cabo tareas del sector financiero. Pudiendo comparar los análisis realizados por un modelo específico y otro genérico e identificando las diferencias de resultados en distintos contextos de predicción.

REFERENCIAS

- Aguado Peña, D. (2018). Fondos de Inversión [Trabajo de fin de grado, Universidad de Valladolid]. UvaDoc. <https://uvadoc.uva.es/handle/10324/34256>
- Alloghani, M.A., Obe, D.A., Mustafina, J., Hussain, A., y Aljaaf, A.J. (2020). A Systematic Review on Supervised and Unsupervised Machine Learning Algorithms for Data Science. *Springer*. <https://doi.org/10.1007/978-3-030-22475-2>
- AWS. (2023). ¿Qué son los modelos de lenguaje de gran tamaño (LLM)? AWS. <https://aws.amazon.com/es/what-is/large-language-model/>
- Bankinter. (2024). ¿Qué es un fondo de inversión monetario? *Preguntas frecuentes inversión*. <https://www.bankinter.com/banca/preguntas-frecuentes/productos-de-inversion/que-es-un-fondo-de-inversion-monetario>
- BBVA. (s.f.) Fondos de Retorno Absoluto. *Educación financiera*. <https://www.bbva.es/finanzas-vistazo/ef/fondos-inversion/retorno-absoluto.html>
- Chicu, D. (2020). La valoración del riesgo financiero. *Universitat Oberta de Catalunya*. <https://openaccess.uoc.edu/bitstream/10609/150126/1/LaValoracionDelRiesgoFinanciero.pdf>
- Comisión Nacional del Mercado de Valores. (s.f.-a). *Fondos de inversión garantizados*. https://cnmv.es/DocPortal/Publicaciones/Fichas/GR09_Fondos_Garantizados.pdf
- Comisión Nacional del Mercado de Valores. (s.f.-b). *El mercado de los valores y los productos de inversión*. <https://www.cnmv.es/DocPortal/Publicaciones/Guias/ManualUniversitarios.pdf>
- Comisión Nacional del Mercado de Valores. (s.f.-c). Glosario financiero. *Comisión Nacional Del Mercado De Valores*. <https://www.cnmv.es/Portal/inversor/Glosario.aspx>
- Esteban, J.A. (6 de enero, 2025). El valor liquidativo de un fondo de inversión. *IRONIA*. <https://www.fondos.com/blog/que-es-el-valor-liquidativo-de-un-fondo-de-inversion>
- Fernández, A. (2019). Inteligencia artificial en los servicios financieros. *Boletín Económico / Banco de España*. <https://www.bde.es/f/webbde/SES/Secciones/Publicaciones/InformesBoletinesRevistas/ArticulosAnaliticos/19/T2/descargar/Fich/be1902-art7.pdf>
- INVERCO. (2024). QUÉ SON LOS FONDOS Y SOCIEDADES DE INVERSIÓN (IIC). *INVERCO*. <https://www.inverco.es/20/0/133>
- Kim, A.G., Muhn, M. y Nikolaev, V.V. (2024). Financial Statement Analysis with large Language Models. *Universidad de Chicago*. <http://dx.doi.org/10.2139/ssrn.4835311>
- Lock, R.H., Lock P.F., Morgan, K.L., Lock E.F., Lock D.F. (2021). *Statistics: Unlocking the Power of Data*. Wiley.

McCarthy, J. (2004). What is Artificial Intelligence? *John McCarthy's Home Page*
<https://www-formal.stanford.edu/jmc/>

Santander. (s.f.-a). ¿Qué es un fondo de inversión de renta fija? *Ahorro inversión*.
<https://www.bancosantander.es/faqs/particulares/ahorro-inversion/que-es-fondo-inversion-renta-fija>

Santander. (s.f.-b). ¿Qué son los fondos de inversión de renta variable? *Ahorro Inversión*
<https://www.bancosantander.es/faqs/particulares/ahorro-inversion/fondos-inversion-renta-variable>

Schwab, K. (2020). La Cuarta Revolución Industrial. *Futuro Hoy*, 1(1), 06-10.
doi:10.5281/zenodo.4299164

Shanahan, M. (25 de Enero de 2024). Talking about Large Language Models. *Communications of the ACM*, 67 (2), 68-79. <https://doi.org/10.1145/3624724>

Statista. (2024). Tamaño del mercado de la inteligencia artificial a nivel mundial de 2020 a 2030 (en miles de millones de dólares). *Statista*: <https://es-statista-com.ponton.uva.es/estadisticas/1139768/inteligencia-artificial-vaolr-de-mercado/>

The Editors of Encyclopaedia Britannica. (20 de Octubre de 2024). John McCarthy. *Brittanica*: <https://www.britannica.com/biography/John-McCarthy>

Universidad de Barcelona. (1 de marzo de 2005). Contrastes de normalidad. *Universidad de Barcelona*. http://www.ub.edu/aplica_infor/spss/cap5-6.htm

ANEXO A: TABLAS

En estas tablas se presentan los valores estadísticos de las categorías de los fondos analizadas en el primer caso. La tabla 1 de los fondos índice, la tabla 2 de los fondos monetarios, la tabla 3 de fondos de renta fija a corto plazo y la tabla 4 de los fondos de renta variable.

Tabla 21. Valores estadísticos para los fondos índice

Fondos Índice	Obs	Media	Desviación estándar	Min	Max
ROCT	48	-1,287	1,69	-3,44	2,44
RNOV	48	2,619	3,581	-0,55	8,81
PREDICRENT	48	-1,542	2,132	-5	3,518
DIFRENTAB	48	4,359	2,734	0,021	9,242
DIFRENTREL	47	1667,492	2928,076	51,722	13655,508
VOCT	48	56,092	81,095	5,855	290,97
VNOV	48	56,797	81,292	5,918	290,422
PREDICV	48	54,602	78,085	5,745	276,621
DIFVAB	48	2,216	3,412	0,002	14,001
DIFVREL	48	4,179	2,523	0,016	8,743

Tabla 22. Valores estadísticos para los fondos monetarios

Fondos Monetarios	Obs	Media	Desviación estándar	Min	Max
ROCT	13	0,253	0,08	0	0,32
RNOV	13	0,248	0,032	0,19	0,3
PREDICRENT	13	-0,221	1,341	-2,291	3,179
DIFRENTAB	13	1,069	0,873	0,003	2,919
DIFRENTREL	13	456,926	397,064	1,054	1305,662
VOCT	13	38,335	48,111	6,004	110,317
VNOV	13	38,44	48,247	6,02	110,649
PREDICV	13	38,292	48,081	5,987	111,033
DIFVAB	13	0,25	0,268	0,001	0,924
DIFVREL	13	1,067	0,871	0,007	2,911

Tabla 23. Valores estadísticos para los fondos de renta fija a corto plazo

Fondos renta fija c-p	Obs	Media	Desviación estándar	Min	Max
ROCT	204	0,232	0,148	-0,29	0,86
RNOV	204	0,333	0,16	-0,02	0,88
PREDICRENT	204	0,184	1,223	-3,905	3,762
DIFRENTAB	204	0,999	0,767	0,001	4,125
DIFRENTREL	201	381,802	567,981	0,512	6633,531

VOCT	204	335,979	1217,817	0	12941,81
VNOV	204	337,722	1221,621	0,187	12973,266
PREDICV	204	337,165	1221,837	0	12928,101
DIFVAB	204	3,791	13,506	0	125,015
DIFVREL	204	1,481	6,974	0	100

Tabla 24. Valores estadísticos para los fondos de renta variable

Fondos renta variable	Obs	Media	Desviación estándar	Min	Max
ROCT	65	-1,948	1,088	-3,81	2,81
RNOV	65	0,148	1,518	-1,03	9,85
PREDICRENT	65	-2,195	1,592	-5	1,13
DIFRENTAB	65	2,447	1,939	0,025	8,72
DIFRENTREL	62	1186,018	2382,914	4,717	16766,667
VOCT	65	186,847	372,003	0	1694,8
VNOV	65	187,323	372,111	8,145	1697,806
PREDICV	65	182,573	364,221	0	1682,099
DIFVAB	65	4,769	9,355	0,009	46,154
DIFVREL	65	3,955	12,24	0,023	100

ANEXO B: GRÁFICOS

En estos gráficos se presentan los gráficos de líneas para diferencias pareadas para la rentabilidad y valor liquidativo de varias categorías de fondos empleados en el primer caso.

Gráfico 19 Gráfico de líneas para diferencias pareadas para la rentabilidad en los fondos índice

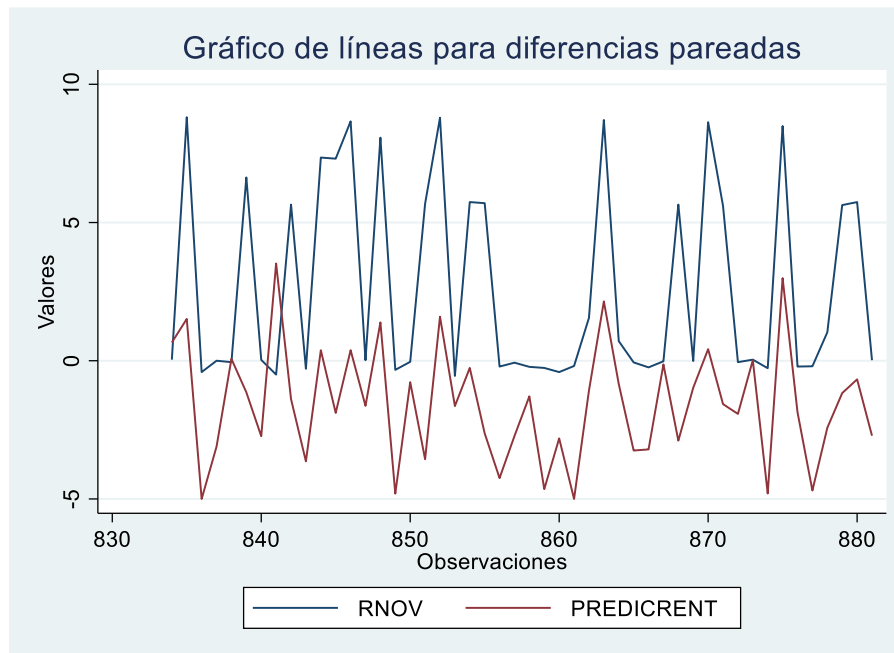


Gráfico 20 Gráfico de líneas para diferencias pareadas para la rentabilidad en los fondos monetarios

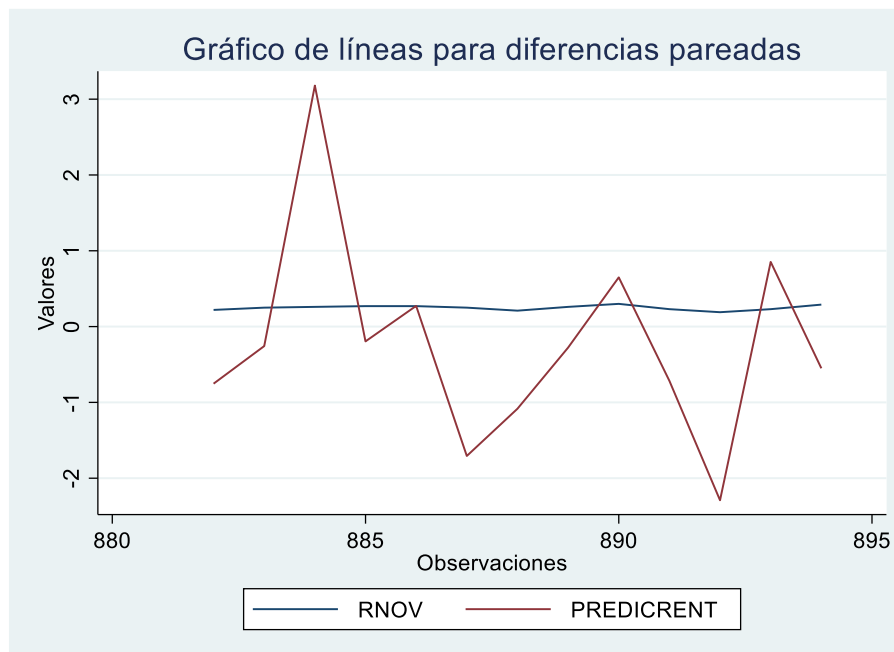


Gráfico 3 Gráfico de líneas para diferencias pareadas para la rentabilidad en los RF c-p

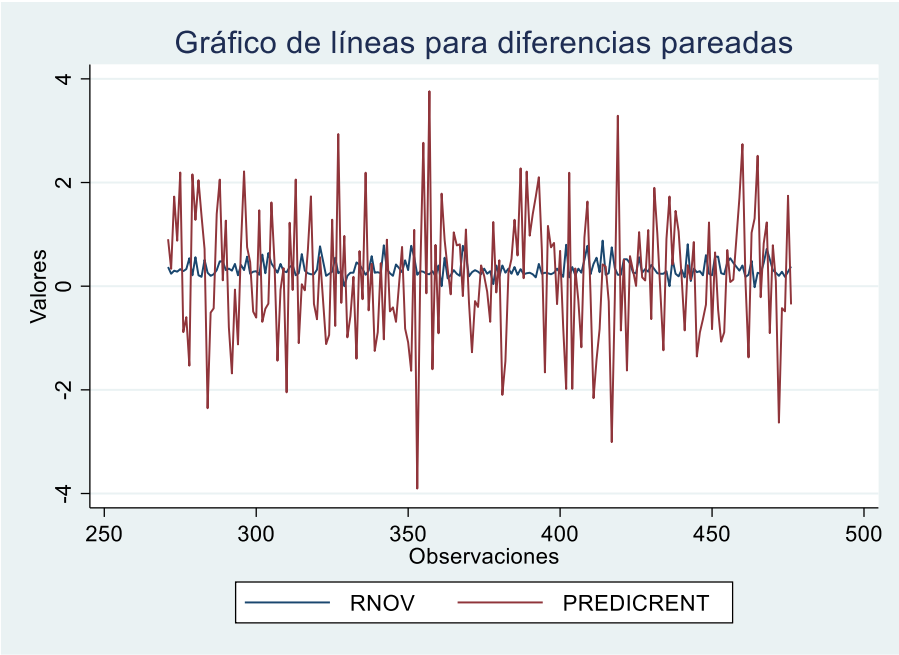


Gráfico 4 Gráfico de líneas para diferencias pareadas para la rentabilidad en los RV

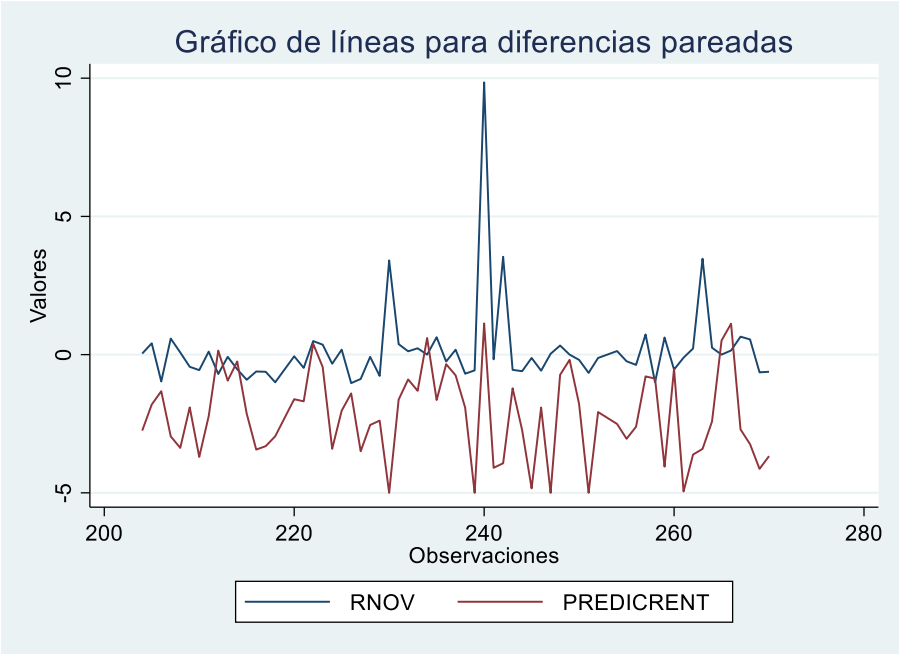


Gráfico 5 Gráfico de líneas para diferencias pareadas para el valor liquidativo en los FIL

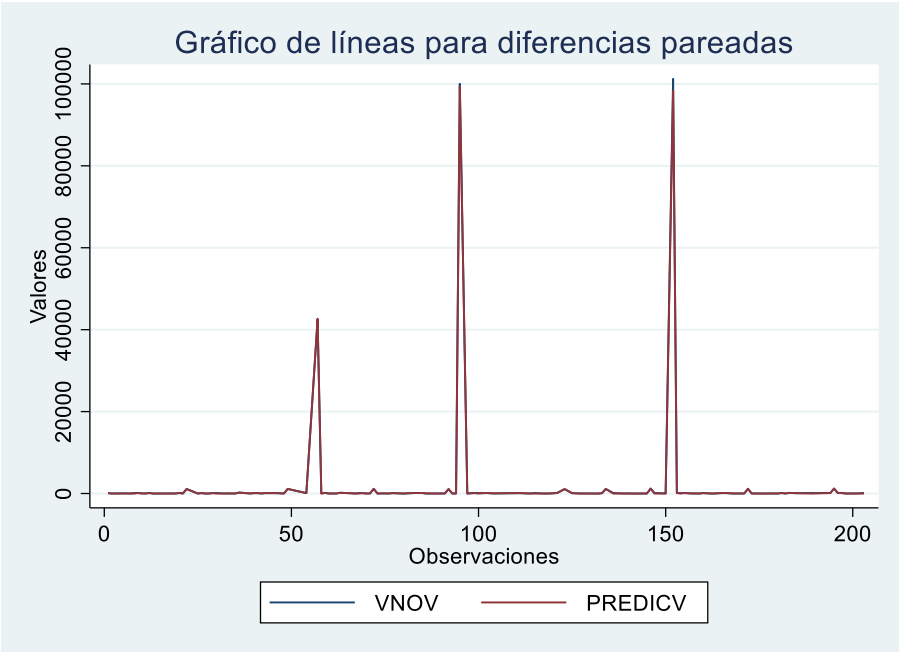


Gráfico 6 Gráfico de líneas para diferencias pareadas para el valor liquidativo en los fondos índice

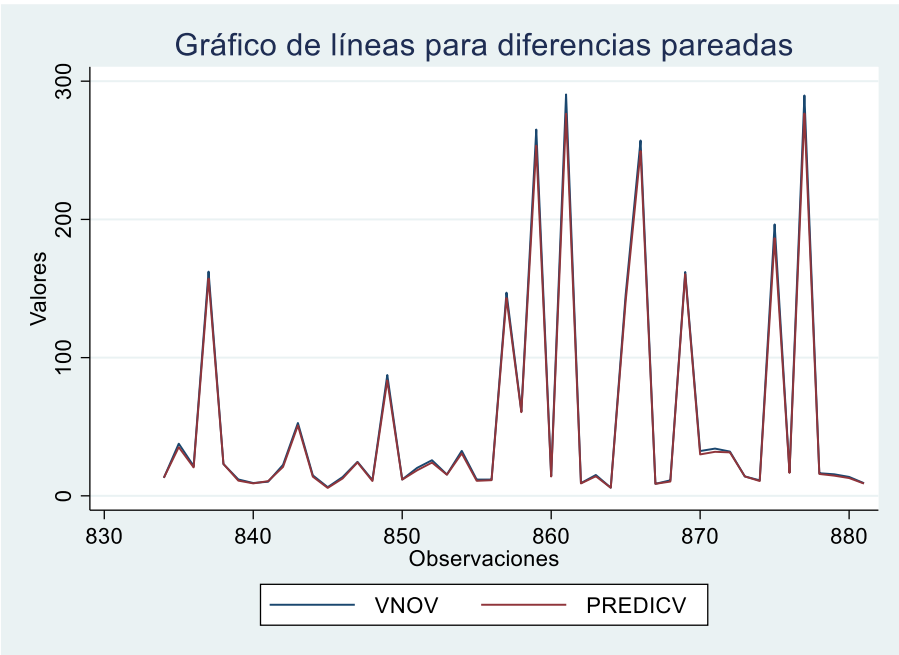


Gráfico 7 Gráfico de líneas para diferencias pareadas para el valor liquidativo en los fondos monetarios

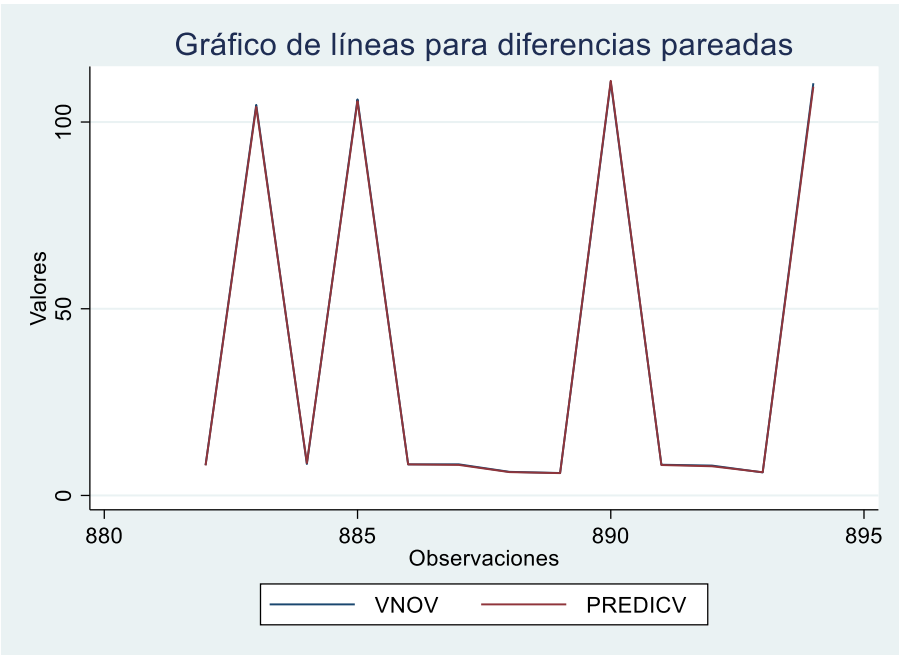


Gráfico 8 Gráfico de líneas para diferencias pareadas para el valor liquidativo en los RF c-p

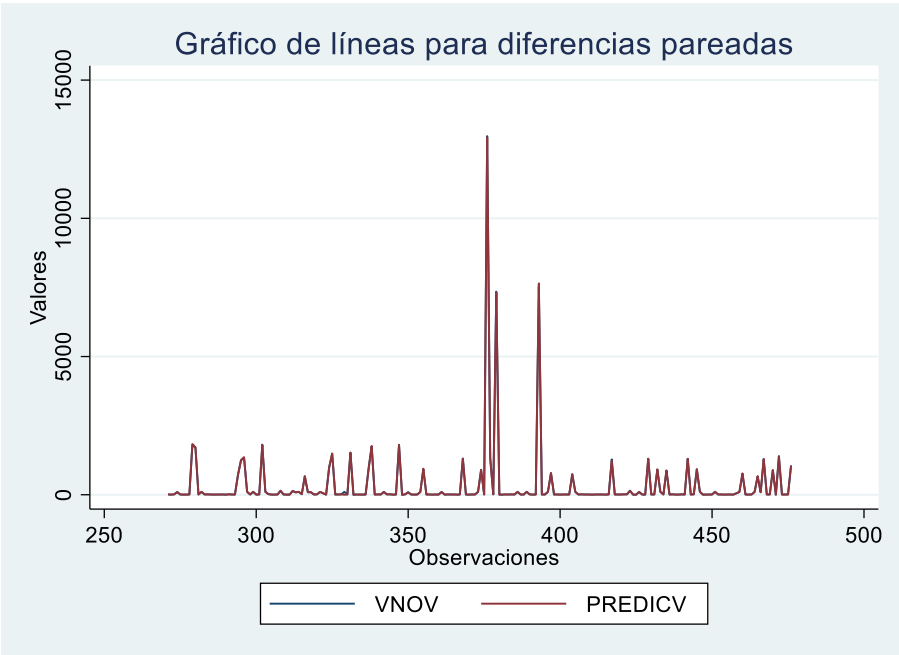


Gráfico 9 Gráfico de líneas para diferencias pareadas para el valor liquidativo en los RF I-p

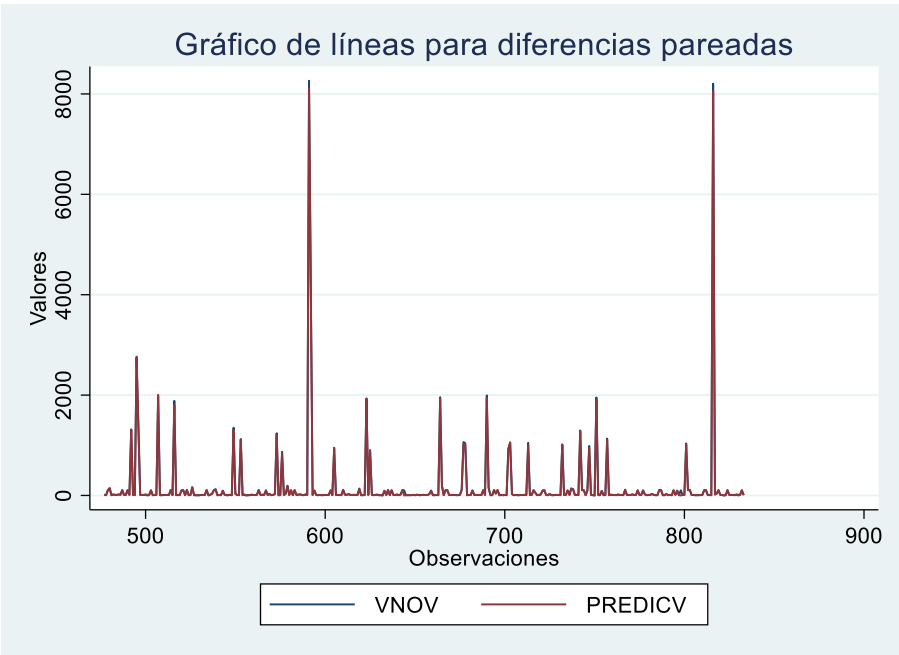


Gráfico 10 Gráfico de líneas para diferencias pareadas para el valor liquidativo en los RV

