Universitat Oberta de Catalunya (UOC)

Máster Universitario en Ciencia de Datos (*Data Science*)

TRABAJO FINAL DE MÁSTER

Área: 002

Agente inversor para acciones de small cap mediante el uso de Reinforcement Learning

---

Autor: Ignacio Such Ballester

Tutor: Rubén Pérez Ibáñez

Profesor: Ismael Benito Altamirano

Madrid, 10 de febrero de 2024

Créditos/Copyright

Declaro que yo, Ignacio Such Ballester, soy el autor de esta tesis titulada "Predicción de los valores de las acciones de empresas small cap a largo plazo mediante el uso del aprendizaje por refuerzoz del trabajo presentado en ella. Confirmo que: Este trabajo se ha realizado total o principalmente durante la candidatura a un máster en esta Universidad. Si alguna parte de esta tesis se ha presentado anteriormente para la obtención de un título o cualquier otra cualificación en esta Universidad o en cualquier otra institución, se ha indicado claramente. En los casos en que he consultado el trabajo publicado de otros, esto siempre es claramente atribuido. Cuando he citado el trabajo de otros, siempre se indica la fuente. Con excepción de dichas citas, esta tesis es enteramente obra mía. He reconocido todas las principales fuentes de ayuda. En los casos en que la tesis se basa en un trabajo realizado por mí conjuntamente con otros, he dejado claro lo que han hecho los demás y lo que he aportado yo.

FICHA DEL TRABAJO FINAL

Título del trabajo:	Agente inversor para acciones de small cap mediante el uso de Reinforcement Learning.
Nombre del autor:	Ignacio Such Ballester
Nombre del colaborador/a docente:	Rubén Pérez Ibáñez
Nombre del PRA:	Ismael Benito Altamirano
Fecha de entrega (mm/aaaa):	02/2024
Titulación o programa:	Máster Ciencia de Datos
Área del Trabajo Final:	Aprendizaje por refuerzo
Idioma del trabajo:	Español
Palabras clave	Deep Reinforcement Learning, Stock market, Proximal Policy Optimization

Dedicatoria/Cita

Quiero agradecer a mi tutor, Rubén, la paciencia y ayuda prestada a lo largo de este tiempo para la redacción y orientación de este trabajo final de máster.

Me gustaría dedicarle este trabajo final de máster a mis padres Ignacio y Amparo, los cuales siempre estuvieron para mí y me motivaron a perseguir mis sueños y orientarme a un campo de estudio, que, a mi parecer, es el más bonito de todos.

Finalmente, y sobre todo, me gustaría dedicárselo a mi pareja, Macarena, por lo que representa para mí, siendo mi mayor motivación y apoyo durante esta etapa.

Por otro lado, quiero dedicar una carta al Ignacio del futuro:

Querido Ignacio del futuro:

Espero que estés bien y que vaya todo fenomenal en nuestra vida. A día de hoy sólo sueño con poder ayudar a las personas mediante el uso de Inteligencia Artificial. Espero que cuando leas esta pequeña dedicatoria, estemos orgullosos tanto del camino como de la meta que hemos tenido siempre.

Fdo: Ignacio del pasado.

Per Aspera Ad Astra

Resumen del trabajo

En el contexto actual, el mercado bursátil se ha convertido en una industria compleja y en constante cambio, donde la toma de decisiones puede ser crucial para obtener beneficios o incurrir en pérdidas significativas. La evolución de los sistemas de inversión ha llevado a la incorporación de técnicas innovadoras de aprendizaje automático, como el Reinforcement Learning, que permiten aprender y adaptarse a los cambios del mercado en tiempo real. El Reinforcement Learning es una rama del aprendizaje automático que se basa en el concepto de premio y castigo, y su objetivo es maximizar la recompensa obtenida a través de un proceso de interacción con un entorno. Esta técnica ha demostrado su eficacia en la resolución de problemas complejos y ha sido aplicada con éxito en entornos como la robótica y los videojuegos. En este contexto, el uso del Reinforcement Learning en el mercado bursátil se presenta como una alternativa prometedora para el diseño de estrategias de inversión óptimas y rentables en un entorno cambiante y altamente competitivo.

El objetivo de esta tesis de fin de Máster es establecer una nueva línea de investigación en la predicción de valores de las acciones de empresas "small cap"mediante el uso de algoritmos de Deep Reinforcement Learning. El primer algoritmo es el Proximal Policy Optimization (PPO) A, donde se utilizará la implementación de Liu et al. [11]. En ella, se muestra como es un buen agente de bolsa en momentos al alza pero son mas vulnerables en etapas de descenso. Por otro lado, se mencionan los algoritmos A2C C y B, los cuales han sido prometedores en la misma obra. Se demuestra que el DDPG, no es tan eficaz como el PPO, aunque sí es mas cauto en etapas de caídas en bolsa.

Palabras clave: Deep Reinforcement Learning, Stock Market, Small Caps

Abstract

In the current context, the stock market has become a complex and ever-changing industry, where decision-making can be crucial for making significant profits or incurring significant losses. The evolution of investment systems has led to the incorporation of innovative machine learning techniques, such as Reinforcement Learning, which allow them to learn and adapt to market changes in real-time. Reinforcement Learning is a branch of machine learning based on the concept of reward and punishment, which aims to maximise the reward obtained through interaction with an environment. This technique has proven its effectiveness in solving complex problems and has been successfully applied in environments such as robotics and video games. In this context, the use of Reinforcement Learning in the stock market presents itself as a promising alternative for the design of optimal and profitable investment strategies in a changing and highly competitive environment.

The objective of this Master's thesis is to establish a new line of research in predicting stock values of "small capçompanies through the use of Deep Reinforcement Learning algorithms. The first algorithm is the Proximal Policy Optimization (PPO), where the implementation of Liu et al. [11] will be used. On the other hand, the A2C and DDPG algorithms will be employed, which have been promising according to this paper Liu et al. [10].

Keywords: Deep Reinforcement Learning, Stock Market, Small Caps

Índice general

Índice de figuras

Índice de cuadros

Capítulo 1

Introducción

1.1. Contexto

Durante la historia de la humanidad, se han ido desarrollando teorías o conceptos como la teoría del caos, el efecto mariposa o la ley de la acción reacción, las cuales demuestran que a partir de una acción actual tiene repercusiones tanto a corto plazo como a largo plazo.

El aprendizaje por refuerzo, o Reinforcement Learning (RL), es una rama del aprendizaje automático que también se relaciona con estas teorías. El RL se enfoca en cómo un agente de software puede aprender a tomar decisiones óptimas en un entorno dado, a través de la interacción con ese entorno y la observación de las recompensas obtenidas. Al igual que en la teoría del caos o en la ley de la acción reacción, las decisiones tomadas por el agente en RL también pueden tener un impacto a largo plazo en el entorno y en las recompensas obtenidas.

Es en este contexto donde se sitúa la presente tesis de máster, que tiene como objetivo explorar el potencial del RL aplicado al campo de las inversiones. En concreto, se buscará desarrollar y analizar algoritmos de RL para la predicción de valores de acciones de empresas "small cap". Para ello, se emplearán modelos de Deep Reinforcement Learning, que permitirán al agente de software aprender patrones complejos en los datos históricos de cotización de las empresas y tomar decisiones de inversión basadas en ellos.

El resultado de esta investigación podría ser de gran interés tanto para inversores como para empresas, ya que la capacidad de predecir los valores de las acciones de empresas "small cap"puede resultar crucial para la toma de decisiones de inversión efectivas y rentables. Además, esta tesis puede contribuir al avance en el campo del RL y su aplicación a problemas del mundo real.

1.2. Justificación

Los algoritmos de aprendizaje por refuerzo han logrado imitar la conducta de los seres vivos en entornos dinámicos y complejos con éxito. En este sentido, el objetivo de esta tesis de máster es mejorar el estado del arte en el campo de la inversión en bolsa, específicamente determinando qué enfoque de aprendizaje por refuerzo brinda los mejores resultados en un ambiente con un número elevado de dimensiones y una alta volatilidad como lo son empresas que acaban de entrar en la bolsa de valores o que actualmente no tienen una base sólida como multinacional.

1.3. Motivación

Desde siempre me ha sorprendido y entusiasmado el dinero, lo veía como el poder total para el mundo en el que vivimos y siempre quise encontrar una mina de oro. Es por ello por lo que trato de emplear un algoritmo de aprendizaje por refuerzo para poder encontrar empresas con mucho potencial que actualmente no están tan valoradas.

Considero que es una buena oportunidad construir y entrenar un modelo que pueda predecir el valor de dichas empresas. Pensemos en lo siguiente:

1. Tenemos una multinacional, como por ejemplo Berkshire Hathaway, la cual sus acciones (Clase B) a día de hoy, 12 de marzo de 2023, valen 303$. Imaginemos que mañana, 13 de marzo, valen 306$.

2. Imaginemos que ahora tenemos una small cap, la cual su valor por acción es de 3$ y mañana valen 6$.

En ambos casos ha aumentado 3$ el valor de la acción. Sin embargo, mientras que Berkshire ha aumentado 0.99 %, la small cap ha aumentado un 100 %. Por ello, si hubiéramos invertido 1000€, con la multinacional hubiéramos ganado apenas 10€, mientras que con la pequeña empresa, tendríamos 1000€ más.

1.4. Objetivo Principal

El objetivo principal de este proyecto es la recreación de un entorno bursátil volátil definida en la obra *Deep Reinforcement Learning for Automated Stock Trading: An Ensemble Strategy* [23] donde se tratará de encontrar el mayor rendimiento a través de un agente de RL con una alta fiabilidad. Para ello, se tratará de adaptar diferentes algoritmos y mostrar mediante gráficas el rendimiento de cada uno de ellos de manera conjunta.

1.5. Objetivos parciales

Los objetivos parciales son los siguientes:

1. Se tratará de alcanzar un nuevo estado del arte en referencia a las empresas "small caps", las cuales podrían ser un nuevo método para que los pequeños inversores tengan la oportunidad de aumentar su capital. Para ello, se entrenarán tres agentes

2. Por otro lado, se enfocará en desarrollar tres agentes capaces de ejecutar estrategias de trading a corto plazo. Este enfoque permitirá al usuario aprovechar las oportunidades de mercado en un marco temporal más inmediato, adaptándose a las actualizaciones de datos cada 5 minutos.

Es importante señalar que esta Tesis se basa en una revisión exhaustiva y crítica de la literatura científica existente en el campo de estudio. A través de esta revisión, se buscará identificar las principales teorías, modelos y enfoques que se han utilizado previamente para abordar el problema de investigación en cuestión. Aunque se utilizarán trabajos de terceros para desarrollar y apoyar las ideas presentadas en esta Tesis, se garantiza que se seguirán los estándares éticos y académicos en cuanto a la correcta atribución de fuentes y el uso de citas y referencias adecuadas. La intención de esta Tesis es construir un análisis crítico y original del tema de investigación, a partir de la combinación de ideas previas y la propuesta de nuevas perspectivas y enfoques que permitan avanzar en el conocimiento de la materia.

1.6. Hipótesis

La hipótesis planteada para este proyecto es que, a pesar de la alta volatilidad de las "small cap" en el mercado, es posible lograr un ROI[1] alto y confiable mediante una estrategia de inversión adecuada. Se espera que esta estrategia permita alcanzar metas parciales de inversión de manera constante y sostenible, lo que puede resultar en un rendimiento financiero satisfactorio.

1.7. Metodología

Para la elaboración de la presente Tesis, se empleará el lenguaje de programación Python, el cual incorpora una gran cantidad de librerías para la elaboración de los agentes del aprendizaje por refuerzo profundo como pueden ser la librería de Tensorflow o Torch.

[1] ROI: Return On Investment

Por otro lado, el entorno que se pretende emplear para el presente proyecto es el FinRL, el cual se ha demostrado en el siguiente paper ([10]) que se ha llegado a obtener un ROI entre el 149 % y 362 %.

En referencia a la información financiera, la librería que se pretende emplear para la extracción de datos es yfinance. Este módulo escrito en Python permite obtener datos financieros de Yahoo! finance para el proyecto.

El estudio se desarrollará con la ayuda de Git y Gitlab, para el control de versiones.

Para la elaboración de la memoria se ha utilizado Latex, un sistema de composición de textos que, junto el entorno de programación proporcionado por Overleaf [16], permite elaborar documentos de alta calidad tipográfica en un entorno colaborativo.

1.8. Competencia de compromiso ético y global (CCEG) y Objetivos de Desarrollo Sostenible (ODS)

El presente proyecto consistirá en la creación de agentes de Aprendizaje por Refuerzo cuyo objetivo es obtener ganancias de capital mediante el uso de trading algorítmico. Solamente se va a tener en cuenta el uso de datos financieros, por lo que no se cometerá ningún crimen ético en referencia a la sostenibilidad, comportamiento ético, responsabilidad social, Diversidad y derechos humanos.

He llegado a la conclusión que no se llega a vulnerar ni los derechos de personas ni cometer un acto para nada ético ya que la misión del presente proyecto es hacer un estudio matemático mediante datos financieros donde solamente se tendrán en cuanta la evolución de los precios de las acciones de hasta 600 empresas. Por otro lado, no está involucrada ninguna persona física en el presente proyecto salvo los autores de los artículos y libros que empleo como apoyo para la realización del proyecto por lo que se descarta cualquier tipo de discriminación tanto racial, como de género, sexual, social, capital, etc.

1.9. Planificación del proyecto o de la investigación

La planificación para el presente proyecto es el siguiente:

- Investigación: Se hará una profunda investigación sobre el estado del arte del aprendizaje por refuerzo, así como los diferentes enfoques para construir motores de negociación automatizados. Esta tarea debería llevar entre dos y tres semanas.

- Extracción de datos: Esta tarea consistirá en extraer los datos a través de la librería yfinance para su futuro procesamiento. Esta tarea tomaría alrededor de uno o dos días.

- Creación de un entorno: Esta tarea consistirá en programar el entorno en el que cada uno de los modelos escogidos pueda interactuar con el mercado de acciones. Esta es una tarea crucial porque debe realizarse con precisión considerando muchos factores. Para mostrar por qué es una tarea tan relevante, recordemos que, por ejemplo, cuando se ejecuta una operación de compra, la cantidad de dinero disponible en la billetera debe disminuir, pero también puede suceder que ese día el agente decida vender otras acciones, lo que significará que el monto de la billetera tendrá que aumentar. Es necesario asegurarse de que este proceso esté bien controlado, que no se pierda información y que se almacene un histórico con cada operación para su futuro análisis. Esta tarea debería llevar dos semanas de trabajo.

- Implementación de los modelos seleccionados. En esta tarea se deben implementar los diferentes modelos seleccionados en la primera tarea con el objetivo de compararlos en el entorno de estudio. Esta tarea debería llevar tres semanas.

- Entrenamiento de los modelos. Esta tarea consistirá en entrenar los diferentes modelos en el entorno creado. Esta tarea debería completarse en tres semanas.

- Elaboración de las conclusiones. Esta tarea consistirá en analizar las diferentes conclusiones obtenidas del estudio realizado. Debería completarse en las dos últimas semanas del proyecto

- Redacción de tesis: esta tarea consistirá en poner por escrito todos los procesos desarrollados y se realizará en paralelo con el resto de tareas. Para ello, se debe de realizar a lo largo del proyecto.

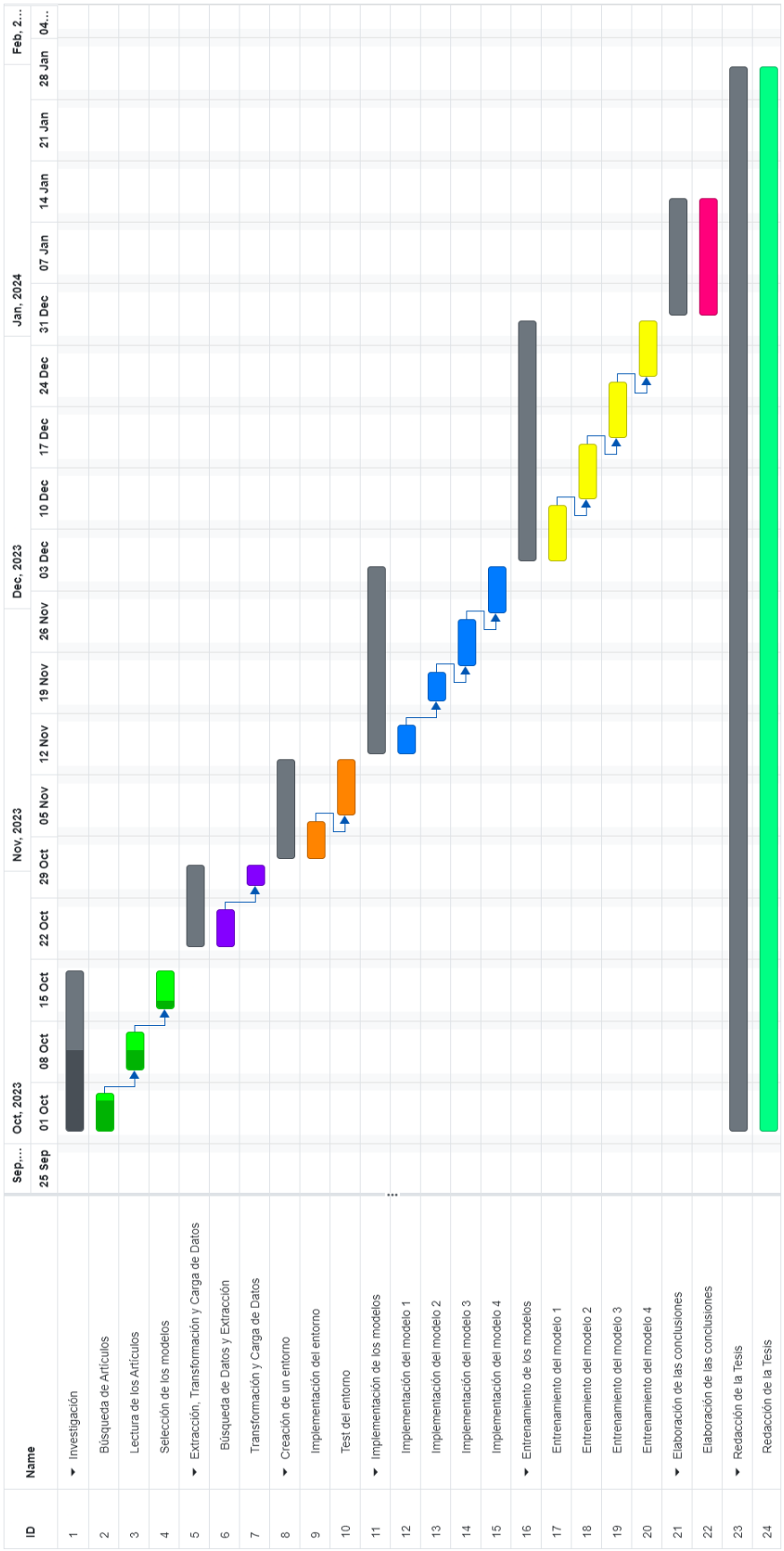


Figura 1.1: Diagrama Gantt. Fuente propia.

Capítulo 2

Estado del arte

2.1. Introducción

En este capítulo se introducirán lo que son los mercados financieros 2.2 así como el mercado de valores 2.3. Asimismo, mencionar que se describirá lo que es el Trading (o venta en corto) y su evolución 2.4. En referencia al aprendizaje automático, se describirá lo que es y como es aplicado al trading 2.5. También se definirán las diferentes vertientes que tiene el aprendizaje automático, como la gestión de carteras, predicción de valor de activos financieros o análisis de sentimientos entre otros 2.6.

Por otro lado, se introducirá el Aprendizaje por Refuerzo con una analogía y su desarrollo en el trading 2.7. Y, al final de este capítulo, se describirán los algoritmos existentes en el campo del Aprendizaje por refuerzo 2.8, así como su justificación de los seleccionados.

2.2. Mercados Financieros

De acuerdo con *Financial Markets and Institutions* [12], un mercado financiero es un mercado en el que se negocian instrumentos financieros como acciones, bonos, divisas y derivados. Estos mercados permiten a compradores y vendedores negociar activos financieros con el objetivo de obtener un beneficio o gestionar un riesgo. La función principal de los activos financieros es canalizar fondos de quienes tienen un excedente de capital a quienes tienen un déficit de los mismos. Por otro lado, los mercados financieros desempeñan un papel fundamental en la economía al permitir compartir y transferir riesgos entre distintas partes. Al comprar y vender activos financieros, los inversores pueden diversificar sus carteras y reducir su exposición a los riesgos asociados a los activos individuales.

2.3. Mercado de valores

En el mercado de valores se comercializan los valores de renta variable, comúnmente conocido como acciones [2]. Según el libro *Paul Wilmott Introduces Quantitative Finance*[22], una acción o *stock* representa una participación en la propiedad de una empresa. Cuando una empresa decide emitir acciones, las vende a los inversores con el fin de recaudar fondos para financiar sus operaciones y proyectos futuros.

El comprador de una acción se convierte en accionista de la empresa y, por lo tanto, tiene derecho a recibir una parte de las ganancias de la empresa en forma de dividendos y a participar en las decisiones importantes de la empresa a través de votaciones en asambleas de accionistas.

El precio de una acción en el mercado es determinado por la oferta y la demanda de los inversores, y puede variar ampliamente en función de la percepción del mercado sobre la salud y el futuro de la empresa [4]. En la práctica, el trading de acciones es una actividad clave en los mercados financieros, ya que ofrece a los inversores la oportunidad de obtener beneficios a través de la compra y venta de acciones en diferentes empresas.

Por otro lado, el libro *Small Stocks, Big Money: Interviews With Microcap Superstars* [26], las small caps son acciones de empresas de baja capitalización, lo que significa que tienen una capitalización de mercado relativamente pequeña en comparación con las grandes empresas que cotizan en bolsa. Por lo general, se considera que las empresas de pequeña capitalización tienen un valor de mercado de entre $300 millones y $2 mil millones.

Las small caps suelen ser empresas jóvenes y en crecimiento que están en una etapa temprana de su ciclo de vida, y por lo tanto, tienen un mayor potencial de crecimiento que las empresas más grandes y establecidas. Debido a su menor capitalización de mercado, las small caps son generalmente menos líquidas que las grandes empresas, lo que significa que pueden ser más volátiles y tener mayores fluctuaciones de precios en un corto período de tiempo.

A pesar de esto, muchos inversores ven a las small caps como una oportunidad de inversión atractiva, ya que tienen el potencial de proporcionar mayores ganancias a largo plazo debido a su capacidad para crecer más rápido que las empresas más grandes y establecidas. [26]

Además de la inversión tradicional en acciones, existe otra forma de obtener beneficios mediante la denominada venta en corto. Esta técnica de inversión se basa en la idea de que el precio de una acción puede disminuir en el futuro y permite a los inversores obtener beneficios mediante la venta de las acciones prestadas que no poseen en el momento de la venta. En el

caso de que el precio de la acción aumenta, el inversor incurre una pérdida. No obstante, si el precio de la acción disminuye, el inversor puede volver a comprar las acciones a un precio menor y devolverlas al prestamista, obteniendo unos beneficios en el proceso.

Por otro lado, al existir la necesidad de una gran cantidad de entidades y personas que desean adquirir acciones de empresas, se crearon las bolsas de valores. Estas bolsas de valores permiten a las empresas obtener capital al emitir acciones al público en general y a su vez, ofrecen una oportunidad de inversión a los inversores interesados en participar en el mercado de valores.

Las bolsas de valores también proporcionan un ambiente seguro y transparente para la negociación de acciones, lo que promueve la confianza en el mercado y aumenta la eficiencia en la asignación de recursos. A medida que las tecnologías avanzan, la negociación de acciones se ha vuelto cada vez más accesible, eficiente y globalizada, lo que ha permitido a más personas y empresas participar en el mercado de valores y obtener los beneficios potenciales que ofrece la inversión en acciones.

2.4. Trading

Tras la evolución de los mercados bursátiles y su globalización, ha permitido que cada vez más entidades y personas físicas puedan interactuar con las acciones de las sociedades.

Un corredor de bolsa, a lo que comúnmente se le denomina *broker*. tiene como función actuar como intermediario entre los inversores y los mercados financieros, facilitando la compra y venta de acciones y otros valores. Los corredores pueden realizan órdenes de compra y venta en el momento en el que un cliente necesite ejecutar la acción u ordenarle que se compre o venda cuando el valor de una acción valga un precio determinado.

Cuando un inversor desea comprar o vender una acción, se comunica con su corredor de bolsa, quien a su vez realiza la operación en su nombre. El corredor de bolsa cobra una comisión por sus servicios, que suele ser un porcentaje del valor de la transacción.

Las órdenes de compra venta de las acciones quedan registradas en el libro de órdenes, el cual muestra la cantidad y precio de las órdenes de compra y venta para un instrumento financiero en particular. El precio de la acción varía debido a la oferta y la demanda, pero otros factores como el análisis técnico y fundamental, también pueden influir en su valor.

- Análisis técnico: Consiste en la recopilación de datos y eventos pasados en el mercado

de valores como el volumen y precio de cada acción para hacer una predicción de cuánto podría valer en un tiempo determinado en el futuro. [22]

- Análisis fundamental: Se enfoca en estudiar los aspectos internos y externos de una empresa con el objetivo de determinar su valor intrínseco. Esto implica identificar los elementos financieros y no financieros que afectan el desempeño de la empresa, como la calidad de su gestión, su estructura de costos y su posición en el mercado. Los inversores que emplean esta estrategia creen que los mercados pueden fijar un precio incorrecto a corto plazo, pero eventualmente se alcanzará el precio correcto. Es por ello por lo que se puede obtener grandes ganancias comprando o vendiendo acciones a precios erróneos. [22]

No solo los mercados bursátiles se digitalizaron, el trading también lo hizo. El trading algorítmico aparece en la década de 1980, aunque comenzó a emplearse de manera más frecuente a principio de los años 2000 con el surgimiento de las plataformas online para el trading.

El trading algorítmico consiste en el uso de programas informáticos, los cuales emplean algoritmos matemáticos, para automatizar la ejecución de órdenes de compra y venta de los mercados financieros. Dichos algoritmos son capaces de procesar grandes cantidades de datos, permitiendo identificar oportunidades de negociación y ejecutar órdenes de manera rápida y precisa.

El trading algorítmico se emplea en diversas estrategias de inversión, desde la ejecución de órdenes sencillas hasta estrategias más complejas que utilizan el análisis técnico, el análisis fundamental y el aprendizaje automático. El giro de los medios, el seguimiento de tendencias y el arbitraje estadístico son algunas de las estrategias más populares. El uso de la negociación automatizada ha reducido los costes y errores asociados a la negociación humana y ha aumentado la eficiencia en la ejecución de órdenes.

2.5. Aprendizaje Automático aplicado al Trading

El trading ha evolucionado hacia el trading algorítmico, y lo mismo ha ocurrido con el campo de la computación y el aprendizaje automático. En este sentido, el aprendizaje automático o ML[1] es un subcampo de la inteligencia artificial (IA) que se centra en el desarrollo de algoritmos y modelos que permiten a las máquinas aprender y mejorar su rendimiento a partir de la experiencia, es decir, de los datos. En lugar de ser programadas explícitamente para realizar una tarea específica, las máquinas que utilizan el aprendizaje automático ajustan y optimizan su rendimiento a medida que se les expone a más datos y situaciones.

[1]ML: Machine Learning (Aprendizaje Automático)

Sin embargo, el uso del Machine Learning no se limita a la inteligencia artificial, sino que también ha revolucionado el mundo del trading al permitir el análisis y la interpretación de grandes cantidades de datos para tomar decisiones de inversión más informadas y precisas. En particular, las técnicas de Machine Learning son cada vez más utilizadas en el mercado financiero para la predicción de precios, la clasificación de patrones, el análisis de sentimiento, la gestión de riesgos, la clasificación de activos y la optimización de carteras. En este contexto, el aprendizaje por refuerzo se ha convertido en una herramienta poderosa para la toma de decisiones en entornos dinámicos y complejos del mercado de valores.

- Predicción de precios: El uso de ML permite la predicción del futuro precio de un activo financiero. Para ello se emplean técnicas de aprendizaje supervisado, donde el modelo se entrena con un conjunto de datos históricos con la finalidad de predecir el comportamiento futuro del activo. Este enfoque puede ser aplicado tanto a activos individuales como a carteras enteras. Entre las técnicas más utilizadas se encuentran la regresión lineal, los árboles de decisión, las redes neuronales y el aprendizaje por refuerzo. [21], [24], [18].

- Clasificación de patrones: El análisis técnico se basa en la identificación de patrones en los movimientos de precios históricos. El ML puede ser empleado para detectar y clasificar dichas tendencias. Esto permite ayudar a los individuos o instituciones a tomar decisiones informadas sobre cuándo comprar o vender un activo. Entre las técnicas de ML más empleadas para la clasificación de patrones se encuentran los árboles de decisión, las redes neuronales y el aprendizaje por refuerzo. [9]

- Análisis de sentimiento: El análisis de sentimiento se utiliza para evaluar noticias, informes financieros y otras fuentes de información con el fin de determinar el sentimiento del mercado. El ML puede ser empleado para realizar esta tarea más eficiente que los humanos, mediante técnicas de procesamiento de lenguaje natural y minería de datos. [13], [15], [17], [14]

- Gestión de riesgos: Los modelos de ML pueden ser empleados para valorar el riesgo asociado con una inversión y con la finalidad de ayudar a los traders a tomar decisiones informadas sobre la gestión de riesgos de sus activos. Por ejemplo, se pueden utilizar modelos de clasificación para evaluar el riesgo de que una empresa sea objeto de una investigación regulatoria. [25]

- Clasificación de activos: Consiste en emplear técnicas de ML para identificar patrones y tendencias en los precios de diferentes tipos de activos financieros, como materias primas, acciones, bonos, etc. Esto permite a los traders diversificar sus carteras de inversión y

reducir el riesgo asociado con la inversión en un solo tipo de activo del que se ha hablado en el apartado 2.2.

- Optimización de carteras: Consiste en emplear el ML para crear carteras de inversión que maximicen los rendimientos y se minimicen el riesgo en función de diferentes objetivos de inversión y restricciones. Esto es posible mediante el análisis de grandes cantidades de datos y la identificación de tendencias y relaciones entre diferentes activos. Con ello, es posible crear carteras personalizadas para cada cliente, teniendo en cuenta sus objetivos de inversión y su tolerancia al riesgo. [7], [1], [5]

2.6. Introducción al Aprendizaje por Refuerzo

Tal y como se ha definido antes, el Machine Learning es una rama de la Inteligencia Artificial, donde un modelo es capaz de aprender de manera autónoma a partir de los datos. Dentro de este campo, existe diversas sub-ramas, entre las cuales, está el Aprendizaje por Refuerzo (*Reinforcement Learning, RL*).

El aprendizaje por refuerzo se puede definir como la ciencia de tomar decisiones a partir de la interacción con el entorno. El RL consta de 5 elementos clave:

- **Agente**: Es el sistema o algoritmo que aprende y toma decisiones en un entorno determinado.
- **Entorno**: Es el sistema en el que el agente interactúa y toma decisiones.
- **Acciones**: Son las acciones que el agente puede realizar en el entorno.
- **Estado**: Es la representación de la situación actual del entorno en un momento dado.
- **Recompensa**: Es la retroalimentación que el agente recibe del entorno en función de las acciones realizadas y que le permite evaluar si una acción es buena o mala para su objetivo.

Un agente interactúa con un entorno y toma una serie de decisiones en función de su estado actual. A medida que el agente toma decisiones, el entorno proporciona una retroalimentación en forma de recompensas o castigos, lo que permite al agente evaluar la bondad de sus acciones.

A través de la repetición de este proceso, el agente aprende a maximizar su recompensa a largo plazo, ajustando continuamente su comportamiento y estrategias. Este proceso de aprendizaje por refuerzo es capaz de adaptarse a diferentes entornos y objetivos, lo que lo hace útil en una amplia variedad de aplicaciones, desde la robótica hasta la toma de decisiones en el

trading financiero.

Política y función de valor: Un agente dispone de:

- **Función de valor**: es una función que nos indica cuán bueno (en términos del posible retorno) es estar en un determinado estado o llevar a cabo una determinada acción.
- **Política**: rige la manera como el agente selecciona las acciones.

Una vez definida la política, es necesario contemplar lo siguiente:

Exploración vs. Explotación: es una dicotomía importante en el aprendizaje por refuerzo. La explotación implica elegir acciones que el agente ya ha utilizado y que se sabe que funcionan bien para maximizar las recompensas a corto plazo. Por otro lado, la exploración se enfoca en probar nuevas acciones que podrían tener un mayor potencial a largo plazo, pero que también implican más riesgo. En general, el objetivo del aprendizaje por refuerzo es encontrar el equilibrio óptimo entre la exploración y la explotación, lo que permite al agente obtener la mayor cantidad de recompensas en el menor tiempo posible, sin comprometer su capacidad para seguir aprendiendo y mejorando su desempeño a largo plazo.

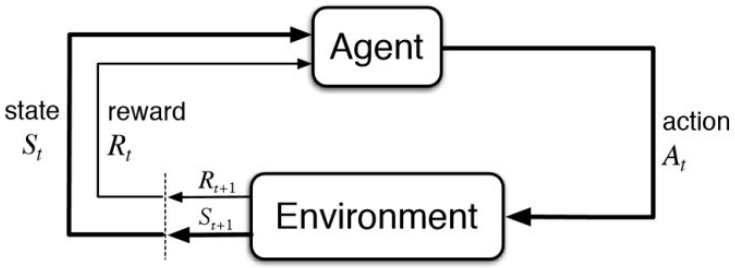


Figura 2.1: Ciclo del aprendizaje por refuerzo. Fuente: https://www.analyticsvidhya.com/

2.7. Aprendizaje por Refuerzo en el Trading.

El RL se ha aplicado con éxito en una amplia variedad de campos, entre ellos la finanza. En particular, el trading algorítmico es un campo en el que el RL ha demostrado ser muy efectivo. En este contexto, el objetivo del agente es aprender una estrategia de trading que maximice el rendimiento de la cartera de inversión a largo plazo.

Una de las principales ventajas del RL en el trading es su capacidad para adaptarse a un

entorno dinámico y desconocido. Los mercados financieros son volátiles y cambiantes, y el comportamiento de los precios de los activos es impredecible. El RL permite a los agentes adaptarse a estos cambios y tomar decisiones óptimas en un entorno altamente incierto.

Otra ventaja del RL en el trading es su capacidad para manejar grandes cantidades de datos y analizar patrones complejos en ellos. Los mercados financieros generan grandes cantidades de datos en tiempo real, y el RL puede procesar esta información para identificar patrones y tendencias en los precios de los activos.

Además de la predicción de precios, el RL también se utiliza en el trading para la gestión de riesgos. Los modelos de RL pueden ser entrenados para evaluar el riesgo asociado con una inversión y para desarrollar estrategias de gestión de riesgos para minimizar las pérdidas potenciales.

2.8. Algoritmos Seleccionados

En el aprendizaje por refuerzo existen numerosos algoritmos, los cuales se pueden clasificar según si se basan en aprendizaje por valores o por política, y si se utilizan técnicas de RL o DeepRL.

	Función de Valor	Política
RL	Q-Learning, Double Q- Learning, Weighted Q-Learning, Speedy Q-Learning, R-Learning, SARSA, Expected SARSA, True Online SARSA, FQI LSPI	REINFORCE, GPOMDP, eNAC, RWR, PGPE, REPS, COPDAC-Q, Stochastic ActorCritic
DRL	DQN, Double-DQN, Averaged DQN, Categorical DQN	A2C, DDPG, TD3, SAC, TRPO, PPO

Cuadro 2.1: Algoritmos de Aprendizaje por Refuerzo. Fuente propia.

Los algoritmos basados en valores aprenden una función de valor óptima, la cual indica el valor esperado de estar en un estado y seguir una política dada. Estos algoritmos estiman la función de valor óptima mediante la iteración de la política y la actualización de los valores de Q (o V) utilizando la regla de Bellman. Aunque estos algoritmos pueden ser efectivos en entornos con un número limitado de acciones, en entornos de alta dimensionalidad, como el trading o entornos financieros, su complejidad aumenta de manera exponencial con la cantidad

de estados y acciones posibles, lo que hace que su coste computacional sea elevado.

El problema principal de los algoritmos por función de valor es que la complejidad aumenta de manera exponencial con la cantidad de estados y acciones posibles, lo que lo hace computacionalmente inviable para un gran número de situaciones. No obstante, aunque se hayan implementado redes neuronales a los algoritmos de Q-valor, convirtiendo así el RL en DRL[2], sigue necesitando un coste computacional elevado.

Es por ello por lo que en esta Tesis se ha decidido emplear algoritmos de DRL de política. Estos algoritmos:

- Permiten trabajar de manera efectiva y óptima con entornos de alta dimensionalidad.
- Son más flexibles que los de la función de valor ya que aprenden directamente una política que maximiza la recompensa.
- En algunos casos, los algoritmos de DRL de política pueden tener un mejor rendimiento que los de aprendizaje por valor en entornos complejos. Esto se debe a que aprenden directamente una política que maximiza la recompensa, en lugar de tener que estimar una función de valor.

Entre las posibilidades dentro de los algoritmos de DRL de política, he decidido emplear los algoritmos PPO [20], DDPG [8] y A2C [23].

La selección del algoritmo PPO se basa en la investigación de Yang et al.[23], donde demuestra que el PPO es el que mayor retorno consigue a lo largo del tiempo en sus experimentos, pero con la mayor volatilidad de todos.

Este algoritmo puede ser interesante puesto que, las small caps son empresas poco afianzadas en el mercado y, con lo que hemos explicado en el punto 2.4., el mercado de valor esta basado en que los precios de una empresa sean incorrectos y se pueda obtener beneficios de ellas. Es por ello que el PPO puede llegar a obtener un mayor retorno de la inversión.

Cabe añadir que, en la obra *FinRL-Podracer: High Performance and Scalable Deep Reinforcement Learning for Quantitative Finance* [6] funciona muy bien en un entorno alcista.

Por otro lado, quiero intentar demostrar que el algoritmo DDPG, ante el entorno de las small

[2]DRL: Deep Reinforcement Learning

caps tiene un mejor resultado en comparación a la obra *Deep Reinforcement Learning for Automated Stock Trading: An Ensemble Strategy* [23] ya que muestra que entre el PPO, A2C y el DDPG, éste último es el que peor resultados tiene.

Antes habíamos definido que los algoritmos por valor daban muy buenos resultados pero a un costo computacional elevado. El DDPG es una combinación entre la función por valor y política, por lo que puede ser un buen candidato a obtener buenos resultados tal y como se comenta en la obra de Liu et al. [11], donde es tiene una gran resiliencia en épocas a la baja.

Asimismo, en referencia al A2C[23], es el que mejor Sharpe Ratio tiene de los 3 mencionados previamente, llegando a obtener un mayor rendimiento a un menor riesgo en comparacion con los otros dos, por lo que se adapta mejor a las tendencias bajista. Por otro lado, basándonos en la obra *FinRL: Deep Reinforcement Learning Framework to Automate Trading in Quantitative Finance* [10] se menciona que el A2C tiene el mejor desempeño mediante el uso de la librería FinRL en comparación con los dos anteriores.

Estos 3 algoritmos se emplearán en el entorno de small caps debido a que son empresas en las que algunas aún están por desarrollarse de manera sólida en el mercado y es una buena oportunidad obtener un mayor rendimiento al comprar acciones a menor precio de otras multinacionales más conocidas y sólidas.

En primer lugar, se realizará una comparación día a día con el artículo *Deep Reinforcement Learning for Automated Stock Trading: An Ensemble Strategy* [23]. Luego, se realizará un segundo experimento actualizando los datos cada 5 minutos, con el objetivo de comparar los resultados con el primer experimento y comprobar si es rentable actualizarlos con mayor frecuencia. Además, se tendrán en cuenta las comisiones de los brokers, el cual comprende entre un 2 y 3 % para cada operación.

Capítulo 3

Implementación

A lo largo de este capítulo, se presentarán los agentes implementados así como el entorno de este trabajo final de máster.

Empezaremos definiendo la estructura del proyecto y su distribución en la sección 3.1. Posteriormente, se empleará el conocimiento obtenido en el apartado 2.4 para presentar un enfoque del Mercado de Valores como un problema de Aprendizaje por Refuerzo según las características presentadas en la sección 2.6. Para ello, dividiremos este enfoque en los siguientes pasos.

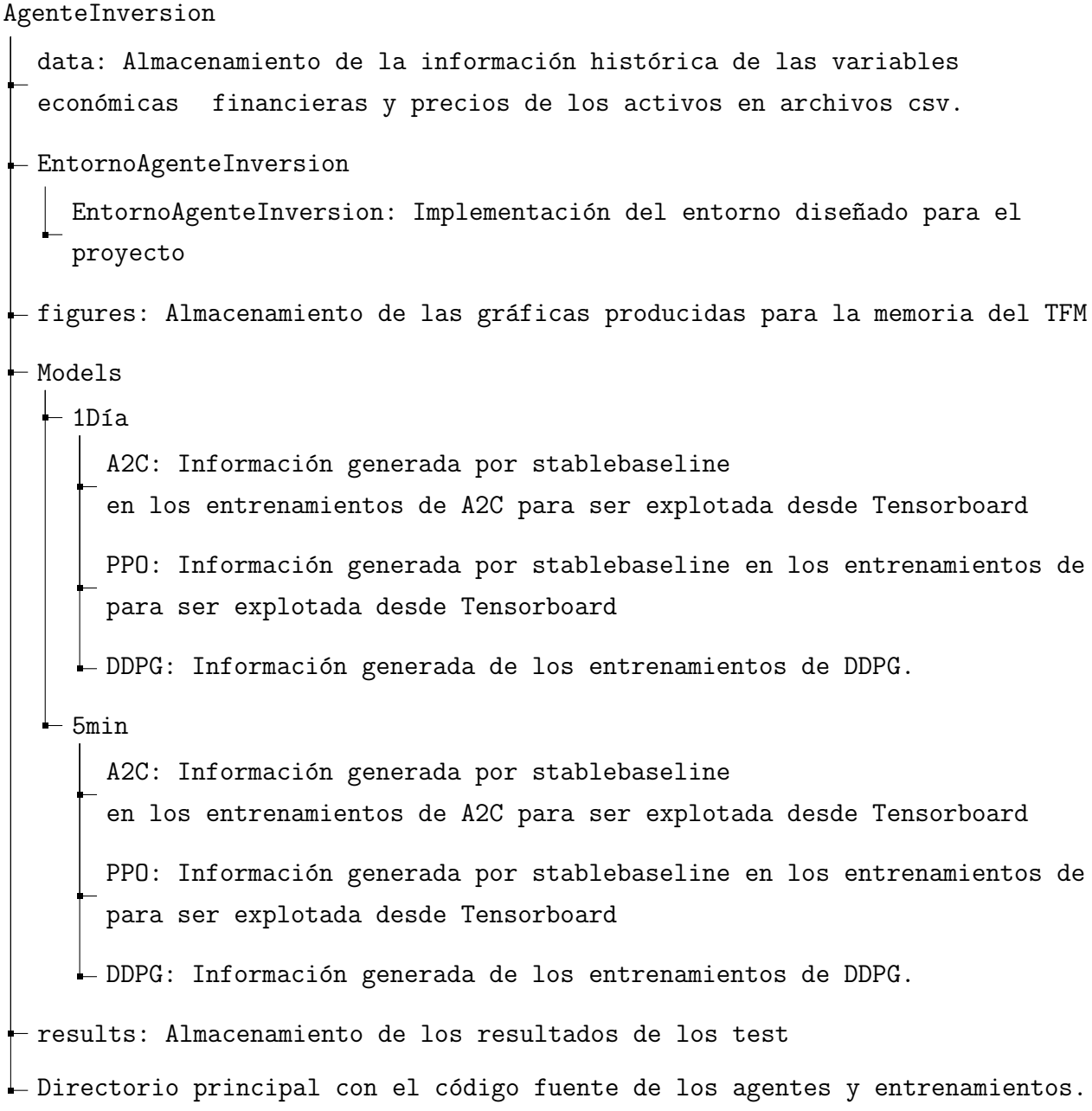
1. El primer paso sería construir un emulador del entorno, la Bolsa de Valores, usando las pautas de OpenAI Gym [3], es decir, el marco más extendido para desarrollar problemas de Aprendizaje por Refuerzo en la sección 3.2.

2. En segundo lugar, en la sección 3.3 tenemos que crear agentes capaces de interactuar con este entorno, utilizando Tensorflow, Pytorch y los frameworks más extendidos para desarrollar aplicaciones de Deep Learning.

3. Finalmente, en la sección 3.4 tenemos que implementar los test destinados a comprobar los agentes construidos en los pasos anteriores para el conjunto de datos que son para cada día y para cada 5 minutos.

3.1. Estructura del proyecto

A continuación se presenta la estructura del proyecto que se ha realizado en este trabajo final de máster. Dicha arquitectura se puede encontrar en el siguiente repositorio GitHub: https://github.com/isuchb/TFM.



3.2. Entorno de aprendizaje

Para poder modelar un agente inversor que, pueda operar en bolsa, hay que definir previamente cómo percibe este entorno. Esto es, tenemos que construir un simulador de la inversión bursátil.

Para ello, el primer paso es analizar cómo una persona realiza la acción de invertir y adaptar este contexto a un problema de aprendizaje por refuerzo. Para crear este simulador debemos definir los siguientes conceptos:

Observaciones

El inversor conoce tanto la evolución histórica de una acción como el valor actual de la misma. Asimismo, también posee información sobre de la cantidad de dinero disponible para invertir y la cantidad de acciones que actualmente posee.

Acciones

Hay tres acciones posibles en cualquier paso de tiempo con respecto a una acción: mantener, comprar o vender.

- **Comprar:** Esta acción supone un aumento de la posesión de acciones. Para ello se realiza una compra de cierta cantidad de valores y se almacena el precio de entrada en ese momento para calcular una futura ganancia o perdida patrimonial.
- **Mantener:** Permite el paso de tiempo sin aumentar o disminuir una posesión. Se mantiene constante en el tiempo.
- **Vender:** Desprenderse de una acción o conjunto de acciones a cambio de capital. El precio obtenido por la venta menos el precio de entrada permite obtener la ganancia o pérdida de la operación realizada.

Recompensa

La forma habitual de analizar la evolución de los precios en Bolsa y construir carteras rentables no es analizando directamente los precios de venta sino calculando las rentabilidades. Este cálculo lo podemos realizar con la siguiente aproximación:

$$Recompensa = Precio_{\text{Venta}} - Precio_{\text{Compra}} \tag{3.1}$$

Una vez mencionadas las acciones disponibles que se pueden realizar, Yang et al. [23] sugieren y disponen de un entorno que permite modelar la naturaleza dinámica del mercado de valores empleando un proceso de decisión de Markov (MDP) del siguiente modo:

- **Estado** $s = [p, h, b]$: consta de un vector con los precios de las acciones $p \in \mathbb{R}_D^+$ otro vector con la distribución de las acciones en la cartera $h \in \mathbb{Z}_D^+$ y por último el saldo restante $b \in \mathbb{R}^+$ donde D denota el número de acciones y $\mathbb{Z}^+$ denota números enteros no negativos.
- **Acción** α: denota un vector con los movimientos realizados sobre las acciones D. Para cada valor está permitido comprar, vender o mantener lo que resulta en una aumento, disminución o ningún cambio en los valores de h respectivamente.

- **Recompensa** $r(s, a, s')$: la recompensa directa de ejecutar a en el estado s y llegar al nuevo estado s'.
- **Política** $\pi(s)$: se define como la estrategia comercial en el estado s. Esta es la distribución de probabilidad de las acciones dado el estado s.
- **Valor** $Q_\pi(s, a)$: la recompensa esperada de tomar la acción α en el estado s siguiendo la política π.

  En contra parte, hay que tener en consideración las siguientes restricciones:

- **Liquidez del mercado:** Las órdenes se pueden ejecutar rápidamente al precio de cierre. En este estudio suponemos que el mercado de valores no se verá afectado por nuestro agente.
- **Saldo no negativo** $b \geq 0$: las acciones permitidas no deben obtener un saldo negativo de la cuenta.

Se puede definir matemáticamente del siguiente método:

Si nos encontramos en el día t con un conjunto de acciones D en el que podemos invertir, se pueden dividir estos valores en un conjunto para vender S, otro para comprar B y otro para mantener Hdonde

$$S \cup B \cup H = \{1, \ldots, D\}$$

y no se superponen.

Una vez descrito, se puede definir $p_t^B = [p_t^i : i \in B]$ y $k_t^B = [k_t^i : i \in B]$ como los vectores de precio y número de acciones los cuales pertenecen al conjunto de compra B. De manera similar, se define p_t^S y k_t^S para conjunto de venta y p_t^H y k_t^H para el conjunto de mantener.

A partir de estas definiciones, la restricción para el saldo no negativo se puede expresar como:

$$b_{t+1} = b_t + (p_t^S)^T k_t^S - (p_t^B)^T k_t^B \geq 0 \tag{3.2}$$

Coste de transacción: Hay muchos tipos de coste de transacción como tarifas de cambio tarifas de ejecución y tarifas de la SEC.

Cada plataforma posee diferentes tarifas y comisiones para las acciones de compra, venta e incluso mantenimiento de la cuenta con acciones. A pesar de estas variaciones asumimos que nuestros costos de transacción son 0.003 % del valor de cada operación (ya sea compra o venta) y lo definimos con la siguiente ecuación:

$$c_t = p^T k_t \cdot 0{,}003\,\% \tag{3.3}$$

Cabe remarcar que hay que añadir este coste a la ecuación (3.2) obteniendo el siguiente resultado:

$$b_{t+1} = b_t + ((p_t^S)^T k_t^S - c_t^S) - ((p_t^B)^T k_t^B + c_t^B) \geq 0 \tag{3.4}$$

Salida de emergencia por exceso de volatilidad: Existen eventos esporádicos (y a veces aleatorios) que pueden causar el colapso del mercado de valores como guerras, burbujas, pandemias, incumplimiento de la deuda soberana y/o crisis financieras. Para controlar el riesgo en los peores casos, como la crisis financiera mundial de 2008, se hace uso de un índice de turbulencia financiera (*turbulencet*) que mide los movimientos extremos de los precios de los activos:

$$\text{turbulence} = (y_t - \mu)(\sum)^{-1}(y_t - \mu) \tag{3.5}$$

donde:

- $y_t \in \mathbb{R}^D$ denota los rendimientos de las acciones para el período t
- $\mu \in \mathbb{R}^D$ denota los rendimientos medios históricos
- $\Sigma \in \mathbb{R}^{D \times D}$ denota la covarianza del rendimiento histórico.

Cuando la turbulencia es superior a un umbral es decir las condiciones de mercado son extremas simplemente se debe dejar de comprar acciones y el agente debería vender toda su cartera. Las operaciones se reanudarán una vez que el índice de turbulencia regresa a un valor estable.

Una vez definidas las bases del entorno el siguiente punto es definir la recompensa del agente. Se establece la formula matemática que represente los beneficios de un trader. La función de recompensa se define como el cambio del valor de la cartera cuando se ejecute la acción a en el estado s y se llega al nuevo estado s'. El objetivo es diseñar una estrategia comercial que maximice el cambio del valor de la cartera:

$$r(s_t, a_t, s_{t+1}) = (b_{t+1} + p_{t+1}^T h_{t+1}) - (b_t + p_t^T h_t) - c_t \tag{3.6}$$

3.2.1. Preparación del entorno

Los indicadores económicos y variables son elementos fundamentales para el análisis y seguimiento de la salud financiera de una economía, empresa o país. Dichas medidas proporcionan una información clave con las cuales se pueden hacer un seguimiento de la salud del mercado laboral, inflación, desempeño económico y el crecimiento, entre otros aspectos. Dicha información es crucial para comprender el rumbo por el que va la economía.

Este trabajo final de máster se centrará en el índice de bolsa Standard & Poors 600 (S&P600) donde engloban el mercado de las empresas Small Cap.

Para nuestro mercado de valores ficticio, se han seleccionado 500 valores de los 600 que se pueden consultar en la tabla que aparece en el Apéndice F de manera aleatoria. Esto es debido a que, al ser empresas con mayor volatilidad que las del S&P500, puede ser que hayan quebrado, hayan ascendido en capitalización o hayan descendido. Todos esos escenarios y algunos otros, pueden hacer que hayan tickers que hayan desaparecido dentro de la tabla o que falte información creando errores dentro del comportamiento del agente. Asimismo, de estas 500 empresas, se han dividido en dos grupos:

1. *Conjunto de Entrenamiento*

2. *Conjunto de Prueba*

donde en ambos casos son de 250 acciones. De este modo, lo que se pretende es que tengamos un balance y evitar el sesgo entre el entrenamiento y el test. Si hubiéramos hecho los 500 para entrenamiento y emplear los mismos para el test, es probable que hubiera aparecido un *overfitting*.

Todos los valores en los que operamos se encuentran dentro del mercado de valores americano, lo que permite reducir el coste de las comisiones.

Con esta información, podemos realizar la configuración del estado de la siguiente forma:

1. **Estado (state):** Empleamos un vector de 1501 dimensiones que consta de siete partes de información para representar el espacio de estado del entorno de negociación de múltiples acciones: $[b_t, p_t, h_t, M_t, R_t, C_t, X_t]$. Cada componente se define de la siguiente manera:

   - $b_t \in \mathbb{R}^+$: saldo disponible en el paso de tiempo actual t.
   - $p_t \in \mathbb{R}_+^{250}$: precio de cierre ajustado de cada acción.
   - $h_t \in \mathbb{Z}_+^{250}$: acciones en cartera.
   - $M_t \in \mathbb{R}^{250}$: La divergencia de convergencia de la media móvil (MACD) para cada valor. El MACD es un indicador de seguimiento de tendencias que muestra la relación

entre dos medias móviles del precio de un activo. Típicamente, se calcula restando la media móvil exponencial (EMA) de 26 períodos de la EMA de 12 períodos. El resultado de esto es la línea MACD. Luego, una EMA de 9 días del MACD llamada la "línea de señal,"se dibuja encima del MACD, que puede funcionar como un disparador para señales de compra o venta. Cuando la línea MACD cruza por encima de la línea de señal, puede ser una señal de compra, y cuando cruza por debajo, una señal de venta.

- $R_t \in \mathbb{R}_+^{250}$: El índice de fuerza relativa (RSI) por valor. El RSI es un oscilador de momento que mide la velocidad y cambio de los movimientos de precios. El RSI oscila entre cero y 100 y se considera sobrecomprado cuando está por encima de 70 y sobrevendido cuando está por debajo de 30. Las señales se pueden generar mediante la búsqueda de divergencias, cruces de línea de centro y lecturas de sobrecompra o sobrevendida.
- $C_t \in \mathbb{R}_+^{250}$: El índice Osciladores de Momento (CCI). El CCI compara el precio actual de un activo con su precio promedio en un período de tiempo. El indicador oscila por encima y por debajo de cero. Valores altos muestran que el precio está inusualmente alto comparado con el precio promedio, y valores bajos indican que el precio es inusualmente bajo. Generalmente, si el CCI es superior a +100, se considera que el activo está en una fuerte tendencia alcista y podría ser una señal de compra. Por el contrario, un CCI por debajo de -100 puede señalar una tendencia bajista y sería una señal de venta.
- $X_t \in \mathbb{R}^{250}$: El índice medio direccional (ADX). El ADX es un indicador de fuerza de tendencia. No indica la dirección de la tendencia, sino que mide su fuerza, ya sea al alza o a la baja. El ADX se calcula basándose en el rango de movimiento de precios en un período de tiempo dado y es una media móvil de la expansión del rango de precios. Un valor ADX por encima de 25 sugiere una tendencia fuerte, mientras que un valor por debajo de este umbral podría indicar una tendencia débil o la ausencia de tendencia.

2. **Espacio de acción (action space):** para una sola acción, el espacio de acción se define como $\{-k, \ldots, -1, 0, 1, \ldots, k\}$ donde k y $-k$ representan el número de acciones que podemos comprar y vender, y $k \leq h_{\max}$, siendo $h_{\max}$ un parámetro predefinido que establece como la cantidad máxima de acciones permitidas por compra.

   Por ello, el tamaño de todo el espacio de acción es $(2k+1)^{250}$. Luego, el espacio de acción se normaliza a $[-1, 1]$, ya que los algoritmos PPO definen la política directamente en una distribución Gaussiana que debe ser normalizada y simétrica [19].

3.2.2. Obtención de los datos de entrenamiento

A continuación se describe la obtención de los datos empleados para este trabajo final de máster, los cuales están descritos en la tabla anexa en el Apéndice F.1.

Prosiguiendo con el análisis del módulo, podemos ver el archivo 1 procesamiento datos extracción_datos.py cuya finalidad es recuperar para cada valor su precio de apertura, cierre, de la compra más cara, la venta más barata y el volumen de acciones comerciadas por sesión desde 16 de octubre de 2017 hasta el 12 de abril de 2023 para 1 día, y desde el 27 de julio de 2023 hasta el 10 de diciembre de 2023 para cada 5 minutos. El resultado de estos ficheros es la creación del documento **1día.csv** y **5min.csv** que contiene todos los datos obtenidos.

Estos datos han sido obtenidos a partir de Yahoo Fin, una librería de Python 3. Dicho paquete fue diseñado con la finalidad de poder extraer datos históricos de precios de acciones, así como para proporcionar información actual sobre límites de mercado, rendimientos de dividendos y qué acciones comprenden las principales bolsas.

En el fichero **postprocessed.ipynb** añadiremos la información económica MACD, RSI, CCI y ADX a cada valor. Adicionalmente, se calculará el índice de volatilidad definido en la ecuación 3.2 El resultado de este proceso es la creación de un fichero done data.csv donde se puede consultar las 5 primeras filas de un fichero de ejemplo:

datadate	tic	adjcp	open	high	low
2017-10-23	ABCB	44.418537	48.599998	47.849998	48.000000
2017-10-23	ABG	56.049999	56.500000	55.549999	56.049999
2017-10-23	ABM	38.011845	42.880001	42.310001	42.389999
2017-10-23	ACLS	33.099998	33.250000	32.650002	33.099998
2017-10-23	ADTN	18.942442	22.100000	21.549999	21.600000

Cuadro 3.1: Ejemplo de las primeras 5 filas del fichero 1día.csv con las primeras 6 columnas. Tabla propia.

volume	macd	rsi	cci	adx	turbulence
126600.0	-0.082124	17.596425	-75.143931	47.518253	0.000000
319400.0	-0.068000	40.709900	-26.832254	8.682980	0.000000
168100.0	-0.008431	31.550213	-92.760045	2.944239	0.000000
616700.0	0.126690	93.749129	94.661890	100.000000	0.000000
386100.0	-0.111675	4.025454	-71.636742	100.000000	0.000000

Cuadro 3.2: Ejemplo de las primeras 5 filas del fichero 1día.csv con las últimas 6 columnas. Tabla propia.

Para obtener la información sobre los índices mencionados previamente, se ha usado la librería stockstats que, a partir de la información obtenida del fichero fichero sin tratar.csv permite calcular estas métricas.

3.2.3. Implementación del entorno

El entorno se define como todo aquello que es independiente del agente, y que dicho agente interactúa con él y aprenda. En este trabajo, se ha empleado una versión del entorno creado en la obra *Deep Reinforcement Learning for Automated Stock Trading: An ensemble strategy* [23]. Este diseño se realiza de acuerdo con las directrices de OpenAI [3], la cual es una de las bibliotecas más importantes y empleadas para el desarrollo de aplicaciones de RL. Para cumplir con los requisitos de Gym y, por ende, utilizarlo, se han definido los siguientes métodos:

1. **init(self)** : En este apartado se declara la información inicial del entorno como el action space, observation space, etc. Recibirá como parámetro el documento **processed_1d.csv** el saldo inicial y los costes por transacción.

2. **reset(self)**: Retorna el entorno al estado inicial.

3. **step(self, action)**: Avanza el entorno en un día. Este método solicita al agente qué operación de compra, mantener o venta va a realizar y devuelve como resultado la siguiente información:

   a) *Observation*: El estado en que se queda el entorno después de la acción.

   b) *Reward*: La recompensa acumulada hasta al momento. En este caso es el valor de la

   c) *Done*: Si ha llegado al último día disponible para realizar trading.

4. **render(self)**: Devolvemos el estado. Aparte de estas especificaciones, la clase StockTradingEnv implementada en el fichero envMultipleStock.py tiene otros atributos y métodos cuyo análisis en profundidad exceden el objetivo de esta memoria. Es por ello por lo que se incita al lector interesado a realizar un análisis del código para obtener más información.

3.3. Implementación de los agentes

En este apartado, se analizarán las bases del algoritmo A2C, DDPG y PPO, siendo este último el actual estado del arte según el artículo [11]

Primeramente, se analizará el algoritmo PPO en la sección 3.3.1. A continuación analizaremos el algoritmo DDPG en la sección 3.3.2 y por último el algoritmo A2C en la sección 3.3.3.

3.3.1. PPO

Descrito en [20], el algoritmo PPO A emerge tras evaluar los puntos débiles y fortalezas de métodos preexistentes. Su principal ventaja radica en la habilidad de conservar proximidad entre políticas nuevas y antiguas, logrando esto a través de un proceso notablemente simplificado.

La capacidad de evitar desvíos erróneos en el entrenamiento de algoritmos convencionales es crucial. Tales desvíos a menudo impiden un avance adecuado en el aprendizaje.

Enfoques tradicionales buscan prevenir cambios abruptos y otorgar consistencia al aprendizaje limitando cuánto puede divergir una nueva política de su predecesora, lo cual requiere cálculos intensivos.

La complejidad de estos cálculos, necesarios para ejecutarse millones de veces, se aborda en PPO a través de una simplificación ingeniosa. Se introduce un elemento de ajuste elemental en la función de objetivo, permitiendo restringir variaciones en las políticas con menor demanda computacional.

Considérese la relación de probabilidades entre políticas viejas y nuevas, descrita por:

$$r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{\text{Old}}}(a_t|s_t)} \tag{3.7}$$

Aquí, $\pi_\theta(a_t|s_t)$ representa la política de la red neuronal. Dicha relación, conocida como muestreo de importancia, puede mitigar actualizaciones excesivas.

La función objetivo alternativa de PPO se define como:

$$J_{\text{CLIP}}(\theta) = E_t\left[\text{mín}\left(r_t(\theta)\hat{A}(s_t,a_t), \text{clip}\left(r_t(\theta), 1-\epsilon, 1+\epsilon\right)\hat{A}(s_t,a_t)\right)\right] \tag{3.8}$$

Donde:

- θ representa la política actual.
- E_t simboliza la expectativa matemática en un tiempo t.
- r_t es la razón de probabilidades entre políticas nuevas y viejas.
- $\hat{A}(s_t,a_t)$ se refiere a la ventaja estimada en el tiempo, un cálculo de la diferencia entre el resultado obtenido y lo esperado. Un valor positivo incrementa la probabilidad de seleccionar la acción (s_t,a_t), mientras que un valor negativo la disminuye.
- $\hat{A}(s_t,a_t) = \text{DiscountedReward} - \text{ValueFunction}$
- ϵ es un hiperparámetro, usualmente establecido en 0.2.

En este contexto, $r_t(\theta)\hat{A}(s_t, a_t)$ es la función objetivo habitual en el gradiente de política, y la función $clip(r_t(\theta), 1-\epsilon, 1+\epsilon)$ limita la relación $r_t(\theta)$ dentro del rango $[1-\epsilon, 1+\epsilon]$. La función objetivo de PPO utiliza el menor valor entre la meta ajustada y la estándar, mejorando así la estabilidad durante la formación de redes al restringir la actualización de la política en cada etapa de entrenamiento.

Las ventajas principales del PPO incluyen:

- Una implementación más directa y eficiente desde el punto de vista computacional.
- Versatilidad y efectividad en una amplia gama de problemas y entornos, independientemente de su complejidad.
- Aplicabilidad en entornos con espacios de acción tanto continuos como discretos.

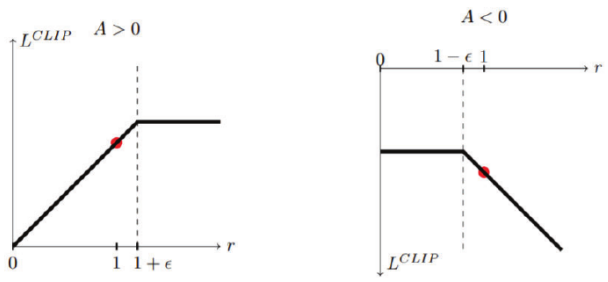


Figura 3.1: Impacto positivo y negativo en la función objetivo de PPO. El punto rojo marca el comienzo de la optimización. Se suavizan las acciones beneficiosas para la política, mientras que se incorporan directamente las perjudiciales. Fuente [20].

Entre las limitaciones de PPO, destaca la necesidad de datos recientes y relevantes para realizar predicciones efectivas. Por ejemplo, el uso de información del periodo 2017 a 2023 para anticipar tendencias del mercado de valores en la década de 2020 podría llevar a una actualización deficiente de la política. Es esencial contar con muestras adecuadas y representativas de la realidad.

3.3.2. DDPG

El algoritmo *Deep Deterministic Policy Gradient* (DDPG) B es una implementación actor-crítico que combina aspectos de los algoritmos de gradiente de política con el clásico DQN (Deep Q-Network). Es especialmente útil en entornos dinámicos como el mercado de valores. DDPG utiliza dos redes neuronales separadas: una para la función de valor y otra para la política.

El proceso de aprendizaje en DDPG se puede describir en los siguientes pasos:

1. Generación de un conjunto de experiencias aleatorias, representadas como (s, a, r, s', d), donde d es 1 si el estado es terminal o 0 en caso contrario.

2. Uso de la ecuación de Bellman para aprender la función Q.

3. Aprendizaje de la política de forma determinista a partir de esta función.

Matemáticamente, el proceso se describe de la siguiente manera:

- La acción óptima para cualquier estado s se determina por:

$$a^*(s) = \arg\max_a Q^*(s, a)$$

- La red neuronal $Q_\phi(s, a)$ se entrena para satisfacer la ecuación de Bellman. El error cuadrático medio de Bellman (MSBE) se utiliza para ajustar $Q_\phi(s, a)$:

$$L(\phi, M) = E_{(s,a,r,s',d)\sim M}\left[Q_\phi(s, a) - \left(r + \gamma(1 - d)\max_{a'} Q_\phi(s', a')\right)\right]^2$$

- El objetivo es aprender una política determinista parametrizada $\mu_\phi(s)$ que devuelva la acción que maximiza $Q_\phi(a, s)$. Los parámetros de la política ϕ se entrenan con un paso de ascenso del gradiente.

DDPG se ha aplicado con éxito en una variedad de entornos, demostrando su versatilidad y eficacia en el aprendizaje por refuerzo.

3.3.3. A2C

El algoritmo *Advantage Actor-Critic* (A2C) C es una variante del enfoque Actor-Critic en aprendizaje por refuerzo. A2C se diferencia de las versiones tradicionales de Actor-Critic por utilizar la ventaja para actualizar las políticas, lo que reduce la varianza en las actualizaciones de la política y mejora la eficiencia del aprendizaje.

En A2C, la política (actor) y la función de valor (crítico) se actualizan simultáneamente. La ventaja, que es la diferencia entre el retorno esperado y el valor estimado, guía el proceso de aprendizaje:

$$A(s, a) = Q(s, a) - V(s) \tag{3.9}$$

donde $A(s, a)$ es la ventaja, $Q(s, a)$ es el valor de acción-estado, y $V(s)$ es el valor del estado.

El algoritmo A2C se puede resumir en los siguientes pasos:

1. Recopilación de experiencias por parte del actor basándose en la política actual.
2. Cálculo de la ventaja $A(s, a)$ utilizando las recompensas y la función de valor $V(s)$.
3. Actualización de la política del actor utilizando la ventaja.
4. Actualización de la función de valor $V(s)$ del crítico basándose en las recompensas y los valores estimados.

La actualización de la política se realiza utilizando el gradiente de la función objetivo con respecto a los parámetros de la política:

$$\nabla_\theta J(\theta) = E_{\pi_\theta} \left[\nabla_\theta \log \pi_\theta(a|s) A(s, a) \right] \tag{3.10}$$

donde π_θ es la política parametrizada por θ, y $J(\theta)$ es la función objetivo.

A2C ha demostrado ser eficaz en una amplia gama de entornos de aprendizaje por refuerzo, tanto en espacios de acción discretos como continuos, facilitando la recopilación eficiente de experiencias y la actualización de políticas y funciones de valor.

3.4. Elaboración de los test

En esta sección, se detallan los procedimientos y metodologías empleadas para realizar las pruebas con el índice S&P600. Se han considerado dos enfoques distintos basados en la frecuencia de los datos: pruebas con datos diarios y pruebas con datos de alta frecuencia (cada 5 minutos). Estos enfoques permiten evaluar el rendimiento de los algoritmos de trading en diferentes condiciones de mercado y con diferentes granularidades de datos.

3.4.1. Cada 1 Día

Las pruebas con datos diarios se realizaron utilizando información del índice S&P600 desde el 16 de octubre de 2017 hasta el 12 de abril de 2023. Este conjunto de datos incluye las fluctuaciones diarias del mercado, lo que permite analizar cómo los algoritmos de trading se adaptan y reaccionan a las tendencias y cambios a largo plazo en el mercado.

Los principales objetivos de estas pruebas son:

- Evaluar la capacidad de los algoritmos para capturar tendencias a largo plazo y generar estrategias de trading consistentes.
- Analizar la adaptabilidad de los algoritmos a las variaciones del mercado a lo largo de un período prolongado.

- Comparar el rendimiento de los algoritmos en términos de retorno ajustado al riesgo y la estabilidad de las ganancias a lo largo del tiempo.

Las métricas clave consideradas en esta evaluación incluyen el retorno total, la volatilidad, el ratio de Sharpe, y la reducción máxima.

3.4.2. Cada 5 minutos

La segunda parte de las pruebas se centró en el análisis de datos de alta frecuencia, concretamente en intervalos de 5 minutos. Estos datos abarcan desde el 27 de julio de 2023 hasta el 10 de diciembre de 2023. El uso de datos de alta frecuencia permite evaluar cómo los algoritmos gestionan la volatilidad a corto plazo y las oportunidades de trading intradía.

Los objetivos específicos de las pruebas con datos cada 5 minutos incluyen:

- Investigar la eficacia de los algoritmos en la identificación de oportunidades de trading a corto plazo y su capacidad para reaccionar rápidamente a los cambios del mercado.
- Evaluar la robustez de los algoritmos frente a la volatilidad intradía y su habilidad para gestionar el riesgo en un entorno de mercado de alta frecuencia.
- Analizar el impacto de los costes de transacción y el deslizamiento en el rendimiento del trading.

En este contexto, además de las métricas tradicionales de rendimiento, se presta especial atención a la frecuencia y eficacia de las operaciones, el impacto del deslizamiento y los costes de transacción en los resultados.

En ambas secciones, se realizaron pruebas exhaustivas para garantizar la validez y fiabilidad de los resultados, proporcionando una comprensión profunda del rendimiento de los algoritmos de trading bajo diversas condiciones de mercado.

Capítulo 4

Análisis y comparativa de los resultados

En este capítulo conclusivo, se aplicarán los entornos y agentes desarrollados en el capítulo previo, tal como se detalla en la Sección 3. Posteriormente, se demostrarán los resultados obtenidos a partir de las pruebas realizadas en el entorno y los agentes.

En la sección 4.1 se expone las métricas empleadas para realizar la comparación entre los agentes tanto para el escenario de 1 día como para 5 minutos. Por último, en la sección 4.3 se exponen las principales conclusiones obtenidas y en la sección 4.4 se detallan los posibles trabajos y líneas de investigación en el futuro.

4.1. Resultados

Para evaluar los modelos entrenados, se realizado una serie de tests en donde se ha empleado un set de datos totalmente diferente al que han sido los modelos. Recordemos que durante el entrenamiento se han empleado los 250 primeras acciones del S&P600. En el test se han usado los siguientes 250. Para cada uno de los 3 agentes (PPO, DDPG y A2C) se han entrenado con la siguiente cantidad de *steps* o pasos:

- 1 Millón
- 2 Millones
- 5 Millones
- 10 Millones

Para evaluar los modelos, han sido testeados habiendo ejecutado los agentes PPO A y A2C C 100 iteraciones debido a que presentan una naturaleza estocástica. No obstante, el agente DDPG B, se ha realizado una sola iteración ya que es un agente con naturaleza determinista:

- **End Total Asset:**
  - PPO: Distribución tipo violín de valores del valor de la cartera al final en millones de euros.
  - A2C: Distribución tipo violín de valores del valor de la cartera al final en millones de euros.
  - DDPG: Valor final de la cartera en millones de euros.
- **Total Reward:**
  - PPO: Distribución tipo violín de valores de la recompensa total final en millones de euros.
  - A2C: Distribución tipo violín de valores de la recompensa total final en millones de euros.
  - DDPG: Recompensa final de la cartera en millones de euros.
- **Total Cost:**
  - PPO: Distribución tipo violín de valores del costo total en miles de euros.
  - A2C: Distribución tipo violín de valores del costo total en miles de euros.
  - DDPG: Costo total final en miles de euros.
- **Total Trades:**
  - PPO: Distribución tipo violín de valores del número total de operaciones realizadas.
  - A2C: Distribución tipo violín de valores del número total de operaciones realizadas.
  - DDPG: Número total final de operaciones realizadas.
- **Sharpe Ratio:**
  - PPO: Distribución tipo violín de valores del ratio de Sharpe.
  - A2C: Distribución tipo violín de valores del ratio de Sharpe.
  - DDPG: Valor final del ratio de Sharpe.

4.1.1. 1 Día

En esta sección, se detallan los resultados obtenidos para los test de los modelos que han sido entrenados durante 1, 2, 5 y 10 Millones de pasos. Se emplaza al lector interesado a realizar una lectura más profunda sobre los resultados obtenidos en el apéndice E.

Para los modelos entrenados con un conjunto de datos que, se han ido actualizando de manera diaria desde el 16 de octubre de 2017 hasta el 12 de abril de 2023, se detallan los resultados a continuación:

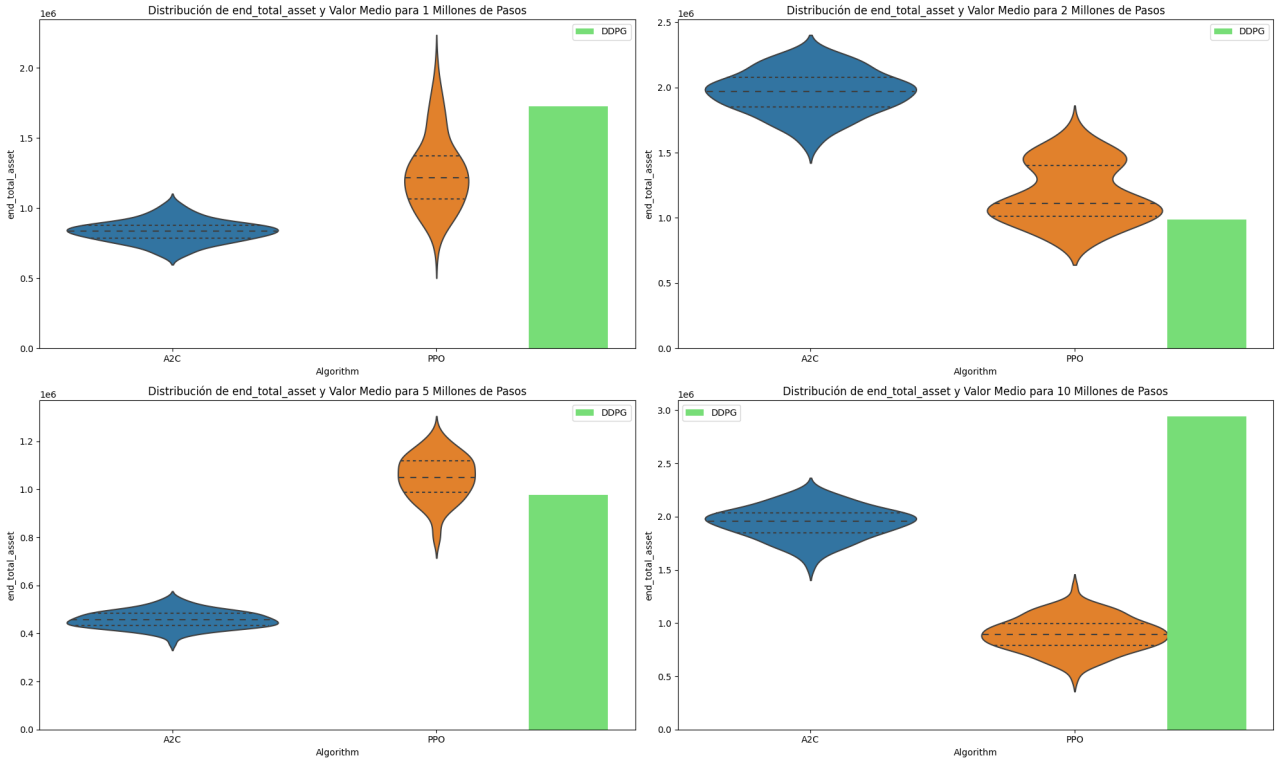


Figura 4.1: Estadística del mínimo, media y máximo para el valor de la cartera el último día para 1, 2, 5 y 10 millones de pasos. Fuente propia.

En esta figura, se han calculado los datos del siguiente modo:

- Para los agentes PPO y A2C, se han realizado 100 iteraciones a modo de test, en los cuales, se han representado los valores mínimos, medios y máximos de los datos recogidos en dichas iteraciones. Los valores recogidos simbolizan el valor final de la cartera (End Total Asset).
- En el contexto del algoritmo DDPG, se realizó únicamente una iteración al tratarse de un gradiente de política con naturaleza determinista, tal y como se indica en el apartado

2.8. Esto se debe a que, a diferencia de los métodos que incorporan actores críticos con un componente de aleatoriedad, el DDPG carece de esta variabilidad en su proceso de toma de decisiones. Gracias a su naturaleza, el DDPG tiende a seleccionar consistentemente los mismos valores, resultando así en una salida predecible y uniforme en cada ejecución.

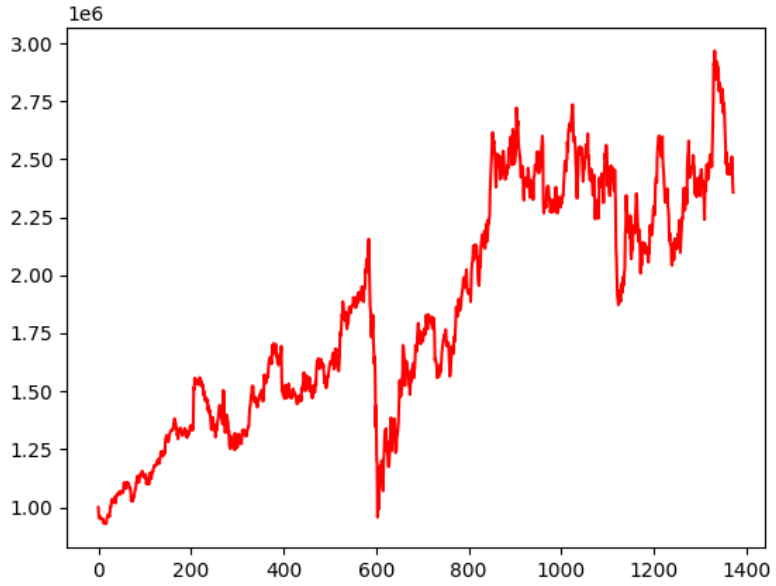


Figura 4.2: Evolución del entrenamiento del agente DDPG en una de sus iteraciones. Fuente propia.

A partir de la figura anterior E.1, podemos observar una línea roja que atraviesa las gráficas. Esta contempla el valor inicial de la cartera, que era de 1 millón de dólares americanos. Además, se pueden obtener las siguientes conclusiones:

- El modelo que mejor rendimiento ha tenido de los 3 ha sido el DDPG, desbancando totalmente al A2C y al PPO. Éste B, ha llegado a aumentar en un 300 % su valor final respecto el inicio. Esto permite que obtenga un rendimiento medio anual del 33,33 %.
- El modelo PPO A, ha obtenido un rendimiento muy por debajo de lo esperado. Ya que, como se menciona en la obra de *Deep reinforcement learning for automatedstock trading: An ensemble strategy* [23] marcaba el estado del arte. En este escenario, ha sido todo lo contrario. Ha sido el que peor rendimiento ha tenido de los 3.
- El A2C C logra posicionarse entre el DDPG y el PPO. Este hito sorprende también ya que, en el artículo mencionado, se encuentra como el que peor rendimiento tiene. No

obstante, en este escenario se ha obtenido un rendimiento medio del 100 %, obteniendo mínimo 500.000 y un máximo de 1.250.000 dólares.

4.1.2. 5 Min

Para este escenario, se han entrenado los agentes con 2,5 Millones de pasos. La diferencia de por qué para un 1 día se ha empleado más cantidad de pasos radica en que el set de datos es 3 veces menor que el de cada 5 minutos. Al estar entrenando un set de datos de 3,5 millones de líneas, obliga a reducir el número de pasos a realizar debido a que no se posee un hardware con prestaciones superiores. En caso contrario, se hubiera realizado para 1, 5 y 10 Millones de pasos.

Para los modelos entrenados con un set de datos que, se han ido actualizando de manera continuada cada 5 minutos desde el 27 de julio de 2023 hasta el 10 de diciembre de 2023, se detallan los resultados a continuación:

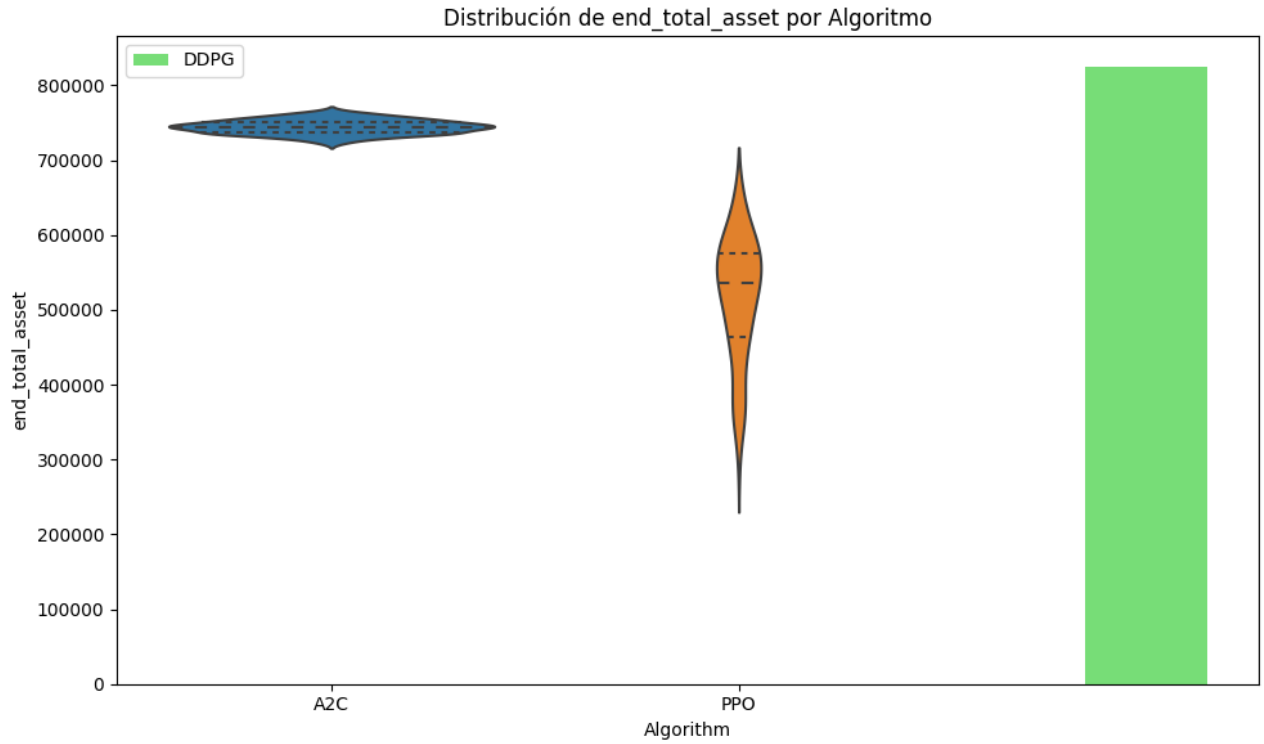


Figura 4.3: Resultado de evaluación de los agentes A2C, PPO y DDPG con actualización cada 5 minutos. Fuente propia.

En esta visualización 4.3, se representan los valores mínimos, medios y máximos de los agentes evaluados a lo largo de este trabajo final de máster. En ella, podemos ver claramente

como el rendimiento negativo de los 3 agentes. Podemos fijarnos en que, aun teniendo un set de datos mayor que el diario, no ha sido suficiente como para poder mejorar, o al menos, superar la barrera del millón de dólares americanos.

Este acontecimiento, puede deberse a que, no ha tenido suficiente lapso de tiempo como para tener un buen rendimiento. Hay que recordar que, en este caso, la diferencia de tiempo entre el inicio y el final del set de datos es de apenas 5 meses, mientras que el de actualizaciones diarias es de 6 años.

Por ello, considero que esta hipótesis puede llevar a dos escenarios diferentes:

1. Podría ser que, el que los datos se actualicen cada 5 minutos sea un escenario que tenga mucho potencial ya que tiene un mayor margen de rectificación que si es de actualización diaria. El problema yace en que es necesario un lapso de tiempo mayor para ver como se puede desarrollar y que pueda tener incluso un mejor rendimiento que el diario, pudiendo crear un nuevo benchmark[1].

2. Otro posible escenario es que, por mucho que se entrene y que tenga un set de datos mayor, no tenga suficiente potencial debido a que, el cambio tan rápido de las condiciones del mercado puede llevar a una volatilidad excesiva, dificultando la capacidad de los agentes para adaptarse y aprender patrones consistentes. Esta alta frecuencia de actualización puede resultar en decisiones basadas en cambios marginales y no en tendencias más significativas, lo cual podría ser contraproducente para el rendimiento a largo plazo. Por tanto, puede ser que un enfoque de actualización más pausado, que permita a los agentes procesar y aprender de tendencias más estables, sea más adecuado para lograr un rendimiento sostenido y fiable en el tiempo.

4.2. Impacto Medioambiental y consumo energético

Por último, en este entrenamiento incluimos también la evaluación del CO2 emitido en el entrenamiento del agente, para lo que utilizamos el programa de Nvidia Geforce Experience. Este software nos proporciona una estimación de consumo energético que varía desde 350 W hasta 9.6 kW, con una media de 300 Wh por modelo entrenado.

Basándonos en estos datos, la generación estimada de CO2 se ajustaría proporcionalmente a este consumo. Si consideramos que 3.28 kW de energía resultan en 0.55 kg de CO2, el consumo medio de 300 Wh implicaría una emisión significativamente menor. Tomando como referencia esta proporción y el número total de pasos de entrenamiento realizados, se podría estimar la huella de carbono correspondiente para la nueva media de consumo energético. Como

[1]benchmark: Prueba para comparar rendimiento tecnológico

comparativa, y ajustando los valores a la nueva media de consumo, la generación de CO2 sería equivalente a la que produce un vehículo promedio de gasolina al recorrer una distancia proporcionalmente menor que los 2.2 km mencionados previamente, adecuándose así a la eficiencia energética del proceso de entrenamiento adoptado.

4.3. Conclusiones

A continuación, se presentan las conclusiones finales de este trabajo final de máster. En ellas, se evalúan los rendimientos generales de los agentes empleados en este proyecto. El criterio de rendimiento radica en el valor medio final de nuestra cartera en función de los agentes empleados. Se emplaza al lector interesado a realizar un análisis más profundo de los resultados obtenidos en el apéndice E.

4.3.1. 1 Día

El artículo en el que se ha basado este trabajo final de máster *Deep reinforcement learning for automatedstock trading: An ensemble strategy* [23], se indica que el PPO era el mejor agente como inversor de bolsa, creando un nuevo estado del arte.

En este trabajo final de máster se ha demostrado que, para un entorno más volátil que las empresas que se detallan en la sección 3.2.1, ha sido más eficaz el DDPG para actualizaciones diarias basándonos en la recompensa media final. Tras el paso de los 6 años entre el inicio y el final del set de datos, se ha obtenido una recompensa bruta (antes de impuestos) de casi 2 Millones de dólares para un agente entrenado con 10 Millones de pasos. Esto es, un incremento patrimonial del **200 %**, o, lo que es equivalente, un **20 %** de media anual. También hay que tener en cuenta el periodo de la pandemia del SaRS-Covid, donde hubo una caída general de los índices bursátiles a nivel mundial.

Por otro lado, el agente PPO, aunque no ha sido tan eficaz en la maximización de la recompensa final como el DDPG o el A2C, ha mostrado un comportamiento distintivo en función de la tendencia del mercado. Se ha observado que el PPO tiende a subperformar durante tendencias bajistas, donde la cautela y la preservación del capital son críticas. Sin embargo, durante las tendencias alcistas, el PPO ha capitalizado eficientemente las oportunidades de crecimiento, lo que sugiere una sensibilidad a la dirección del mercado que podría ser explotada en una estrategia de inversión combinada.

Por último, el A2C ha obtenido una recompensa mayor de la esperada. A pesar de enfrentarse a un mercado con tendencias impredecibles y periodos de alta volatilidad, el modelo A2C ha demostrado una capacidad notable para adaptarse y generar resultados positivos E. Este

desempeño resalta la eficacia de la metodología de ventaja de actores-críticos en entornos financieros complejos y sugiere que, con la adecuada configuración y entrenamiento, modelos como A2C pueden sobrepasar las expectativas y ofrecer estrategias de inversión viables y rentables.

Como conclusión, el agente con mejor rendimiento ha sido el DDPG, ya que, de los 4 escenarios posibles, 2 de ellos ha quedado en primera posición E.1, y 1 vez en segunda. Esto permite y asegura al usuario, que, en caso de que se emplee con fines con ánimo de lucro, obtenga una buena recompensa y aumente el valor de la cartera.

Si definimos el benchmark como superar el porcentaje del IPC, se puede decir que ha tenido un muy buen rendimiento. Para 1 Millón y 10 Millones, ha sido capaz de superar el benchmark del IPC con buenos beneficios.

4.3.2. 5 Minutos

En referencia a los agentes entrenados para el escenario donde se actualizan los datos en periodos más cortos, han obtenido resultados peores que los de actualizaciones diarias.

Los resultados obtenidos plantean reflexiones importantes sobre la eficacia de los modelos de aprendizaje por refuerzo en mercados de alta volatilidad y con actualizaciones frecuentes. Contrario a lo que se podría esperar de sus sofisticadas arquitecturas y su adaptabilidad a las condiciones del mercado, todos los modelos han demostrado un rendimiento insatisfactorio, lo que sugiere limitaciones inherentes en su capacidad para navegar y aprovechar las dinámicas de mercado de alta frecuencia.

Este hallazgo cuestiona la robustez de las estrategias de aprendizaje por refuerzo en contextos donde la información cambia rápidamente y donde las oportunidades de arbitraje son fugaces y complejas. Parece ser que, más allá de la capacidad de identificar patrones en los datos, existe un desafío significativo en la aplicación efectiva de estos patrones en decisiones de trading en tiempo real.

Además, la uniformidad en los resultados de los diferentes modelos indica que, a pesar de sus distintas metodologías, ninguno ha logrado una ventaja competitiva clara. Este fenómeno podría apuntar a una saturación en la capacidad de innovación y diferenciación dentro del marco del aprendizaje por refuerzo aplicado a la negociación algorítmica.

En resumen, estos resultados subrayan la necesidad de una evaluación más crítica de los modelos de aprendizaje por refuerzo en el trading algorítmico y sugieren un camino hacia estrategias más complejas y holísticas que puedan adaptarse eficazmente a las rápidas variaciones del mercado.

4.4. Trabajos Futuros

Esta sección aborda las posibles direcciones de investigación y mejora que se derivan de este trabajo final de máster.

Los hallazgos de nuestras investigaciones sugieren dos líneas principales de trabajo futuro:

1. ***1 Día***: Los resultados obtenidos para el intervalo de 1 día han sido satisfactorios. Esto abre la posibilidad de desarrollar un bot que realice análisis diarios y ayude en la gestión de carteras de activos. En investigaciones futuras, se debería profundizar en la optimización de este modelo, explorando cómo las variaciones en las configuraciones del agente y la introducción de nuevas características de los datos pueden mejorar aún más su eficacia. Esto incluiría la incorporación de análisis de sentimiento de mercado y tendencias económicas globales, con el objetivo de aumentar la precisión y la rentabilidad del bot en diversos escenarios de mercado.

2. ***5 Minutos***: Los resultados menos alentadores obtenidos en este intervalo indican la necesidad de una investigación más profunda. Se planea obtener un conjunto de datos más amplio para determinar si los pobres resultados se deben a una insuficiencia de datos o a limitaciones en la capacidad de los agentes para operar en intervalos de tiempo tan cortos. Además, como se sugiere en tu párrafo final, la investigación futura debería centrarse en identificar las limitaciones de los modelos actuales en estos entornos de mercado altamente dinámicos.

Adicionalmente, se propone entrenar los agentes con diferentes configuraciones y redes neuronales alternativas para encontrar resultados óptimos. La exploración de diferentes arquitecturas de red y parámetros de entrenamiento puede proporcionar información valiosa sobre cómo mejorar el rendimiento en ambos intervalos de tiempo.

En resumen, el trabajo futuro se centrará en la optimización de estrategias de trading algorítmico para intervalos de tiempo variados, utilizando técnicas avanzadas de aprendizaje automático y considerando una gama más amplia de datos y factores del mercado.

Apéndice A

Implementación PPO con Stable Baselines

Esta implementación está basada en la que se ha empleado en el artículo *Deep reinforcement learning for automatedstock trading: An ensemble strategy* [23]

La configuración empleada consta de la siguiente adaptación con respecto a la configuración por defecto de la librería:

1. n_steps = 1371 El número de pasos que se ha de ejecutar para cada entorno por actualización.

2. learning_rate = 0.0007 Este hiperparámetro controla cuánto debe cambiar el modelo en respuesta al error estimado cada vez que se actualizan los pesos del gradiente.

3. mini_batches = 128 Número de entrenamientos simultáneos ejecutados.

4. ent_coef = 0.01 Coeficiente de entropía para el cálculo de pérdidas.

Como se detalla en el artículo, se decidió utilizar una política llamada MlpPolicy, construida como una red neuronal multicapa (MLP) D con dos capas, cada una de las cuales consta de 64 neuronas.

Como parte de nuestro trabajo, las redes se enfrentan a la tarea de gestionar secuencias de tiempo que no son más que conjuntos de datos continuos. Esta secuencia se trata como un vector de información rica que se ingresa en el MLP. Estos enfoques suelen servir como puntos de referencia para evaluar y comparar el rendimiento de diferentes modelos en tareas específicas.

La implementación del PPO se puede consultar en el siguiente repositorio de GitHub:

https://github.com/Stable-Baselines-Team/stable-baselines/blob/master/stable_baselines/ppo2

Apéndice B

Implementación DDPG con Stable Baselines

La implementación del Deep Deterministic Policy Gradient (DDPG) que se describe a continuación, se inspira en la metodología presentada en el artículo *Deep reinforcement learning for automated stock trading: An ensemble strategy* [23], y se ha adaptado a partir de la configuración predeterminada proporcionada por la biblioteca Stable Baselines.

A continuación, se detalla la configuración específica que se ha adoptado:

1. n_steps = 1371: Determina la cantidad de pasos ejecutados en cada actualización para cada entorno.

2. learning_rate = 0.0007: Este hiperparámetro regula la magnitud del ajuste en los pesos del modelo en respuesta al error estimado en cada actualización del gradiente.

3. mini_batches = 128: Define el número de lotes de entrenamiento que se procesan simultáneamente.

4. ent_coef = 0.01: Este coeficiente regula el término de entropía en el cálculo de la función de pérdida.

A diferencia de los enfoques PPO y A2C que utilizan políticas de selección de acciones probabilísticas, los agentes DDPG se basan en políticas deterministas. Esto significa que el modelo siempre sugiere la misma acción específica para una condición observada determinada. Esta característica es muy importante para entornos como los sistemas comerciales automatizados donde se desea una toma de decisiones consistente y la previsibilidad de las acciones es beneficiosa.

La política determinista de DDPG elimina la variabilidad en la selección de acciones inherente a las políticas estocásticas y puede resultar beneficiosa en entornos estables y continuos.

Sin embargo, esta característica también puede ser una limitación en entornos con requisitos de exploración más complejos, donde la aleatoriedad en la selección de acciones puede ayudar a descubrir nuevas estrategias y soluciones.

La implementación de DDPG con **stable baselines** proporciona un marco sólido para desarrollar y evaluar algoritmos de aprendizaje por refuerzo. Con una arquitectura basada en redes neuronales profundas y la integración de técnicas avanzadas como la iteración de experiencias y el suavizado de objetivos, DDPG se establece como una poderosa herramienta para aprender políticas complejas en el espacio de acción continua.

El algoritmo DDPG ha demostrado ser eficaz en una amplia gama de tareas, desde el control de robots hasta la gestión de inversiones, y sigue siendo referenciado en la comunidad de aprendizaje por refuerzo por su capacidad para aprender políticas efectivas y su flexibilidad para adaptarse a una variedad de aplicaciones.

Como se detalla en el artículo, se decidió utilizar una política llamada MlpPolicy, construida como una red neuronal multicapa (MLP) D con dos capas, cada una de las cuales consta de 64 neuronas.

Como parte de nuestro trabajo, las redes se enfrentan a la tarea de gestionar secuencias de tiempo que no son más que conjuntos de datos continuos. Esta secuencia se trata como un vector de información rica que se ingresa en el MLP. Estos enfoques suelen servir como puntos de referencia para evaluar y comparar el rendimiento de diferentes modelos en tareas específicas.

Para más información sobre la implementación específica y configuraciones del DDPG en Stable Baselines, se recomienda revisar el código fuente y la documentación disponible en el repositorio de GitHub.

https://github.com/Stable-Baselines-Team/stable-baselines/tree/master/stable_baselines/ddpg

Apéndice C

Implementación A2C con Stable Baselines

La implementación de A2C (Advantage Actor-Critic) que presentamos aquí está inspirada en la metodología utilizada en el artículo *Deep reinforcement learning for automated stock trading: An ensemble strategy* [23], adaptada a partir de las configuraciones por defecto de la biblioteca Stable Baselines.

La configuración específica utilizada es la siguiente:

1. n_steps = 1371: Se refiere al número de pasos ejecutados en cada actualización por cada entorno.

2. learning_rate = 0.0007: Este parámetro ajusta la tasa a la que el modelo aprende, controlando el tamaño de los ajustes de los pesos del modelo con cada actualización del gradiente.

3. mini_batches = 128: Cantidad de lotes de entrenamiento que se procesan de forma simultánea.

4. ent_coef = 0.01: Coeficiente de entropía utilizado en la función de pérdida para fomentar la exploración.

Siguiendo la estructura propuesta en el artículo, se eligió una política llamada MlpPolicy, que consiste en una red neuronal multicapa (MLP) con dos capas de 64 neuronas cada una. Las redes MLP están formadas por múltiples capas: una capa de entrada, varias capas ocultas para procesar diferentes niveles de abstracción y una capa de salida para realizar las predicciones. La Figura D.1 proporciona una representación visual de esta estructura de red.

Este tipo de redes es ideal para resolver tareas de clasificación, asignando entradas a categorías específicas, y también para construir modelos de regresión que predicen valores continuos a partir de un conjunto de variables de entrada.

En nuestra investigación, el modelo A2C se encarga de gestionar secuencias temporales, tratándolas como vectores de información que se procesan a través de la red MLP. Estos modelos suelen establecerse como benchmarks para comparar la efectividad de diferentes algoritmos en tareas determinadas.

Como se detalla en el artículo, se decidió utilizar una política llamada MlpPolicy, construida como una red neuronal multicapa (MLP) D con dos capas, cada una de las cuales consta de 64 neuronas.

Como parte de nuestro trabajo, las redes se enfrentan a la tarea de gestionar secuencias de tiempo que no son más que conjuntos de datos continuos. Esta secuencia se trata como un vector de información rica que se ingresa en el MLP. Estos enfoques suelen servir como puntos de referencia para evaluar y comparar el rendimiento de diferentes modelos en tareas específicas.

Para acceder a la implementación específica del A2C y a sus configuraciones, se puede visitar el repositorio de GitHub en el siguiente enlace:

https://github.com/Stable-Baselines-Team/stable-baselines/tree/master/stable_baselines/a2c

Apéndice D

MLPPolicy

Como se detalla en el artículo, se decidió utilizar una política llamada MlpPolicy, construida como una red neuronal multicapa (MLP) con dos capas, cada una de las cuales consta de 64 neuronas. Las redes neuronales del tipo MLP se caracterizan por tener una gran cantidad de capas: una capa de entrada, varias capas ocultas que permiten diferentes niveles de abstracción y una capa de salida donde se realizan las predicciones. En la Figura D.1 se muestra un diagrama visual típico de dicha red.

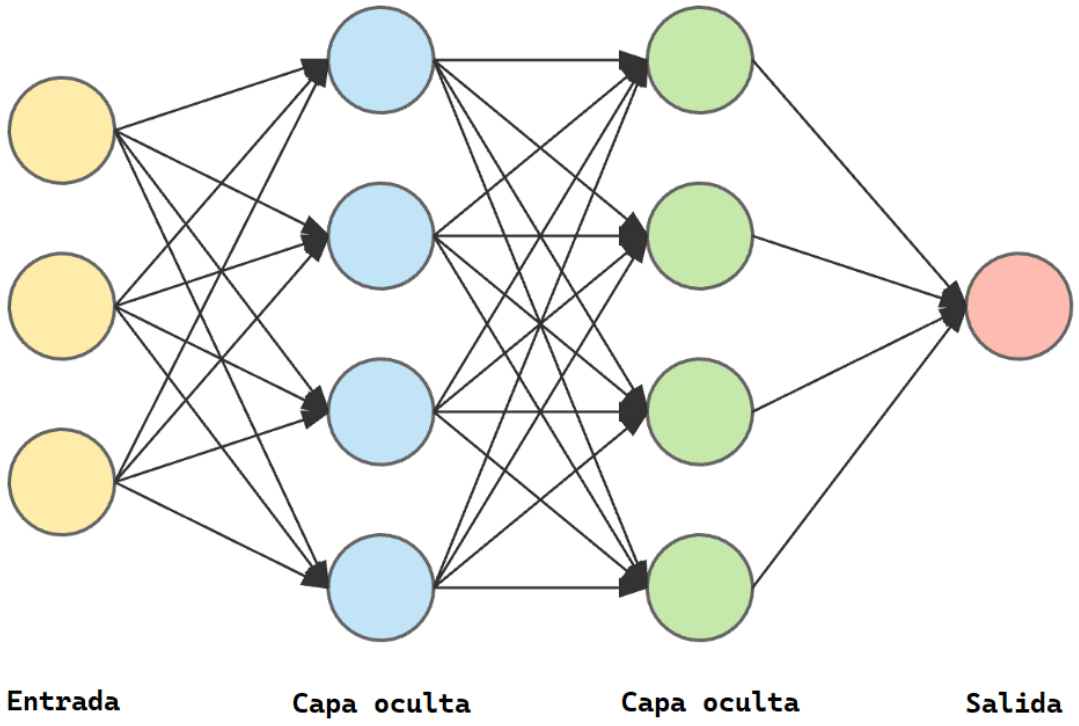


Figura D.1: Ejemplo de una MLP con dos capas ocultas de 4 neuronas. Fuente propia.

Estas arquitecturas neuronales son adecuadas para tareas de clasificación que asignan entradas a categorías específicas, así como modelos de regresión que estiman números continuos a partir de un conjunto de variables de entrada.

Como parte de nuestro trabajo, las redes se enfrentan a la tarea de gestionar secuencias de tiempo que no son más que conjuntos de datos continuos. Esta secuencia se trata como un vector de información rica que se ingresa en el MLP. Estos enfoques suelen servir como puntos de referencia para evaluar y comparar el rendimiento de diferentes modelos en tareas específicas.

Apéndice E

Ampliación de los resultados del rendimiento de los agentes

A lo largo de este Apéndice, se relatan los resultados obtenidos en los test realizados para los agentes entrenados con 1, 2 y 5 millones de pasos. Por otro lado, también se ampliarán los resultados expuestos en el capítulo E tanto 10 Millones de pasos como actualizaciones cada 5 minutos.

E.1. Resultados

A lo largo de esta sección, se presentarán los resultados obtenidos teniendo en cuenta la siguiente jerarquía:

1. 1 Día
    a) 1 Millón
    b) 2 Millones
    c) 5 Millones
    d) 10 Millones
2. 5 Minutos

donde cada una engloba los diferentes test realizados para los 3 agentes A2C, PPO y DDPG:

- Valor de la cartera al final de cada ciclo (End Total Asset).
- Ganancias generadas al final de cada ciclo (Total Reward).

- La cantidad total de las compras realizadas a lo largo del ciclo (Total Trades).
- El coste total al comprar las acciones por cada ciclo (Total Cost)
- El ratio de Sharpe del ciclo (Sharpe Ratio)

E.1.1. 1 Día

En este apartado se contemplan los resultados obtenidos para una actualización diaria.

E.1.1.1. End Total Asset (Valor final de la cartera)

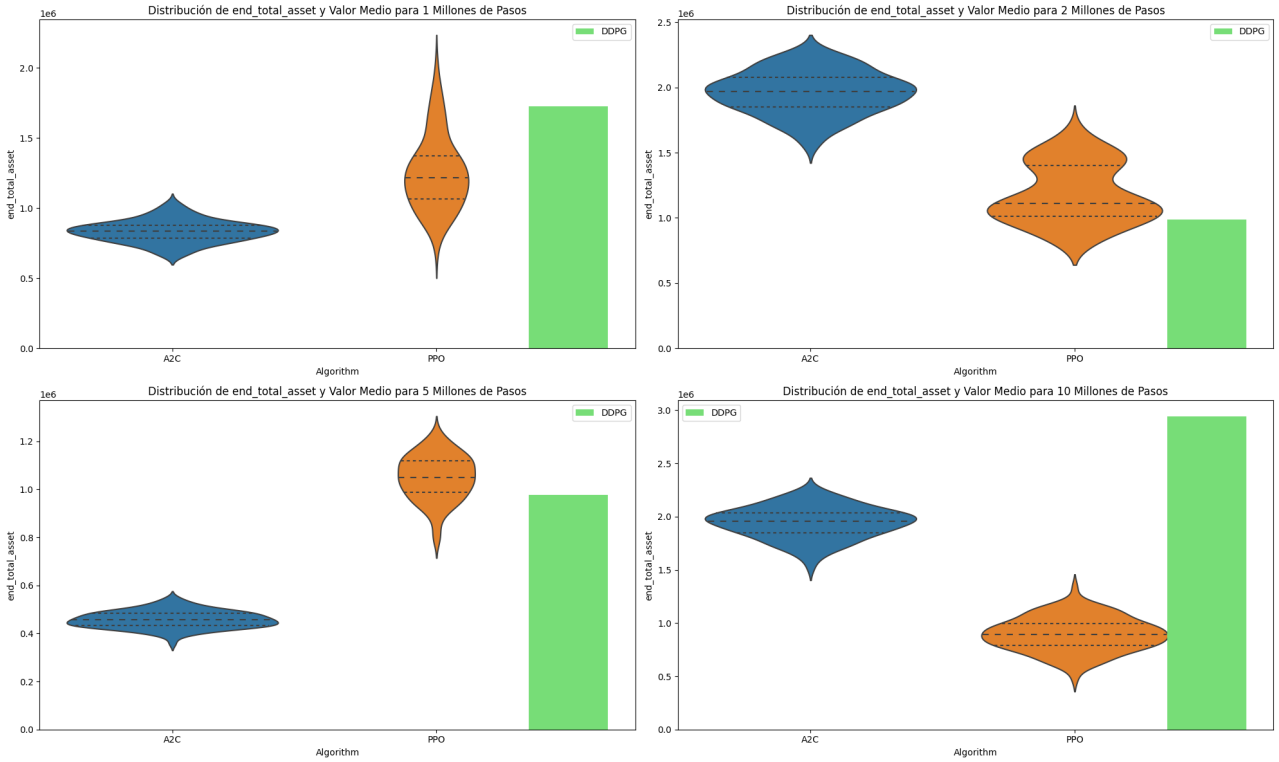


Figura E.1: Resultado de evaluación para 1, 2, 5 y 10 Millones de pasos de los agentes A2C, PPO y DDPG. Fuente propia.

Los resultados de los agentes A2C, PPO y DDPG variaron significativamente a lo largo de diferentes etapas, reflejando la complejidad y la variabilidad en el comportamiento de cada algoritmo.

Para 1 Millón de pasos:

- *A2C*: Mostró un bajo rendimiento, con pérdidas casi aseguradas, sugiriendo una estrategia de trading poco efectiva en este escenario.

- *PPO*: Exhibió una amplia variabilidad entre sus resultados mínimos y máximos, con una media superando el millón de dólares. Esto indica una incertidumbre significativa en los resultados, lo que podría ser riesgoso en una implementación real.

- *DDPG*: Se destacó con un rendimiento sobresaliente, generando casi 750,000 dólares de beneficio, lo que lo posiciona como una opción prometedora para este intervalo de tiempo.

Para 2 Millones de pasos:

- *A2C*: Mejoró su rendimiento, alcanzando un promedio ligeramente superior al 100 %.

- *PPO*: Mantuvo un rendimiento notable, aunque con un riesgo potencial de pérdidas.

- *DDPG*: Mostró un equilibrio entre ganancias y pérdidas, sin un rendimiento definitivamente positivo o negativo.

Para 5 Millones de pasos:

- *A2C*: Registró pérdidas significativas, promediando medio millón de dólares en el negativo.

- *PPO*: Alcanzó un ligero beneficio, aunque cerca del umbral de pérdidas.

- *DDPG*: Se mantuvo por debajo del millón de dólares, indicando pérdidas moderadas pero controladas.

Para 10 Millones de pasos:

- El *DDPG* demostró un rendimiento excepcional, logrando una recompensa del 200 % después de 6 años de pruebas, destacando su potencial en escenarios de largo plazo.

- Los resultados de *A2C* y *PPO* mostraron una variabilidad considerable, con el *A2C* teniendo un mejor rendimiento que el *PPO*.

Agente	Nº Pasos	Posición
DDPG	**1M**	**1º**
PPO	1M	2º
A2C	1M	3º
A2C	**2M**	**1º**
PPO	2M	2º
DDPG	2M	3º
PPO	**5M**	**1º**
DDPG	5M	2º
A2C	5M	3º
DDPG	**10M**	**1º**
A2C	10M	2º
PPO	10M	3º

Cuadro E.1: Clasificación de los modelos según el valor medio final de las carteras. Fuente Propia.

Como podemos ver en la tabla anterior E.1, se demuestra que el DDPG ha sido el que mejor rendimiento ha tenido de los 3 de manera general. Ha sido el más conservador de todos y el que mejores resultados ha obtenido. El segundo mejor agente ha sido el PPO, y por último el A2C.

E.1.1.2. Total Reward (Recompensa final)

A continuación se muestran la recompensa final de los mismos agentes y mismos pasos que en la sección anterior E.1.2.1:

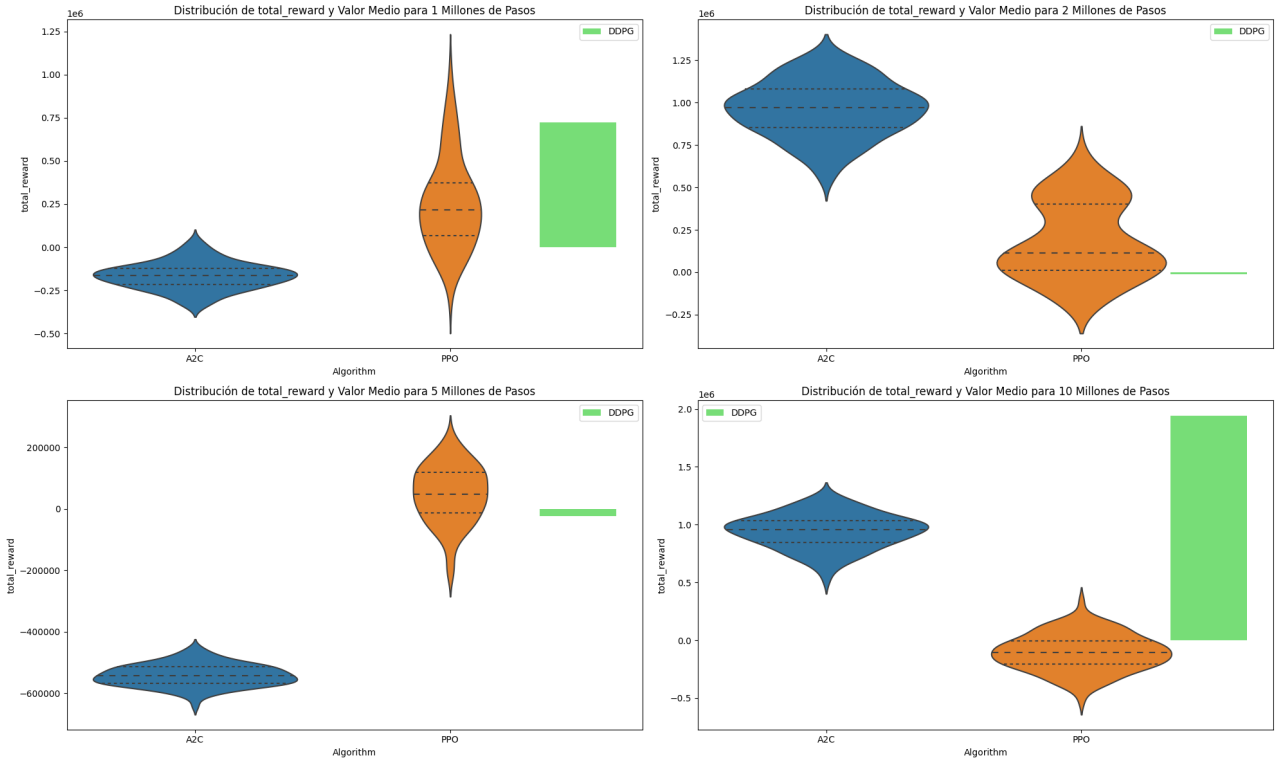


Figura E.2: Recompensa final para 1, 2, 5 y 10 Millones de pasos de los agentes A2C, PPO y DDPG. Fuente propia.

E.1.1.3. Total Cost (Coste Total)

A continuación se muestran el coste total de los mismos agentes y mismos pasos que en la sección anterior E.1.2.1:

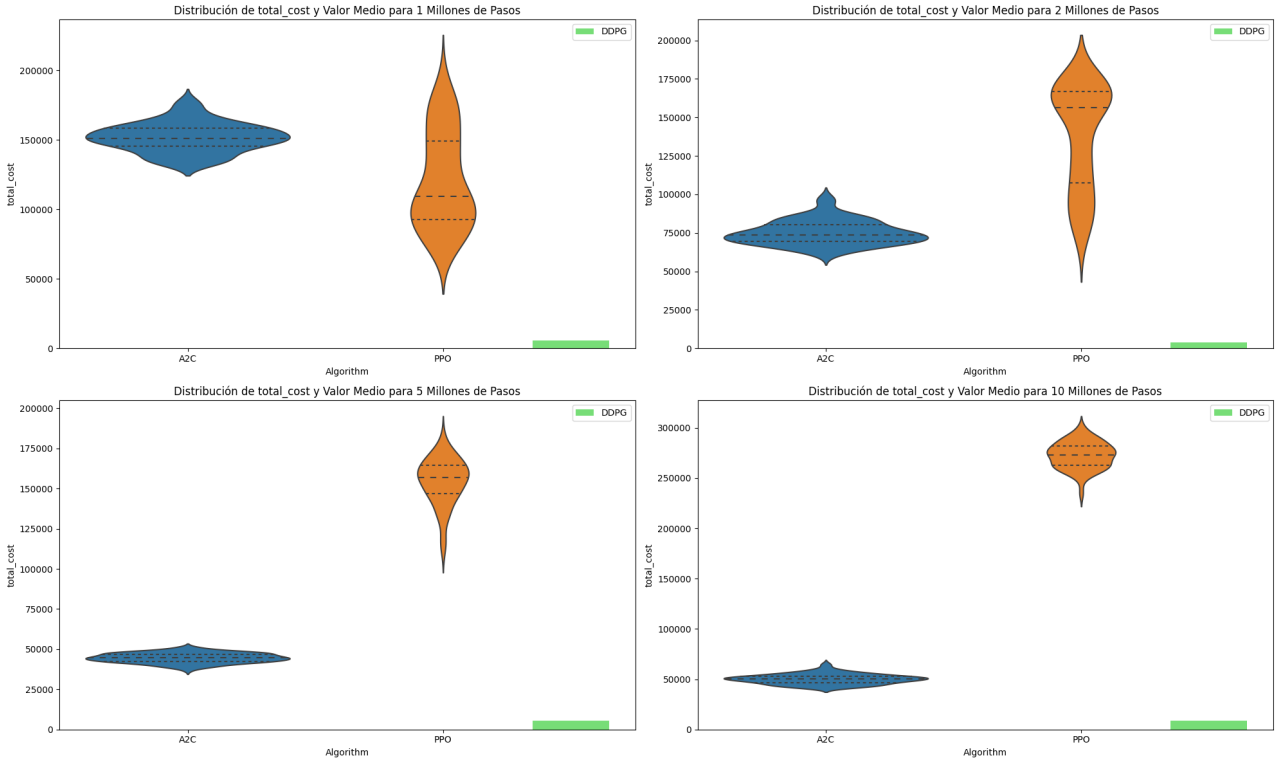


Figura E.3: Coste total para 1, 2, 5 y 10 Millones de pasos de los agentes A2C, PPO y DDPG. Fuente propia.

E.1.1.4. Total Trades (Intercambios totales)

A continuación se muestran los intercambios totales de los mismos agentes y mismos pasos que en la sección anterior E.1.2.1:

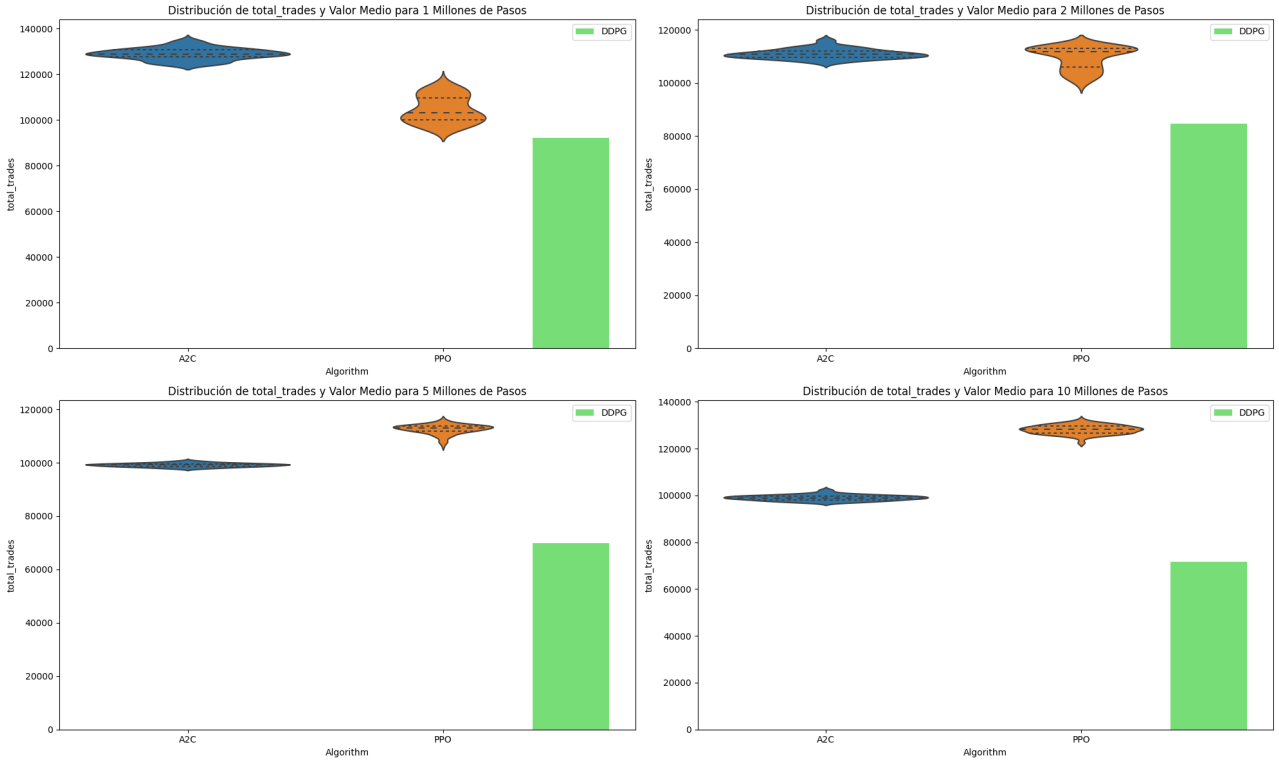


Figura E.4: Recompensa final para 1, 2, 5 y 10 Millones de pasos de los agentes A2C, PPO y DDPG. Fuente propia.

E.1.1.5. Ratio Sharpe

A continuación se muestran el ratio de Sharpe de los mismos agentes y mismos pasos que en la sección anterior E.1.2.1:

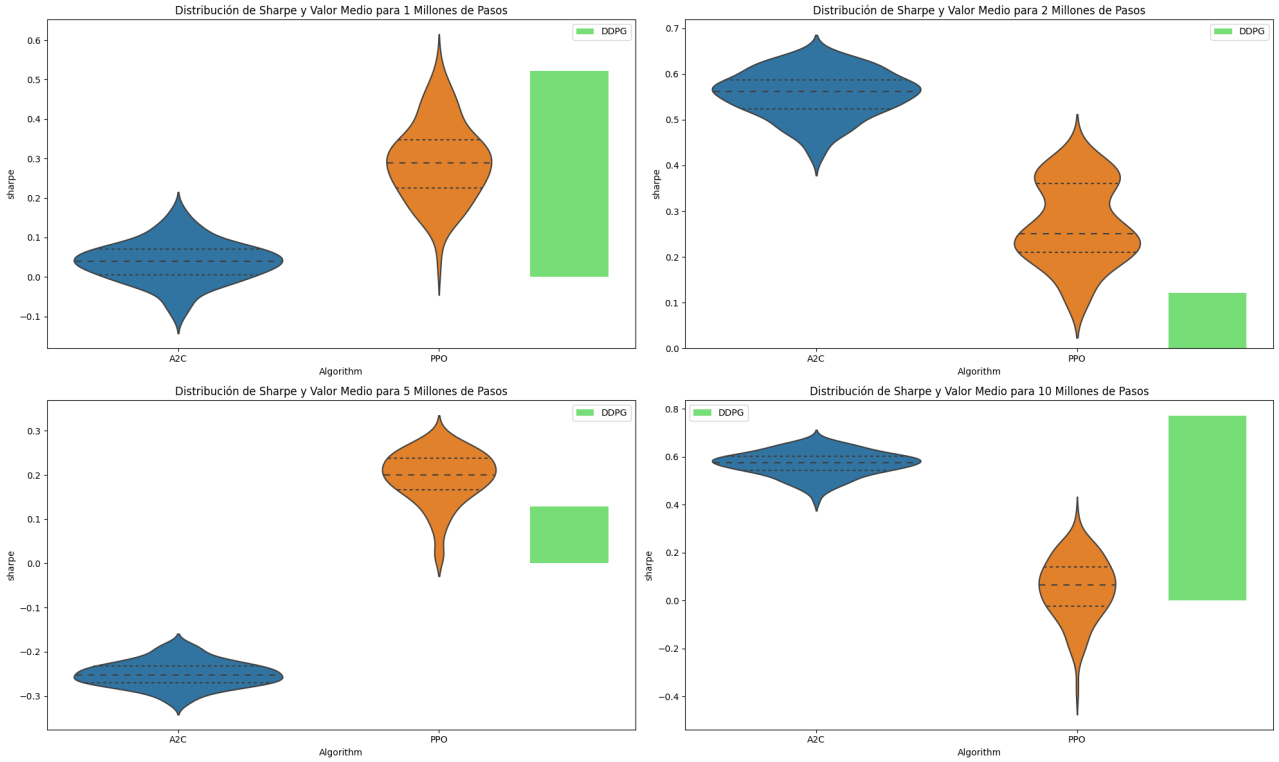


Figura E.5: Recompensa final para 1, 2, 5 y 10 Millones de pasos de los agentes A2C, PPO y DDPG. Fuente propia.

E.1.2. 5 minutos

Al igual que en el subsección anterior, se detallan los resultados obtenidos para aquellos ciclos con actualización de 5 minutos:

E.1.2.1. End Total Asset (Valor final de la cartera)

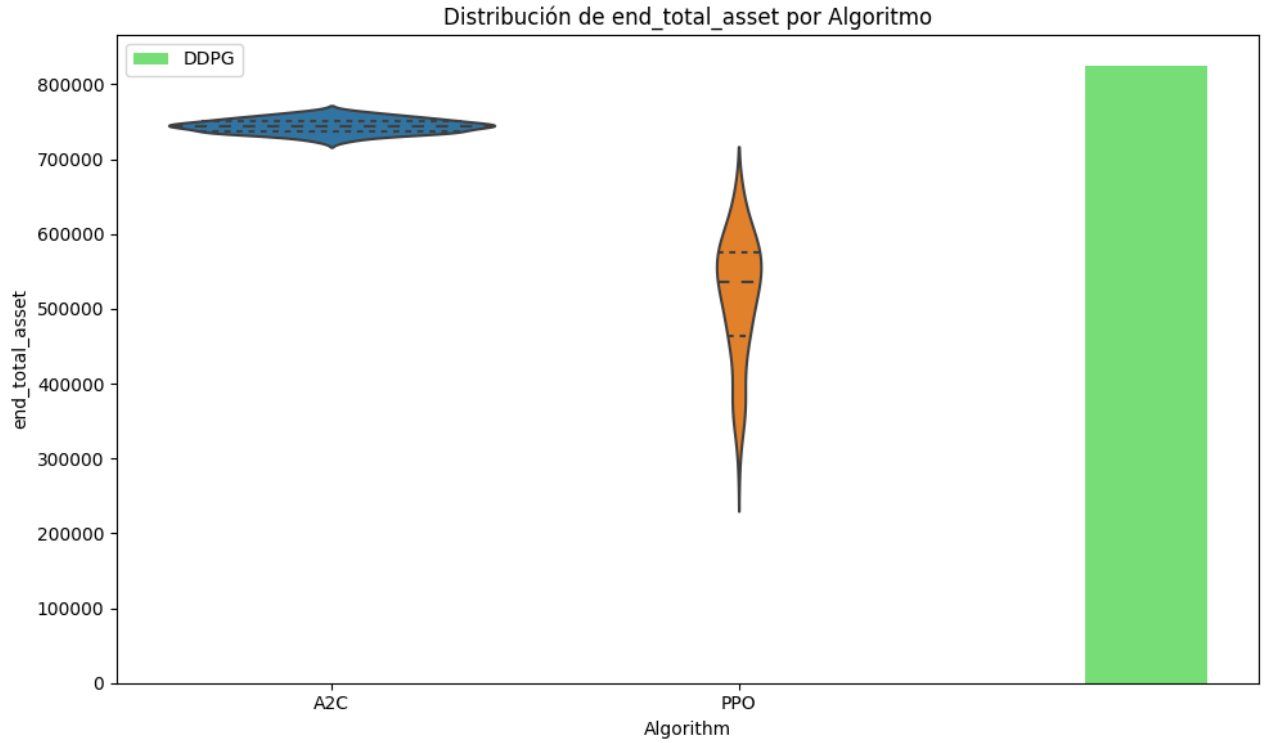


Figura E.6: Valor final de la cartera mediante el uso de los agentes A2C, PPO y DDPG con actualización cada 5 minutos. Fuente propia.

Como se ha podido corroborar, todos los agentes han tenido un rendimiento bajo, haciendo que el valor de la cartera quede por debajo de su estado inicial. Este acontecimiento, puede deberse a que, no ha tenido suficiente lapso de tiempo como para tener un buen rendimiento. Hay que recordar que, en este caso, la diferencia de tiempo entre el inicio y el final del dataset es de apenas 5 meses, mientras que el de actualizaciones diarias es de 6 años.

Por ello, considero que esta hipótesis puede llevar a dos escenarios diferentes:

1. Podría ser que, el que los datos se actualicen cada 5 minutos sea un escenario que tenga mucho potencial ya que tiene un mayor margen de rectificación que si es de actualización diaria. El problema yace en que es necesario un lapso de tiempo mayor para ver como se puede desarrollar y que pueda tener incluso un mejor rendimiento que el diario, pudiendo crear un nuevo benchmark.

2. Otro posible escenario es que, por mucho que se entrene y que tenga un set de datos mayor, no tenga suficiente potencial debido a que, el cambio tan rápido de las condicio-

nes del mercado puede llevar a una volatilidad excesiva, dificultando la capacidad de los agentes para adaptarse y aprender patrones consistentes. Esta alta frecuencia de actualización puede resultar en decisiones basadas en cambios marginales y no en tendencias más significativas, lo cual podría ser contraproducente para el rendimiento a largo plazo. Por tanto, puede ser que un enfoque de actualización más pausado, que permita a los agentes procesar y aprender de tendencias más estables, sea más adecuado para lograr un rendimiento sostenido y fiable en el tiempo.

Agente	Posición
DDPG	**1**
A2C	2
PPO	3

Cuadro E.2: Clasificación de los modelos según el valor medio final de las carteras. Fuente Propia.

Como podemos ver en la tabla anterior E.1, se demuestra que el DDPG ha sido el que mejor rendimiento ha tenido de los 3 de manera general. Ha sido el más conservador de todos y el que mejores resultados ha obtenido. El segundo mejor agente ha sido el PPO, y por último el A2C.

E.1.2.2. Total Reward (Recompensa final)

A continuación se muestran la recompensa final de los mismos agentes y mismos pasos que en la sección anterior E.1.2.1:

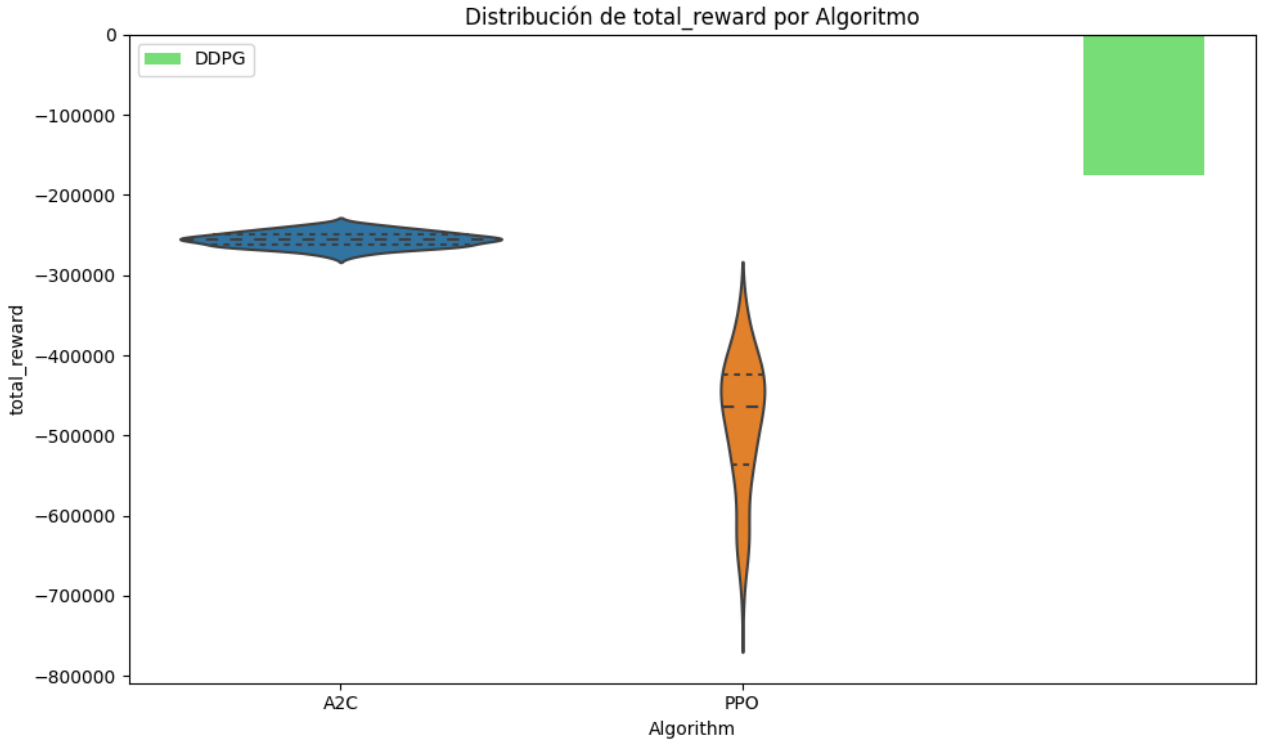


Figura E.7: Recompensa final de los agentes A2C, PPO y DDPG con actualización cada 5 minutos. Fuente propia.

E.1.2.3. Total Cost (Coste Total)

En la siguiente figura E.3 representa el coste total de los agentesE.1.2.1:

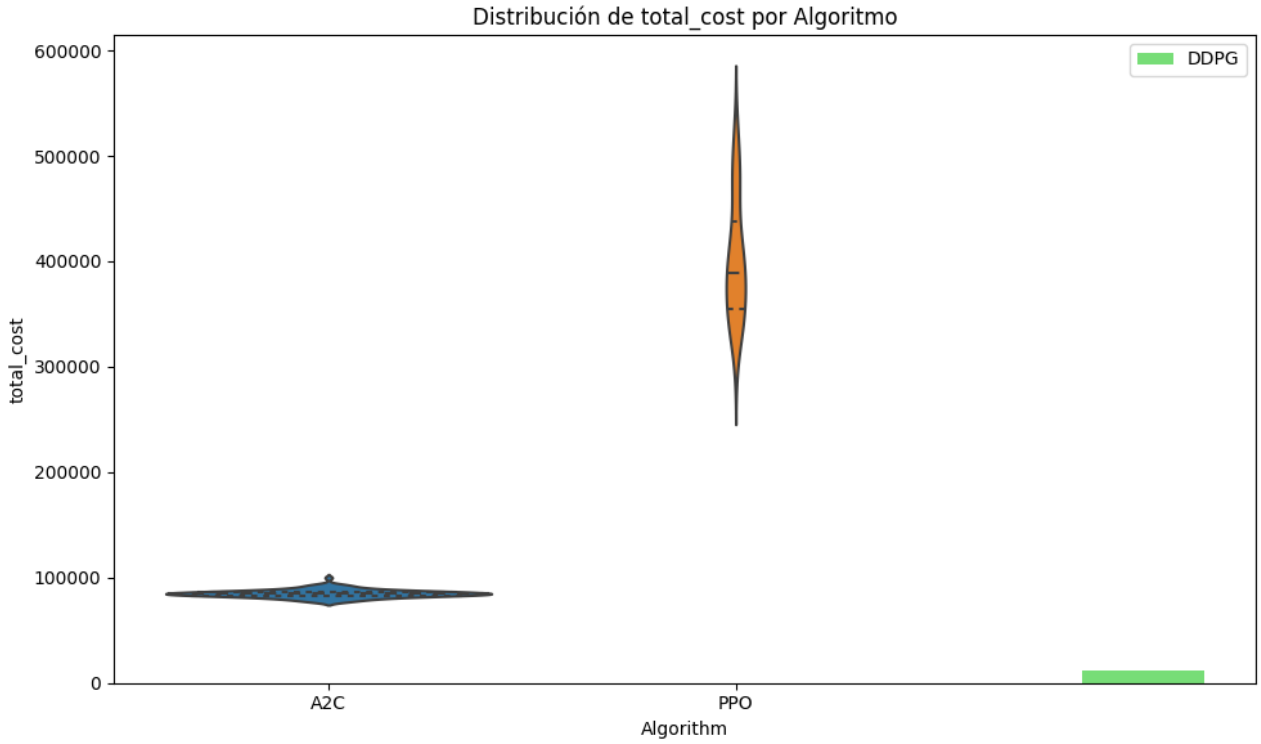


Figura E.8: Coste total de los agentes A2C, PPO y DDPG con actualización cada 5 minutos. Fuente propia.

E.1.2.4. Total Trades (Intercambios totales)

Ahora se visualizan los intercambios totales para cada uno de los agentes que en la sección anterior E.1.2.1:

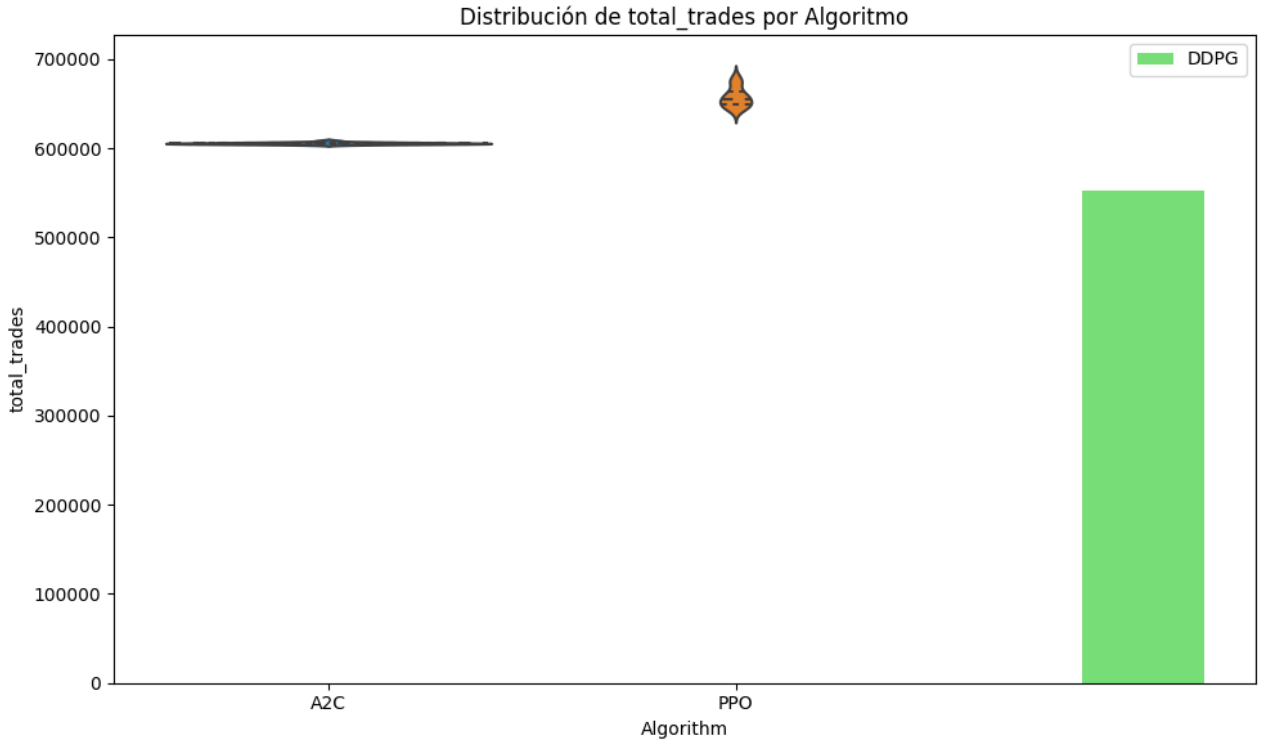


Figura E.9: Intercambios totales de los agentes A2C, PPO y DDPG con actualización cada 5 minutos. Fuente propia.

E.1.2.5. Ratio Sharpe

Finalmente, se ha elaborado una gráfica que representa las estadísticas del ratio de Sharpe de los agentes que en la sección anterior E.1.2.1:

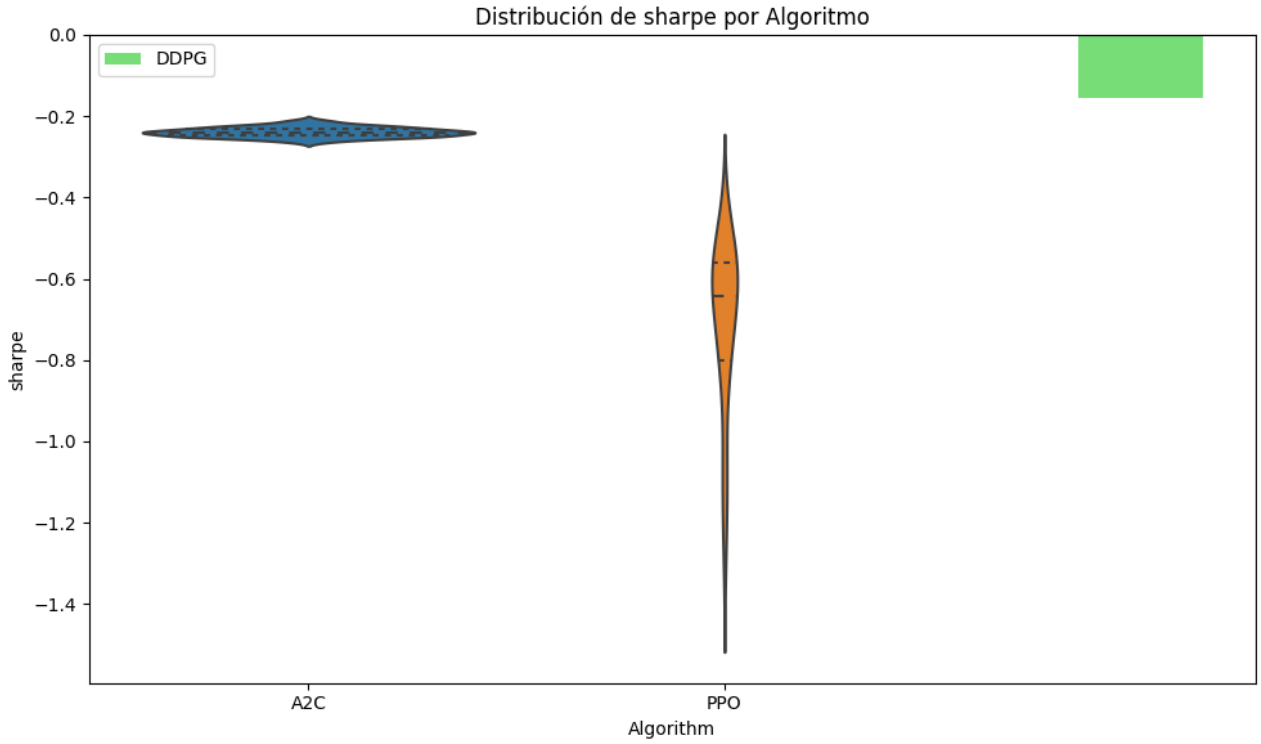


Figura E.10: Ratio de Sharp de los agentes A2C, PPO y DDPG con actualización cada 5 minutos. Fuente propia.

Apéndice F

Tabla de las empresas que comprenden el mercado de valores S&P600

En el presente apéndice, se introducen en una tabla las compañías que comprenden el índice de bolsa Standard & Poors 600, así como su símbolo, País de su residencia fiscal, Sector y Mercado. Esta tabla se ha extraído del siguiente enlace: https://en.wikipedia.org/wiki/List_of_S%26P_600_companies

Símbolo	**Nombre de la Compañía**	**País**	**Sector**	**Mercado**
AAON	AAON, Inc.	Estados Unidos	Industrials	us market
AAP	Advance Auto Parts, Inc.	Estados Unidos	Consumer Discretionary	us market
AAT	American Assets Trust	Estados Unidos	Real Estate	us market
ABCB	Ameris Bancorp	Estados Unidos	Financials	us market
ABG	Asbury Automotive Group	Estados Unidos	Consumer Discretionary	us market
ABM	ABM Industries, Inc.	Estados Unidos	Industrials	us market
ABR	Arbor Realty Trust	Estados Unidos	Financials	us market
ACA	Arcosa, Inc.	Estados Unidos	Industrials	us market
ACIW	ACI Worldwide	Estados Unidos	Information Technology	us market
ACLS	Axcelis Technologies, Inc.	Estados Unidos	Information Technology	us market
ADEA	Adeia, Inc.	Estados Unidos	Information Technology	us market

Continúa en la siguiente página

Cuadro F.1 – Continúa de la página anterior

Símbolo	Nombre de la Compañía	País	Sector	Mercado
ADTN	Adtran, Inc.	Estados Unidos	Information Technology	us market
ADUS	Addus HomeCare Corp.	Estados Unidos	Health Care	us market
AEIS	Advanced Energy	Estados Unidos	Information Technology	us market
AEL	American Equity Investment Life Holding Co.	Estados Unidos	Financials	us market
AEO	American Eagle Outfitters	Estados Unidos	Consumer Discretionary	us market
AGO	Assured Guaranty Ltd.	Estados Unidos	Financials	us market
AGTI	Agiliti, Inc.	Estados Unidos	Health Care	us market
AGYS	Agilysys, Inc.	Estados Unidos	Information Technology	us market
AHCO	AdaptHealth Corp.	Estados Unidos	Health Care	us market
AHH	Armada Hoffler Properties, Inc.	Estados Unidos	Real Estate	us market
AIN	Albany International Corp.	Estados Unidos	Industrials	us market
AIR	AAR CORP.	Estados Unidos	Industrials	us market
AIT	Applied Industrial Technologies	Estados Unidos	Industrials	us market
AKR	Acadia Realty Trust	Estados Unidos	Real Estate	us market
ALEX	Alexander & Baldwin	Estados Unidos	Real Estate	us market
ALG	Alamo Group	Estados Unidos	Industrials	us market
ALGT	Allegiant Travel Company	Estados Unidos	Industrials	us market
ALRM	Alarm.Com, Inc.	Estados Unidos	Financials	us market
AMBC	Ambac Financial Group	Estados Unidos	Financials	us market
AMCX	AMC Networks	Estados Unidos	Consumer Discretionary	us market
AMEH	Apollo Medical Holdings, Inc.	Estados Unidos	Health Care	us market
AMN	Amn Healthcare Services, Inc.	Estados Unidos	Health Care	us market

Continúa en la siguiente página

Cuadro F.1 – Continúa de la página anterior

Símbolo	Nombre de la Compañía	País	Sector	Mercado
AMPH	Amphstar Pharmaceuticals, Inc.	Estados Unidos	Health Care	us market
AMR	Alpha Metallurgical Resources, Inc.	Estados Unidos	Materials	us market
AMSF	Amerisafe, Inc.	Estados Unidos	Financials	us market
AMWD	American Woodmark Corp.	Estados Unidos	Industrials	us market
ANDE	The Andersons, Inc.	Estados Unidos	Consumer Staples	us market
ANF	Abercrombie & Fitch Company	Estados Unidos	Consumer Discretionary	us market
ANIP	ANI Pharmaceuticals, Inc.	Estados Unidos	Health Care	us market
AORT	Artivion	Estados Unidos	Health Care	us market
AOSL	Alpha and Omega Semiconductor, Ltd.	Estados Unidos	Information Technology	us market
APAM	Artisan Partners Asset Management, Inc.	Estados Unidos	Financials	us market
APLE	Apple Hospitality REIT, Inc.	Estados Unidos	Real Estate	us market
APOG	Apogee Enterprises, Inc.	Estados Unidos	Industrials	us market
APPS	Digital Turbine	Estados Unidos	Information Technology	us market
ARCB	ArcBest Corp.	Estados Unidos	Industrials	us market
ARI	Apollo Commercial Real Estate Finance	Estados Unidos	Real Estate	us market
ARLO	Arlo Technologies	Estados Unidos	Information Technology	us market
AROC	Archrock, Inc.	Estados Unidos	Energy	us market
ARR	Armour Residential REIT	Estados Unidos	Real Estate	us market
ASIX	Advansix, Inc.	Estados Unidos	Materials	us market
ASO	Academy Sports + Outdoors	Estados Unidos	Consumer Discretionary	us market
ASTE	Astec Industries, Inc.	Estados Unidos	Industrials	us market

Continúa en la siguiente página

Cuadro F.1 – Continúa de la página anterior

Símbolo	Nombre de la Compañía	País	Sector	Mercado
ATEN	A10 Networks, Inc.	Estados Unidos	Information Technology	us market
ATGE	Adtalem Global Education	Estados Unidos	Consumer Discretionary	us market
ATI	ATI Inc.	Estados Unidos	Materials	us market
ATNI	ATN International, Inc.	Estados Unidos	Communication Services	us market
AUB	Atlantic Union Bankshares, Corp.	Estados Unidos	Financials	us market
AVA	Avista Corporation	Estados Unidos	Utilities	us market
AVAV	AeroVironment, Inc.	Estados Unidos	Industrials	us market
AVNS	Avanos Medical, Inc.	Estados Unidos	Health Care	us market
AWR	American States Water Company	Estados Unidos	Utilities	us market
AX	Axos Financial, Inc.	Estados Unidos	Financials	us market
AXL	American Axle & Manufacturing, Inc.	Estados Unidos	Consumer Discretionary	us market
AZZ	AZZ, Inc.	Estados Unidos	Industrials	us market
B	Barnes Group, Inc.	Estados Unidos	Industrials	us market
BANC	Banc Of California, Inc.	Estados Unidos	Financials	us market
BANF	Bancfirst Corp	Estados Unidos	Financials	us market
BANR	Banner Corporation	Estados Unidos	Financials	us market
BCC	Boise Cascade	Estados Unidos	Industrials	us market
BCPC	Balchem Corporation	Estados Unidos	Materials	us market
BDN	Brandywine Realty Trust	Estados Unidos	Real Estate	us market
BFH	Bread Financial Holdings, Inc.	Estados Unidos	Financials	us market
BFS	Saul Centers, Inc.	Estados Unidos	Real Estate	us market
BGS	B&G Foods	Estados Unidos	Consumer Staples	us market
BHE	Benchmark Electronics, Inc.	Estados Unidos	Information Technology	us market

Continúa en la siguiente página

Cuadro F.1 – Continúa de la página anterior

Símbolo	Nombre de la Compañía	País	Sector	Mercado
BHLB	Berkshire Hills Bancorp, Inc.	Estados Unidos	Financials	us market
BJRI	BJ's Restaurants, Inc.	Estados Unidos	Consumer Discretionary	us market
BKE	The Buckle, Inc.	Estados Unidos	Consumer Discretionary	us market
BKU	BankUnited, Inc.	Estados Unidos	Financials	us market
BLFS	BioLife Solutions, Inc.	Estados Unidos	Health Care	us market
BLMN	Bloomin' Brands, Inc.	Estados Unidos	Consumer Discretionary	us market
BMI	Badger Meter, Inc.	Estados Unidos	Information Technology	us market
BOH	Bank of Hawaii	Estados Unidos	Financials	us market
BOOT	Boot Barn Holdings, Inc.	Estados Unidos	Consumer Discretionary	us market
BRC	Brady Corporation	Estados Unidos	Industrials	us market
BRKL	Brookline Bancorp, Inc.	Estados Unidos	Financials	us market
BSIG	BrightSphere Investment Group, Inc.	Estados Unidos	Financials	us market
BXMT	Blackstone Mortgage Trust, Inc.	Estados Unidos	Financials	us market
CAKE	The Cheesecake Factory, Inc.	Estados Unidos	Consumer Discretionary	us market
CAL	Caleres, Inc.	Estados Unidos	Consumer Discretionary	us market
CALM	Cal-Maine Foods, Inc.	Estados Unidos	Consumer Staples	us market
CARG	CarGurus	Estados Unidos	Communication Services	us market
CARS	Cars.com	Estados Unidos	Communication Services	us market
CASH	Pathward Financial, Inc.	Estados Unidos	Financials	us market

Continúa en la siguiente página

Cuadro F.1 – Continúa de la página anterior

Símbolo	Nombre de la Compañía	País	Sector	Mercado
CATY	Cathay General Bancorp	Estados Unidos	Financials	us market
CBRL	Cracker Barrel	Estados Unidos	Consumer Discretionary	us market
CBU	Community Bank System, Inc.	Estados Unidos	Financials	us market
CCOI	Cogent Communications Holdings, Inc.	Estados Unidos	Communication Services	us market
CCRN	Cross Country Healthcare, Inc.	Estados Unidos	Health Care	us market
CCS	Century Communities, Inc.	Estados Unidos	Consumer Discretionary	us market
CCSI	Consensus Cloud Solutions, Inc.	Estados Unidos	Information Technology	us market
CDMO	Avid Bioservices, Inc.	Estados Unidos	Health Care	us market
CEIX	CONSOL Energy, Inc.	Estados Unidos	Energy	us market
CENT	Central Garden & Pet Company	Estados Unidos	Consumer Staples	us market
CENTA	Central Garden & Pet Company (Class A)	Estados Unidos	Consumer Staples	us market
CENX	Century Aluminum Company	Estados Unidos	Materials	us market
CERT	Certara, Inc.	Estados Unidos	Health Care	us market
CEVA	CEVA, Inc.	Estados Unidos	Information Technology	us market
CFFN	Capitol Federal Savings Bank	Estados Unidos	Financials	us market
CHCO	City Holding Company	Estados Unidos	Financials	us market
CHCT	Community Healthcare Trust, Inc.	Estados Unidos	Health Care	us market
CHEF	Chefs' Warehouse, Inc.	Estados Unidos	Consumer Staples	us market

Continúa en la siguiente página

Cuadro F.1 – Continúa de la página anterior

Símbolo	Nombre de la Compañía	País	Sector	Mercado
CHS	Chico's FAS, Inc.	Estados Unidos	Consumer Discretionary	us market
CHUY	Chuy's Holdings, Inc.	Estados Unidos	Consumer Discretionary	us market
CLB	Core Laboratories	Estados Unidos	Energy	us market
CLDT	Chatham Lodging Trust	Estados Unidos	Real Estate	us market
CLFD	Clearfield, Inc.	Estados Unidos	Information Technology	us market
CLW	Clearwater Paper Corporation	Estados Unidos	Materials	us market
CMP	Compass Minerals International, Inc.	Estados Unidos	Materials	us market
CNK	Cinemark Holdings, Inc.	Estados Unidos	Communication Services	us market
CNMD	CONMED Corporation	Estados Unidos	Health Care	us market
CNSL	Consolidated Communications Holdings, Inc.	Estados Unidos	Communication Services	us market
CNXN	PC Connection, Inc.	Estados Unidos	Information Technology	us market
COHU	Cohu, Inc.	Estados Unidos	Information Technology	us market
COLL	Collegium Pharmaceutical, Inc.	Estados Unidos	Health Care	us market
COOP	Mr. Cooper Group, Inc.	Estados Unidos	Financials	us market
CORT	Corcept Therapeutics Incorporated	Estados Unidos	Health Care	us market
CPE	Callon Petroleum	Estados Unidos	Energy	us market
CPF	Central Pacific Financial Corp.	Estados Unidos	Financials	us market
CPK	Chesapeake Utilities Corp.	Estados Unidos	Utilities	us market
CPRX	Catalyst Pharmaceuticals Partners, Inc.	Estados Unidos	Health Care	us market

Continúa en la siguiente página

Cuadro F.1 – Continúa de la página anterior

Símbolo	Nombre de la Compañía	País	Sector	Mercado
CRC	California Resources Corporation	Estados Unidos	Energy	us market
CRK	Comstock Resources, Inc.	Estados Unidos	Energy	us market
CRMT	Americas Carmart, Inc.	Estados Unidos	Consumer Discretionary	us market
CRNC	Cerence, Inc.	Estados Unidos	Information Technology	us market
CRS	Carpenter Technology	Estados Unidos	Materials	us market
CRSR	Corsair Gaming	Estados Unidos	Information Technology	us market
CRVL	CorVel Corporation	Estados Unidos	Health Care	us market
CSGS	CSG Systems International, Inc.	Estados Unidos	Industrials	us market
CSR	Centerspace Trust	Estados Unidos	Real Estate	us market
CTKB	Cytek Biosciences, Inc.	Estados Unidos	Health Care	us market
CTRE	CareTrust REIT, Inc.	Estados Unidos	Real Estate	us market
CTS	CTS Corporation	Estados Unidos	Information Technology	us market
CUBI	Customers Bancorp, Inc.	Estados Unidos	Financials	us market
CVBF	CVB Financial Corp.	Estados Unidos	Financials	us market
CVCO	Cavco Industries, Inc.	Estados Unidos	Consumer Discretionary	us market
CVGW	Calavo Growers, Inc.	Estados Unidos	Consumer Staples	us market
CVI	CVR Energy, Inc.	Estados Unidos	Energy	us market
CWEN	Clearway Energy, Inc.	Estados Unidos	Utilities	us market
CWEN.A	Clearway Energy, Inc. (Class A)	Estados Unidos	Utilities	us market
CWK	Cushman & Wakefield plc	Estados Unidos	Real Estate	us market
CWT	California Water Service Group	Estados Unidos	Utilities	us market
CXW	CoreCivic	Estados Unidos	Industrials	us market

Continúa en la siguiente página

Cuadro F.1 – Continúa de la página anterior

Símbolo	Nombre de la Compañía	País	Sector	Mercado
CYH	Community Health Systems, Inc.	Estados Unidos	Health Care	us market
CYTK	Cytokinetics, Incorporated	Estados Unidos	Health Care	us market
DAN	Dana Incorporated	Estados Unidos	Consumer Discretionary	us market
DBI	Designer Brands, Inc.	Estados Unidos	Consumer Discretionary	us market
DCOM	Dime Community Bancshares, Inc.	Estados Unidos	Financials	us market
DDD	3D Systems Corporation	Estados Unidos	Industrials	us market
DEA	Easterly Government Properties, Inc.	Estados Unidos	Real Estate	us market
DEI	Douglas Emmett	Estados Unidos	Real Estate	us market
DFIN	Donnelley Financial Solutions, Inc.	Estados Unidos	Financials	us market
DGII	Digi International Inc.	Estados Unidos	Information Technology	us market
DIN	Dine Brands Global, Inc.	Estados Unidos	Consumer Discretionary	us market
DIOD	Diodes Incorporated	Estados Unidos	Information Technology	us market
DISH	Dish Network	Estados Unidos	Communication Services	us market
DLX	Deluxe Corporation	Estados Unidos	Industrials	us market
DNOW	NOW Inc	Estados Unidos	Industrials	us market
DORM	Dorman Products, Inc.	Estados Unidos	Consumer Discretionary	us market
DRH	DiamondRock Hospitality Company	Estados Unidos	Real Estate	us market
DRQ	Dril-Quip Inc.	Estados Unidos	Energy	us market
DV	DoubleVerify Holdings, Inc.	Estados Unidos	Information Technology	us market

Continúa en la siguiente página

Cuadro F.1 – Continúa de la página anterior

Símbolo	Nombre de la Compañía	País	Sector	Mercado
DVAX	Dynavax Technologies Corp.	Estados Unidos	Health Care	us market
DXC	DXC Technology	Estados Unidos	Information Technology	us market
DXPE	DXP Enterprises, Inc.	Estados Unidos	Industrials	us market
DY	Dycom Industries, Inc.	Estados Unidos	Industrials	us market
EAT	Brinker International, Inc.	Estados Unidos	Consumer Discretionary	us market
ECPG	Encore Capital Group, Inc.	Estados Unidos	Financials	us market
EFC	Ellington Financial, Inc.	Estados Unidos	Financials	us market
EGBN	Eagle Bancorp Inc	Estados Unidos	Financials	us market
EHAB	Enhabit, Inc.	Estados Unidos	Health Care	us market
EIG	Employers Holdings, Inc.	Estados Unidos	Financials	us market
ELF	e.l.f. Beauty, Inc.	Estados Unidos	Consumer Staples	us market
ELME	Elme Communities	Estados Unidos	Real Estate	us market
EMBC	Embecta Corp.	Estados Unidos	Health Care	us market
ENR	Energizer	Estados Unidos	Consumer Staples	us market
ENSG	Ensign Group, Inc.	Estados Unidos	Health Care	us market
ENV	Envestnet, Inc.	Estados Unidos	Information Technology	us market
ENVA	Enova International, Inc.	Estados Unidos	Financials	us market
EPAC	Enerpac Tool Group	Estados Unidos	Industrials	us market
EPC	Edgewell Personal Care	Estados Unidos	Consumer Staples	us market
EPRT	Essential Properties Realty Trust, Inc.	Estados Unidos	Real Estate	us market
ESE	ESCO Technologies Inc.	Estados Unidos	Industrials	us market
ETD	Ethan Allen Interiors, Inc.	Estados Unidos	Consumer Discretionary	us market
EVTC	EVERTEC, Inc.	Estados Unidos	Financials	us market

Continúa en la siguiente página

Cuadro F.1 – Continúa de la página anterior

Símbolo	Nombre de la Compañía	País	Sector	Mercado
EXPI	eXp World Holdings, Inc.	Estados Unidos	Real Estate	us market
EXTR	Extreme Networks, Inc.	Estados Unidos	Information Technology	us market
EYE	National Vision Holdings	Estados Unidos	Consumer Discretionary	us market
EZPW	EZCORP, Inc.	Estados Unidos	Financials	us market
FBK	FB Financial Corp.	Estados Unidos	Financials	us market
FBNC	First Bancorp (Southern Pines NC)	Estados Unidos	Financials	us market
FBP	First BanCorp (Puerto Rico)	Estados Unidos	Financials	us market
FBRT	Franklin BSP Realty Trust, Inc.	Estados Unidos	Real Estate	us market
FCF	First Commonwealth Financial, Corp.	Estados Unidos	Financials	us market
FCPT	Four Corners Property Trust, Inc.	Estados Unidos	Real Estate	us market
FDP	Fresh Del Monte Produce, Inc.	Estados Unidos	Consumer Staples	us market
FELE	Franklin Electric	Estados Unidos	Industrials	us market
FFBC	First Financial Bancorp.	Estados Unidos	Financials	us market
FHB	First Hawaiian, Inc.	Estados Unidos	Financials	us market
FIX	Comfort Systems USA, Inc.	Estados Unidos	Industrials	us market
FIZZ	National Beverage Corp.	Estados Unidos	Consumer Staples	us market
FL	Foot Locker	Estados Unidos	Consumer Discretionary	us market
FLGT	Fulgent Genetics, Inc.	Estados Unidos	Health Care	us market
FN	Fabrinet	Estados Unidos	Information Technology	us market
FORM	FormFactor, Inc.	Estados Unidos	Information Technology	us market

Continúa en la siguiente página

Cuadro F.1 – Continúa de la página anterior

Símbolo	Nombre de la Compañía	País	Sector	Mercado
FORR	Forrester Research, Inc.	Estados Unidos	Industrials	us market
FSS	Federal Signal Corporation	Estados Unidos	Industrials	us market
FTDR	Frontdoor, Inc.	Estados Unidos	Consumer Discretionary	us market
FTRE	Fortrea Holdings, Inc.	Estados Unidos	Health Care	us market
FUL	H.B. Fuller Company	Estados Unidos	Materials	us market
FULT	Fulton Financial Corporation	Estados Unidos	Financials	us market
FWRD	Forward Air Corp.	Estados Unidos	Industrials	us market
GBX	The Greenbrier Companies, Inc.	Estados Unidos	Industrials	us market
GDEN	Golden Entertainment	Estados Unidos	Consumer Discretionary	us market
GDOT	Green Dot Corporation	Estados Unidos	Financials	us market
GEO	GEO Group, Inc.	Estados Unidos	Industrials	us market
GES	Guess, Inc.	Estados Unidos	Consumer Discretionary	us market
GFF	Griffon Corporation	Estados Unidos	Industrials	us market
GIII	G-III Apparel Group, Ltd.	Estados Unidos	Consumer Discretionary	us market
GKOS	Glaukos Corp.	Estados Unidos	Health Care	us market
GMS	GMS, Inc.	Estados Unidos	Industrials	us market
GNL	Global Net Lease, Inc.	Estados Unidos	Real Estate	us market
GNW	Genworth Financial, Inc.	Estados Unidos	Financials	us market
GOGO	Gogo, Inc.	Estados Unidos	Communication Services	us market
GPI	Group 1 Automotive, Inc.	Estados Unidos	Consumer Discretionary	us market
GPRE	Green Plains, Inc.	Estados Unidos	Energy	us market
GRBK	Green Brick Partners, Inc.	Estados Unidos	Consumer Discretionary	us market
GSHD	Goosehead Insurance, Inc.	Estados Unidos	Financials	us market

Continúa en la siguiente página

Cuadro F.1 – Continúa de la página anterior

Símbolo	Nombre de la Compañía	País	Sector	Mercado
GTY	Getty Realty Corp.	Estados Unidos	Real Estate	us market
GVA	Granite Construction, Inc.	Estados Unidos	Industrials	us market
HAFC	Hanmi Financial Corporation	Estados Unidos	Financials	us market
HAIN	Hain Celestial Group	Estados Unidos	Consumer Staples	us market
HASI	Hannon Armstrong Sustainable Infrastructure Capital, Inc.	Estados Unidos	Financials	us market
HAYN	Haynes International, Inc.	Estados Unidos	Materials	us market
HAYW	Hayward Holdings, Inc.	Estados Unidos	Industrials	us market
HBI	Hanesbrands, Inc.	Estados Unidos	Consumer Discretionary	us market
HCC	Warrior Met Coal, Inc.	Estados Unidos	Materials	us market
HCI	HCI Group, Inc.	Estados Unidos	Financials	us market
HCSG	Healthcare Services Group, Inc.	Estados Unidos	Industrials	us market
HFWA	Heritage Financial Corporation	Estados Unidos	Health Care	us market
HI	Hillenbrand, Inc.	Estados Unidos	Industrials	us market
HIBB	Hibbett Sports, Inc.	Estados Unidos	Consumer Discretionary	us market
HIW	Highwoods Properties]	Estados Unidos	Real Estate	us market
HLIT	Harmonic Inc.	Estados Unidos	Information Technology	us market
HLX	Helix Energy Solutions Group, Inc.	Estados Unidos	Energy	us market
HMN	Horace Mann Educators Corporation	Estados Unidos	Financials	us market
HNI	HNI Corporation	Estados Unidos	Industrials	us market
HOPE	Hope Bancorp, Inc.	Estados Unidos	Financials	us market
HOUS	Anywhere Real Estate	Estados Unidos	Real Estate	us market

Continúa en la siguiente página

Cuadro F.1 – Continúa de la página anterior

Símbolo	Nombre de la Compañía	País	Sector	Mercado
HP	Helmerich & Payne, Inc.	Estados Unidos	Energy	us market
HPP	Hudson Pacific Properties	Estados Unidos	Financials	us market
HRMY	Harmony Biosciences Holdings, Inc.	Estados Unidos	Health Care	us market
HSII	Heidrick & Struggles International, Inc.	Estados Unidos	Industrials	us market
HSTM	HealthStream, Inc.	Estados Unidos	Health Care	us market
HTH	Hilltop Holdings Inc.	Estados Unidos	Financials	us market
HTLD	Heartland Express, Inc.	Estados Unidos	Industrials	us market
HUBG	Hub Group, Inc.	Estados Unidos	Industrials	us market
HVT	Haverty Furniture Companies, Inc.	Estados Unidos	Consumer Discretionary	us market
HWKN	Hawkins, Inc.	Estados Unidos	Materials	us market
HZO	MarineMax, Inc.	Estados Unidos	Consumer Discretionary	us market
IBP	Installed Building Products, Inc.	Estados Unidos	Consumer Discretionary	us market
IBTX	Independent Bank Group, Inc.	Estados Unidos	Financials	us market
ICHR	Ichor Holdings, Ltd.	Estados Unidos	Information Technology	us market
ICUI	ICU Medical	Estados Unidos	Health Care	us market
IDCC	InterDigital, Inc.	Estados Unidos	Information Technology	us market
IIIN	Insteel Industries, Inc.	Estados Unidos	Industrials	us market
IIPR	Innovative Industrial Properties, Inc.	Estados Unidos	Real Estate	us market
INDB	Independent Bank Corp.	Estados Unidos	Financials	us market
INN	Summit Hotel Properties, Inc.	Estados Unidos	Real Estate	us market
INVA	Innoviva, Inc.	Estados Unidos	Health Care	us market
IOSP	Innospec, Inc.	Estados Unidos	Materials	us market

Continúa en la siguiente página

Cuadro F.1 – Continúa de la página anterior

Símbolo	Nombre de la Compañía	País	Sector	Mercado
IPAR	Inter Parfums, Inc.	Estados Unidos	Consumer Staples	us market
IRBT	iRobot Corporation	Estados Unidos	Consumer Discretionary	us market
IRWD	Ironwood Pharmaceuticals, Inc.	Estados Unidos	Health Care	us market
ITGR	Integer Holdings Corporation	Estados Unidos	Health Care	us market
ITRI	Itron, Inc.	Estados Unidos	Information Technology	us market
IVR	Invesco Mortgage Capital, Inc.	Estados Unidos	Real Estate	us market
JACK	Jack in the Box	Estados Unidos	Consumer Discretionary	us market
JBGS	JBG Smith	Estados Unidos	Real Estate	us market
JBLU	JetBlue	Estados Unidos	Industrials	us market
JBSS	John B. Sanfilippo & Son, Inc.	Estados Unidos	Consumer Staples	us market
JBT	John Bean Technologies Corporation	Estados Unidos	Industrials	us market
JJSF	J&J Snack Foods Corp.	Estados Unidos	Consumer Staples	us market
JOE	St. Joe Company	Estados Unidos	Real Estate	us market
JRVR	James River Group Holdings, Ltd.	Estados Unidos	Financials	us market
JXN	Jackson Financial, Inc.	Estados Unidos	Financials	us market
KALU	Kaiser Aluminum Corporation	Estados Unidos	Materials	us market
KAMN	Kaman Corporation	Estados Unidos	Industrials	us market
KAR	OPENLANE, Inc.	Estados Unidos	Industrials	us market
KELYA	Kelly Services, Inc.	Estados Unidos	Industrials	us market
KFY	Korn/Ferry International	Estados Unidos	Industrials	us market

Continúa en la siguiente página

Cuadro F.1 – Continúa de la página anterior

Símbolo	Nombre de la Compañía	País	Sector	Mercado
KLG	WK Kellogg Co	Estados Unidos	Consumer Staples	us market
KLIC	Kulicke and Soffa Industries, Inc.	Estados Unidos	Information Technology	us market
KMT	Kennametal	Estados Unidos	Industrials	us market
KN	Knowles Corporation	Estados Unidos	Information Technology	us market
KOP	Koppers Holdings, Inc.	Estados Unidos	Materials	us market
KREF	KKR Real Estate Finance Trust, Inc.	Estados Unidos	Financials	us market
KSS	Kohl's Corp.	Estados Unidos	Consumer Discretionary	us market
KTB	Kontoor Brands	Estados Unidos	Consumer Discretionary	us market
KW	Kennedy-Wilson Holdings, Inc.	Estados Unidos	Real Estate	us market
KWR	Quaker Chemical Corporation	Estados Unidos	Materials	us market
LBRT	Liberty Energy, Inc.	Estados Unidos	Energy	us market
LCII	LCI Industries	Estados Unidos	Consumer Discretionary	us market
LESL	Leslie's, Inc.	Estados Unidos	Consumer Discretionary	us market
LGIH	LGI Homes	Estados Unidos	Consumer Discretionary	us market
LGND	Ligand Pharmaceuticals, Inc.	Estados Unidos	Health Care	us market
LKFN	Lakeland Financial	Estados Unidos	Financials	us market
LMAT	LeMaitre Vascular	Estados Unidos	Health Care	us market
LNC	Lincoln Financial	Estados Unidos	Financials	us market
LNN	Lindsay Corporation	Estados Unidos	Materials	us market
LPG	Dorian LPG Ltd.	Estados Unidos	Energy	us market

Continúa en la siguiente página

Cuadro F.1 – Continúa de la página anterior

Símbolo	Nombre de la Compañía	País	Sector	Mercado
LQDT	Liquidity Services, Inc.	Estados Unidos	Industrials	us market
LRN	Stride, Inc.	Estados Unidos	Consumer Discretionary	us market
LTC	LTC Properties, Inc.	Estados Unidos	Real Estate	us market
LTHM	Livent Corp.	Estados Unidos	Materials	us market
LUMN	Lumen Technologies	Estados Unidos	Communication Services	us market
LXP	Lexington Realty Trust	Estados Unidos	Real Estate	us market
LZB	La-Z-Boy, Inc.	Estados Unidos	Consumer Discretionary	us market
MAC	Macerich	Estados Unidos	Real Estate	us market
MATV	Mativ Holdings, Inc.	Estados Unidos	Materials	us market
MATW	Matthews International Corporation	Estados Unidos	Industrials	us market
MATX	Matson, Inc.	Estados Unidos	Industrials	us market
MBC	MasterBrand, Inc.	Estados Unidos	Industrials	us market
MC	Moelis & Company	Estados Unidos	Financials	us market
MCRI	Monarch Casino & Resort, Inc.	Estados Unidos	Consumer Discretionary	us market
MCS	The Marcus Corporation	Estados Unidos	Consumer Discretionary	us market
MCW	Mister Car Wash, Inc.	Estados Unidos	Consumer Discretionary	us market
MCY	Mercury General	Estados Unidos	Financials	us market
MD	Mednax	Estados Unidos	Health Care	us market
MDC	M.D.C. Holdings, Inc.	Estados Unidos	Consumer Discretionary	us market
MDRX	Veradigm	Estados Unidos	Health Care	us market
MED	Medifast, Inc.	Estados Unidos	Consumer Staples	us market
MEI	Methode Electronics, Inc.	Estados Unidos	Information Technology	us market

Continúa en la siguiente página

Cuadro F.1 – Continúa de la página anterior

Símbolo	Nombre de la Compañía	País	Sector	Mercado
MERC	Mercer International, Inc.	Estados Unidos	Materials	us market
MGPI	MGP Ingredients, Inc.	Estados Unidos	Consumer Staples	us market
MHO	M/I Homes, Inc.	Estados Unidos	Consumer Discretionary	us market
MLAB	Mesa Laboratories Inc	Estados Unidos	Health Care	us market
MLKN	MillerKnoll, Inc.	Estados Unidos	Industrials	us market
MLI	Mueller Industries, Inc.	Estados Unidos	Industrials	us market
MMI	Marcus & Millichap, Inc.	Estados Unidos	Real Estate	us market
MMSI	Merit Medical Systems, Inc.	Estados Unidos	Health Care	us market
MNRO	Monro, Inc.	Estados Unidos	Consumer Discretionary	us market
MODV	ModivCare, Inc.	Estados Unidos	Health Care	us market
MOG-A	Moog Inc.	Estados Unidos	Industrials	us market
MOV	Movado Group, Inc.	Estados Unidos	Consumer Discretionary	us market
MRCY	Mercury Systems	Estados Unidos	Industrials	us market
MRTN	Marten Transport Ltd.	Estados Unidos	Industrials	us market
MSEX	Middlesex Water Company	Estados Unidos	Utilities	us market
MSGS	Madison Square Garden Sports Corp.	Estados Unidos	Communication Services	us market
MTH	Meritage Homes Corporation	Estados Unidos	Consumer Discretionary	us market
MTRN	Materion Corp.	Estados Unidos	Materials	us market
MTX	Minerals Technologies	Estados Unidos	Materials	us market
MXL	MaxLinear, Inc.	Estados Unidos	Information Technology	us market
MYE	Myers Industries, Inc.	Estados Unidos	Materials	us market
MYGN	Myriad Genetics, Inc.	Estados Unidos	Health Care	us market
MYRG	MYR Group, Inc.	Estados Unidos	Industrials	us market
NABL	N-able, Inc.	Estados Unidos	Information Technology	us market

Continúa en la siguiente página

Cuadro F.1 – Continúa de la página anterior

Símbolo	Nombre de la Compañía	País	Sector	Mercado
NATL	NCR Atleos	Estados Unidos	Financials	us market
NAVI	Navient	Estados Unidos	Financials	us market
NBHC	National Bank Holdings Corporation	Estados Unidos	Financials	us market
NBR	Nabors Industries, Ltd.	Estados Unidos	Energy	us market
NBTB	NBT Bancorp, Inc.	Estados Unidos	Financials	us market
NEO	NeoGenomics Laboratories, Inc.	Estados Unidos	Health Care	us market
NFBK	Northfield Bancorp, Inc. (Staten Island, NY)	Estados Unidos	Financials	us market
NGVT	Ingevity, Corp.	Estados Unidos	Materials	us market
NMIH	NMI Holdings, Inc.	Estados Unidos	Financials	us market
NOG	Northern Oil and Gas, Inc.	Estados Unidos	Energy	us market
NPK	National Presto Industries, Inc.	Estados Unidos	Industrials	us market
NPO	EnPro Industries, Inc.	Estados Unidos	Industrials	us market
NSIT	Insight Enterprises, Inc.	Estados Unidos	Industrials	us market
NTCT	NETSCOUT Systems, Inc.	Estados Unidos	Information Technology	us market
NUS	Nu Skin Enterprises	Estados Unidos	Consumer Staples	us market
NVEE	NV5 Global	Estados Unidos	Industrials	us market
NVRI	Enviri Corporation	Estados Unidos	Industrials	us market
NWBI	Northwest Bancshares, Inc.	Estados Unidos	Financials	us market
NWL	Newell Brands	Estados Unidos	Consumer Discretionary	us market
NWN	NW Natural	Estados Unidos	Utilities	us market
NX	Quanex Building Products Corporation	Estados Unidos	Industrials	us market
NXRT	NexPoint Residential Trust, Inc.	Estados Unidos	Real Estate	us market

Continúa en la siguiente página

Cuadro F.1 – Continúa de la página anterior

Símbolo	Nombre de la Compañía	País	Sector	Mercado
NYMT	New York Mortgage Trust, Inc.	Estados Unidos	Real Estate	us market
ODP	The ODP Corporation	Estados Unidos	Consumer Discretionary	us market
OFG	OFG Bancorp	Estados Unidos	Financials	us market
OFIX	Orthofix Medical, Inc.	Estados Unidos	Health Care	us market
OGN	Organon & Co.	Estados Unidos	Health Care	us market
OI	O-I Glass, Inc.	Estados Unidos	Materials	us market
OII	Oceaneering International, Inc.	Estados Unidos	Energy	us market
OIS	Oil States International, Inc.	Estados Unidos	Energy	us market
OMCL	Omnicell	Estados Unidos	Health Care	us market
OMI	Owens & Minor, Inc.	Estados Unidos	Health Care	us market
OSIS	OSI Systems, Inc.	Estados Unidos	Information Technology	us market
OSPN	OneSpan, Inc.	Estados Unidos	Information Technology	us market
OSUR	OraSure Technologies, Inc.	Estados Unidos	Health Care	us market
OTTR	Otter Tail Corporation	Estados Unidos	Utilities	us market
OUT	Outfront Media	Estados Unidos	Real Estate	us market
OXM	Oxford Industries, Inc.	Estados Unidos	Consumer Discretionary	us market
PAHC	Phibro Animal Health	Estados Unidos	Health Care	us market
PARR	Par Pacific Holdings, Inc.	Estados Unidos	Energy	us market
PAYO	Payoneer Global Inc.	Estados Unidos	Financials	us market
PATK	Patrick Industries, Inc.	Estados Unidos	Consumer Discretionary	us market
PBH	Prestige Consumer Healthcare	Estados Unidos	Health Care	us market
PBI	Pitney Bowes, Inc.	Estados Unidos	Industrials	us market
PCRX	Pacira BioSciences, Inc.	Estados Unidos	Health Care	us market

Continúa en la siguiente página

Cuadro F.1 – Continúa de la página anterior

Símbolo	Nombre de la Compañía	País	Sector	Mercado
PDFS	PDF Solutions, Inc.	Estados Unidos	Information Technology	us market
PEB	Pebblebrook Hotel Trust	Estados Unidos	Real Estate	us market
PECO	Phillips Edison & Company	Estados Unidos	Real Estate	us market
PFBC	Preferred Bank	Estados Unidos	Financials	us market
PFS	Provident Financial Services, Inc.	Estados Unidos	Financials	us market
PGTI	PGT Innovations, Inc.	Estados Unidos	Industrials	us market
PHIN	PHINIA, Inc.	Estados Unidos	Consumer Discretionary	us market
PINC	Premier, Inc.	Estados Unidos	Health Care	us market
PIPR	Piper Sandler Companies	Estados Unidos	Financials	us market
PLAB	Photronics, Inc.	Estados Unidos	Information Technology	us market
PLAY	Dave & Buster's Entertainment, Inc.	Estados Unidos	Consumer Discretionary	us market
PLMR	Palomar Holdings, Inc.	Estados Unidos	Financials	us market
PLUS	ePlus, Inc.	Estados Unidos	Information Technology	us market
PLXS	Plexus Corp.	Estados Unidos	Information Technology	us market
PMT	PennyMac Mortgage Investment Trust	Estados Unidos	Real Estate	us market
POWL	Powell Industries, Inc.	Estados Unidos	Industrials	us market
PPBI	Pacific Premier Bancorp, Inc.	Estados Unidos	Financials	us market
PRA	ProAssurance Corporation	Estados Unidos	Financials	us market
PRAA	PRA Group, Inc.	Estados Unidos	Financials	us market
PRDO	Perdoceo Education Corp.	Estados Unidos	Consumer Discretionary	us market
PRFT	Perficient, Inc.	Estados Unidos	Information Technology	us market

Continúa en la siguiente página

Cuadro F.1 – Continúa de la página anterior

Símbolo	Nombre de la Compañía	País	Sector	Mercado
PRG	PROG Holdings, Inc.	Estados Unidos	Financials	us market
PRGS	Progress Software Corporation	Estados Unidos	Information Technology	us market
PRK	Park National Corp.	Estados Unidos	Financials	us market
PRLB	Protolabs	Estados Unidos	Industrials	us market
PRVA	Privia Health Group, Inc.	Estados Unidos	Health Care	us market
PSMT	PriceSmart	Estados Unidos	Consumer Staples	us market
PTEN	Patterson-UTI Energy, Inc.	Estados Unidos	Energy	us market
PUMP	ProPetro Holding Corp.	Estados Unidos	Energy	us market
PZZA	Papa John's Pizza	Estados Unidos	Consumer Discretionary	us market
QNST	QuinStreet, Inc.	Estados Unidos	Communication Services	us market
RAMP	LiveRamp Holdings, Inc.	Estados Unidos	Information Technology	us market
RC	Ready Capital Corp	Estados Unidos	Financials	us market
RCUS	Arcus Biosciences, Inc.	Estados Unidos	Health Care	us market
RDN	Radian Group, Inc.	Estados Unidos	Financials	us market
RDNT	RadNet, Inc.	Estados Unidos	Health Care	us market
RES	RPC, Inc.	Estados Unidos	Energy	us market
REX	REX American Resources Corporation	Estados Unidos	Energy	us market
REZI	Resideo Technologies, Inc.	Estados Unidos	Industrials	us market
RGNX	REGENXBIO Inc.	Estados Unidos	Health Care	us market
RGP	Resources Connection, Inc.	Estados Unidos	Industrials	us market
RGR	Sturm, Ruger & Company, Inc.	Estados Unidos	Consumer Discretionary	us market
RILY	B. Riley Financial	Estados Unidos	Financials	us market
RMBS	Rambus, Inc.	Estados Unidos	Information Technology	us market
RNST	Renasant Corp.	Estados Unidos	Financials	us market

Continúa en la siguiente página

Cuadro F.1 – Continúa de la página anterior

Símbolo	Nombre de la Compañía	País	Sector	Mercado
ROCK	Gibraltar Industries, Inc.	Estados Unidos	Industrials	us market
ROG	Rogers Corporation	Estados Unidos	Information Technology	us market
ROIC	Retail Opportunity Investments Corp.	Estados Unidos	Real Estate	us market
RPT	Ramco-Gershenson Properties Trust	Estados Unidos	Real Estate	us market
RWT	Redwood Trust, Inc.	Estados Unidos	Real Estate	us market
RXO	RXO, Inc.	Estados Unidos	Industrials	us market
SAFE	Safehold, Inc.	Estados Unidos	Real Estate	us market
SABR	Sabre	Estados Unidos	Consumer Discretionary	us market
SAFT	Safety Insurance Group, Inc.	Estados Unidos	Financials	us market
SAH	Sonic Automotive, Inc.	Estados Unidos	Consumer Discretionary	us market
SANM	Sanmina Corporation	Estados Unidos	Information Technology	us market
SBCF	Seacoast Banking Corporation of Florida	Estados Unidos	Financials	us market
SBH	Sally Beauty Holdings, Inc.	Estados Unidos	Consumer Discretionary	us market
SBSI	Southside Bancshares, Inc.	Estados Unidos	Financials	us market
SCHL	Scholastic Corporation	Estados Unidos	Consumer Discretionary	us market
SCL	Stepan Company	Estados Unidos	Materials	us market
SCSC	ScanSource, Inc.	Estados Unidos	Information Technology	us market
SCVL	Shoe Carnival, Inc.	Estados Unidos	Consumer Discretionary	us market
SDGR	Schrödinger, Inc.	Estados Unidos	Health Care	us market

Continúa en la siguiente página

Cuadro F.1 – Continúa de la página anterior

Símbolo	Nombre de la Compañía	País	Sector	Mercado
SEM	Select Medical Holdings, Corp.	Estados Unidos	Health Care	us market
SFBS	ServisFirst Bancshares, Inc.	Estados Unidos	Financials	us market
SFNC	Simmons First National Corporation	Estados Unidos	Financials	us market
SGH	SMART Global Holdings, Inc.	Estados Unidos	Information Technology	us market
SHAK	Shake Shack, Inc.	Estados Unidos	Consumer Discretionary	us market
SHEN	Shenandoah Telecommunications Co	Estados Unidos	Communication Services	us market
SHO	Sunstone Hotel Investors, Inc.	Estados Unidos	Real Estate	us market
SHOO	Steven Madden, Ltd.	Estados Unidos	Consumer Discretionary	us market
SIG	Signet Jewelers	Estados Unidos	Consumer Discretionary	us market
SITC	SITE Centers Corp.	Estados Unidos	Real Estate	us market
SITM	SiTime	Estados Unidos	Information Technology	us market
SIX	Six Flags	Estados Unidos	Consumer Discretionary	us market
SJW	SJW Group	Estados Unidos	Utilities	us market
SKT	Tanger Factory Outlet Centers, Inc.	Estados Unidos	Real Estate	us market
SKYW	SkyWest, Inc.	Estados Unidos	Industrials	us market
SLCA	U.S. Silica Holdings, Inc.	Estados Unidos	Energy	us market
SLG	SL Green Realty	Estados Unidos	Real Estate	us market
SLP	Simulations Plus, Inc.	Estados Unidos	Health Care	us market
SLVM	Sylvamo Corp.	Estados Unidos	Materials	us market
SM	SM Energy Company	Estados Unidos	Energy	us market

Continúa en la siguiente página

Cuadro F.1 – Continúa de la página anterior

Símbolo	Nombre de la Compañía	País	Sector	Mercado
SMP	Standard Motor Products, Inc.	Estados Unidos	Consumer Discretionary	us market
SMPL	Simply Good Foods Company	Estados Unidos	Consumer Staples	us market
SMTC	Semtech Corporation	Estados Unidos	Information Technology	us market
SNCY	Sun Country Airlines	Estados Unidos	Industrials	us market
SNEX	StoneX Group Inc.	Estados Unidos	Financials	us market
SONO	Sonos, Inc.	Estados Unidos	Consumer Discretionary	us market
SPNT	SiriusPoint Ltd.	Estados Unidos	Financials	us market
SPSC	SPS Commerce, Inc.	Estados Unidos	Information Technology	us market
SPTN	SpartanNash Company	Estados Unidos	Consumer Staples	us market
SPWR	SunPower Corporation	Estados Unidos	Industrials	us market
SPXC	SPX Corporation	Estados Unidos	Industrials	us market
SSP	The E.W. Scripps Company	Estados Unidos	Consumer Discretionary	us market
SSTK	Shutterstock, Inc.	Estados Unidos	Communication Services	us market
STAA	STAAR Surgical Company	Estados Unidos	Health Care	us market
STBA	S&T Bancorp, Inc.	Estados Unidos	Financials	us market
STC	Stewart Information Services Corporation	Estados Unidos	Financials	us market
STEL	Stellar Bancorp, Inc.	Estados Unidos	Financials	us market
STRA	Strategic Education, Inc.	Estados Unidos	Consumer Discretionary	us market
SUPN	Supernus Pharmaceuticals, Inc.	Estados Unidos	Health Care	us market
SVC	Service Properties Trust	Estados Unidos	Real Estate	us market
SXC	SunCoke Energy, Inc.	Estados Unidos	Materials	us market

Continúa en la siguiente página

Cuadro F.1 – Continúa de la página anterior

Símbolo	Nombre de la Compañía	País	Sector	Mercado
SXI	Standex International Corporation	Estados Unidos	Industrials	us market
SXT	Sensient Technologies	Estados Unidos	Materials	us market
TALO	Talos Energy, Inc.	Estados Unidos	Energy	us market
TBBK	The Bancorp, Inc.	Estados Unidos	Financials	us market
TBI	TrueBlue, Inc.	Estados Unidos	Industrials	us market
TDS	Telephone and Data Systems, Inc.	Estados Unidos	Communication Services	us market
TFIN	Triumph Bancorp, Inc.	Estados Unidos	Financials	us market
TGI	Triumph Group, Inc.	Estados Unidos	Industrials	us market
THRM	Gentherm Incorporated	Estados Unidos	Consumer Discretionary	us market
THRY	Thryv Holdings, Inc.	Estados Unidos	Communication Services	us market
THS	TreeHouse Foods, Inc.	Estados Unidos	Consumer Staples	us market
TILE	Interface, Inc.	Estados Unidos	Industrials	us market
TMP	Tompkins Financial Corporation	Estados Unidos	Financials	us market
TMST	TimkenSteel Corporation	Estados Unidos	Materials	us market
TNC	Tennant Company	Estados Unidos	Industrials	us market
TNDM	Tandem Diabetes Care	Estados Unidos	Health Care	us market
TPH	Tri Pointe Homes, Inc.	Estados Unidos	Consumer Discretionary	us market
TR	Tootsie Roll Industries, Inc.	Estados Unidos	Consumer Staples	us market
TRIP	TripAdvisor	Estados Unidos	Communication Services	us market
TRMK	Trustmark Corp.	Estados Unidos	Financials	us market
TRN	Trinity Industries, Inc.	Estados Unidos	Industrials	us market
TRST	TrustCo Bank Corp NY	Estados Unidos	Financials	us market
TRUP	Trupanion	Estados Unidos	Financials	us market

Continúa en la siguiente página

Cuadro F.1 – Continúa de la página anterior

Símbolo	Nombre de la Compañía	País	Sector	Mercado
TTEC	TTEC Holdings, Inc.	Estados Unidos	Industrials	us market
TTGT	TechTarget	Estados Unidos	Communication Services	us market
TTMI	TTM Technologies, Inc.	Estados Unidos	Information Technology	us market
TWI	Titan International, Inc.	Estados Unidos	Materials	us market
TWO	Two Harbors Investment Corp.	Estados Unidos	Financials	us market
UCBI	United Community Banks, Inc.	Estados Unidos	Financials	us market
UCTT	Ultra Clean Holdings, Inc.	Estados Unidos	Information Technology	us market
UE	Urban Edge Properties	Estados Unidos	Real Estate	us market
UFCS	United Fire Group, Inc.	Estados Unidos	Financials	us market
UFPT	UFP Technologies, Inc.	Estados Unidos	Health Care	us market
UHT	Universal Health Realty Income Trust	Estados Unidos	Real Estate	us market
UNF	UniFirst Corporation	Estados Unidos	Industrials	us market
UNFI	United Natural Foods Inc	Estados Unidos	Consumer Staples	us market
UNIT	Uniti Group	Estados Unidos	Real Estate	us market
UPBD	Upbound Group, Inc.	Estados Unidos	Consumer Discretionary	us market
URBN	Urban Outfitters, Inc.	Estados Unidos	Consumer Discretionary	us market
USNA	Usana Health Sciences, Inc.	Estados Unidos	Consumer Staples	us market
USPH	U.S. Physical Therapy, Inc.	Estados Unidos	Health Care	us market
UTL	Unitil Corporation	Estados Unidos	Utilities	us market
UVV	Universal Corporation	Estados Unidos	Consumer Staples	us market
VBTX	Veritex Holdings, Inc.	Estados Unidos	Financials	us market

Continúa en la siguiente página

Cuadro F.1 – Continúa de la página anterior

Símbolo	Nombre de la Compañía	País	Sector	Mercado
VCEL	Vericel	Estados Unidos	Health Care	us market
VECO	Veeco Instruments Inc.	Estados Unidos	Information Technology	us market
VGR	Vector Group	Estados Unidos	Consumer Staples	us market
VIAV	VIAVI Solutions	Estados Unidos	Information Technology	us market
VICR	Vicor Corporation	Estados Unidos	Industrials	us market
VIR	Vir Biotechnology, Inc.	Estados Unidos	Health Care	us market
VRE	Veris Residential, Inc.	Estados Unidos	Real Estate	us market
VREX	Varex Imaging Corporation	Estados Unidos	Health Care	us market
VRRM	Verra Mobility Corporation	Estados Unidos	Industrials	us market
VRTS	Virtus Investment Partners, Inc.	Estados Unidos	Financials	us market
VSAT	Viasat, Inc.	Estados Unidos	Information Technology	us market
VSCO	Victoria's Secret	Estados Unidos	Consumer Discretionary	us market
VSTO	Vista Outdoor Inc.	Estados Unidos	Consumer Discretionary	us market
VTOL	Bristow Group Inc.	Estados Unidos	Energy	us market
VTLE	Vital Energy, Inc.	Estados Unidos	Energy	us market
VVI	Viad Corp	Estados Unidos	Industrials	us market
WABC	Westamerica Bancorporation	Estados Unidos	Financials	us market
WAFD	Washington Federal	Estados Unidos	Financials	us market
WD	Walker & Dunlop, Inc.	Estados Unidos	Financials	us market
WDFC	WD-40 Company	Estados Unidos	Consumer Staples	us market
WGO	Winnebago Industries, Inc.	Estados Unidos	Consumer Discretionary	us market
WIRE	Encore Wire Corporation	Estados Unidos	Industrials	us market

Continúa en la siguiente página

Cuadro F.1 – Continúa de la página anterior

Símbolo	Nombre de la Compañía	País	Sector	Mercado
WKC	World Kinect Corporation	Estados Unidos	Energy	us market
WLY	John Wiley & Sons	Estados Unidos	Communication Services	us market
WNC	Wabash National	Estados Unidos	Industrials	us market
WOR	Worthington Enterprises	Estados Unidos	Consumer Discretionary	us market
WRLD	World Acceptance Corporation	Estados Unidos	Financials	us market
WS	Worthington Steel	Estados Unidos	Materials	us market
WSFS	WSFS Financial Corporation	Estados Unidos	Financials	us market
WSR	Whitestone REIT	Estados Unidos	Real Estate	us market
WT	WisdomTree Investments, Inc.	Estados Unidos	Financials	us market
WWW	Wolverine World Wide, Inc.	Estados Unidos	Consumer Discretionary	us market
XHR	Xenia Hotels& Resorts, Inc.	Estados Unidos	Real Estate	us market
XNCR	Xencor Inc	Estados Unidos	Health Care	us market
XPEL	XPEL, Inc.	Estados Unidos	Consumer Discretionary	us market
XPER	Xperi, Inc.	Estados Unidos	Information Technology	us market
XRX	Xerox	Estados Unidos	Information Technology	us market
YELP	Yelp, Inc.	Estados Unidos	Communication Services	us market

Cuadro F.1: Tabla de las acciones que engloba el mercado bursátil S&P600. Fuente externa.

Bibliografía

[1] Söhnke M. Bartram, Jürgen Branke, Giuliano De Rossi, and Mehrshad Motahari. Machine learning for active portfolio management. *The Journal of Financial Data Science*, 3(3):9–30, 2021.

[2] Zvi Bodie, Alex Kane, and Alan J Marcus. *Essentials of Investments: Global Edition*. McGraw Hill, 2013.

[3] G. Brockman, V. Cheung, L. Pettersson, J. Schneider, J. Schulman, J. Tang, and W. Zaremba. Openai gym. *arXiv preprint arXiv:1606.01540*, 2016.

[4] Eugene F Fama and Kenneth R French. The cross-section of expected stock returns. *Journal of Finance*, 47(2):427–465, 1992.

[5] Yuh-Jong Hu and Shang-Jen Lin. Deep reinforcement learning for optimizing finance portfolio management. In *2019 Amity International Conference on Artificial Intelligence (AICAI)*, pages 83–87. IEEE, 2019.

[6] Zechu Li, Xiao-Yang Liu, Jiahao Zheng, Zhaoran Wang, Anwar Walid, and Jian Guo. Finrl-podracer: High performance and scalable deep reinforcement learning for quantitative finance. *arXiv preprint arXiv:2104.05480*, 2021.

[7] Zhipeng Liang, Hao Chen, Junhao Zhu, Kangkang Jiang, and Yanran Li. Adversarial deep reinforcement learning in portfolio management. *arXiv preprint arXiv:1808.09940*, 2018.

[8] Timothy P Lillicrap, Jonathan J Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*, 2015.

[9] Yaohu Lin, Shancun Liu, Haijun Yang, Harris Wu, and Bingbing Jiang. Improving stock trading decisions based on pattern recognition using machine learning technology. *PloS one*, 16(8):e0255558, 2021.

[10] Xiao-Yang Liu, Hongyang Yang, Jiechao Gao, and Christina Dan Wang. Finrl: Deep reinforcement learning framework to automate trading in quantitative finance. *Journal of Open Source Software*, 6(57):2937, 2021.

[11] Yifeng Liu, Cong Ma, Yiding Liu, Guanghui Wang, and Xin Wang. Deep reinforcement learning for automated stock trading: An ensemble strategy. *Applied Soft Computing*, 81:105511, 2019.

[12] Frederic S. Mishkin and Stanley G. Eakins. *Financial Markets and Institutions*. Pearson, 10th edition, 2021.

[13] Anshul Mittal and Arpit Goel. Stock prediction using twitter sentiment analysis. *Stanford University*, 2012.

[14] Saloni Mohan, Sahitya Mullapudi, Sudheer Sammeta, Parag Vijayvergia, and David C. Anastasiu. Stock price prediction using news sentiment analysis. In *2019 IEEE Fifth International Conference on Big Data Computing Service and Applications (BigDataService)*, pages 214–221. IEEE, 2019.

[15] Thien Hai Nguyen, Kiyoaki Shirai, and Julien Velcin. Sentiment analysis on social media for stock movement prediction. *Expert Systems with Applications*, 42(24):9603–9611, 2015.

[16] Overleaf. Online latex editor - overleaf. https://www.overleaf.com/, 2023.

[17] Venkata Sasank Pagolu, Kamal Nayan Reddy Challa, Ganapati Panda, and Babita Majhi. Sentiment analysis of twitter data for predicting stock market movements. *IEEE Transactions on Computational Social Systems*, 3(1):1–12, 2016.

[18] J Patel, S Shah, P Thakkar, and K Kotecha. Predicting stock market index using fusion of machine learning techniques. *Expert Systems with Applications*, 2015.

[19] A. Raffin, A. Hill, M. Ernestus, A. Gleave, A. Kanervisto, and N. Dormann. Stable baselines3. https://github.com/DLR-RM/stable-baselines3, 2019.

[20] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.

[21] Mehak Usmani, Syed Hasan Adil, Kamran Raza, and Syed Saad Azhar Ali. Stock market prediction using machine learning techniques. In *2017 IEEE 13th International Conference on Emerging Technologies (ICET)*, pages 1–6. IEEE, 2017.

[22] Paul Wilmott. *Quantitative Finance*. John Wiley & Sons, 2007.

[23] Hongyang Yang, Xiao-Yang Liu, Shan Zhong, and Anwar Walid. Deep reinforcement learning for automated stock trading: An ensemble strategy. *arXiv preprint arXiv:1912.05372*, 2019.

[24] Hongyang (Bruce) Yang, Xiao-Yang Liu, and Qingwei Wu. A practical machine learning approach for dynamic stock recommendation. *Department of Statistics, Columbia University*, 2018.

[25] Zheng Zhongbin, Fang Jinwu, and Fu Tao. Stock market risk measurement based on qgarch and machine learning algorithm. *2019 International Conference on Economic Management and Model Engineering (ICEMME)*, 2019.

[26] Dave Zimmerman. *Small Stocks, Big Money: Interviews With Microcap Superstars*. John Wiley & Sons, 2012.