

# PAP - PROGRAMA DE MODELACIÓN MATEMÁTICA PARA EL DESARROLLO DE PLANES Y PROYECTOS DE NEGOCIO

*“Estrategias de Rotación Sectorial”*

## **Integrantes:**

Alvarado Garnica Óscar Uriel - **734194**

Enriquez Nares Diego Emilio - **728356**

Martínez Ramírez José Alfonso - **734272**

Mugica Liparoli Juan Antonio - **728370**



*Prof. Sean Nicolás  
González Vázquez*

# ÍNDICE

<b>INTRODUCCIÓN.....</b>	<b>2</b>
<b>FLUJO DE TRABAJO .....</b>	<b>3</b>
OBJETIVO DEL PROYECTO .....	3
DEFINICIÓN DEL PROYECTO .....	3
DIAGRAMA DE FLUJO DEL PROYECTO .....	4
<b>FUNDAMENTACIÓN TEÓRICA DEL PROYECTO.....</b>	<b>5</b>
ROTACIÓN SECTORIAL .....	5
NATURALEZA DE LOS ACTIVOS .....	6
INDICADORES ECONÓMICOS .....	6
FASES DE LA ECONOMÍA .....	7
ASSET ROTATION .....	8
METODOLOGÍA GENERAL .....	9
<b>ANÁLISIS EXPLORATORIO DE LOS DATOS .....</b>	<b>11</b>
RESUMEN ESTADÍSTICO HISTÓRICO .....	11
COMPORTAMIENTO HISTÓRICO .....	12
DISTRIBUCIÓN HISTÓRICA.....	15
CUARTILES HISTÓRICOS .....	16
CORRELACIÓN HISTÓRICA .....	17
<b>MODELO .....</b>	<b>19</b>
DISEÑO Y CONCEPTUALIZACIÓN .....	19
ENTRENAMIENTO Y VALIDACIÓN .....	22
OPTIMIZACIÓN Y GENERALIZACIÓN DE PARÁMETROS E HIPER-PARÁMETROS .....	23
<i>OPTIMIZACIÓN DEL MODELO XGBOOST.....</i>	<i>24</i>
<i>OPTIMIZACIÓN DE LA RED NEURONAL MULTICAPA (MLP).....</i>	<i>25</i>
<i>GENERALIZACIÓN DE LOS MODELOS.....</i>	<i>26</i>
<i>RESULTADOS DE LOS MODELOS.....</i>	<i>26</i>
<i>RESULTADOS DE LA REGRESIÓN LOGÍSTICA (BENCHMARK): .....</i>	<i>27</i>
<i>RESULTADOS DEL MODELO XGBOOST: .....</i>	<i>30</i>
<i>RESULTADOS DE LA RED NEURONAL MULTICAPA (MLP) con ReLU .....</i>	<i>32</i>
<i>CONCLUSIONES, RESULTADOS Y SELECCIÓN DEL MODELO .....</i>	<i>34</i>
IDENTIFICACIÓN DE TENDENCIA ESPERADA .....	35
SELECCIÓN DE ESTRATEGIA DE INVERSIÓN.....	36
IMPLEMENTACIÓN .....	37
<b>BACKTESTING DINÁMICO .....</b>	<b>37</b>
FUNDAMENTACIÓN .....	37
MÉTRICAS DE DESEMPEÑO .....	38
COMPARACIÓN CON EL ‘BENCHMARK’ .....	45
<b>CONCLUSIONES.....</b>	<b>47</b>
<b>ANEXOS.....</b>	<b>49</b>
SOFTWARE .....	49
GLOSARIO .....	49
<b>BIBLIOGRAFIA.....</b>	<b>53</b>

# ESTRATEGIAS DE ROTACIÓN SECTORIAL

## INTRODUCCIÓN

En el dinámico y complejo entorno de los mercados financieros, los gestores de portafolios se enfrentan al desafío de anticipar los movimientos del mercado y ajustar sus estrategias de inversión de acuerdo con las fases del ciclo económico. Este proyecto se centra en desarrollar un modelo matemático basado en el concepto de la rotación sectorial, utilizando indicadores económicos clave, como: CLI, BCI, GDP y CCI, para guiar esta toma de decisiones de inversión. La teoría fundamental es que los diferentes sectores económicos responden de manera distinta a las fluctuaciones económicas, lo que ofrece oportunidades para optimizar la asignación de activos en función de la fase del ciclo económico del momento.

Mediante la recopilación y el análisis de datos históricos, este proyecto busca identificar patrones y relaciones entre los indicadores económicos propuestos y el rendimiento del S&P 500 (índice de mercado benchmark), con el objetivo de construir un modelo matemático que no solo anticipe los movimientos del mercado financiero, sino que también ofrezca estrategias cuantitativas de cuándo tener una postura overweight (sobre-ponderar), underweight (sub-ponderar) o neutral en algún sector económico. La implementación de estas técnicas y del modelo permitirán evaluar su efectividad, validando mediante backtesting dinámico y comparando con un modelo benchmark.

# FLUJO DE TRABAJO

## OBJETIVO DEL PROYECTO

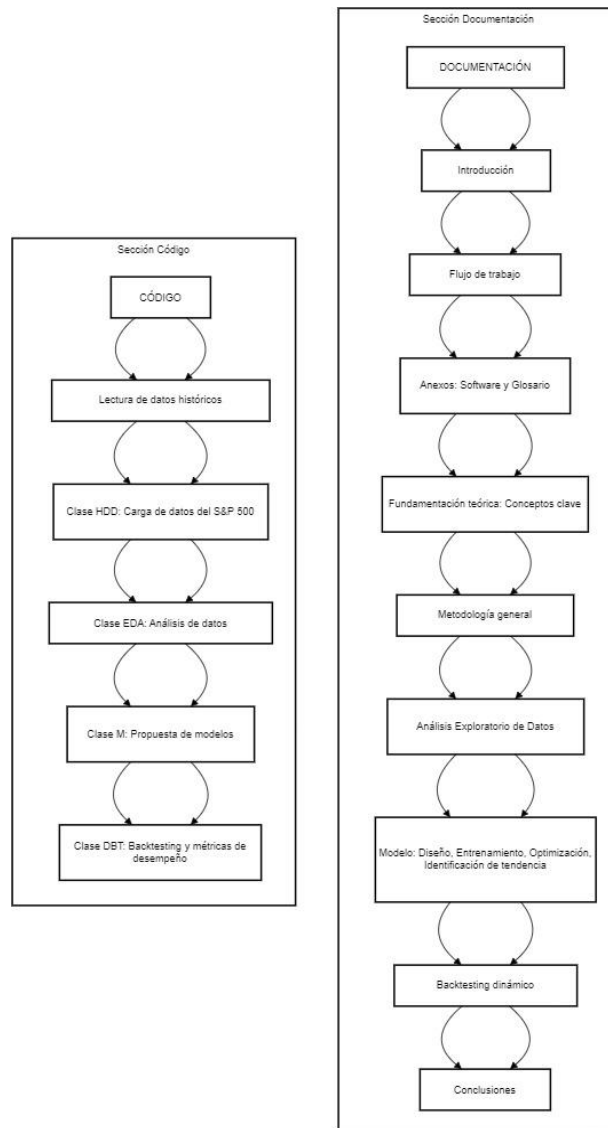
El propósito de este proyecto es desarrollar un modelo matemático que permita anticipar las fases del ciclo económico y, a partir de esta predicción, diseñar una estrategia de inversión cuantitativa que optimice la toma de decisiones en función de las condiciones económicas del momento. Se buscan aplicar técnicas de asset allocation para ajustar la exposición a diferentes sectores económico. Además, se llevará a cabo una comparación del desempeño de la estrategia frente a un modelo benchmark, con el objetivo de validar su efectividad.

## DEFINICIÓN DEL PROYECTO

Se realizará un análisis descriptivo y una comparación gráfica de los indicadores económicos (mencionados anteriormente) contra un índice de mercado benchmark (S&P 500). Luego, se descargarán los datos históricos del índice y de sus respectivos activos financieros y se clasificarán según su naturaleza (beta): pro-cíclico (1) o anti-cíclico (0); después, se propondrán distintos modelos matemáticos con el objetivo de clasificar la postura (overweight, underweight o neutral) de la estrategia de rotación sectorial seleccionada. Posteriormente, se entrenará y validará el mejor modelo hallado, se optimizarán y generalizarán sus respectivos parámetros e hiperparámetros y se evaluará la estrategia de inversión. Finalmente, se implementará dicha estrategia con el modelo seleccionado, se realizará un backtesting dinámico (el cual validará la implementación de dicho modelo) con una selección aleatoria de activos financieros, se identificarán sus respectivas métricas de desempeño.

## DIAGRAMA DE FLUJO DEL PROYECTO

El siguiente diagrama de flujo (**imagen 1**), representa el proceso a seguir tanto en la parte de la documentación (teoría) como en la parte del código (aplicación práctica) que respalda el desarrollo e implementación del proyecto.



*Imagen 8. Diagrama de flujo del proyecto.*

# FUNDAMENTACIÓN TEÓRICA DEL PROYECTO

## ROTACIÓN SECTORIAL

La rotación sectorial es una estrategia de inversión que consiste en ajustar portafolios cambiando la asignación de activos financieros entre sectores económicos según la fase del ciclo económico del momento. Imagina que eres un gestor de portafolios y observas que hay cambios en la economía, ya sea por cambios en las tasas de interés, nuevas políticas monetarias de diferentes bancos centrales o también por ser un gestor de portafolios que revisa de forma recurrente los indicadores económicos para entender qué es lo próximo que puede ocurrir en el mercado financiero en base a distintas métricas económicas. En base a un análisis, prevés que en los próximos meses es probable que la economía entre en recesión o que se experimente una expansión económica. Dependiendo de estas condiciones económicas, debes ajustar tu portafolio para maximizar los rendimientos, minimizar las pérdidas y ajustar el riesgo en base a la aversión al riesgo que tienen tus clientes.

Dicha estrategia implica seleccionar y ponderar sectores económicos específicos en el portafolio según la fase del ciclo económico en la que se encuentra la economía. Los mercados financieros tienen tendencias alcistas y bajistas dependiendo las expectativas económicas que tienen los participantes. En base a esto, es importante tener en cuenta que hay momentos donde conviene estar overweight en activos financieros agresivos (pro-cíclicos) y underweight en activos financieros defensivos (anti-cíclicos). Una métrica interesante para saber cuándo un activo financiero es pro-cíclico o anti-cíclico se conoce, como: 'beta' ( $\beta$ ), el cual es un indicador que nos alerta cuánto se mueve un activo financiero con respecto al mercado financiero (índice), el significado de la beta varía dependiendo su valor:

- $\beta = 1$ : El activo se mueve igual que el mercado financiero.
- $\beta < 1$ : El activo tiene una relación inversa al mercado financiero. Por ejemplo, una beta de 0.40 indica que, si el mercado sube 1, nuestro activo sube menos (0.40), lo que implica menor riesgo sistemático.

- $\beta > 1$ : El activo se mueve en la misma dirección que el mercado financiero, pero con mayor fuerza, lo que conlleva mayor riesgo sistemático.

Conociendo el ciclo económico, podemos seleccionar activos financieros con alguna beta específica. En periodos de recesión económica, es conveniente optar por activos defensivos con una beta inferior a uno. En tiempos de expansión económica, es preferible elegir activos con una beta igual o superior a uno. Es por eso por lo que un gestor de portafolios puede superar al mercado financiero alternando entre posturas overweight, underweight o neutral en activos agresivos y defensivos, según el ciclo económico en el que se encuentre la economía.

## NATURALEZA DE LOS ACTIVOS

- *Pro-cíclica*: Activo o estrategia financiera que tiende a beneficiarse durante periodos de expansión económica, como el sector industrial o tecnológico. Ponderación alta en activos pro-cíclicos. Este tipo de activos son elásticos, es decir, tienen un gran impacto respecto al cambio de precios o condiciones económicas; presentan una beta ( $\beta$ )  $> 0.7$ . Algunos ejemplos de sectores: Tech, Consumer Discretionary, Financial y Real State.
- *Anti-cíclica*: Activo o estrategia financiera que tiende a tener un buen desempeño durante periodos de recesión económica, como el sector de consumo básico. Ponderación alta en activos anti-cíclicos. Este tipo de activos son inelásticos, es decir, no se ven afectados por cambios o alteraciones generales del mercado; presentan un rango de beta  $0 > \beta > 0.70$ . Algunos ejemplos de sectores: Consumer Staples, Healthcare, Utilities y Safe-Haven Commodities.

## INDICADORES ECONÓMICOS

- 1- *Business Confidence Index (BCI)*: Este indicador económico mide la confianza de las empresas en la economía, lo que refleja su disposición a invertir en nuevos proyectos, contratar empleados, y expandir operaciones. Si las empresas son optimistas sobre el futuro, comenzarán a realizar inversiones antes de que se materialice el crecimiento económico, lo que puede hacer que el BCI se adelante al ciclo económico. Un aumento en la confianza empresarial sugiere un

incremento en la actividad económica futura, mientras que una disminución podría indicar una desaceleración.

- 2- *Consumer Confidence Index (CCI)*: Este indicador económico mide la confianza de los consumidores en la economía, lo que afecta directamente sus decisiones de gasto y ahorro. Si los consumidores están confiados en su situación financiera y en el estado de la economía, es probable que aumenten su consumo, impulsando la demanda agregada antes de que se refleje en el crecimiento económico. Por otro lado, si la confianza disminuye, los consumidores pueden recortar sus gastos, lo que podría anticipar una recesión económica.
- 3- *Composite Leading Indicator (CLI)*: Este indicador económico es un índice compuesto que incluye varios indicadores económicos adelantados, como el empleo, los pedidos de manufactura, y los permisos de construcción, entre otros. Estos componentes están seleccionados porque tienden a moverse antes que el ciclo económico general. Al agrupar múltiples indicadores adelantados, el CLI proporciona una señal temprana de cambios en la actividad económica, permitiendo prever fases de expansión o recesión en la economía.
- 4- *Gross Domestic Product (GDP)*: Este indicador económico mide el valor total de los bienes y servicios producidos en una economía durante un período específico. Sin embargo, a diferencia de los otros indicadores, el GDP es un indicador rezagado, ya que refleja el resultado final de la actividad económica en lugar de preverla. Dicho esto, algunos componentes del GDP, como el gasto en inversión y el gasto de los consumidores, pueden contener información sobre tendencias futuras.

## FASES DE LA ECONOMÍA

- 1- *Macro Bottoming*: En esta fase existe un incremento en las ganancias o utilidades de una empresa o del mercado en general también conocido como ‘earnings expansion’. También en el periodo de acumulación el mercado financiero está en una tendencia ascendente, donde los precios de las acciones están subiendo. Esto generalmente ocurre en respuesta a factores positivos como un crecimiento económico sólido, buenas noticias corporativas o políticas económicas favorables, a este término se le conoce como ‘rising equity markets’. En esta temporada las condiciones macroeconómicas son desfavorables, existe un sentimiento positivo



en el mercado financiero y es momento de inversión en activos financieros con beta alta (pro-cíclicos).

- 2- *Macro Improving*: La economía está en su mejor momento, altas ganancias corporativas, los bancos centrales comienzan a subir las tasas de interés, valuaciones elevadas, las acciones alcanzan su techo, también existe un sentimiento eufórico en el mercado financiero donde todos recomiendan comprar. Se recomienda comenzar la transición de activos financieros pro-cíclicos a anti-cíclicos.
- 3- *Macro Peaking*: La economía tiene un crecimiento económico insostenible y hay una contracción de la demanda. Existe una disminución de las ganancias empresariales, las tasas de interés se encuentran en máximos históricos, caen los mercados financieros, los inversionistas tienen un sentimiento pesimista del mercado. Se recomienda tener el portafolio compuesto por activos financieros defensivos, es decir, un portafolio compuesto con activos de beta baja y empresas en sectores inelásticos.
- 4- *Macro Declining*: En esta fase económica comienza un período de recesión, caída de ganancias empresariales, se crean políticas económicas expansivas donde inician recortes en las tasas de interés, el mercado financiero alcanza su punto más bajo, las métricas fundamentales de las empresas comienzan a mejorar, el sentimiento en el mercado es mixto y se recomienda que el portafolio pase por una transición hacia activos más agresivos.

## ASSET ROTATION

La rotación de activos, conocida como 'asset rotation', es una estrategia de inversión que busca maximizar los rendimientos mediante la diversificación sectorial económica. Esta estrategia se basa en la naturaleza pro-cíclica o anti-cíclica de los activos financieros, aprovechando las distintas fases del ciclo económico. En este proyecto, se propone una estrategia de rotación sectorial en la que se seleccionarán activos financieros de manera aleatoria y se ajustarán sus ponderaciones en el portafolio de inversión, de acuerdo con las características de cada activo y la postura determinada.

Posturas:

- a) *Overweight*: Ponderación alta en activos pro-cíclicos.

- b) *Neutral*: Ponderación combinada y/o equilibrada en activos pro-cíclicos y anti-cíclicos.
- c) *Underweight*: Ponderación alta en activos anti-cíclicos.

Para identificar si un activo es pro-cíclico o anti-cíclico, se utilizará la beta como medida de referencia.

$$\beta = \frac{Cov(R_i, R_m)}{Var(R_m)}$$

Donde:

- $R_i$ : rendimiento del activo.
- $R_m$ : rendimiento del mercado.
- $Cov(R_i, R_m)$ : covarianza entre los rendimientos del activo y los rendimientos del mercado.
- $Var(R_m)$ : varianza de los rendimientos del mercado.

## METODOLOGÍA GENERAL

Implementar un modelo matemático para predecir el ciclo económico, teniendo como base los indicadores económicos previamente mencionados. Aunque tienen distintos enfoques, metodologías y componentes, son indicadores seleccionados con el objetivo de anticiparse al mercado financiero con distintos factores económicos, como los datos de construcción de viviendas, expectativas empresariales, confianza del consumidor, expectativas de producción, entre otros. Debido a esto, pueden llegar a ir adelantados al índice seleccionado. Los indicadores representaran X en el modelo, mientras que se construirá Y, identificando con los siguientes valores: 1, 0 y -1, a cada registro, en función de los rendimientos históricos de dicho índice.

- a) *Overweight*: 1.
- b) *Neutral*: 0.
- c) *Underweight*: -1.

Posteriormente, se entrenará y testeará el modelo matemático inicialmente utilizando una regresión logística (modelo benchmark) para establecer una modelo de base que nos permita evaluar el rendimiento de este en la predicción del ciclo económico y en la clasificación de activos financieros pro-cíclicos o anti-cíclicos. Una vez obtenido este modelo benchmark, se comparará

con un par de modelos, como: redes neuronales multicapa y XGBoost. Evaluando su desempeño y optimizando sus parámetros e hiper-parámetros solo si ofrecen una mejora significativa en las predicciones en relación con la regresión logística. La elección final del modelo se realizará en función de su capacidad para predecir con mayor precisión las fases del ciclo económico, así como su desempeño en el backtesting dinámico. Una vez identificada la postura económica esperada se seleccionará una estrategia de inversión, la cual será implementada siguiendo el concepto de rotación sectorial (una vez teniendo el modelo previamente entrenado y validado).

En este punto, se obtendrán los precios de los activos financieros que conforman el índice seleccionado, clasificándolos en pro-cíclicos o anti-cíclicos (según su beta). Se seleccionarán aquellos que mejor se ajusten a la fase del ciclo económico. Una vez seleccionados, la optimización se llevará a cabo utilizando la estrategia de asset allocation propuesta. Posteriormente, se evaluarán distintas métricas de desempeño para compararlos con el índice.

# ANÁLISIS EXPLORATORIO DE LOS DATOS

Mediante un análisis exploratorio de datos (EDA), se identificarán patrones, anomalías y relaciones en los indicadores económicos seleccionados y en el índice.

## RESUMEN ESTADÍSTICO HISTÓRICO

```
Summary statistics:
^GSPC CLOSE      CLI      BCI      GDP      CCI
count    289.000000  289.000000  289.000000  289.000000  289.000000
mean    1987.169547  99.779276  99.937007  99.900981  99.696149
std     1078.264977  1.271074  1.151665  1.206805  1.526347
min      735.090027  93.483700  95.797870  92.026080  96.201320
25%     1191.329956  99.188310  99.288160  99.556070  98.560120
50%     1454.599976  99.910780  99.956170  99.936840  99.904290
75%     2640.870117 100.657200 100.735500 100.521100 100.915900
max     4845.649902 101.954200 102.119800 101.872500 102.848900

Missing values:
^GSPC CLOSE      0
CLI              0
BCI              0
GDP              0
CCI              0
dtype: int64

Outliers
(array([106, 107, 108, 108, 109, 109, 110, 110, 111, 112, 242, 243, 243,
        244, 244, 245, 246]), array([2, 2, 1, 2, 1, 2, 1, 2, 1, 1, 3, 1, 3, 1, 3, 3, 3]))
```

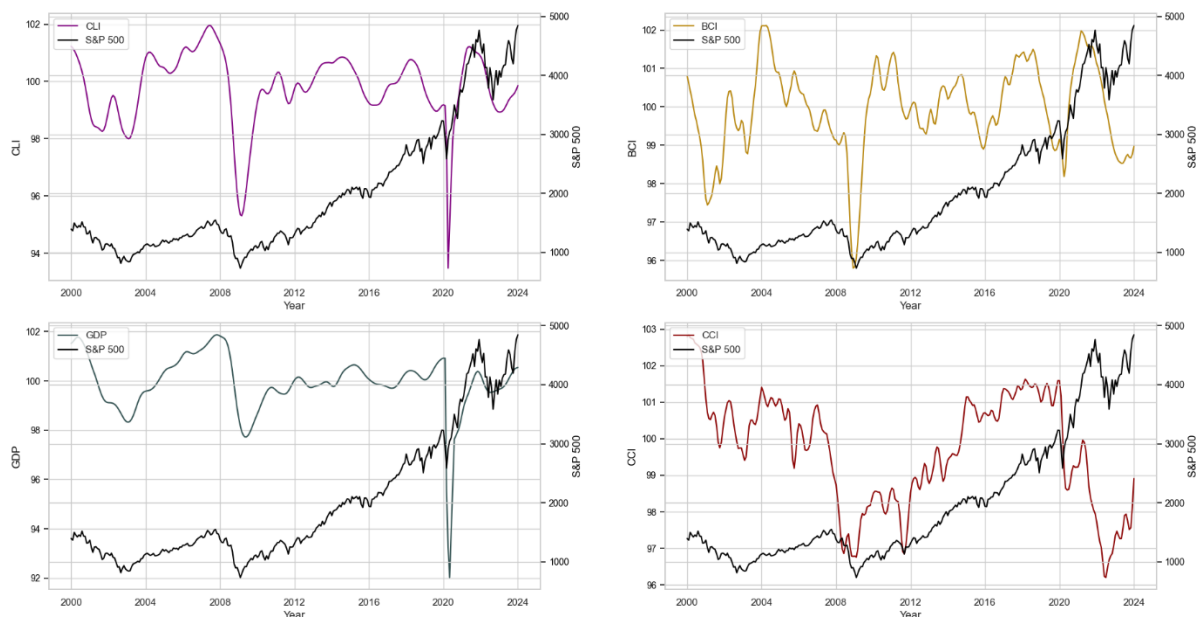
*Imagen 2. Comparativa del resumen estadístico, valores faltantes y outliers: indicadores económicos VS índice.*

En primera instancia, se calculó un conjunto de estadísticas descriptivas (**imagen 2**), lo que permitió observar que los indicadores económicos mantienen una estabilidad relativa con medias cercanas a 100, lo cual está alineado con la normalización base 100 aplicada previamente. Por otro lado, el índice mostró una mayor volatilidad, con una desviación estándar considerablemente alta, indicando fluctuaciones significativas en el mercado durante el período analizado. Adicionalmente, se identificaron outliers en los indicadores económicos, específicamente en BCI, GDP y CCI, lo que sugiere la presencia de eventos atípicos o inusuales que podrían afectar la correlación entre estos indicadores y el comportamiento del mercado.

Además, la mediana de los indicadores es mayor que la media, lo que sugiere una asimetría positiva o una distribución sesgada a la izquierda. Esto implica que, aunque la mayoría de los valores tienden a ser altos, hubo algunos eventos negativos significativos (outliers) que redujeron las medias de los indicadores. Estos outliers negativos son cruciales para entender los períodos de recesión económica, que pueden haber afectado los niveles de confianza tanto empresarial como del consumidor.

Por otro lado, se identificaron outliers en algunos de los indicadores económicos, particularmente en el BCI, GDP y CCI. Estos outliers indican la presencia de eventos económicos inusuales que, en algunos casos, podrían estar relacionados con crisis económicas o períodos de recesión. Estos eventos pueden haber afectado significativamente la relación entre los indicadores económicos y el comportamiento del mercado financiero, alterando las expectativas de los inversores y generando una mayor volatilidad.

## COMPORTAMIENTO HISTÓRICO

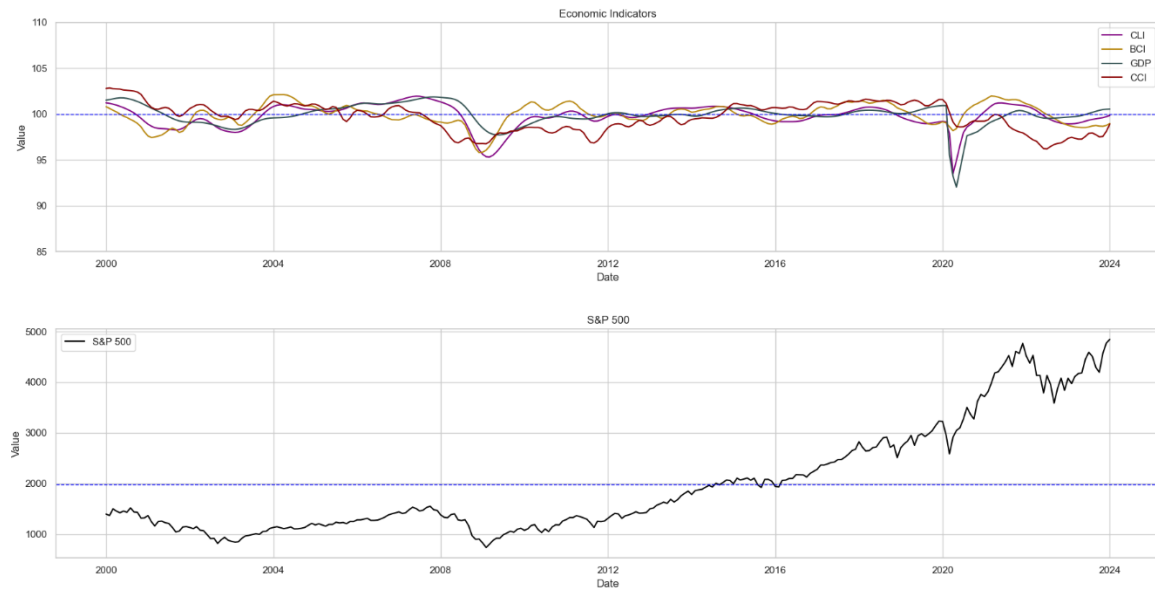


*Imagen 3. Comparativa individual del comportamiento histórico: indicadores económicos VS índice.*

El análisis gráfico (**imagen 3**) de la evolución del índice en comparación con los indicadores económicos seleccionados revela dinámicas interesantes que refuerzan la complejidad de la relación entre el mercado financiero y el ciclo económico. En particular, se observa que el índice muestra un patrón de crecimiento sostenido, especialmente a partir de la década de 2010, mientras que los indicadores económicos como el CLI y el BCI presentan fluctuaciones más volátiles. Este comportamiento sugiere que, aunque los indicadores económicos capturan las condiciones macroeconómicas en diferentes fases del ciclo, el mercado financiero parece estar más influenciado por las expectativas de futuro y no solo por la situación económica actual. Esto puede explicar por qué el mercado financiero anticipa movimientos importantes antes de que los indicadores reflejen cambios significativos.

En el caso del GDP, un indicador de naturaleza más rezagada, se observa una correlación menos directa con las fluctuaciones del índice. Esta falta de sincronización puede atribuirse a que el GDP refleja la actividad económica pasada, mientras que el mercado financiero generalmente reacciona a expectativas de la población. Este desfase temporal es clave para entender cómo los inversores anticipan los movimientos del mercado, basándose en otros indicadores económicos más adelantados, como el CLI y el BCI, para tomar decisiones informadas.

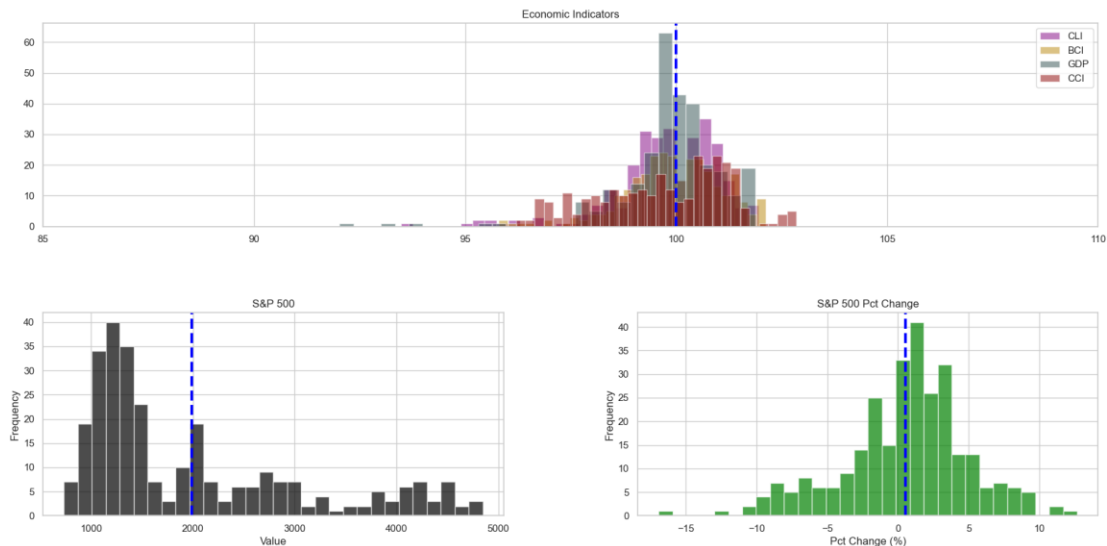
Por otro lado, la relación entre el CCI y el índice, aunque es más consistente, presenta momentos de divergencia. Esto indica que la confianza del consumidor puede ser más volátil y estar sujeta a cambios abruptos en el entorno macroeconómico, como eventos políticos o globales, que afectan la percepción del mercado financiero. La volatilidad observada en el CCI puede también reflejar una mayor sensibilidad de los consumidores a factores externos, lo que genera oscilaciones en sus expectativas de gasto.



*Imagen 4. Comparativa colectiva del comportamiento histórico: indicadores económicos VS índice.*

Estos gráficos (**imagen 3 e imagen 4**) ponen como tema la necesidad de un enfoque más sofisticado para captar las relaciones temporales y la causalidad entre los indicadores económicos y el rendimiento del mercado financiero. La presencia de estas divergencias sugiere que el modelo deberá considerar la posibilidad de diferentes ventanas temporales y técnicas de suavizado que permitan capturar mejor la dinámica entre los indicadores adelantados y el índice.

## DISTRIBUCIÓN HISTÓRICA



*Imagen 5. Comparativa colectiva de las distribuciones históricas: indicadores económicos VS índice.*

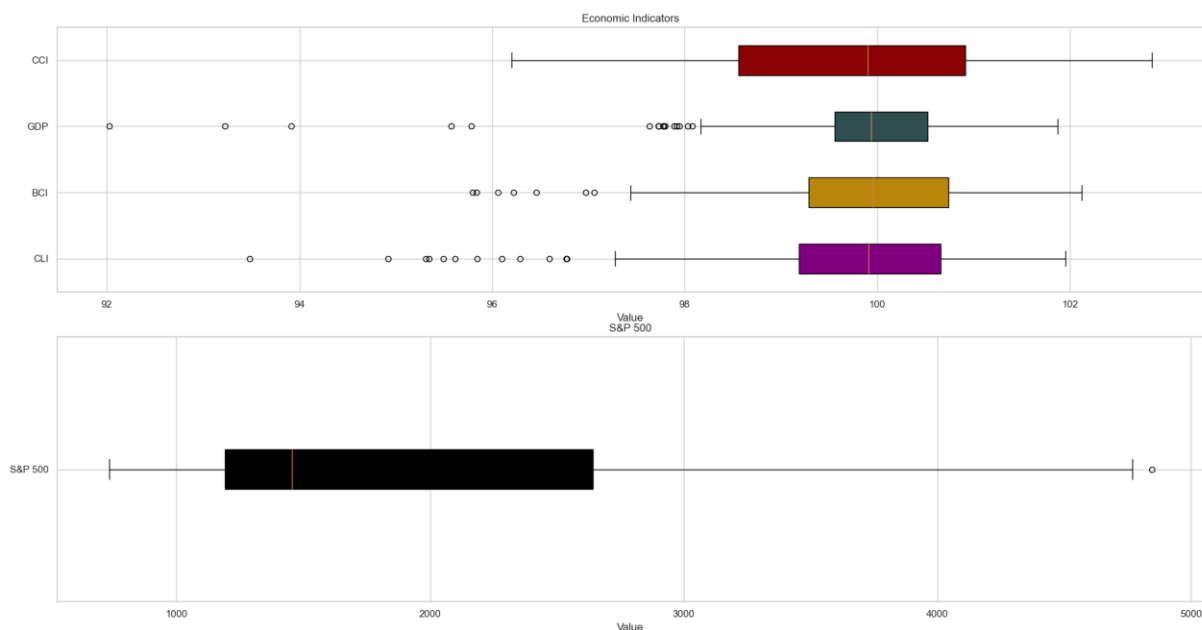
La distribución de los indicadores económicos (**imagen 5**) en comparación con el índice revela una notable diferencia en la variabilidad de los datos. Todos los indicadores económico se agrupan estrechamente alrededor del valor de 100, lo que indica una estabilidad con fluctuaciones limitadas a lo largo del tiempo. Esta concentración en torno a 100 sugiere que estos tienden a mantenerse relativamente constantes, sin grandes variaciones extremas, lo que contrasta fuertemente con la mayor volatilidad del índice.

Por otro lado, el índice muestra una dispersión más amplia, lo cual es evidente en su distribución más extendida hacia valores más bajos. Este comportamiento refleja su alta volatilidad, que es característica de los mercados financieros, donde las fluctuaciones son más pronunciadas. La gráfica que muestra los cambios porcentuales en el índice refuerza esta idea, evidenciando desviaciones significativas tanto positivas como negativas, lo que subraya la naturaleza impredecible y a menudo brusca del mercado financiero en comparación con los indicadores económicos que son más estables.



Esta divergencia entre la estabilidad de los indicadores económicos y la volatilidad del índice destaca cómo el mercado financiero reacciona no solo a las condiciones económicas actuales, sino también a las expectativas futuras y a eventos imprevistos, lo que provoca fluctuaciones más pronunciadas. En conjunto, la comparación refuerza la idea de que los mercados financieros tienden a ser más reactivos, mientras que los indicadores económicos reflejan una imagen más estable y moderada de la economía.

## CUARTILES HISTÓRICOS

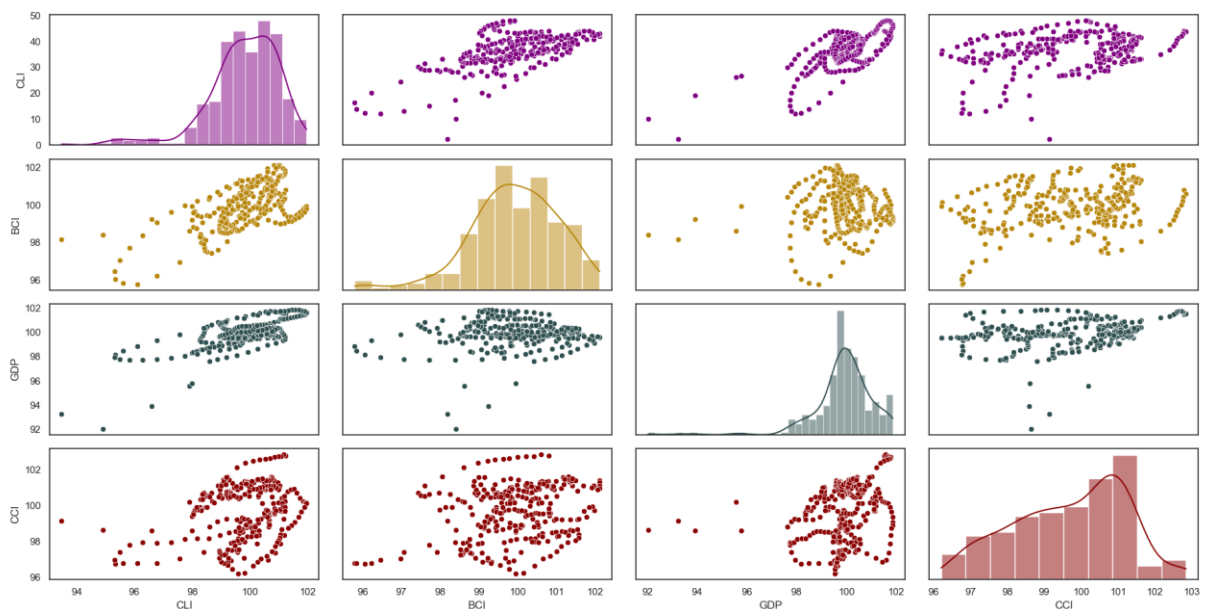


*Imagen 6. Comparativa colectiva de los cuartiles históricos: indicadores económicos VS índice.*

El análisis mediante los diagramas de caja (**imagen 6**) pone de relieve las diferencias en la dispersión y la presencia de outliers entre los indicadores económicos y el índice. En los indicadores económicos CCI, BCI y GDP, se observa una mayor concentración de los datos alrededor de sus medianas, lo que refleja una estabilidad relativa. Sin embargo, se identifican outliers, especialmente en el GDP y el BCI, lo cual sugiere episodios de variabilidad anómala que pueden estar asociados a eventos económicos puntuales que afectaron de manera significativa estos, lo cual se ha venido comentando a lo largo de esta exploración.

Por otro lado, el índice muestra una mayor extensión en su rango intercuartílico, lo que evidencia una mayor volatilidad en comparación con los indicadores económicos. El amplio rango en el índice indica que el mercado financiero experimenta fluctuaciones mucho más extremas. Además, la presencia de un outlier significativo en el índice resalta la naturaleza impredecible del mercado financiero, subrayando que este puede estar influenciado por eventos extremos que afectan su comportamiento de manera considerable, contrastando con la estabilidad relativa de los indicadores económicos.

## CORRELACIÓN HISTÓRICA



*Imagen 7. Análisis colectivo de la dispersión histórica: indicadores económicos.*

En el gráfico de pares (**imagen 7**), se puede observar que los indicadores económicos muestran una correlación positiva en la mayoría de los casos, lo que sugiere que en determinados contextos económicos tienden a reaccionar de manera similar. Sin embargo, no se aprecia una relación lineal clara, ya que la dispersión de los puntos y la estructura de los datos indican que las relaciones entre estos son más complejas de lo que una correlación lineal simple podría describir. También se identifican outliers en algunas de las relaciones, lo que podría deberse a eventos

económicos atípicos que alteraron las dinámicas entre los indicadores. Además, se puede notar la formación de clústeres en algunos gráficos, lo que podría indicar comportamientos diferenciados en ciertos períodos o bajo condiciones económicas específicas.



*Imagen 8. Análisis colectivo de la correlación histórica: indicadores económicos.*

La matriz de correlación (**imagen 8**) entre los indicadores económicos revela una relación moderada a fuerte entre la mayoría de ellos. En particular, destaca la alta correlación entre el CLI y el GDP (72%), lo que sugiere que estos indicadores suelen moverse en la misma dirección, aunque con distinta magnitud. Este nivel de correlación indica que, si bien estos pueden estar relacionados, la magnitud y el momento en que afectan la actividad económica pueden variar. No obstante, aún no se puede concluir si estas correlaciones serán útiles para anticipar cambios en el ciclo económico, ya que esto dependerá del comportamiento del modelo predictivo en etapas posteriores del análisis.

Por otro lado, la relación más débil se observa entre el BCI y los otros indicadores, especialmente con el GDP (0.084), lo que indica que la confianza empresarial puede estar influenciada por factores independientes que no están necesariamente relacionados de manera directa con el

crecimiento económico general. Esta falta de correlación sugiere que el BCI podría ser más volátil o responder a expectativas de corto plazo, en lugar de cambios estructurales en la economía.

En conjunto, estos patrones de correlación subrayan la importancia de considerar múltiples indicadores económicos para obtener una visión más completa y sus posibles impactos en el mercado financiero, ya que cada uno puede capturar diferentes aspectos del ciclo económico.

En resumen, el análisis exploratorio de datos ha revelado que mientras que los indicadores económicos muestran una estabilidad relativa y una agrupación cercana a sus valores normalizados, el índice refleja una mayor volatilidad y una sensibilidad a eventos extremos. Las divergencias observadas en los gráficos y la identificación de outliers subrayan la importancia de capturar las dinámicas temporales entre los indicadores económicos adelantados y los rezagados, lo que será crucial para desarrollar un modelo matemático predictivo complejo. Adicionalmente, las correlaciones entre indicadores económicos como el CLI y el GDP refuerzan su valor como señales complementarias en la anticipación del ciclo económico, mientras que otros, como el BCI, pueden estar influenciados por factores de corto plazo. Estos hallazgos proporcionan una base sólida para ajustar los parámetros del modelo matemático y asegurar que capture tanto la estabilidad económica subyacente como la volatilidad y las fluctuaciones del mercado financiero.

## MODELO

### DISEÑO Y CONCEPTUALIZACIÓN

En esta fase del proyecto, se procede a la descarga de datos históricos del índice (S&P 500) y de los activos financieros que lo componen. Se obtienen los precios de cierre ajustados de cada uno de estos activos, descargando datos mensuales en el rango de fechas especificado (desde enero del 2020 hasta junio del 2024). Posteriormente, se consolida el precio de cierre ajustado de cada activo en particular. Esta recopilación de datos no solo proporciona una base sólida para el análisis y modelado matemático posterior, sino que también permite comparar el rendimiento de los activos individuales con el índice, facilitando la evaluación de la estrategia de rotación sectorial futura.

Para este punto, los datos de nuestros indicadores económicos funcionarían como las X (input) del modelo matemático. En el caso del output, es necesario la generación de nuestras Y, para contar con los datos completos y así entrenar el modelo. Estas Y, serán generadas a partir del comportamiento del precio ajustado de cierre del S&P500. Se calcula el rendimiento diario de este índice utilizando el método de cambio porcentual y se define en función de este rendimiento. Cabe recalcar que el 2% (umbral), es una variable posible para optimizar más adelante en el proceso de este modelo:

- $r_{S\&P\ 500} > 2\%$ : Overweight.
- $2\% > r_{S\&P\ 500} > -2\%$ : Neutral.
- $r_{S\&P\ 500} < -2\%$ : Underweight.

Una vez definido los rangos para el rendimiento del S&P 500, se pueden generar los valores de las Y, los cuales serán:

- a) *Overweight*: 1.
- b) *Neutral*: 0.
- c) *Underweight*: -1.

Este enfoque permite categorizar el comportamiento del S&P 500 en tres posturas distintas, facilitando así la predicción del ciclo económico. Es importante mencionar que se ajustan los datos para que la Y refleje el rendimiento en el periodo que ocurrió, utilizando 1 rezago (shift), desplazando los valores hacia arriba, asegurando que el modelo matemático pueda anticipar cambios futuros en el mercado financiero. En esta parte del proyecto, así como se puede optimizar el umbral para el rendimiento del S&P 500, también se puede aplicar esta técnica de optimización para el número de rezagos por utilizar.

Una vez definidos los indicadores económicos adelantados al ciclo económico, así como la realización del análisis EDA de estos, la descarga de nuestros datos y la generación de nuestro output, completando así nuestra base de datos, se puede pasar a la propuesta de un modelo matemático que ayude a determinar a un gestor de portafolios, cuando es conveniente estar overweight, underweight o neutral, en activos financieros pro-cíclicos o anti-cíclicos. Los modelos propuestos son:

1. *Regresión Logística*: Este modelo matemático será empleado como modelo benchmark a superar por los demás modelos propuestos cuyo objetivo principal es analizar la relación entre los indicadores económicos adelantados y las decisiones de inversión en activos financieros pro-cíclicos o anti-cíclicos según la fase del ciclo económico. Este permite estimar la probabilidad de que un activo financiero se clasifique como overweight, underweight o neutral en función de las variables predictivas.
2. *XGBoost*: Algoritmo de aprendizaje automático basado en árboles de decisión que ha demostrado ser altamente eficaz en tareas de predicción y clasificación. XGBoost ofrece una serie de ventajas, como la capacidad de manejar datos faltantes y su eficiencia en términos de tiempo de entrenamiento y rendimiento. En el desarrollo de la estrategia de rotación sectorial, este modelo permitirá optimizar la selección de activos financieros pro-cíclicos o anti-cíclicos mediante la integración de múltiples indicadores económicos, proporcionando así una base sólida para tomar decisiones informadas y mejorar el rendimiento ajustado al riesgo frente al benchmark.
3. *Red Neuronal*: Esta herramienta es particularmente útil para modelar matemáticamente relaciones complejas y no lineales entre los indicadores económicos y las decisiones de inversión. Su capacidad para identificar patrones en grandes volúmenes de datos la convierte en una opción ideal para predecir la dinámica del ciclo económico a partir de múltiples variables. En el contexto de la rotación sectorial, las redes neuronales son capaces de capturar interacciones que otros modelos podrían pasar por alto, permitiendo una mayor precisión en las predicciones. Su flexibilidad en la arquitectura y la capacidad de ajustar hiper-parámetros contribuyen a la optimización del rendimiento del modelo, lo que puede mejorar la efectividad de la estrategia de inversión.

En esta etapa del proyecto, los activos financieros se clasifican como pro-cíclicos o anti-cíclicos en función de su beta ( $\beta$ ), que mide la sensibilidad de un activo frente a los movimientos del mercado financiero. Esta clasificación es fundamental para guiar las decisiones de inversión en la estrategia de rotación sectorial, permitiendo ajustar la exposición a los sectores económicos según las condiciones económicas previstas.

- $\beta > 0.7$ : Pro-cíclico.
- $\beta \leq 0.7$ : Anti-cíclico.

Donde, un activo financiero pro-cíclico, tiende a amplificar los movimientos del mercado financiero, en cambio, un activo financiero anti-cíclico, se mueve en la misma dirección que el mercado financiero, pero con menor intensidad. Con esta clasificación, más adelante, se podrá establecer una combinación óptima en función de la tendencia económica identificada.

## ENTRENAMIENTO Y VALIDACIÓN

Se pondrá en marcha el modelo matemático para predecir el valor de la variable objetivo (Y) con respecto al comportamiento de las variables predictoras (X), las cuales son los indicadores económicos definidos anteriormente (CLI, BCI, GDP, CCI); así como también, el comportamiento pasado de los rendimientos del S&P 500 (índice). La Y en base al modelo nos regresará una predicción donde puede almacenar alguno de los siguientes valores.

- a) *Overweight*: 1.
- b) *Neutral*: 0.
- c) *Underweight*: -1.

Para entrenar el modelo matemático, se dividirá el conjunto de datos en dos partes: 80% para el conjunto de entrenamiento y 20% para el conjunto de prueba. El entrenamiento se llevó a cabo utilizando distintos modelos de clasificación, nuestro modelo benchmark es una regresión logística, que por más que sea un modelo clásico puede ajustarse de buena manera y no requiere de mucho poder computacional para arrojar un resultado interesante. Los otros dos modelos seleccionados para obtener posibles mejores resultados son: XGBoost y redes neuronales multicapa con una función de activación ReLU.

Posteriormente, se entrenaron estos otros dos modelos matemáticos con los datos de entrada (X), y la salida (Y). Cada modelo fue ajustado con sus respectivos parámetros e hiper-parámetros, y, luego, se realizó una optimización mediante búsquedas aleatorias y el uso de la librería “optuna” para encontrar los respectivos hiper-parámetros óptimos, mejorando así el accuracy de dichos modelos. Después, para validar el rendimiento de los modelos, se utilizó el conjunto de prueba previamente reservado. En esta fase, se calcularon métricas clave, como: precision (proporción de predicciones correctas entre todas las predicciones positivas realizadas por el modelo), recall (proporción de verdaderos positivos detectados por el modelo respecto al total de casos realmente

positivos en los datos) y F1-Score; con el objetivo de evaluar la capacidad de cada modelo para clasificar correctamente los valores de salida.

A su vez, para garantizar la capacidad del modelo matemático y que pueda generalizarse (sin llegar a un sobreajuste), es decir, que pueda arrojar un buen accuracy para futuros datos aunque sean nuevos, se realizó la validación cruzada o mejor conocida como “cross-validation” con k-folds, donde el conjunto de datos se divide en 5 partes, cada modelo se entrena y valida en diferentes combinaciones de estos subconjuntos, mitigando así la variabilidad en los resultados y proporcionando una evaluación con mayor capacidad sobre el accuracy.

## OPTIMIZACIÓN Y GENERALIZACIÓN DE PARÁMETROS E HIPER-PARÁMETROS

Se aborda el proceso de optimización de los parámetros e hiper-parámetros utilizados en los distintos modelos matemáticos. La optimización es un paso crucial para mejorar el rendimiento predictivo de cada modelo, ya que, los parámetros e hiper-parámetros determinan el comportamiento de los modelos de Machine Learning, como: la complejidad, la regularización y la tasa de aprendizaje.

Por su parte, indagando un poco sobre el concepto de la regularización, se refiere a la evaluación mediante métricas, como: el F1-Score y el F2-Score; los cuales son conceptos muy importantes para asegurar la precisión y la sensibilidad de la clasificación. El F1-Score (combinación de la precisión y el recall), se usa para medir la eficiencia del modelo al identificar correctamente las fases del ciclo económico, garantizando un equilibrio entre la detección de verdaderos positivos y la minimización de falsos positivos. Por otro lado, el F2-Score (pondera más el recall que la precisión), resulta ser una métrica valiosa en situaciones donde los costos de omitir casos positivos superan los de reconocer falsos positivos.



## OPTIMIZACIÓN DEL MODELO XGBOOST

El modelo XGBoost se centró en la optimización de hiper-parámetros para mejorar de manera significativa el accuracy en comparación con el modelo benchmark. Se utilizaron técnicas avanzadas como la búsqueda aleatoria y una herramienta de optimización (librería optuna), la cual permitió una exploración eficiente y extensa de los hiper-parámetros, los cuales fueron ajustados debido a que influyen directamente en la capacidad del modelo para generalizar de mejor manera y aprender de forma efectiva. Algunos de los hiper-parámetros, son:

- *n\_estimators*: Determina el número de árboles utilizados, afectando directamente la complejidad del modelo.
- *learning\_rate*: Ajusta la contribución de cada árbol al resultado final, esencial para evitar el sobreajuste.
- *max\_depth*: Limita la profundidad de cada árbol, proporcionando un control sobre la especialización del modelo.
- *subsample* y *colsample\_bytree*: Estos parámetros controlan la fracción de muestras y características utilizadas por cada árbol, promoviendo la diversidad entre los árboles y evitando el sobreajuste.
- *gamma*, *reg\_alpha*, y *reg\_lambda*: Estos controlan la regularización del modelo, penalizando la complejidad para mejorar la generalización en datos no vistos.
- *min\_child\_weight* y *max\_delta\_step*: Ajustan el umbral para la creación de nodos adicionales y la estabilidad en la actualización de los pesos, respectivamente, lo que puede ayudar a prevenir cambios bruscos en las predicciones del modelo, favoreciendo un aprendizaje más suave y estable.

Se ejecutaron 100 iteraciones de búsqueda con la librería optuna para ajustar los hiper-parámetros anteriormente enlistados, en donde cada iteración se diseñó para maximizar el accuracy del modelo matemático, probando diferentes combinaciones y adaptando dinámicamente los hiper-parámetros basados en el rendimiento observado. El resultado fue un modelo XGBoost optimizado que demostró una precisión superior al modelo benchmark.

## OPTIMIZACIÓN DE LA RED NEURONAL MULTICAPA (MLP)

Al elaborar la Red Neuronal Multicapa (MLP), se optimizó tanto la arquitectura del modelo como sus hiper-parámetros mediante el uso de la librería *optuna*. Dichos hiper-parámetros ajustados incluyen el número de unidades en cada una de las capas ocultas (*hidden\_layer\_sizes*), el parámetro de regularización L2 (*alpha*), y la tasa de aprendizaje inicial (*learning\_rate\_init*). A través de esta estructura de optimización, se definió la estructura óptima de la red, experimentando con diversas configuraciones de capas y unidades para cada capa. Este proceso implicó la evaluación de múltiples combinaciones, desde unidades mínimas hasta configuraciones más complejas en tres capas ocultas, con el objetivo de maximizar la precisión en la clasificación.

También, se utilizó la validación cruzada y una serie de 50 pruebas para ajustar dichos hiper-parámetros, maximizando el *accuracy* del modelo matemático en el conjunto de validación. La configuración final de la red demostró una mejora considerable comparada con los otros dos modelos anteriores. Además, se incorporó un análisis por medio de la gráfica AUC-ROC (también implementado para los otros dos modelos) para cada clase, lo que nos permitió evaluar la capacidad de este para manejar clasificaciones multiclase de manera correcta.

Además, acerca de la optimización de nuestro modelo de Red Neuronal Multicapa (MLP), una consideración fundamental fue la definición de la arquitectura de la red, específicamente el número de capas ocultas y la cantidad de neuronas por capa. Esta estructura óptima encontrada incluye 3 capas ocultas, en donde, a lo largo del proceso de optimización por medio de la librería *optuna*, se experimentaron diferentes configuraciones, variando así el número de neuronas en cada capa (de 16 a 128). Esta diferencia de cantidad de neuronas permitió explorar de manera más compleja a la red, en donde influyó su capacidad de aprender y generalizar de mejor manera. Es por eso por lo que, la configuración final quedó determinada finalmente por la maximización del *accuracy* del modelo matemático, consistente en una primera capa con 64 neuronas, seguida por una segunda capa con 32 neuronas, y una tercera y última capa oculta con 32 neuronas.

## GENERALIZACIÓN DE LOS MODELOS

Uno de los mayores retos en la construcción de modelos de Machine Learning es evitar el sobreajuste (overfitting). Para garantizar la capacidad de generalización del modelo matemático, se aplicaron técnicas de validación cruzada y estratificación en los conjuntos de entrenamiento y prueba. De esta manera, se asegura que cada modelo no solo se ajuste a los datos de entrenamiento, sino que también sea capaz de predecir correctamente datos no vistos.

Además, se utilizó regularización (L1-Score y L2-Score en la regresión logística, y regularización en XGBoost y MLP) para controlar la complejidad de cada modelo y prevenir el ajuste excesivo a los datos de entrenamiento. También se ajustaron las tasas de aprendizaje y otros hiper-parámetros para equilibrar la precisión de cada modelo y su capacidad de generalización.

## RESULTADOS DE LOS MODELOS

En esta sección, vamos a detallar el desempeño de los distintos modelos matemáticos, evaluados principalmente por sus curvas AUC-ROC que muestran la capacidad de cada uno de estos para distinguir entre las tres clases de predicción: overweight, underweight y neutral. El gráfico AUC-ROC es una herramienta visual y analítica en la evaluación de modelos de clasificación, donde: AUC se refiere al Área Bajo la Curva, y ROC se refiere a la Curva de Característica Operativa del Receptor. Este gráfico representa la capacidad de un modelo para discriminar entre clases positivas y negativas a diversos umbrales de clasificación. La curva ROC traza la tasa de verdaderos positivos (sensibilidad) en el eje Y contra la tasa de falsos positivos (1 - especificidad) en el eje X para diferentes puntos de corte. Un AUC cercana a 1 indica un modelo altamente eficaz que distingue perfectamente entre las clases, mientras que un AUC de 0.5 sugiere un rendimiento no mejor que el azar.

RESULTADOS DE LA REGRESIÓN LOGÍSTICA (BENCHMARK):

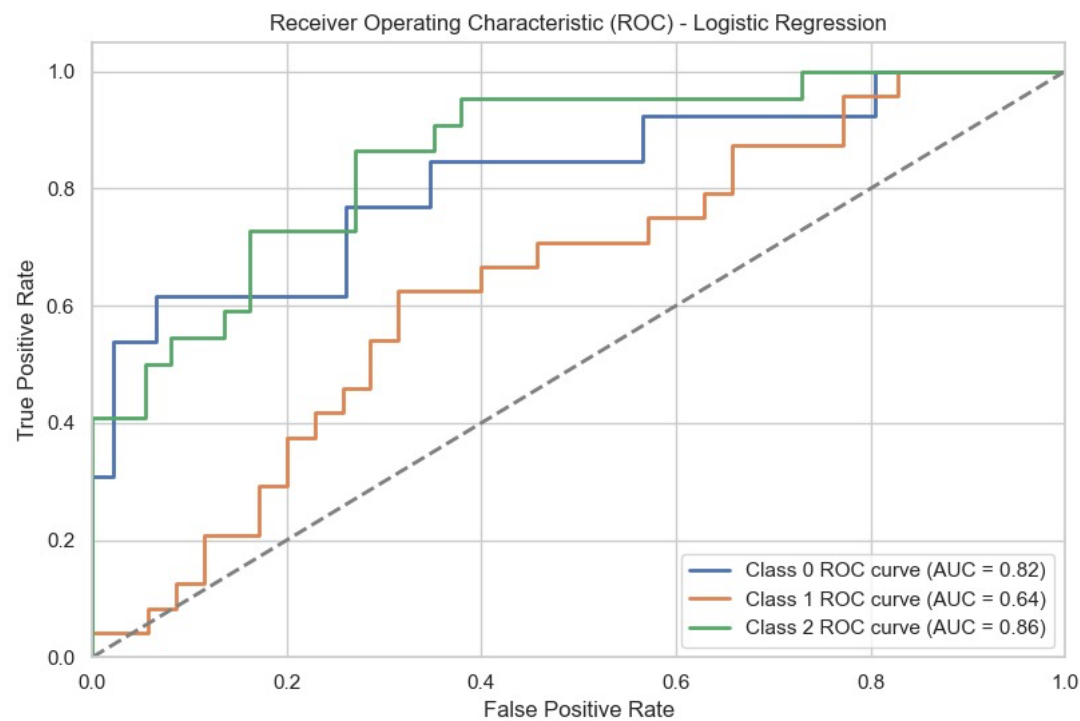


Imagen 9. Gráfica AUC-ROC para el modelo de Regresión Logística.

Accuracy: 0.1695

-----

Classification Report:

	precision	recall	f1-score	support
0	1.00	0.23	0.38	13
1	0.50	0.21	0.29	24
2	1.00	0.09	0.17	22
micro avg	0.67	0.17	0.27	59
macro avg	0.83	0.18	0.28	59
weighted avg	0.80	0.17	0.26	59
samples avg	0.17	0.17	0.17	59

Imagen 10. Reporte de clasificación para el modelo de Regresión Logística.

- *Gráfica AUC-ROC*: Como se puede observar en la **imagen 9**, la curva ROC muestra un AUC 82%, 64% y 86% para las clases 0, 1 y 2 respectivamente. El modelo de regresión logística predice y clasifica de forma adecuada para la clase underweight ( $y = -1$ ) y para la clase overweight ( $y = 1$ ), lo esperado por nuestro modelo es que todas las clases tengan un AUC lo más cercano a 1 donde 1 significa que predice correctamente los verdaderos positivos con respecto los falsos positivos que presenta por cada predicción. En la gráfica el valor optimo es llegar al extremo izquierdo, es decir (0,1) en el plano cartesiano el cual representa que por cada verdadero positivo no tenemos ningun falso positivo. La regresión logística es un modelo clásico el cual es nuestro benchmark y es un modelo eficiente computacionalmente para predecir comportamientos sin embargo para este caso de clasificacion de 3 clases observamos que ninguna de las clases se aproxima al resultado deseado.
- En la **imagen 10**, se puede ver el desempeño detallado del modelo con las métricas de precisión, recall y F1-score. A continuación, un resumen de las métricas clave:
  - *Clase -1 (underweight)*: La clase identifica de manera correcta todas las predicciones que realmente fueron correctas, es decir los verdaderos positivos (Precision = 100%), sin embargo son 13 casos los que predice de forma correcta, la sensibilidad de la clase (Recall) tiene un bajo desempeño de 23% , este indicador nos dice que solo el 23% de los verdaderos positivos fueron identificados de forma correcta, en cuanto el f1-score que hace referencia al promedio entre precision y recall, donde un valor más alto nos indica que para ambos indicadores el modelo clasifico de forma correcta, para nuestra clase de recesión economica el f1-score es de 0.38 lo que indica que nuestra clase tiene un bajo desempeño tanto para identificar verdaderos positivos y consigue identificar pocos casos que fueron verdaderos positivos.
  - *Clase 0 (neutral)*: Siguiendo con el analisis anterior en este caso tenemos una precisión de 50% es decir que la mitad de los casos que identifica la clase fueron verdaderos positivos en los datos reales, al igual que la anterior clase aquí conseguimos un bajo desempeño de 21%, viendo el desempeño de estos dos indicadores, podemos inferir que el f1-score es bajo, debido a que el f1-score es una media armónica o promedio entre los otros dos indicadores pasados. Nuestra clase tiene 29% de f1-score.

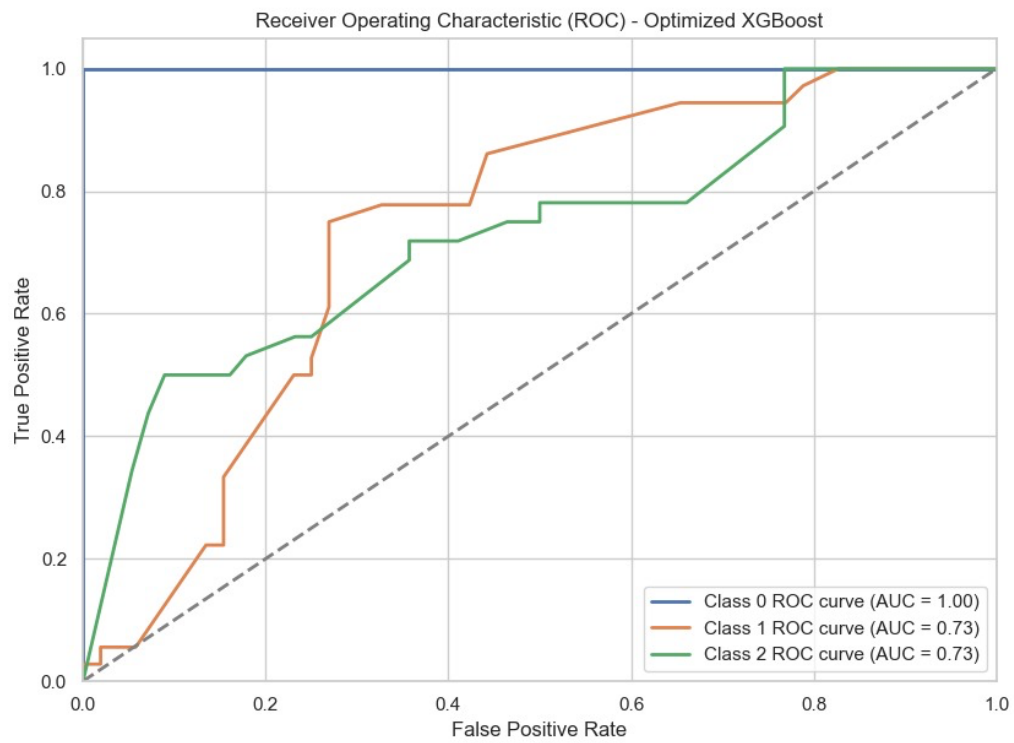
- *Clase 1 (overweight)*: La clase identifica de manera adecuada los verdaderos positivos teniendo una precisión del 100%, el modelo tiene una precisión de 9% lo que indica que comete muchos errores a la hora de predecir algo como positivo, en cuanto al f1-score obtuvo un 17%, este indicador nos dice el rendimiento del modelo considerando tanto los falsos positivos como los falsos negativos, observamos que consiguió un valor con bajo desempeño.

### **Promedios (micro, macro, weighted, samples):**

- **Micro avg:**
  - **Precisión, Recall, F1-Score: 0.67, 0.17, 0.27**: Aquí se suman los verdaderos positivos, falsos negativos y falsos positivos para todas las clases. La precisión es moderada (67%), pero el recall es muy bajo (17%), lo que indica que el modelo está fallando en capturar la mayoría de los ejemplos positivos.
- **Macro avg:**
  - **Precisión, Recall, F1-Score: 0.83, 0.18, 0.28**: Este promedio da igual peso a cada clase. La precisión es alta, lo que significa que en promedio el modelo es preciso en sus predicciones. Sin embargo, el recall es bajo, ya que el modelo tiene problemas para identificar correctamente ejemplos de cada clase.
- **Weighted avg:**
  - **Precisión, Recall, F1-Score: 0.80, 0.17, 0.26**: Este promedio está ponderado por el número de instancias de cada clase. Es similar al promedio macro, pero toma en cuenta que hay más ejemplos de algunas clases (como la clase 1).
- **Samples avg:**
  - **Precisión, Recall, F1-Score: 0.17, 0.17, 0.17**: Esta métrica es útil en casos de clasificación multietiqueta, donde una muestra puede pertenecer a más de una clase a la vez. En este caso, parece reflejar el rendimiento general bastante bajo del modelo.

En general, en base al anterior análisis, el modelo de regresión logística presenta un bajo rendimiento. Esto sugiere que un modelo más avanzado, capaz de capturar interacciones más complejas entre los indicadores económicos, sería necesario para mejorar la clasificación.

## RESULTADOS DEL MODELO XGBOOST:



*Imagen 11.* Gráfica AUC-ROC para el modelo de XGBoost.

```

Accuracy: 0.6705
-----

Classification Report:

              precision    recall  f1-score   support

   -1.0         1.00        0.60        0.75         20
    0.0         0.57        0.83        0.67         36
    1.0         0.74        0.53        0.62         32

 accuracy          0.67         88
  macro avg         0.77         88
 weighted avg         0.73         88

```

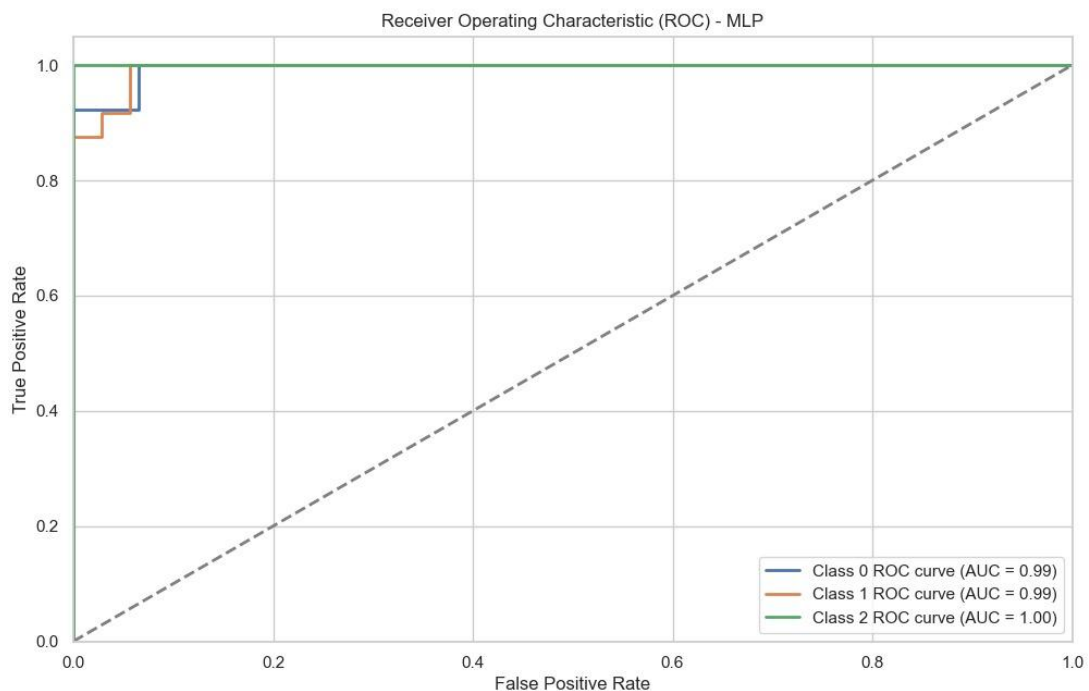
*Imagen 12. Reporte de clasificación para el modelo de XGBoost.*

- *Gráfica AUC-ROC:* El modelo XGBoost (**imagen 11**) muestra una mejora significativa, con un AUC de 0.82 para cada clase. Esto refleja una mejora en la capacidad de predicción para las tres categorías, con un aumento notable en la detección de fases económicas.
- En la **imagen 12**, podemos observar algunas de las métricas clave:
  - *Clase -1 (underweight):* F1-Score y Recall de 1.00, lo que indica que el modelo clasifica de buena manera las fases de la economía con un comportamiento anticíclico.
  - *Clase 0 (neutral):* F1-Score de 0.68, con un Recall de 0.75. Esto refleja que el modelo es relativamente eficaz en predecir fases neutrales, aunque con cierto margen para mejorar su precisión. El desempeño del modelo en esta clase es bueno, pero no sólido como en la clase recesiva, lo que sugiere que podría haber una mejora en la clasificación en estas fases.
  - *Clase 1 (overweight):* F1-Score de 0.56, con un recall de 0.50. El modelo tiene dificultades para predecir con exactitud las fases expansivas, lo cual es un aspecto para mejorar para lograr mejores predicciones en tiempos de crecimiento económico. Si bien el modelo tiene un desempeño razonable, no logra captar completamente las señales de expansión, lo que podría limitar su utilidad para identificar oportunidades de inversión.



El AUC de 0.82 indica que XGBoost tiene un mejor desempeño en la clasificación, logrando una mayor precisión en comparación con la regresión logística. En general, presenta un buen rendimiento, con una precisión global del 71.59% y un F1-Score ponderado de 0.71. Aunque el modelo logra identificar con gran precisión las fases recesivas, su capacidad para diferenciar entre fases neutrales y expansivas aún tiene espacio para mejorar. Esto sugiere que, aunque el modelo es una herramienta robusta para decisiones económicas, una mayor optimización podría mejorar su utilidad en contextos de crecimiento económico.

## RESULTADOS DE LA RED NEURONAL MULTICAPA (MLP) con ReLU



*Imagen 13. Gráfica AUC-ROC para la Red Neuronal Multicapa (MLP).*

MLP Neural Network (ReLU) Accuracy: 0.8813559322033898				
	precision	recall	f1-score	support
-1.0	0.80	0.92	0.86	13
0.0	0.87	0.83	0.85	24
1.0	0.95	0.91	0.93	22
accuracy			0.88	59
macro avg	0.87	0.89	0.88	59
weighted avg	0.89	0.88	0.88	59

*Imagen 14. Reporte de clasificación para la Red Neuronal Multicapa.*

- *Gráfica AUC-ROC:* La red neuronal multicapa (MLP) con activación ReLU (**imagen 13**) muestra un AUC de 0.99 para cada una de las clases (Clase 0: Underweight, Clase 1: Neutral, Clase 2: Overweight), un desempeño bastante bueno. Esto indica que el MLP es muy preciso para clasificar las diferentes fases del ciclo económico, superando con creces a los modelos previos como la regresión logística y XGBoost. El modelo tiene una capacidad elevada para diferenciar entre las clases.
- *Reporte:* En la **imagen 14**, se pueden observar las métricas detalladas de precisión, recall y F1-score para cada clase. El modelo MLP ha alcanzado una precisión general del 88% y un F1-Score ponderado de 0.88, mostrando un gran desempeño en la clasificación de todas las fases económicas:
  - *Clase 1 (underweight):* F1-Score de 0.86, con un recall de 0.92. El modelo predice de forma precisa las fases anticíclicas.
  - *Clase 0 (neutral):* F1-Score de 0.85, con un recall de 0.83. El modelo identifica con buen desempeño las fases neutrales.
  - *Clase 1 (overweight):* F1-Score de 0.93, con un recall de 0.91. El rendimiento en la identificación de fases expansivas es sobresaliente, siendo el mejor entre las tres clases, lo que sugiere que el modelo tiene un mejor rendimiento en las fases procíclicas del mercado.

El MLP con activación ReLU mantiene un gran equilibrio en las métricas de precisión, recall y F1-Score. La capacidad del MLP para capturar relaciones no lineales complejas entre los indicadores económicos permite realizar predicciones confiables.

## CONCLUSIONES Y SELECCIÓN DEL MODELO

**Desempeño del Modelo MLP con ReLU:** La Red Neuronal Multicapa (MLP) con función de activación ReLU tiene un rendimiento notable entre los modelos analizados, como evidencian los resultados obtenidos. Este modelo alcanzó una precisión global del 88%, un F1-Score ponderado de 0.88 y un buen desempeño en cada una de las tres categorías:

**Underweight (-1):** Logró un F1-Score de 0.86, indicando una alta precisión en la predicción de situaciones de tendencia anticíclica.

**Neutral (0):** Obtuvo un F1-Score de 0.85, mostrando un buen desempeño en la predicción de fases neutras, fundamental para una diversificación óptima del portafolio.

**Overweight (1):** Con un F1-Score de 0.93, el modelo demostró gran precisión al anticipar expansiones económicas, facilitando decisiones estratégicas en tendencias procíclicas.

El modelo MLP tiene el AUC más alto, y maneja un equilibrio robusto en las métricas clave de precisión, recall y F1-Score en todas las categorías. Esto implica que es el modelo que mejor generaliza y maneja las complejidades de los datos económicos y de mercado.

**Regresión Logística y XGBoost:** Tanto la Regresión Logística como XGBoost mostraron un rendimiento aceptable, sus AUC más bajos (0.73 y 0.82 respectivamente) demuestran sus limitaciones para capturar patrones no lineales y complejos en los indicadores económicos. La Regresión Logística tuvo dificultades para diferenciar entre las tres clases, lo que la convierte en una opción menos eficiente. Aunque XGBoost mostró mejoras significativas, no alcanzó el nivel de precisión y estabilidad logrado por el MLP, especialmente en la predicción de fases expansivas.

**Selección del Modelo MLP con ReLU:** Dado que el MLP con activación ReLU ha demostrado su capacidad para capturar las interacciones complejas entre los indicadores económicos y el mercado, es la opción más sólida para guiar las decisiones de inversión. Este modelo fue seleccionado por su capacidad para predecir con mayor exactitud los movimientos del ciclo económico, lo cual es crucial para implementar una estrategia de rotación sectorial exitosa.

## IDENTIFICACIÓN DE TENDENCIA ESPERADA

El modelo multiclase desarrollado en este proyecto permite identificar la tendencia económica y ajustar la composición del portafolio con base en las predicciones del modelo. El objetivo es optimizar la ponderación de los activos en función de las fases del ciclo económico, dividiendo las acciones en dos categorías:

- *Pro-cíclicas:* Empresas que tienden a tener un mejor desempeño en fases expansivas de la economía. Estos activos tienen una beta mayor a 0.7, lo que implica que tienden a amplificar los movimientos del mercado.
- *Anti-cíclicas:* Empresas que se comportan mejor en periodos de recesión, con una beta entre 0 y 0.7, lo que indica que su rendimiento es menos sensible a las fluctuaciones del mercado.

La decisión del modelo influye directamente en la ponderación de estos tipos de activos, lo cual se puede observar en la composición del portafolio en función de tres estados:

1. *Underweighted Portfolio:* Cuando el modelo predice una recesión económica (valoración -1), el portafolio se pondera en un 75% hacia activos anti-cíclicos, priorizando la estabilidad y la menor exposición a riesgo. Mientras que el 25% del portafolio está asignado a activos pro-cíclicos.
2. *Neutral Portfolio:* En una fase neutral del ciclo económico (valoración 0), el portafolio se distribuye equitativamente, con un 50% en activos pro-cíclicos y 50% en anti-cíclicos. Esto refleja un equilibrio entre sectores que podrían beneficiarse de condiciones económicas moderadas o volátiles.
3. *Overweighted Portfolio:* Durante una fase expansiva del ciclo económico (valoración 1), el modelo indica que se debería ponderar un 75% de los activos en empresas pro-cíclicas,

buscando aprovechar el crecimiento económico. Mientras que un 25% del portafolio se mantiene en sectores defensivos o anti-cíclicos.

## SELECCIÓN DE ESTRATEGIA DE INVERSIÓN

Una vez que el modelo identifica la tendencia esperada del mercado, es posible construir un portafolio otorgando mayor peso a los activos que se beneficien más de la fase económica en la que se encuentre. Existen diversas estrategias de inversión que permiten optimizar un portafolio, como el Ratio de Sharpe, la Mínima Varianza o el Ratio de Sortino. Sin embargo, dado el enfoque de este proyecto, basado en la rotación sectorial a partir de indicadores adelantados del mercado, se considera que el Ratio de Sharpe es la estrategia más adecuada. A continuación, se define la estrategia:

Un portafolio con el Ratio de Sharpe es una estrategia que, mide el rendimiento de una inversión, ajustando el riesgo y comparándola con la rentabilidad de activo libre de riesgo. Esta razón se cuestiona si el rendimiento adicional de una inversión compensa lo suficiente al riesgo adicional que se asume. Históricamente, el Máximo Ratio de Sharpe ha ganado mucha popularidad, hasta convertirse en una de las herramientas más comunes para la evaluación de portafolios formando parte de los conceptos básicos de todo inversor. Fue desarrollado por el Premio Nobel William F. Sharpe de la Universidad de Stanford. Introducido en 1966 por Sharpe en un artículo publicado en el Journal of Business, “Mutual Fund Performance”, con el objetivo principal de proporcionar una estrategia que ayudara a los inversionistas a evaluar la rentabilidad de un activo en relación con el riesgo asumido. Sharpe desarrolló esta estrategia como parte de su trabajo en el campo de la teoría moderna de portafolios, que incluye otros modelos como CAPM (Modelo de Valoración de Activos Financieros).

## IMPLEMENTACIÓN

Para la implementación de la estrategia de inversión se utilizó el algoritmo planteado para predecir la tendencia esperada, este algoritmo está diseñado para evaluar y optimizar la rotación sectorial de un portafolio de inversión utilizando datos históricos. La rotación sectorial se basa en adaptar la selección de activos a las condiciones económicas, clasificando los activos como “anticiclicos” o “prociclicos” según su comportamiento frente a los ciclos economicos que vimos anteriormente.

## BACKTESTING DINÁMICO

### FUNDAMENTACIÓN

El backtesting es un método de simulación de portafolios que permite evaluar de forma estadística la efectividad de una estrategia de inversión. Este proceso implica simular la estrategia (En este caso de rotación sectorial) utilizando los pesos y datos históricos del portafolio, para, en función de un horizonte temporal, analizar cómo se habría comportado.

Realizaremos un backtesting dinamico donde se actualizan las ponderaciones periodicamente mediante la re-optimización del portafolio. Para este proyecto se descargaron los precios de cierre de las 500 empresas que forman parte del S&P500 y el precio de cierre del índice, desde 01/01/2000 hasta 31/05/2024. Para analizar el desempeño de nuestra estrategia contra el benchmark (S&P500). A continuación se desglosará el flujo de trabajo que se realizó para el backtesting:

1. Con nuestra base de datos, seleccionamos aleatoriamente 20 acciones, donde en el primer periodo a optimizar serán 10 prociclicas y 10 anticiclicas. Con una asignación equitativa del 50% para cada tipo, esta estructura busca evitar sesgos hacia una sola categoría de activos, estableciendo un punto de referencia neutral que permita analizar de manera objetiva el desempeño de ambos segmentos en el contexto del ciclo económico.
2. Se define la periodicidad del rebalanceo, la cual es anual.
3. Se parte la base de datos en función de las ventanas de tiempo y de los años de información disponibles.

4. Se optimizan las ponderaciones utilizando la primer ventana de tiempo  $t_0$  y se realiza la simulación para el periodo posterior, es decir de  $t_1$  hasta  $t_2$ , una vez finalizada la ventana, se optimizan de nuevo las ponderaciones, ahora usando la información de  $t_1$  hasta  $t_2$  y se simula utilizando desde  $t_2$  a  $t_3$ , este proceso se repite iterativamente hasta llegar a  $t_n$ .
5. Con la evolucion historica del portafolio, obtenemos las metricas de desempeño.

De esta manera, pondremos a prueba el modelo generado anteriormente, el cual predice la tendencia esperada del ciclo económico. Un buen desempeño del modelo , en teoría, ayudará a tener mejores resultados en el backtest, ya que los activos seleccionados se ajustarán de mejor manera a cada momento del mercado.

El backtesting dinámico permitirá evaluar el modelo predictivo del mercado, así como generar datos y resultados que permitan mostrar de una mejor manera, si el portafolio gestionado con la estrategia de rotación sectorial es viable, cumple con los objetivos y supera al benchmark. Para esto se generarán 2000 simulaciones, y posteriormente se utilizarán estos escenarios para calcular métricas de desempeño del portafolio.

## MÉTRICAS DE DESEMPEÑO

Para evaluar el desempeño de estrategias de inversión mediante backtesting dinámico, es fundamental contar con métricas que ofrezcan una evaluación objetiva y completa del rendimiento del portafolio en comparación con un benchmark de referenciam como el índice S&P 500. Entre las métricas más relevantes para este tipo de análisis se encuentran Ratio de Sharpe y la comparación del rendimiento del portafolio vs el benchmark.

- I. Ratio de Sharpe: Mide la eficiencia de una estrategia de inversión en relación con el riesgo asumido, se define como el retorno promedio excedente del portafolio (el retorno por encima de la tasa libre de riesgo) dividido por la desviación estándar de estos retornos. Un ratio de sharpe alto indica que la estrategia está generando mayores retornos ajustados por riesgo. Esta métrica permite a los inversores evaluar si un rendimiento elevado se justifica por el nivel de riesgo asumido.

$$\text{Ratio de Sharpe} = \frac{E(R_p) - R_f}{\sigma_p}$$

Dónde:

- $E(R_p)$ : Rendimiento promedio del portafolio.
- $R_f$ : Tasa libre de riesgo.
- $\sigma_p$ : Desviación estándar de los rendimientos del portafolio.

- II. Rendimiento del portafolio vs Benchmark (S&P500): Comparar directamente el rendimiento del portafolio contra el benchmark permite evaluar si la estrategia cumple o supera el comportamiento de un índice de referencia. Esta métrica se basa en analizar el rendimiento acumulado del portafolio frente al rendimiento acumulado del benchmark, observando si la estrategia logra añadir valor consistente a lo largo del tiempo.

$$\text{Exceso de Rendimiento} = E(R_p) - E(R_b)$$

Dónde:

- $E(R_p)$ : Rendimiento promedio del portafolio.
- $E(R_b)$ : Rendimiento promedio del benchmark.

- III. Upside Risk: El riesgo al alza mide la variabilidad de los rendimientos de un portafolio que exceden un umbral de referencia. Este indicador se centra en las ganancias, proporcionando información sobre el potencial positivo de un portafolio. Un upside elevado indica que el portafolio tiene una mayor capacidad de generar rendimientos significativamente superiores al umbral definido.

$$\text{UpsideRisk} = \sqrt{\frac{\sum_{i=1}^n (R_i - T)^2 * I(R_i > T)}{n}}$$

Dónde:

- $(R_i)$ : Rendimiento observado del portafolio en el período  $i$ .



- $T$ : Umbral de referencia (por ejemplo,  $R_f$ ).
- $I(R_i > T)$ : Indicador que toma el valor 1 si  $(R_i > T)$  y 0 en caso contrario.
- $n$ : Número total de observaciones.

IV. Downside Risk: El riesgo a la baja mide la variabilidad de los rendimientos de un portafolio que se encuentran por debajo de un umbral de referencia. Este indicador se enfoca únicamente en las pérdidas potenciales, ignorando los rendimientos positivos, lo que lo convierte en una métrica relevante para los inversionistas aversos al riesgo. Un downside elevado indica que el portafolio tiene una mayor exposición a rendimientos inferiores al umbral definido, reflejando un mayor riesgo de pérdida.

$$DownsideRisk = \sqrt{\frac{\sum_{i=1}^n (R_i - T)^2 * I(R_i < T)}{n}}$$

Dónde:

- $(R_i)$ : Rendimiento observado del portafolio en el período  $i$ .
- $T$ : Umbral de referencia (por ejemplo,  $R_f$ ).
- $I(R_i < T)$ : Indicador que toma el valor 1 si  $(R_i < T)$  y 0 en caso contrario.
- $n$ : Número total de observaciones.

V. Ratio de Sortino: El ratio de sortino mide la eficiencia de una estrategia de inversión considerando solo el riesgo a la baja, es decir, el downside risk. A diferencia de sharpe, este indicador excluye la variabilidad causada por los rendimientos positivos, centrándose únicamente en las pérdidas potenciales. Un ratio de sortino alto indica que el portafolio está generando mayores rendimientos ajustados por riesgo a la baja, siendo especialmente útil para inversionistas que buscan minimizar las pérdidas mientras maximizan los retornos.

$$Sortino\ Ratio = \frac{R_p - R_f}{\sigma_d}$$

Dónde:

- $(R_p)$ : Rendimiento promedio del portafolio.
- $(R_f)$ : Tasa libre de riesgo.
- $(\sigma_d)$ : Desviación estándar a la baja.

VI. Alpha de Jensen: El alpha de jensen mide la capacidad de un portafolio para generar rendimientos superiores a los esperados en función de su nivel de riesgo sistemático, según lo definido por el modelo CAPM. Este indicador representa el retorno adicional obtenido por el portafolio después de ajustar por el riesgo asumido frente al benchmark. Un alpha de jensen positivo indica que el portafolio ha superado el rendimiento ajustado por riesgo del benchmark, lo que sugiere un desempeño superior. Por otro lado, un alpha negativo refleja un rendimiento inferior al esperado dado el nivel de riesgo.

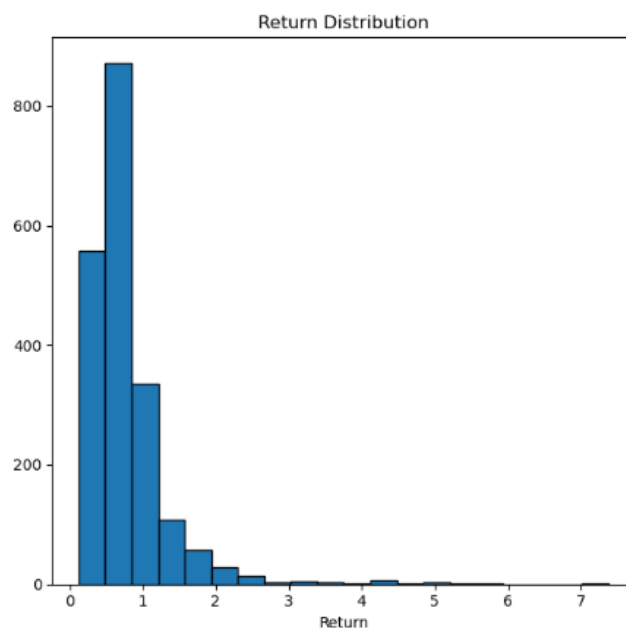
$$\text{Alpha de Jensen} = R_p - [R_f + \beta (R_m - R_f)]$$

Dónde:

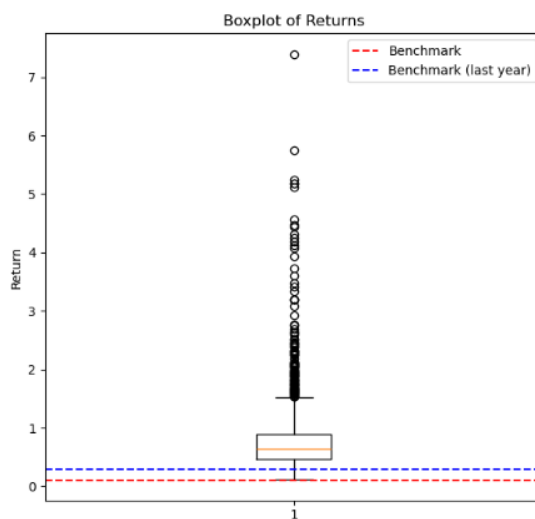
- $E(R_p)$ : Rendimiento promedio del portafolio.
- $E(R_m)$ : Rendimiento promedio del benchmark.
- $E(R_f)$ : Tasa libre de riesgo.
- $\beta$ : Sensibilidad del portafolio a los movimientos del mercado.

## RESULTADOS DE MÉTRICAS DE DESEMPEÑO

El análisis de desempeño de la estrategia de rotación sectorial se basó en la simulación de 20,000 portafolios, cuyos resultados se presentan a continuación. En la Imagen 15, que corresponde al histograma de retornos de las simulaciones, se observa una distribución con un marcado sesgo positivo hacia la derecha. Este comportamiento indica una tendencia a obtener resultados superiores al valor esperado, con un rendimiento promedio de 77.55% y una mediana de 63.89%. Sin embargo, como se muestra en la Imagen 16 (*boxplot de los retornos*), se identificaron múltiples datos fuera del rango intercuartílico. Esto motivó un análisis adicional para excluir estos "outliers" y obtener una representación más precisa del desempeño de la estrategia.

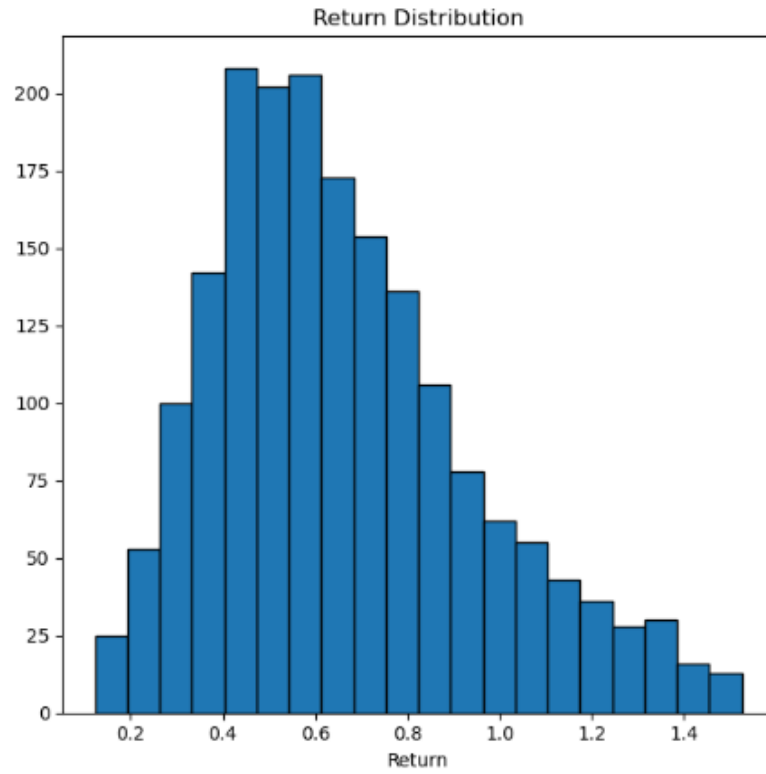


*Imagen 15. Histograma de Retornos de Simulaciones.*

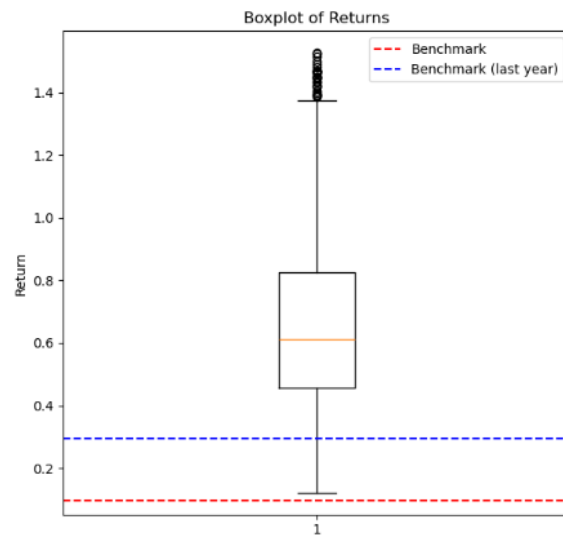


*Imagen 16. Boxplot de Retornos de Simulaciones.*

Tras aplicar el criterio del rango intercuartílico, considerando un factor de 1.5 para los límites superior e inferior, los datos se ajustaron para eliminar los valores atípicos. Los resultados refinados se resumen en la Imagen 17 (*histograma ajustado*) y la Imagen 18 (*boxplot ajustado*). Sin los outliers, el rendimiento esperado disminuyó a un 66.20%, con una mediana de 61.27%. Además, se determinó que este rendimiento esperado se supera en un 42.93% de las simulaciones, lo que refuerza la confiabilidad de la estrategia.



*Imagen 17. Histograma de Retornos de Simulaciones sin Atípicos.*



*Imagen 18. Boxplot de Retornos de Simulaciones sin Atípicos.*

La estrategia de rotación sectorial presentó un desempeño sólido, respaldado por métricas clave que evidencian su capacidad para generar rendimientos ajustados por riesgo de manera eficiente. En términos de Sharpe Ratio, se obtuvo un valor de 3.90, lo cual es significativamente superior al umbral considerado como bueno (mayor a 1). Este resultado sugiere que el portafolio se gestiona de manera efectiva, logrando rendimientos atractivos en relación con el riesgo asumido, un aspecto esencial en cualquier estrategia de inversión bien estructurada.

- Métrica	Rotación Sectorial
Sharpe	3.90
Volatility	15.51%
Upside	11.93%
Downside	10.15%
Sortino	5.97
Jensen Alpha	0.025

*Tabla 1. Métricas de desempeño.*

La volatilidad del portafolio, que mide la variabilidad de los rendimientos, alcanzó un nivel moderado de 15.51%. Este grado de riesgo es razonable, especialmente considerando los rendimientos obtenidos, lo que lo hace aceptable para inversionistas que buscan un equilibrio entre riesgo y recompensa. Adicionalmente, el riesgo al alza (Upside Risk) del portafolio fue de 11.93%, lo que indica un potencial positivo significativo para generar rendimientos superiores, un aspecto sumamente atractivo para los inversionistas.

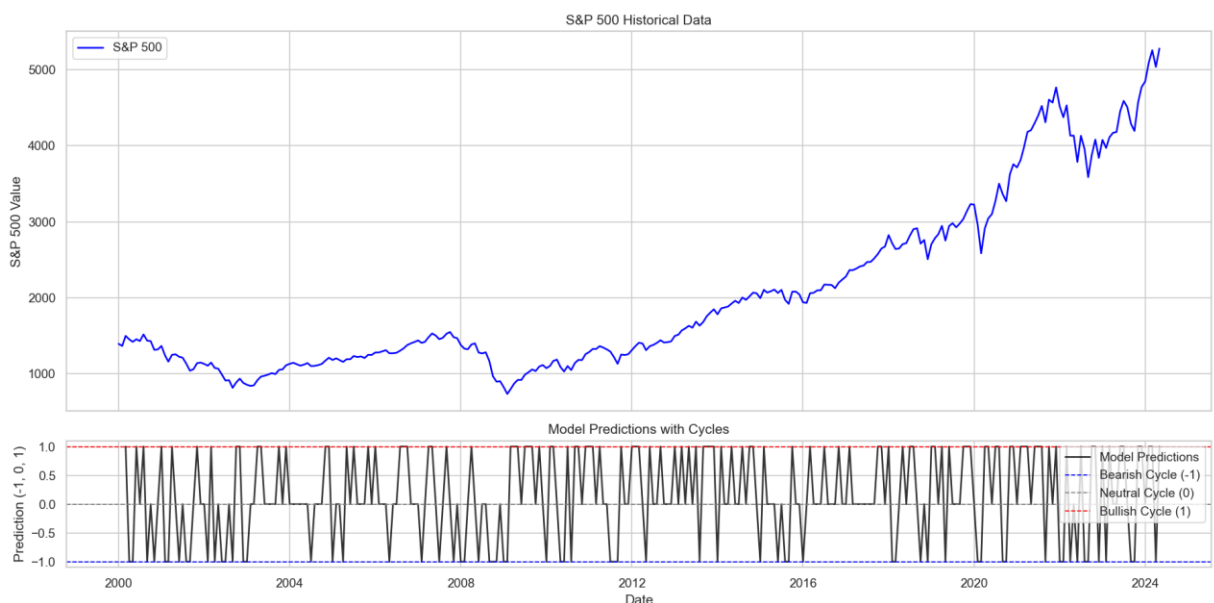
Por otro lado, el riesgo a la baja (Downside Risk) del portafolio se situó en 10.15%, evidenciando una exposición controlada a las pérdidas. Este resultado muestra que el portafolio logra mantener una mayor inclinación hacia el potencial de ganancias, con una exposición limitada a resultados negativos, algo crucial en la construcción de estrategias conservadoras pero rentables.

El Sortino Ratio alcanzó un nivel de 5.97, un indicador que refuerza el desempeño positivo de la estrategia al considerar únicamente el riesgo a la baja. Esto demuestra que el portafolio no solo está diseñado para mitigar pérdidas potenciales, sino que también maximiza los rendimientos ajustados por este tipo de riesgo. Finalmente, el Alpha de Jensen, con un valor de 0.025, confirma que la estrategia genera rendimientos por encima de lo esperado, considerando el riesgo sistemático frente al benchmark. Este resultado resalta la capacidad del portafolio para agregar valor más allá de las expectativas del mercado.

En conjunto, el portafolio generado mediante la estrategia de rotación sectorial demuestra un desempeño notable. Su capacidad para gestionar eficientemente las pérdidas potenciales, junto con un sólido potencial de ganancias y un Alpha positivo, lo posiciona como una alternativa interesante para los inversionistas que buscan maximizar retornos sin asumir riesgos excesivos. Esto lo convierte en una herramienta valiosa para quienes desean combinar estabilidad con resultados superiores frente al benchmark.

## COMPARACIÓN CON EL ‘BENCHMARK’

Posteriormente, al comparar los resultados obtenidos con un benchmark, se logran desarrollar múltiples observaciones, las cuales van desde los resultados del modelo entrenado y el comportamiento del mercado, hasta comprender cual fue el desempeño de la estrategia de inversión y generación de portafolios.



*Imagen 19. Relación entre Predicciones del Modelo y Comportamiento del S&P 500.*

Al generar las gráficas de los resultados del modelo y del benchmark, se permite entender la interacción entre la predicción del ciclo económico (según el modelo) y el comportamiento del S&P 500, en donde se observa:

1. **Comportamiento histórico del Benchmark:** La primera gráfica muestra cómo se ha comportado a lo largo del periodo establecido el valor del mercado (S&P500), en donde se aprecia un comportamiento alcista en la mayoría del periodo.

2. **Predicciones del Modelo y Ciclos Económico:** La segunda gráfica muestra cómo el modelo clasifica los ciclos económicos en categorías de tendencia: -1 (recesión), 0 (neutral), y 1 (expansión), reflejando una correlación entre el estado anticipado del mercado y la categoría económica estimada.

Estas gráficas permiten analizar las similitudes entre las predicciones del modelo y el comportamiento del S&P 500, mostrando que las concentraciones en las predicciones tienden a coincidir con movimientos importantes en el índice, se puede apreciar cómo en etapas donde el índice tiende a subir se concentran las predicciones en 1.

Mientras que en el análisis comparativo del rendimiento del portafolio frente al benchmark, el S&P 500, revela diferencias significativas en términos de retorno ajustado al riesgo y estabilidad. El S&P 500 ha registrado un rendimiento total del 278.46%, lo que evidencia un crecimiento sustancial durante el período evaluado. Con un Sharpe Ratio de 1.84, se demuestra una gestión eficiente del riesgo, generando retornos sólidos por cada unidad de volatilidad. Adicionalmente, el Omega Ratio de 1.37 destaca una proporción alta de rendimientos positivos frente a los negativos, subrayando la fiabilidad de este índice como punto de referencia.

En contraste, los portafolios desarrollados presentan un valor final promedio de \$19,610,992.08, con un rendimiento promedio de 18.61%. El Sharpe Ratio general del portafolio es de 2.64, lo cual indica que, en promedio, los rendimientos son suficientes para compensar el riesgo asumido, lo que puede traducirse en ganancias ajustadas al riesgo. El Omega Ratio general de 2.35 mejor que en algunos escenarios anteriores, señala un desequilibrio considerable en la distribución de rendimientos, con una mayor probabilidad de resultados positivos.

En los extremos, los resultados del portafolio varían ampliamente. El mejor caso, con un valor final de \$178,181,409.00 y un rendimiento máximo del 177%, muestra un desempeño destacable. Este escenario tiene un Sharpe Ratio de 3.62, lo que indica una gestión más eficiente del riesgo cuando las condiciones del mercado son favorables, y un Omega Ratio de 2.55, reflejando una buena proporción de rendimientos positivos. No obstante, tales resultados no son la norma y dependen de circunstancias específicas.

Por otro lado, el peor escenario del portafolio termina con un valor final de \$3,932,604.78, equivalente a un rendimiento mínimo de 2.9%. En este caso, el Sharpe Ratio cae a un alarmante -

1.40, lo que señala un manejo del riesgo deficiente. El Omega Ratio es de 0.20, reafirmando el desequilibrio en los rendimientos, con una probabilidad de pérdidas significativas en situaciones adversas.

En resumen, el portafolio muestra un rango amplio de resultados, desde rendimientos excepcionalmente altos hasta mínimos debajo del benchmark. Aunque el S&P 500 mantiene un desempeño consistente y ajustado al riesgo, la estrategia del portafolio generó una mejora en su eficiencia y fiabilidad, especialmente en términos de la relación riesgo-retorno.

## CONCLUSIONES

La metodología implementada en este proyecto logró cumplir los objetivos planteados, inicialmente el desarrollar un modelo capaz de anticipar las fases del ciclo económico y emplear esta predicción para guiar decisiones de inversión basadas en una estrategia de rotación sectorial. A lo largo del proceso, se demostró la utilidad de herramientas de aprendizaje automático como la regresión logística, XGBoost y redes neuronales multicapa (MLP). De manera destacable, el modelo MLP con activación ReLU superó a las demás alternativas, mostrando un equilibrio robusto entre precisión, recall y F1-Score en todas las categorías. Este desempeño refuerza su potencial para convertirse en una herramienta prescriptiva efectiva en la gestión de portafolios.

El análisis del backtesting dinámico confirmó que la estrategia de rotación sectorial tiene la capacidad de mejorar el rendimiento ajustado al riesgo en comparación con el benchmark (S&P 500). Sin embargo, este enfoque no está libre de desafíos. La presencia de un sesgo positivo en los histogramas de simulaciones sugiere que la estrategia puede ser sensible a condiciones económicas específicas. Para minimizar este riesgo, se realizó una comparación con un portafolio exclusivamente compuesto por el benchmark, lo cual permitió validar que los rendimientos adicionales generados por nuestra estrategia no son productos aleatorios, sino del impacto positivo del modelo predictivo.

En cuanto a las métricas de desempeño, se incluyeron indicadores clave como el Ratio de Sharpe, Downside Risk, Upside Risk, Sortino y Alpha de Jensen, lo que permitió una evaluación más profunda del comportamiento de nuestro modelo de rotación sectorial. El Alpha de Jensen positivo destaca particularmente, al evidenciar que el portafolio gestionado con la estrategia de rotación



sectorial genera rendimientos superiores ajustados por riesgo en comparación con el benchmark. Esta métrica subraya el valor agregado de la estrategia y su capacidad para vencer al mercado bajo ciertas condiciones.

Es importante mencionar que la metodología no está exenta de limitaciones. Si bien el modelo mostró un desempeño sobresaliente en datos históricos, su utilización implica dependencias de la estabilidad de las relaciones identificadas en el pasado. Cambios en el mercado, como la irrupción de nuevas tecnologías o políticas económicas atípicas, podrían disminuir su eficacia. Por tanto, es **crucial reentrenar el modelo de manera periódica**, idealmente cada trimestre o tras eventos económicos significativos, para garantizar que las predicciones se mantengan relevantes.

En términos prácticos, se considera que la estrategia propuesta es viable, siempre y cuando se combine con un monitoreo constante de los datos de entrada y un ajuste flexible a las dinámicas del mercado. Adicionalmente, los resultados del proyecto destacan la importancia de ampliar el conjunto de métricas de desempeño y utilizar enfoques avanzados como la estimación de densidad Kernel (KDE) para entender mejor los riesgos asociados con los rendimientos negativos.

Finalmente, aunque la metodología implementada cumplió los objetivos establecidos, se identificaron áreas de oportunidad. Por ejemplo, sería valioso explorar modelos adicionales que integren interpretabilidad, como Random Forests, para enriquecer la toma de decisiones. También se recomienda realizar estudios más profundos sobre el impacto de los umbrales de rendimiento utilizados en la clasificación de posturas económicas, con el fin de ajustar aún más la estrategia de inversión.

En resumen, este proyecto representa un paso significativo hacia la implementación de estrategias de inversión basadas en datos y aprendizaje automático. Los resultados son prometedores y demuestran que es posible superar el benchmark bajo ciertas condiciones, aunque también queda claro que el éxito a largo plazo dependerá de la adaptabilidad del modelo a un entorno económico en constante cambio.

# ANEXOS

## SOFTWARE

1. *Python – Vscode (librerías: pandas, matplotlib, numpy, tensorflow, keras, scipy, yfinance)*: Lenguaje y entorno, respectivamente, de programación para desarrollar y ejecutar el código del proyecto. Las librerías enlistadas son las principales y cumplirán con el objetivo de manipular y analizar datos (pandas), visualizar gráficas (matplotlib), realizar cálculos numéricos (numpy, scipy), implementar modelos (tensorflow, keras), obtener datos financieros históricos (yfinance), entre otras más.
2. *GitHub*: Aplicación donde se almacenará y gestionará el código fuente del proyecto (repositorio), el cual servirá como evidencia del trabajo realizado, permitiendo un control de versiones y facilitando la colaboración entre los integrantes del equipo.

## GLOSARIO

1. *Acción financiera*: Título que representa la propiedad parcial de una empresa y otorga el derecho sobre su beneficio.
2. *Accionista*: Persona o entidad que posee acciones de una empresa y tiene derechos sobre sus ganancias y decisiones corporativas.
3. *Actividad económica*: Conjunto de acciones relacionadas con la producción, distribución y consumo de bienes y servicios en una economía.
4. *Activo financiero*: Bien intangible que tiene valor económico, como acciones, bonos o derivados.
5. *Análisis estadístico y descriptivo (EDA)*: Proceso de analizar y visualizar datos para descubrir patrones, anomalías y relaciones, antes de aplicar modelos predictivos.
6. *Asignación de activos (asset allocation)*: Estrategia de distribución del capital en diferentes clases de activos (acciones, bonos, efectivo) para equilibrar riesgo y rendimiento.
7. *Aversión al riesgo*: Preferencia de los inversores por inversiones seguras, con menor volatilidad, frente a aquellas con mayor riesgo.

8. *Backtesting dinámico*: Proceso de probar una estrategia de inversión utilizando datos históricos, ajustando sus parámetros a lo largo del tiempo para simular condiciones reales.
9. *Banco central*: Institución que gestiona la política monetaria de un país, emite moneda, y regula el sistema bancario.
10. *Benchmark financiero*: Índice o referencia que se utiliza para comparar el rendimiento de una cartera de inversión.
11. *Beta financiera*: Medida de la volatilidad de un activo en relación con el mercado general. Beta mayor a 1 indica mayor volatilidad que el mercado.
12. *Ciclo económico*: Fluctuaciones en la actividad económica de un país, que incluyen fases de expansión, pico, recesión y recuperación.
13. *Clúster*: Agrupamiento de datos con características similares, utilizado en técnicas de análisis como el clustering.
14. *Consumidor*: Persona o entidad que compra y utiliza bienes y servicios.
15. *Correlación (matriz de correlación)*: Medida estadística que indica la relación entre dos o más variables, representada en una matriz para visualizar las interacciones entre múltiples variables.
16. *Covarianza*: Medida que indica el grado en que dos variables cambian juntas.
17. *Demanda agregada*: Total de bienes y servicios demandados en una economía a un nivel dado de precios y durante un periodo de tiempo específico.
18. *Desviación estándar*: Medida de dispersión que indica cuánto varían los datos respecto a la media.
19. *Distribución (normal, anormal, sesgada)*: Representación gráfica de cómo se distribuyen los datos. La distribución normal es simétrica, mientras que las distribuciones sesgadas muestran asimetrías hacia uno de los lados.
20. *Economía*: Ciencia que estudia la producción, distribución y consumo de bienes y servicios.
21. *Elasticidad financiera*: Sensibilidad de una variable financiera (como la demanda) en respuesta a cambios en otra variable, como el precio.
22. *Estrategia de inversión*: Plan estructurado que guía las decisiones sobre la compra, venta o mantenimiento de activos financieros para cumplir con objetivos financieros específicos.
23. *Expansión económica*: Fase del ciclo económico caracterizada por el crecimiento del PIB, aumento en el empleo y la inversión.

- 24. *Fluctuación económica:*** Cambios en el nivel de actividad económica, reflejados en el crecimiento o decrecimiento del PIB, empleo, etc.
- 25. *Gestor de portafolios:*** Profesional encargado de administrar y tomar decisiones sobre la composición y gestión de una cartera de inversiones para maximizar el rendimiento ajustado al riesgo.
- 26. *Indicador o métrica económica:*** Medida estadística que evalúa el estado de la economía (PIB, tasa de desempleo, inflación).
- 27. *Inversión financiera:*** Asignación de capital a activos con el objetivo de obtener un rendimiento en el futuro.
- 28. *Inversionista:*** Persona o entidad que destina capital a activos financieros con la expectativa de obtener rendimientos.
- 29. *Librería:*** Conjunto de funciones y herramientas predefinidas en un lenguaje de programación que facilita la realización de tareas específicas, como análisis de datos, visualización o cálculos matemáticos, sin necesidad de escribir código desde cero.
- 30. *Macroeconomía:*** Rama de la economía que estudia el comportamiento global de una economía, incluyendo el PIB, la inflación y el desempleo.
- 31. *Media:*** Promedio aritmético de un conjunto de datos.
- 32. *Mediana:*** Valor central en un conjunto de datos ordenados.
- 33. *Mercado financiero:*** Espacio donde se intercambian activos financieros como acciones, bonos, divisas, y derivados. Facilita la captación de capital y permite a los inversores comprar y vender activos.
- 34. *Microeconomía:*** Rama de la economía que estudia las decisiones individuales de consumidores y empresas y su interacción en los mercados.
- 35. *Moda:*** Valor que aparece con mayor frecuencia en un conjunto de datos.
- 36. *Modelo matemático (parámetros e hiper-parámetros; X-input, y-output):*** Representación matemática que usa variables de entrada (X) y salida (y) para predecir o modelar el comportamiento de un sistema, donde los parámetros controlan el modelo y los hiper-parámetros ajustan el rendimiento.
- 37. *Política económica:*** Conjunto de medidas y acciones adoptadas por un gobierno para influir en la economía nacional.

38. *Política monetaria*: Estrategias implementadas por un banco central para controlar la oferta de dinero y las tasas de interés con el objetivo de influir en la inflación y el crecimiento económico.
39. *POO (programación orientada a objetos)*: Paradigma de programación que organiza el código en "objetos" que contienen datos y métodos, facilitando la reutilización y modularidad.
40. *Portafolio de inversión*: Conjunto de activos financieros mantenidos por un inversor con el fin de diversificar el riesgo y maximizar el rendimiento.
41. *Precio de cierre*: Último precio al que se negoció un activo en una jornada de mercado.
42. *Recesión económica*: Periodo de contracción en la actividad económica, caracterizado por una disminución en el PIB y un aumento en el desempleo.
43. *Rendimiento financiero*: Retorno generado por una inversión, expresado generalmente como porcentaje.
44. *Repositorio*: Espacio de almacenamiento en una plataforma de control de versiones (como GitHub) donde se guarda el código fuente de un proyecto, junto con su historial de cambios y versiones. Facilita la colaboración y el seguimiento de las modificaciones realizadas.
45. *Rezago (shift)*: Diferencia de tiempo entre el momento en que ocurre un evento económico y cuando se observa su impacto en los datos.
46. *Riesgo financiero (volatilidad financiera)*: Posibilidad de que una inversión pierda valor debido a fluctuaciones en los precios del mercado.
47. *Riesgo sistémico*: Riesgo que afecta a todo el mercado financiero o sistema económico, no se puede eliminar mediante diversificación.
48. *Sector económico*: División de la economía que agrupa a empresas que ofrecen productos o servicios similares (tecnología, salud, energía).
49. *Situación financiera*: Estado económico actual de una persona, empresa o gobierno, basado en sus activos, pasivos y capacidad para generar ingresos.
50. *Tasa de interés*: Costo del dinero prestado o el rendimiento de una inversión, generalmente expresado como porcentaje anual.
51. *Tendencia alcista*: Periodo prolongado en el que los precios de los activos financieros tienden a subir.
52. *Tendencia bajista*: Periodo prolongado en el que los precios de los activos financieros tienden a bajar.

53. *Valor atípico (outlier)*: Punto de datos que se encuentra muy alejado de los demás en un conjunto, lo que podría indicar una anomalía o error.
54. *Valor máximo*: El valor más grande en un conjunto de datos.
55. *Valor mínimo*: El valor más pequeño en un conjunto de datos.
56. *Valuación financiera*: Proceso de determinar el valor justo de un activo financiero, generalmente utilizando métodos como el flujo de caja descontado o comparables de mercado.
57. *Varianza*: Promedio de las diferencias al cuadrado respecto a la media, utilizado para medir la dispersión de los datos.
58. *S&P 500*: Índice que representa el rendimiento de las 500 empresas más grandes de EE. UU., usado como indicador del mercado de acciones.
59. *Función de activación ReLU*: Función no lineal en redes neuronales que pasa valores positivos y convierte negativos en cero, acelerando la convergencia del entrenamiento.
60. *Tasa de aprendizaje*: Parámetro que determina el tamaño del paso en cada iteración del entrenamiento de modelos, crucial para la convergencia del algoritmo.
61. *Validación cruzada*: Método para evaluar la generalización de modelos predictivos, alternando conjuntos de entrenamiento y prueba para estimar el rendimiento.
62. *Especificidad (AUC-ROC)*: Métrica que mide la capacidad del modelo para identificar correctamente casos negativos en análisis de clasificación.

## BIBLIOGRAFIA

- Beers, Brian. Investopedia. (07 de julio del 2024). “*What Is Sector Rotation? How It Works and Importance in Investing*”. Recuperado el 18 de agosto del 2024, de: <https://www.investopedia.com/articles/trading/05/020305.asp#:~:text=What%20Is%20Sector%20Rotation%3F%20Sector%20rotation%20is%20the,them%20thrive%20or%20languish%20depending%20on%20the%20cycle>
- DISFOLD. (25 de agosto del 2024). “*Rotación de Sectores en Trading: Técnicas, Tendencias e Indicadores*”. Recuperado el 31 de agosto del 2024, de: <https://es.disfold.com/rotacion-sectores-trading/>
- Financial Engineers. GitHub. (2024). “*PAP-ERS*”. Recuperado el 18 de agosto del 2024, de: <https://github.com/diegotita4/PAP-ERS>
- González, Sean. (2024). “*Rotación Sectorial*”. Presentación de PowerPoint. Recuperado el 07 de septiembre del 2024.

- IBM. (s. f.). “¿Qué es el análisis exploratorio de datos (EDA)?”. Recuperado el 14 de septiembre del 2024, de: <https://www.ibm.com/mx-es/topics/exploratory-data-analysis>
- IBM. (s. f.). “¿Qué es la regresión logística?”. Recuperado el 21 de septiembre del 2024, de: <https://www.ibm.com/es-es/topics/logistic-regression>
- IBM. (s. f.). “¿Qué son las redes neuronales?”. Recuperado el 21 de septiembre del 2024, de: <https://www.ibm.com/es-es/topics/neural-networks>
- Organisation for Economic Co-operation and Development. OECD. (2024). “Composite leading indicator (CLI)”. Recuperado el 18 de agosto del 2024, de: <https://www.oecd.org/en/data/indicators/composite-leading-indicator-cli.html>
- Organisation for Economic Co-operation and Development. OECD. (2024). “Business confidence index (BCI)”. Recuperado el 18 de agosto del 2024, de: <https://www.oecd.org/en/data/indicators/business-confidence-index-bci.html>
- Organisation for Economic Co-operation and Development. OECD. (2024). “Real Gross Domestic Product (GDP)”. Recuperado el 18 de agosto del 2024, de: <https://www.oecd.org/en/data/indicators/real-gross-domestic-product-gdp.html>
- Organisation for Economic Co-operation and Development. OECD. (2024). “Consumer confidence index (CCI)”. Recuperado el 18 de agosto del 2024, de: <https://www.oecd.org/en/data/indicators/consumer-confidence-index-cci.html>
- PyPI. (01 de septiembre del 2024). “Optuna 4.0.0”. Recuperado el 21 de septiembre del 2024, de: <https://pypi.org/project/optuna/>
- PyPI. (11 de septiembre del 2024). “Scikit-learn 1.5.2”. Recuperado el 21 de septiembre del 2024, de: <https://pypi.org/project/scikit-learn/>
- PyPI. (16 de septiembre del 2024). “Statsmodels 0.14.3”. Recuperado el 14 de septiembre del 2024, de: <https://pypi.org/project/statsmodels/>
- PyPI. (20 de agosto del 2024). “Scipy 1.14.1”. Recuperado el 21 de septiembre del 2024, de: <https://pypi.org/project/scipy/>
- PyPI. (24 de agosto del 2024). “Yfinance 0.2.43”. Recuperado el 07 de septiembre del 2024, de: <https://pypi.org/project/yfinance/>
- Sanz, Francisco. (2024). “Cómo funciona el algoritmo XGBoost en Python”. Recuperado el 21 de septiembre del 2024, de: <https://www.themachinelearners.com/xgboost-python/>
- Tamplin, T. (05 de julio del 2023). “Omega Ratio | Definition, Components, Advantages & Limitations. Finance Strategists”. Recuperado el 11 de octubre del 2024, de: <https://www.financestrategists.com/wealthmanagement/financial-ratios/omega-ratio/>
- Yahoo! Finance. (2024). “S&P 500 (^GSPC)”. Recuperado el 18 de agosto del 2024, de: <https://finance.yahoo.com/quote/%5EGSPC/>
- Zhou, X. (15 de enero del 2023). “Ratio Omega. Rankia”. Recuperado el 11 de octubre del 2024, de: <https://www.rankia.com/diccionario/fondos-inversion/ratio-omega>