



Maestría en

SISTEMAS DE INFORMACIÓN



Mención **Inteligencia de Negocios y Analítica de Datos Masivos.**

Tesis previa a la obtención del título de Magíster en Sistemas de Información mención Inteligencia de Negocios y Analítica de Datos Masivos.

AUTORES:

Jimmy Vladimir Tarira Pérez

Rubén Marcelo Recalde Baquero

Marlon Andrés Cumbal Cumbal

Sebastián Vaca Alarcón

TUTOR:

Ms. Iván Galo Reyes Chacón

TEMA:

Evaluación de la eficacia de modelos de aprendizaje automático en la predicción de movimientos del mercado bursátil y generación de alertas para compra o ventas de acciones

APROBACIÓN DEL TUTOR

Yo, Iván Reyes, certifico que conozco los autores/as del presente trabajo siendo los responsables exclusivos tanto de su originalidad y autenticidad, como de su contenido.

A handwritten signature in blue ink, appearing to read 'Iván Reyes', positioned above a horizontal line.

Ing. Iván Reyes Mgtr.
DIRECTOR DE TESIS

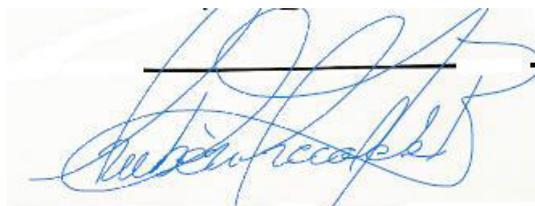
Certificación de autoría

Nosotros, **Jimmy Vladimir Tarira Pérez, Rubén Marcelo Recalde Baquero, Marlon Andrés Cumbal Cumbal y Sebastián Vaca Alarcón**, declaramos bajo juramento que el trabajo aquí descrito es de nuestra autoría; que no ha sido presentado anteriormente para ningún grado o calificación profesional y que se ha consultado la bibliografía detallada.

Cedemos nuestros derechos de propiedad intelectual a la Universidad Internacional del Ecuador (UIDE), para que sea publicado y divulgado en internet, según lo establecido en la Ley de Propiedad Intelectual, su reglamento y demás disposiciones legales.



Firma del graduando
Jimmy Vladimir Tarira Pérez



Firma del graduando
Rubén Marcelo Recalde Baquero



Firma del graduando
Marlon Andrés Cumbal Cumbal



Firma del graduando
Sebastián Vaca Alarcón

Autorización de Derechos de Propiedad Intelectual

Nosotros, **Jimmy Vladimir Tarira Pérez, Rubén Marcelo Recalde Baquero, Marlon Andrés Cumbal Cumbal y Sebastián Vaca Alarcón**, en calidad de autores del trabajo de investigación titulado ***Evaluación de la eficacia de modelos de aprendizaje automático en la predicción de movimientos del mercado bursátil y generación de alertas para compra o ventas de acciones***, autorizamos a la Universidad Internacional del Ecuador (UIDE) para hacer uso de todos los contenidos que nos pertenecen o de parte de los que contiene esta obra, con fines estrictamente académicos o de investigación. Los derechos que como autores nos corresponden, lo establecido en los artículos 5, 6, 8, 19 y demás pertinentes de la Ley de Propiedad Intelectual y su Reglamento en Ecuador.

D. M. Quito, (septiembre 2024)



Firma del graduando
Jimmy Vladimir Tarira Pérez



Firma del graduando
Rubén Marcelo Recalde Baquero



Firma del graduando
Marlon Andrés Cumbal Cumbal



Firma del graduando
Sebastián Vaca Alarcón

DEDICATORIA

En primer lugar, se lo dedico a Dios, quién me ha dado las fuerzas para seguir adelante. Se lo dedico a mi esposa y mi hija porque han estado a mi lado motivándome constantemente. A mis padres a quién les debo la educación escolar, secundaria y universitaria, que sin esos pasos no hubiera podido estar aquí.

Todos ellos han sido un pilar fundamental en este tiempo para sostenerme y poder culminar con la meta trazada.

Jimmy Tarira Pérez

AGRADECIMIENTOS

Agradecemos a todos los profesionales que nos impartieron sus conocimientos y experiencias durante todo este tiempo en que cursamos la maestría.

Agradecemos a nuestros tutores por habernos direccionado para poder lograr este proyecto bajo los lineamientos necesarios.

Jimmy Vladimir Tarira Pérez

Rubén Marcelo Recalde Baquero

Marlon Andrés Cumbal Cumbal

Sebastián Vaca Alarcón

RESUMEN

La alta incertidumbre y el riesgo en la toma de decisiones de inversión en el mercado bursátil es el problema principal al que tratamos de ayudar con el presente proyecto. Los inversionistas a menudo se enfrentan a esto debido a la complejidad y volatilidad del mercado. Para mitigar estos riesgos, el proyecto ofrece señales de compra o venta basadas en predicciones generadas por modelos avanzados de Machine Learning. Utilizando Redes Neuronales Recurrentes como LSTM y GRU, entrenadas con hasta 10 años de datos históricos, el sistema proporciona recomendaciones fundamentadas sobre las mejores oportunidades de inversión y los momentos óptimos para ejecutar transacciones.

Al integrar estos modelos de Machine Learning, el proyecto permite una evaluación más precisa de los patrones y tendencias del mercado, reduciendo la incertidumbre inherente en las decisiones de inversión. Esto no solo ayuda a los inversionistas a evitar decisiones impulsivas o mal informadas, sino que también mejora la capacidad de anticipar movimientos adversos del mercado, disminuyendo así el riesgo general y aumentando las posibilidades de obtener ganancias consistentes.

Palabras Claves:

Bolsa de Valores, acciones, Machine Learning, Red Neuronal, predicción

ABSTRACT

The high uncertainty and risk in making investment decisions in the stock market is the main problem that we try to help with this project. Investors often face this due to the complexity and volatility of the market. To mitigate these risks, the project offers buy or sell signals based on predictions generated by advanced Machine Learning models. Using Recurrent Neural Networks as LSTM y GRU, trained with up to 10 years of historical data, the system provides accurate and informed recommendations on the best investment opportunities and optimal times to execute transactions.

By integrating these models of Machine Learning, the project allows for a more accurate assessment of market patterns and trends, reducing the uncertainty inherent in investment decisions. This not only helps investors avoid impulsive or ill-informed decisions, but also improves the ability to anticipate adverse market movements, thereby decreasing overall risk and increasing the chances of making consistent profits.

Keywords:

Stock Market, Shares, Machine Learning, Neural Network, Prediction

TABLA DE CONTENIDOS (Índice)

Resumen	ix
Abstract	xi
Capítulo I	1
Introducción	1
Naturaleza o tipo de proyecto	2
Objetivo General	2
Objetivos específicos	2
Justificación e importancia del trabajo de investigación	3
Capítulo II	4
Necesidades, stakeholders y áreas del negocio involucradas	4
Fuentes de Información	5
Arquitectura del modelo	5
KPIs	6
Regulación y protección de datos	7
Escalabilidad	7
Capítulo III	10
Análisis PESTEL	10
Metodología	18
Planificación de Recursos	21
Capítulo IV	23
Desarrollo	23
Modelado	26
Gráfico del modelo	28
Resultado	28
Capítulo V	31
Conclusiones generales	31
Conclusiones específicas	31
Limitaciones a la investigación	36

Recomendación	36
Bibliografía	38
Glosario y Acrónimos	40
Apéndice	42

LISTA DE TABLAS (Índice de tablas)

Tabla 1. Planificación de Recursos	22
Tabla 2. Lista de Empresas	23
Tabla 3. Data obtenida del api	25
Tabla 4. Resultados del parámetro “mse” de los modelos	29
Tabla 5. Rentabilidad por distintos períodos de tiempo	35

LISTA DE FIGURAS (Índice de figuras)

Figura 1. Arquitectura del modelo	5
Figura 2. Opciones de suscripción de servicios streaming en AWS	8
Figura 3. Cálculos de los servicios consultados en AWS	8
Figura 4. Arquitectura escalable a servicios en AWS	9
Figura 5. Consulta al api Yahoo Finance para train	24
Figura 6. Consulta al api Yahoo Finance para test	24
Figura 7. Consulta a la api polygon.io para train	25
Figura 8. Consulta a la api polygon.io para test	26
Figura 9. Normalización de datos y obtención de x_train y y_train	27
Figura 10. Obtención del mse	27
Figura 11. Gráfico del modelo	28
Figura 12. Gráfico de predicción de la Red Neuronal LSTM	29
Figura 13. Función para la obtención de señales	30
Figura 14. Gráfica histórica de señales de compra y venta	30
Figura 15. Función para obtener la rentabilidad	32
Figura 16. Gráfico de porcentaje acumulado de rentabilidad	32
Figura 17. Gráfico de porcentaje acumulado de rentabilidad anual	33
Figura 18. Gráfico de porcentaje acumulado de rentabilidad semestral	34
Figura 19. Gráfico de porcentaje acumulado de rentabilidad trimestral	34

CAPITULO 1

INTRODUCCION

1. PLANTEAMIENTO DEL PROBLEMA E IMPORTANCIA DEL ESTUDIO

1.1. Definición del proyecto

El análisis de patrones repetitivos en el comportamiento del mercado bursátil tradicionalmente ha requerido una extensa observación y acumulación de experiencia. A pesar de estos esfuerzos, la predicción de movimientos futuros con total seguridad sigue siendo una tarea compleja y arriesgada. La inteligencia artificial (IA) promete reducir significativamente este tiempo de aprendizaje. Pero ¿hasta qué punto es realmente efectiva en este contexto? Este es el interrogante que nuestro estudio busca responder.

Utilizando datos históricos de hasta 10 años proporcionados por Yahoo Finance y la api polygon.io disponibles libremente en internet, haremos un análisis detallado. Los datos obtenidos incluyen la fecha, precio de apertura, de cierre, precio más bajo, el más alto y volumen y estos datos de forma diaria. El objetivo principal es predecir el momento adecuado para la compra y la venta de las acciones analizadas.

Para este propósito, se aplicarán y evaluarán varios modelos de aprendizaje automático supervisado, entre ellos regresión lineal, bosques aleatorios, gradientes y redes neuronales.

El estudio comparará la precisión de estos modelos y recomendará el más adecuado para predecir movimientos futuros en el mercado bursátil, generando señales para momentos de compra y venta de acciones.

Este análisis es especialmente relevante para los traders principiantes, ya que les permitirá acelerar su curva de aprendizaje en el apasionante mundo del mercado bursátil.

Este proyecto no solo busca evaluar la precisión de diferentes modelos de IA, sino también proporcionar una herramienta práctica para aquellos que buscan incursionar en el trading con un soporte tecnológico avanzado.

1.2. Naturaleza o tipo de proyecto

Este estudio se centrará en el análisis de datos históricos durante un período de hasta diez años de empresas relevantes a nivel mundial para el entrenamiento de los modelos.

Se implementarán y compararán varios modelos de aprendizaje automático supervisado para predecir los movimientos del mercado bursátil. La investigación evaluará la precisión de cada modelo y propondrá el más adecuado para generar señales de compra y venta de acciones.

Aunque el enfoque principal es el análisis técnico a corto plazo, los resultados obtenidos también podrían tener implicaciones para estrategias de inversión a largo plazo.

Se harán comparativas con el histórico para ver la evolución. Se medirá la calidad de las actuaciones para minimizar catalogaciones erróneas.

1.3. Objetivos

1.3.1. Objetivo general

Evaluar la eficacia de diferentes modelos de aprendizaje automático en la predicción de movimientos del mercado bursátil para optimizar las estrategias de compra y venta de acciones.

1.3.2. Objetivos específicos

- Recolectar y procesar datos históricos de hasta diez años de empresas relevantes a nivel mundial para el entrenamiento de los modelos.
- Desarrollar y entrenar varios modelos de aprendizaje automático supervisado como

regresión lineal, bosques aleatorios, gradientes y redes neuronales, utilizando los datos recopilados para optimizar la precisión en la predicción de movimientos de mercado bursátil.

- Determinar el modelo de aprendizaje automático más eficaz para generar señales de compra y venta de acciones, mejorando así las decisiones de inversión.

1.4. Justificación e importancia del trabajo de investigación

En un mundo cada vez más globalizado y dinámico, el mercado bursátil presenta una de las oportunidades más atractivas y, al mismo tiempo, desafiantes para los inversionistas. La capacidad de predecir los movimientos bursátiles con precisión puede marcar la diferencia entre el éxito y el fracaso en el trading. Tradicionalmente, esta predicción ha dependido de la experiencia y el juicio humano, factores que pueden estar sujetos a sesgos y limitaciones. Aquí es donde la inteligencia artificial (IA) puede revolucionar el campo del trading ofreciendo más herramientas para una toma de decisión más informada, herramientas como:

- Generación de análisis más precisos.
- Automatización de las operaciones.
- Pronósticos más precisos.

Si se desea estar preparado para realizar un trabajo más efectivo, es importante que se conozca bien el funcionamiento de esta tecnología y cómo puede impactar el mundo de las inversiones.

CAPITULO 2

2. APLICACIONES ANALÍICAS

2.1. Necesidades, stakeholders y áreas del negocio involucradas

La predicción del precio de las acciones en la bolsa de valores es la búsqueda constante de los inversionistas. Esto es muy complejo debido a que esta variable no es lineal, sino que está sujeta a múltiples factores y condiciones como lo social, político, económico y muchos otros. Como concesión de los integrantes, el proyecto planteado es de interés personal. En este caso somos los involucrados quienes vamos a manejar todos los aspectos del proyecto desde la planificación hasta la ejecución y uso de la herramienta.

Con los avances que se ha tenido en el Machine Learning en su rama del Aprendizaje Automático y utilizando data histórica de Yahoo Finance y polygon.io; y con la presunción de que las futuras variaciones del mercado pueden determinarse a partir de examinar esos datos, se pretende desarrollar una herramienta que proporcione una sugerencia de comprar o de vender las acciones en la bolsa de valores.

Impacto de negocio

A la hora de tomar una decisión de invertir o no en la bolsa, mientras más herramientas a disposición se tenga es mucho mejor.

El valor otorgado de lo que se va a desarrollar se detalla a continuación:

- Reducción de riesgos en nuestras inversiones mediante decisiones informadas.
- Potencial aumento en la rentabilidad de nuestras inversiones.
- Adquisición de conocimientos valiosos en un campo de alta demanda.
- Mayor confianza en nuestras decisiones financieras.

2.2. Fuentes de información

La recopilación de los datos se lo hará a través de peticiones a la api *Polygon.io* y *Yahoo Finance*. Los datos obtenidos son semi-estructurados en formato JSON. Estos datos obtenidos se los importará en un archivo Python con las herramientas proporcionadas por la librería *Pandas*.

Esta información está disponible de forma gratuita y expuesta libremente para quiénes la quieran consultar, por tanto, no incurre en ningún incumplimiento de la ley de protección de datos.

2.3. Arquitectura del modelo

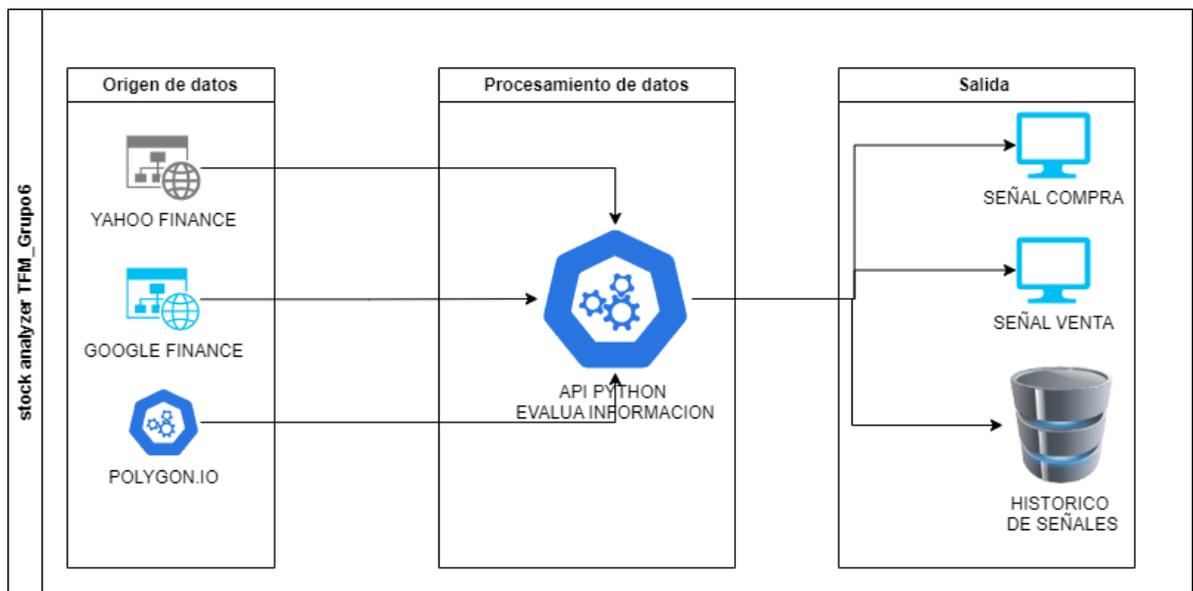


Figura 1. Arquitectura del modelo

Para el presente proyecto, haremos uso de la siguiente arquitectura:

Origen de los datos: estos serán obtenidos de varias fuentes de webs de finanzas y además de la aplicación Polygon.IO en su versión gratuita.

Procesamiento de datos: Estos serán procesados por medio de una API en Python, a ser realizada tentativamente en Visual Studio Code, en esta API los datos serán organizados y

transformados a un estándar que nos permita realizar un análisis por medio de series temporales.

SALIDA: Como salida tendremos la emisión de señales para el operador, que le indique si genera una compra o venta de acciones y adicionalmente se mostrará una gráfica histórica de las señales emitidas dentro de la línea de tiempo.

2.4. KPIs

2.4.1. Precisión de las Señales de Trading:

- Definición: Porcentaje de precisión.
- Fórmula: Error cuadrático medio * 100

2.4.2 Tasa de Retorno Anualizada:

- Definición: Rentabilidad anualizada de la cartera de inversiones basada en las señales generadas.

2.4.3. Cobertura de Fuentes de Datos:

- Definición: Número de fuentes de datos financieras integradas y utilizadas para generar señales.
- Fórmula: Número de fuentes de datos integradas.
- Objetivo: Integrar al menos 2 fuentes de datos relevantes (por ejemplo, Yahoo Finance, Google Finance, Polygon.io, etc.).

2.4.4. Actualización y Frecuencia de Señales:

- Objetivo: Generar y actualizar señales al menos una vez al día.

2.5. Regulación y protección de datos

Aunque aparentemente la información obtenida para el funcionamiento de la aplicación no cae dentro de las leyes de protección de datos, los datos generados si lo hacen y deben de ser guardados y respaldados ya que un operador para cubrir errores puede querer tergiversar las señales obtenidas y la base de datos histórica de recomendaciones generadas será un respaldo para el aplicativo.

Para dicho fin se realizará el entrenamiento siempre con consulta directa a las apis, y de esta forma se minimiza la manipulación de los datos.

2.6. Escalabilidad

En un principio hemos considerado:

- Obtención de datos por medio del API polygon.io
- Obtención de datos por medio de diferentes fuentes libres en línea como lo son Yahoo Finance y Google Finance.

Pero se puede proyectar a ampliar la data. Una de las formas en que la data de nuestro proyecto puede crecer es por medio de la adquisición de nuevas fuentes de información que, tras un análisis exploratorio, demuestren ser un aporte a nuestro proyecto.

Ante tal caso que la ingesta de datos en tiempo real llegase a necesitar de mayores recursos informáticos, hemos considerado suscribirnos a los servicios en la nube de AWS, debido a la valoración de costo / beneficio, ya que después de consulta de precios en diferentes proveedores, nos ofrece un precio accesible frente al respaldo en infraestructura con la que cuenta.

Los costos consultados a la fecha de los servicios requeridos en AWS suman \$38.64 mensuales. Lo considerado para el cálculo de estos costos se muestra en las siguientes capturas de pantalla hechas a la calculadora de la plataforma:

Payment options

Estimated commitment price based on the following selections:
Instance type: **t3.medium** Operating system: **Linux**

Select the container and options to find your best price

<input checked="" type="radio"/> Compute Savings Plans One plan that automatically applies to all usage on EC2, Fargate, and Lambda. Up to 66% discount. Learn more Reservation term <input type="radio"/> 1 year <input checked="" type="radio"/> 3 year Payment Options <input checked="" type="radio"/> No upfront <input type="radio"/> Partial upfront <input type="radio"/> All upfront <hr/> <p style="text-align: center;">Upfront: 0.00 Monthly: 30.66/Month</p>	<input type="radio"/> EC2 Instance Savings Plans Get deeper discount when you only need one instance family and region. Up to 72% discount. Learn more Reservation term <input type="radio"/> 1 year <input checked="" type="radio"/> 3 year Payment Options <input checked="" type="radio"/> No upfront <input type="radio"/> Partial upfront <input type="radio"/> All upfront <hr/> <p style="text-align: center;">Upfront: 0.00 Monthly: 26.50/Month</p>	<input type="radio"/> On-Demand Maximize flexibility. Learn more Expected utilization Enter the expected usage of Amazon EC2 Instances Usage <input type="text" value="100"/> Usage type <input type="text" value="Utilization percent per month"/> <hr/> <p style="text-align: center;">Instance: 0.084/Hour Monthly: 61.32/Month</p>
---	--	--

Figura 2. Opciones de suscripción de servicios streaming en AWS

▼ Show calculations

[Break-even analysis](#)

A cost-optimized strategy for your utilization is found by calculating the breakeven point when Compute Savings Plans instances are more cost effective to use than On-Demand Instances.

Compute Savings Plans rate for t3.medium in the Peru (Lima) for 3 Year term and No Upfront is 0.042 USD
 Hours in the commitment: 365 days * 24 hours * 3 year = 26280.0000 hours
 Total Commitment: 0.042 USD * 26280 hours = 1103.7600 USD
 Upfront: No Upfront (0% of 1103.76) = 0.0000 USD
 Hourly cost for Compute Savings Plans = (Total Commitment - Upfront cost)/Hours in the term: (1103.76 - 0.00)/26280 = 0.0420 USD
 Normalized Compute Savings Plans monthly price: (0.000000 USD / 36 months) + (0.042000 USD x 730 hours in a month) = 30.660000 USD
 On-Demand hourly price: 0.084000 USD
 Normalized On-Demand monthly price: 0.084000 USD x 730 hours in a month = 61.320000 USD
 Breakeven percentage: 30.660000 USD / 61.320000 USD = 0.5
 Breakeven point: 0.5 x 730 hours in month = 365.000000 hours

[Utilization summary](#)

For instance utilization over the breakeven point, 365.000000 hours, it is more cost effective to choose Compute Savings Plans instances than On-Demand Instances.

1 Compute Savings Plans instances x 0.000000 upfront cost = 0.000000 USD

Compute Savings Plans instances (upfront): 0.000000 USD

1 instances x 730 hours in a month = 730 Compute Savings Plans instance hours per month
 730 Compute Savings Plans instance hours per month x 0.042000 USD = 30.660000 USD

Normalized Compute Savings Plans instances (monthly): 30.660000 USD

0 On-Demand instance hours per month x 0.084000 USD = 0.000000 USD

On-Demand (monthly): 0.000000 USD

0.000000 USD On-Demand (monthly) + 30.660000 USD Normalized Compute Savings Plans instances (monthly) = 30.660000 USD

Total cost (monthly): 30.660000 USD

**Please note that you will pay an hourly commitment for Savings Plans and your usage will be accrued at a discounted rate against this commitment.*

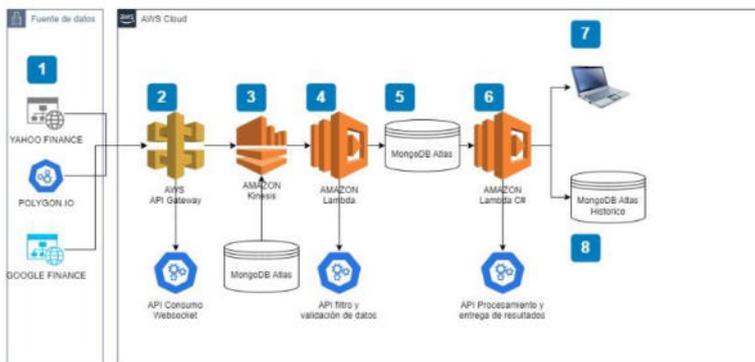
Figura 3. Cálculos de los servicios consultados en AWS

Nuestro proyecto en efecto tiene un componente de análisis de datos en tiempo real que por el momento está limitado a lo que nos ofrece de forma gratuita el api *polygon.io* y *Yahoo Finance*, por lo tanto, el costo inicial es cero, pero si en un futuro se hace necesario adquirir alguna suscripción, a la fecha sería de un costo de \$26 mensuales.

La escalabilidad de la infraestructura de nuestro proyecto migrado a los servicios de AWS, quedaría de la siguiente forma:

Arquitectura en la nube V2

Diagrama de servicios



Ventajas

SIN SERVIDOR: No se necesita manejo de los servidores, solo se paga por consumo
Procesamiento en tiempo real: Manejo de los datos en tiempo real con API Gateway y Kinesis, los cuales son más eficientes que Hadoop por ejemplo Servicios Administrados. Ninguno de los componentes de la arquitectura propuesta requiere de administración directa, lo cual economiza en recursos humanos y económicos
Escalabilidad: Los servicios de AWS son escalables de manera sencilla
Integración: Todos los servicios pertenecen al mismo ecosistema, lo cual facilita enormemente su administración

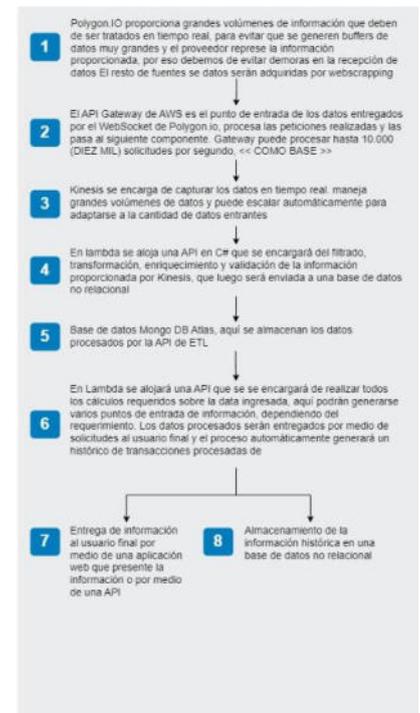


Figura 4. Arquitectura escalable a servicios en AWS

CAPITULO 3

3. INTELIGENCIA DE NEGOCIO

3.1. Análisis PESTEL

3.1.1. Político

A continuación, se detallan los principales elementos políticos que pueden influir en el proyecto.

En 2024, el presidente Daniel Noboa convocó una consulta popular y referéndum ¹para abordar temas de seguridad, institucionalidad y economía. Estos eventos, programados para el pasado marzo de 2024, buscaron, entre otras cosas, reformar la Constitución para permitir que las Fuerzas Armadas apoyen a la Policía Nacional en la lucha contra el crimen organizado. La implementación de estas consultas y referendos podría generar cambios en la política económica que impacten la confianza de los inversionistas y la estabilidad del mercado bursátil.

Además, la crisis política y social en Ecuador, caracterizada por un aumento en la delincuencia y la violencia, ha afectado la percepción de estabilidad en el país. El gobierno ha decretado varios estados de excepción y movilizó a las fuerzas armadas para enfrentar la inseguridad, lo que ha generado preocupación entre los inversionistas sobre la estabilidad política a largo plazo. De acuerdo al sitio web de noticias primicias.ec, entre febrero y junio de 2024 las muertes violentas subieron un 61%. Las tensiones internas y las luchas de poder entre diferentes facciones políticas complican aún más el panorama, aumentando la volatilidad en los mercados financiero.

¹ (Wikipedia, 2024)

La agenda política del país también incluye la preparación para las elecciones generales de 2025, con procesos internos en las organizaciones políticas para definir candidatos y cumplir con los requisitos del Consejo Nacional Electoral (CNE). La capacidad de anticipar y adaptarse a los cambios políticos y geopolíticos, especialmente en contextos de incertidumbre electoral y tensiones diplomáticas, es crucial para mantener la relevancia y eficacia del sistema en un entorno global dinámico.

Las medidas económicas adoptadas por el Gobierno, tales como la disminución de los subsidios y la optimización del uso de fondos públicos, podrían impactar de manera importante la percepción de los inversionistas. Estas acciones podrían ser bien recibidas si se interpretan como avances hacia una gestión más eficaz y responsable de las finanzas.

Asimismo, medidas que fomenten el desarrollo económico, tales como estímulos tributarios para inversores recientes y respaldo a sectores clave, pueden crear un entorno más atractivo para los inversionistas, lo que su vez potenciará el rendimiento del mercado bursátil.

Implementar estas reformas de manera eficaz puede mejorar la estabilidad económica, impulsar la inversión y tener un efecto beneficioso en el ámbito financiero.

Por otro lado, Ecuador ha firmado varios tratados comerciales, como el acuerdo comercial con la Unión Europea ² y el acuerdo de comercio con Estados Unidos. Estos tratados facilitan el comercio y la inversión, abriendo nuevas oportunidades para las empresas ecuatorianas. La apertura de mercados internacionales puede aumentar la demanda de acciones de empresas ecuatorianas, mejorando la liquidez y el rendimiento del mercado bursátil.

3.1.2. Económico

En 2024, la economía ecuatoriana enfrenta una serie de dificultades, como un crecimiento económico mínimo, un déficit fiscal significativo, y una caída en los ingresos petroleros.

² (Ecuador, 2021) (Ecuador U. E., 2020)

Además, la situación fiscal se complica debido a la deuda del sector público, mientras que las inversiones, tanto públicas como privadas, muestran signos de estancamiento. Sumado a esto, la inseguridad jurídica y la falta de incentivos fiscales, como la eliminación del impuesto a la salida de divisas, limitan el atractivo del país para los inversionistas internacionales.³

En la página del Banco Central del Ecuador encontramos que en el primer trimestre de 2024 el Producto Interno Bruto del Ecuador creció 1.2% comparado con el mismo período del año 2023. El crecimiento del PIB en Ecuador ha sido variable, influenciado por factores como los precios del petróleo y las políticas gubernamentales. Un crecimiento sostenido puede aumentar la confianza de los inversores y estimular el mercado bursátil. La economía del Ecuador se ha enfrentado a retos importantes, como la caída en los precios del petróleo, uno de sus principales productos de exportación, en mayo de 2024 se ubicó en \$69.7 por barril y para junio se ubicó en \$65.1, después de cinco meses continuos de ascenso que había tenido, como lo menciona el sitio web primicias.ec. Esto crea la necesidad de implementar políticas fiscales más estrictas para gestionar la deuda pública. Estas medidas son cruciales para mantener la confianza de los inversionistas y minimizar la volatilidad del mercado, evitando que estos factores adversos impacten negativamente en la estabilidad económica del país.⁴

En Ecuador, la adopción del dólar como moneda oficial significa que las estrategias financieras están íntimamente ligadas a las determinaciones de la Reserva Federal estadounidense. Esto implica que las variaciones en las tasas de interés de Estados Unidos pueden influir directamente en la economía del Ecuador. Por esta razón, es fundamental que el sistema de pronóstico pueda fusionar esta información y adaptar las tácticas de inversión en consecuencia.⁵

³ (Group, 2024)

⁴ (economics, 2023)

⁵ (Affairs, 2016)

La inflación es un factor económico crucial a considerar en el análisis PESTEL. Un aumento en la tasa de inflación puede debilitar el poder adquisitivo de los consumidores, reduciendo así las compras y afectando negativamente los ingresos y beneficios de las empresas, lo que a su vez podría resultar en una caída en los precios de las acciones. Durante los años 2023 y 2024, la tasa de inflación del Ecuador ha experimentado variaciones, siendo impactada por diversos factores mencionados. De acuerdo con el Banco Central del Ecuador, la tasa de inflación anual ha variado entre un margen de 2% y 4%. Actualmente la tasa de inflación se encuentra en un nivel bastante moderado, según el banco central a junio de 2024 la inflación acumulada alcanzó el 0.70%, lo cual indica que se ha llevado a cabo una gestión cuidadosa de las políticas monetarias y fiscales. ⁶

3.1.3. Social

En el ámbito del factor social, los ecuatorianos han mostrado históricamente una preferencia por invertir en bienes raíces y negocios familiares. Cambiar esta preferencia hacia el mercado bursátil requiere una adecuada educación y la creación de una cultura de inversión.

Los ecuatorianos tenemos una baja educación financiera con respecto a la región como nos indica un artículo del diario “la hora”, el nivel de educación financiera es hasta 2 veces más bajo que la media regional; y llega hasta 3 veces si se analiza el promedio mundial. En un estudio de la Organización para la Cooperación y el Desarrollo Económico (OCDE), el promedio de puntuación a una prueba simple de conocimientos financieros la puntuación en Ecuador no supera el 7 sobre 21. Las mujeres sacan un mejor resultado que los hombres (8 sobre 21); pero son más relegadas en la llamada inclusión financiera.

En este escenario, la sociedad ecuatoriana no solo tiene grandes vacíos en educación financiera, sino también profundos problemas de acceso que, entre otras cosas, hacen que los

⁶ (Economics, 2024)

ciudadanos sean más propensos a caer en estafas, acudir al chulco, sobre endeudarse, no tener ahorros para imprevistos y la vez, entre muchos otros.

Por ejemplo, a pesar que el interés del chulco puede llegar a más de 1.200%, el 65% de los que acuden a este financiamiento ilegal creen que están pagando menos que en un banco o cooperativa.

A partir del año escolar 2024-2025, Ecuador implementará la educación financiera como una asignatura obligatoria para todos los estudiantes, desde la etapa preescolar hasta la secundaria. Esta modificación en el plan de estudios, impulsada por el Ministerio de Educación, busca potenciar las habilidades de los alumnos en gestión financiera, organización de ahorros, acceso a préstamos e inversiones. La implementación de esta asignatura brindará a las próximas generaciones la oportunidad de desarrollar destrezas fundamentales para la gestión financiera, fomentando una comprensión más profunda y la incorporación de herramientas sofisticadas de evaluación en el mercado de valores.⁷

En paralelo el crecimiento de la clase media está generando un mayor interés en productos financieros sofisticados, lo cual podría favorecer la implementación de sistemas de pronóstico en el mercado de valores. Por lo tanto, una generación joven que cuenta con una educación financiera y digital más sólida muestra mayor disposición para emplear tecnologías innovadoras en la administración de sus inversiones.

La mentalidad de invertir está cambiando, poniendo más énfasis en la transparencia y en garantizar la seguridad de quienes invierten. La adopción de sistemas de predicción bursátil depende en gran medida de la fe depositada en los mercados financieros y en las tecnologías de vanguardia.

⁷ (Educacion, 2023)

Promover la educación financiera y la innovación tecnológica puede fortalecer la confianza, haciendo más sencillo incorporar herramientas de inteligencia artificial en la toma de decisiones relacionadas con las finanzas.

3.1.4. Tecnológico

El marco tecnológico de la Bolsa de Valores en Ecuador está siendo influenciado por varios desarrollos recientes:

Ecuador ha mostrado un avance significativo en su transformación digital y tecnológica. Un estudio reciente revela que el 82% de las empresas en el país planean iniciar o continuar su transformación digital en 2024, aunque el 64% enfrenta desafíos debido a la falta de visión estratégica y capacitación interna adecuada. Esto presenta tanto oportunidades como retos para la implementación de tecnologías avanzadas. La inteligencia artificial (IA) es especialmente prometedora, con empresas ecuatorianas adoptando soluciones de IA en sectores como banca, finanzas, retail y telecomunicaciones para mejorar el servicio al cliente, automatizar procesos y fortalecer la ciberseguridad.⁸

Conjuntamente, el comercio electrónico en Ecuador también está en auge, con proyecciones que indican que alcanzará los 3,19 billones de dólares en transacciones para 2024. Este crecimiento es impulsado por la adopción de tecnologías digitales avanzadas y estrategias de marketing digital. El gobierno ha lanzado iniciativas para fomentar la ciencia, la tecnología y la innovación, como el programa "Idearium", que financia proyectos para resolver necesidades sociales, ambientales y económicas. Además, la política para la transformación digital del Ecuador (2022-2025) incluye componentes clave como la ciberseguridad, la computación en la nube y la infraestructura digital, con un enfoque en la ciber resiliencia y la automatización de la ciberdefensa. También incluye inversiones en tecnología como inteligencia artificial, big

⁸ (Global, 2024)

data y blockchain. ⁹Estos avances están destinados a modernizar y hacer más eficiente el mercado de valores, aumentando la seguridad y la transparencia en las transacciones.

En este contexto, se están utilizando tecnologías modernas para mejorar la infraestructura de las bolsas de valores de Quito y Guayaquil. Esto incluye la digitalización de procesos y la utilización de plataformas en línea para facilitar el comercio y la inversión, mejorando así la accesibilidad para los inversores nacionales e internacionales. Asimismo, el fortalecimiento de la infraestructura tecnológica también busca aumentar la confianza de los inversionistas mediante la implementación de mejores sistemas de supervisión y regulación, esto incluye mejorar las calificaciones de riesgo y aumentar la transparencia de las operaciones en el mercado.

Estos esfuerzos son parte de un objetivo más amplio de Ecuador para integrarse con otros mercados financieros en la región, como el Mercado Integrado Latinoamericano (MILA), y mejorar la participación de inversionistas extranjeros en el país.

3.1.5. Ecológico

El marco ecológico en la Bolsa de Valores de Ecuador está empezando a cobrar relevancia a través de iniciativas de finanzas sostenibles. Las bolsas de valores ecuatorianas están trabajando para posicionarse como líderes en el financiamiento verde, impulsando proyectos de inversión que cumplen con criterios ambientales, sociales y de gobernanza (ESG). Se están promoviendo bonos verdes y otros instrumentos financieros para atraer inversiones que apoyen proyectos ecológicamente responsables. ¹⁰

⁹ (Ecdb, 2024)

¹⁰ (Initiative, 2019)

Además, la implementación de una taxonomía verde en el sistema financiero del país busca incentivar proyectos que contribuyan a la sostenibilidad y al desarrollo de infraestructura ecológica.

3.1.6. Legal

La Ley Orgánica de Protección de Datos Personales establece requisitos estrictos para la recopilación, almacenamiento y procesamiento de datos personales. Las empresas deben cumplir con esta normativa para proteger la información de los usuarios y evitar sanciones legales. Esto es particularmente relevante para proyectos que manejan grandes volúmenes de datos, como los sistemas de predicción de mercado basados en aprendizaje automático.¹¹

Además, las empresas deben adoptar medidas de seguridad cibernética para proteger la información de los inversores y cumplir con las regulaciones de protección de datos. La Superintendencia de Compañías, Valores y Seguros, entidad encargada de la regulación y supervisión del mercado de valores en Ecuador, supervisa a las entidades emisoras, intermediarios y demás participantes del mercado.¹²

Asimismo, Ecuador también está influenciado por normas y prácticas internacionales en materia de mercados financieros, como las recomendaciones de la Organización Internacional de Comisiones de Valores (IOSCO), que promueven estándares globales para la regulación de mercados financieros. Estos estándares ayudan a garantizar la integridad y transparencia del mercado, proporcionando una base sólida para la confianza de los inversores.

Estas recomendaciones se las puede obtener en su totalidad en su sitio web iosco.org. Estos principios enmarcan la regulación, supervisión, sanciones entre otros, hacia los emisores, agencias y proveedores.

¹¹ (Univero, 2021)

¹² (legal, s.f.)

3.2. Metodología

Es conveniente el uso de una metodología AGILE para el presente proyecto ya que contempla etapas de desarrollo e interoperabilidad entre entidades del proyecto que se monitorean de mejor manera con uso de SCRUM, proponemos el siguiente planteamiento para el proyecto:

3.2.1. Necesidades y objetivos

Objetivo del proyecto: Generar señales de compra/venta de acciones en el mercado bursátil a partir del análisis de datos históricos.

3.2.2. Roles en Scrum

Product Owner: Define las características del producto y prioriza el backlog.

Scrum Master: Facilita las reuniones y asegura que el equipo sigue los principios de Scrum.

Equipo de Desarrollo: Desarrolla las funcionalidades del proyecto, compuesto por desarrolladores, analistas de datos y expertos en finanzas.

3.2.3. Backlog del Producto

Epics y User Stories

Epic 1: Recopilar datos

Historia 1.1: Como desarrollador, quiero obtener datos de la API de Polygon.io para tener información bursátil histórica.

Historia 1.2: Como desarrollador, quiero realizar web scraping de páginas financieras para complementar los datos de Polygon.io.

Epic 2: Tratamiento de datos

Historia 2.1: Como analista de datos, quiero uniformar los datos a un formato JSON para facilitar su tratamiento.

Historia 2.2: Como analista de datos, quiero almacenar los datos en un repositorio común para su posterior análisis.

Epic 3: Modelos y algoritmos

Historia 3.1: Como desarrollador, quiero implementar un modelo de series temporales para analizar las tendencias del mercado.

Historia 3.2: Como desarrollador, quiero implementar un modelo de choques de MACD para identificar oportunidades de compra/venta.

Epic 4: Evaluación del modelo

Historia 4.1: Como analista de datos, quiero evaluar el modelo usando MSE para medir su precisión.

Historia 4.2: Como analista de datos, quiero comparar diferentes modelos para seleccionar el más efectivo.

Epic 5: Presentación de resultados

Historia 5.1: Como usuario, quiero recibir señales de compra/venta de acciones basadas en el análisis de datos.

Historia 5.2: Como usuario, quiero ver una gráfica de las señales generadas para entender mejor las recomendaciones.

Epic 6: Despliegue

Historia 6.1: Como desarrollador, quiero desplegar la aplicación en un entorno de desarrollo para pruebas iniciales.

Historia 6.2: Como desarrollador, quiero desplegar la aplicación en un entorno de preproducción para pruebas más avanzadas.

Historia 6.3: Como desarrollador, quiero desplegar la aplicación en un entorno de producción para su uso real.

Epic 7: Puesta en valor

Historia 7.1: Como usuario, quiero que las señales de compra/venta se generen con la data más reciente posible.

Historia 7.2: Como usuario, quiero tomar decisiones informadas basadas en las señales generadas por el modelo.

Epic 8: Seguimiento

Historia 8.1: Como analista de datos, quiero evaluar el performance de los modelos para mejorar su precisión.

Historia 8.2: Como analista de datos, quiero refinar la aplicación en base a las ganancias o pérdidas generadas por las señales.

3.2.4. Sprints y Reuniones Scrum

Sprint Planning: Definir las historias de usuario que se completarán en el próximo sprint.

Daily Standup: Reuniones diarias para revisar el progreso y los impedimentos.

Sprint Review: Revisar el trabajo completado al final del sprint y obtener feedback.

Sprint Retrospective: Evaluar el sprint pasado para identificar mejoras en el proceso.

3.2.5. Incrementos y Entregables

Sprint 1: Configuración inicial del entorno de desarrollo y obtención de datos de Polygon.io.

Sprint 2: Implementación de consulta a Yahoo Finance y uniformización de datos.

Sprint 3: Desarrollo de los modelos de series temporales y MACD.

Sprint 4: Evaluación y comparación de modelos.

Sprint 5: Generación de señales de compra/venta y creación de informes.

Sprint 6: Despliegue en entornos de desarrollo, preproducción y producción.

Sprint 7: Seguimiento y refinamiento del modelo basado en el performance.

3.2.6. Herramientas y Tecnologías

Lenguaje de Programación: Python

API de Datos: Polygon.io y Yahoo Finance

Evaluación de Modelos: Métrica MSE

Plataforma de Despliegue: Entornos de desarrollo, preproducción y producción

3.3. Planificación de Recursos

A continuación, se presenta la tabla de planificación de recursos donde se especifican las tareas y la duración en horas requerida para completarlas.

PLANIFICACIÓN DE RECURSOS	Estimación	Ingeniero Datos	Analista Datos	Gobierno Datos	DevOps
	<i>Total Horas</i>	<i>Horas</i>	<i>Horas</i>	<i>Horas</i>	<i>Horas</i>
DISEÑO TÉCNICO	50				
Estrategia de ingesta	12	12	-	-	-
Modelado	8	-	8	-	-
Seguridad	6	-	-	6	-
Calidad	6	6	-	-	-
KPIs	8	-	8	-	-
Visualización	10	-	-	-	10
IMPLEMENTACIÓN	65				
Carga de datos	16	16	-	-	-
Procesamiento - Enriquecimiento	12	-	12	-	-
Modelado de datos	10	-	10	-	-
KPIs	15	-	-	-	15
Visualización	12	-	12	-	-
IMPLANTACIÓN	110				
Pruebas	40	10	10	10	10
Documentación	40	10	10	10	10
Formación	30	-	15	-	15
TOTAL	225				

Tabla 1. Planificación de recursos

CAPITULO 4

4. APRENDIZAJE AUTOMÁTICO

4.1. Desarrollo

El conjunto de datos inicial sobre el cuál se va a basar la predicción es tomado de la api de *yahoo finance* de un total de 10 años histórico desde el día en que se hace la consulta hacia atrás y 5 años de la api *polygon.io* que es lo que permite de forma gratuita. Se ha elegido para el entrenamiento de los modelos, los datos históricos de algunas de las empresas que son más influyentes en la bolsa de valores de New York, entre las cuales tenemos:

Empresa	Ticker
NVIDIA	NVDA
TESLA	TSLA
MICROSOFT	MSFT
AMAZON	AMZN
INTEL	INTC
AMD	AMD
JOHNSON Y JOHNSON	JNJ
ALI BABA	BABA
GOOGLE	GOOGL
QUALCOMM	QCOM
APPLE	AAPL
META	META
EQUIFAX	EFX
AMERICAN AIRLINES	AAL
DELL	DELL

Tabla 2. Lista de Empresas

Se ha elegido como editor de código Visual Studio Code que es una herramienta gratuita y multiplataforma.

Se ha definido que se considerará un rango de 8 años para el conjunto de entrenamiento y 2 años para el conjunto de validación.

La consulta de la data histórica al api *Yahoo Finance* de la empresa TESLA para obtener datos para el conjunto de entrenamiento, queda de la siguiente manera:

```
company = 'TSLA'  
ticker = yf.Ticker(company)  
df = ticker.history(start = '2014-1-1', end='2022-12-31')  
  
df
```

Figura 5. Consulta al api Yahoo Finance para train

Para el conjunto de validación

```
#Cargar los datos del test  
hist_test = ticker.history(start = '2023-1-1', end='2024-9-25')  
actual_prices = hist_test["close"].values  
  
total_dataset = pd.concat((df['close'], hist_test['close']), axis=0)  
model_inputs = total_dataset[len(total_dataset)-len(hist_test)-prediction_days:].values  
model_inputs = scaler.transform(model_inputs.reshape(-1,1))
```

Figura 6. Consulta al api Yahoo Finance para test

La data obtenida después de la consulta viene de la siguiente manera:

Date	Open	High	Low	Close	Adj Close	Volume
2012-01-03	26.549999	26.959999	26.389999	26.770000	21.200518	64731500
2012-01-04	26.820000	27.469999	26.780001	27.400000	21.699440	80516100
2012-01-05	27.379999	27.730000	27.290001	27.680000	21.921188	56081400
2012-01-06	27.530001	28.190001	27.530001	28.110001	22.261728	99455500
2012-01-09	28.049999	28.100000	27.719999	27.740000	21.968699	59706800
...
2024-07-15	453.299988	457.260010	451.429993	453.959991	453.959991	14429400
2024-07-16	454.220001	454.299988	446.660004	449.519989	449.519989	17175700
2024-07-17	442.589996	444.850006	439.179993	443.519989	443.519989	21778000
2024-07-18	444.339996	444.649994	434.399994	440.369995	440.369995	20794800
2024-07-19	433.100006	441.140015	432.000000	437.109985	437.109985	20862400

Tabla 3. Data obtenida del api.

En una segunda fase se obtiene datos de la api *polygon.io*. Esta api permite la obtención de datos de hasta 5 años de forma gratuita.

A continuación, el código empleado para la obtención de la data para el entrenamiento:

```
# Key para API de Polygon.io
API_KEY = 'QHq8TBx82bQk6U5Lk2btKpESp6GmfHIS'

client = RESTClient(API_KEY)

for a in client.list_aggs(ticker=tickers[i], multiplier=1, timespan="day", from_="2019-09-24", to="2022-09-23", limit=50000):
    lista_polygon.append(a)

fechas = []
cierres = []

for i in range(len(lista_polygon)):
    fecha = pd.to_datetime(lista_polygon[i].timestamp, unit="ms")
    fechas.append(fecha)
    cierres.append(lista_polygon[i].close)

df = pd.DataFrame()
df["Date"] = fechas
df["Close"] = cierres

#df.set_index('Date', inplace=True) # Usar 'Date' como índice

df
```

Figura 7. Consulta a la api *polygon.io* para train

Para el conjunto de validación

```

from polygon import RESTClient

# Key para API de Polygon.io
API_KEY = 'iYrWjtstbjV7E6w6Gd0B1hFIY6EZ814p'

# Lista de tickers (limitado para simplificar pruebas)
tickers = ['AAPL', 'MSFT', 'NVDA', 'TSLA']
i = 0

# Listas para almacenar los datos de Polygon.io y yfinance
lista_polygon = []

client = RESTClient(API_KEY)

for a in client.list_aggs(ticker=tickers[i], multiplier=1, timespan="day", from_="2022-09-24", to="2024-09-23", limit=50000):
    lista_polygon.append(a)

print(len(lista_polygon))

```

Figura 8. Consulta a la api polygon.io para test

Como podemos observar en los gráficos, los datos obtenidos son guardados en DataFrames para su manipulación posterior.

4.2. Modelado

De la data en bruto que nos devuelven las *apis* vamos a trabajar con las columnas “Date” y “Close”. Es decir que se va a trabajar con el precio de la acción con el que cerró en cada día.

Después de filtrar las columnas con las que se quiere trabajar hacemos una limpieza de los datos, es decir, verificando que no existan nulos, que no hayan duplicados, que el tipo de dato sea *date* y *float64* correspondientemente.

Luego de la limpieza de los datos procedemos a hacer una normalización de los datos, es decir pasar los precios a un rango entre 0 y 1, aparte de que evita que los valores extremos tengan mucha más influencia, después de hacer pruebas se obtuvieron mejores resultados de esta forma.

La estrategia a emplear consiste en introducir series de 60 valores consecutivos en fecha y que el modelo aprenda a predecir el siguiente. Para esto empleamos una función que nos permite separar los 60 valores como variables y el siguiente valor como variable objetivo, a lo que hemos denominado *x_train* y *y_train*.

Repetimos lo mismo para la data de validación y obtenemos lo que hemos denominado x_{test} y y_{test} para realizar las validaciones a los modelos entrenados.

```

from sklearn.preprocessing import MinMaxScaler

scaler = MinMaxScaler(feature_range=(0,1))
scaled_data = scaler.fit_transform(df['Close'].values.reshape(-1,1))

prediction_days = 60

x_train = []
y_train = []

for x in range(prediction_days, len(scaled_data)):
    x_train.append(scaled_data[x-prediction_days:x,0])
    y_train.append(scaled_data[x,0])

x_train, y_train = np.array(x_train), np.array(y_train)
#x_train = np.reshape(x_train, (x_train.shape[0], x_train.shape[1], 1))

x_train.shape

```

Figura 9. Normalización de datos y obtención de x_{train} y y_{train}

Con los sets de datos trabajados, empezamos a modelar, primero con una regresión lineal, la definimos, entrenamos y luego la evaluamos, con lo cuál obtenemos el error cuadrático medio (mse), que es la medida de precisión que hemos elegido para saber cuál modelo predice mejor.

```

y_pred = model_lr.predict(x_test)

mse_lr = mean_squared_error(y_test, y_pred)

print("Error cuadrático medio: ", mse_lr)

```

```
Error cuadrático medio: 0.24214997186859333
```

Figura 10. Obtención del mse

Lo mismos pasos los repetimos para los modelos RandomForest de regresión, GradientBoosting de Regresión, XGBoost de Regresión, Red Neuronal Recurrente GRU y para la Red Neuronal LSTM y así obtener el mse de cada modelo para saber qué modelo tiene un mejor desempeño en la predicción.

4.3. Gráfico del modelo

Además de obtener el mse , también se muestra el resultado de las predicciones mediante un gráfico, a continuación, uno de los ejemplos



Figura 11. Gráfico del modelo

4.4. Resultado

Con un conjunto de datos para entrenamiento de 8 años y un conjunto de datos para pruebas de 2 años, se obtuvieron los siguientes resultados con respecto al error cuadrático medio en los modelos de regresión descritos:

Modelo	MSE
Regresión Lineal	12.0001
Random Forest de Regresión	6.93738
XGBoost de Regresión	6.3449
Red Neuronal LSTM (2 capas)	2.16E-04
Red neuronal LSTM (6 capas)	4.73E-04

Tabla 4. Los mse de cada modelo

De esta forma con los datos en mano podemos notar que el mejor modelo para predecir es la Red neuronal LSTM de dos capas.

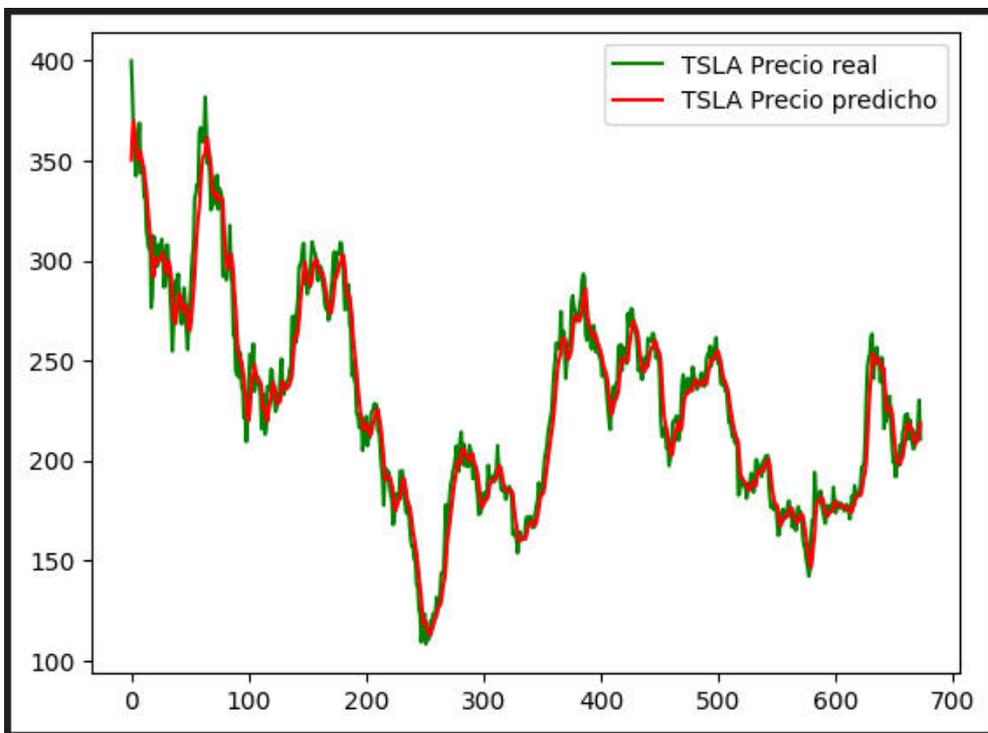


Figura 12. Gráfico de predicción de la Red Neuronal LSTM

Mediante una función preparamos las señales de compra y venta

```

buy_signals = []
sell_signals = []
for i in range(1, len(predicted_prices)):
    if predicted_prices[i] > actual_prices[i-1]:
        buy_signals.append((i, actual_prices[i])) # Alerta de compra
    elif predicted_prices[i] < actual_prices[i-1]:
        sell_signals.append((i, actual_prices[i])) # Alerta de venta

```

Figura 13. Función para la obtención de señales

Con estos sets de datos obtenidos hacemos la gráfica histórica de las señales

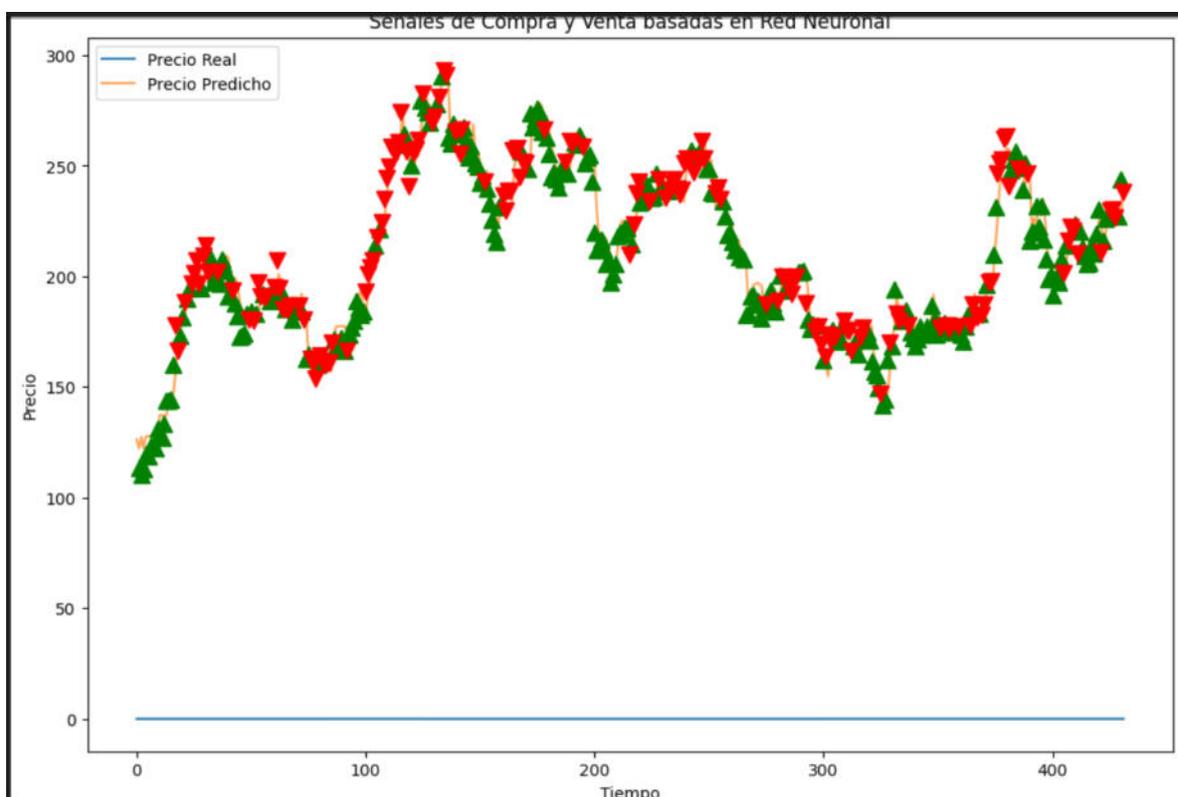


Figura 14. Gráfica histórica de señales de compra y venta

CAPITULO 5

5. CONCLUSIONES Y APLICACIONES

5.1. Conclusiones generales

Después de probar algunos modelos de Machine Learning se concluye como se lo esperaba que las redes neuronales dan mayor acercamiento a los precios reales de las acciones, aunque en la vida real no es suficiente.

Con la ayuda de las herramientas modernas como lo son las redes neuronales se tiene un acercamiento a la realidad más sin embargo aún no se puede predecir con exactitud, por eso es que el planteamiento de este proyecto es de poder proveer una herramienta que proporcione información que ayude a una mejor visualización de ella, sin embargo, se necesita de intervención humana para la toma de decisión final.

Se ha tenido un avance con respecto a la década anterior pero aún sigue siendo un campo que requiere también de análisis de otros agentes externos para poder acertar y poder obtener ganancias con el mínimo de riesgo.

Lo desarrollado sirve para mejorar nuestras inversiones y reducir el riesgo.

Es difícil cuantificar el impacto positivo porque efectivamente existen tantas variables externas que afectan a las acciones que resulta imposible predecir con un alto porcentaje de confiabilidad.

De alguna manera podemos concluir que no es un fenómeno que podamos modelizar con una gran fiabilidad y mucho menos a medio o largo plazo.

5.2. Conclusiones específicas

5.2.1. Análisis del cumplimiento de los objetivos de la investigación

Aunque hemos obtenido el mejor resultado entre los modelos de regresión, al evaluarlo con una función de rentabilidad en un lapso de dos años obtenemos como resultado una pérdida del 9.69%, porque no dependemos sólo de los datos históricos de los precios, sino que otros factores tienen bastante peso sobre el comportamiento del mercado.

```

rentability = 1
for i in range(1,len(actual_prices)):
    if predicted_prices[i] > actual_prices[i-1]:
        rentability*= actual_prices[i]/actual_prices[i-1]

print((rentability-1)*100,"%")

```

✓ 0.0s

-9.695792926384673 %

Figura 15. Función para obtener la rentabilidad

Y al hacerlo por el mes de agosto de 2024 se obtuvo una pérdida del 1.46%, es decir, si se invertía \$100 habríamos perdido \$14,60.

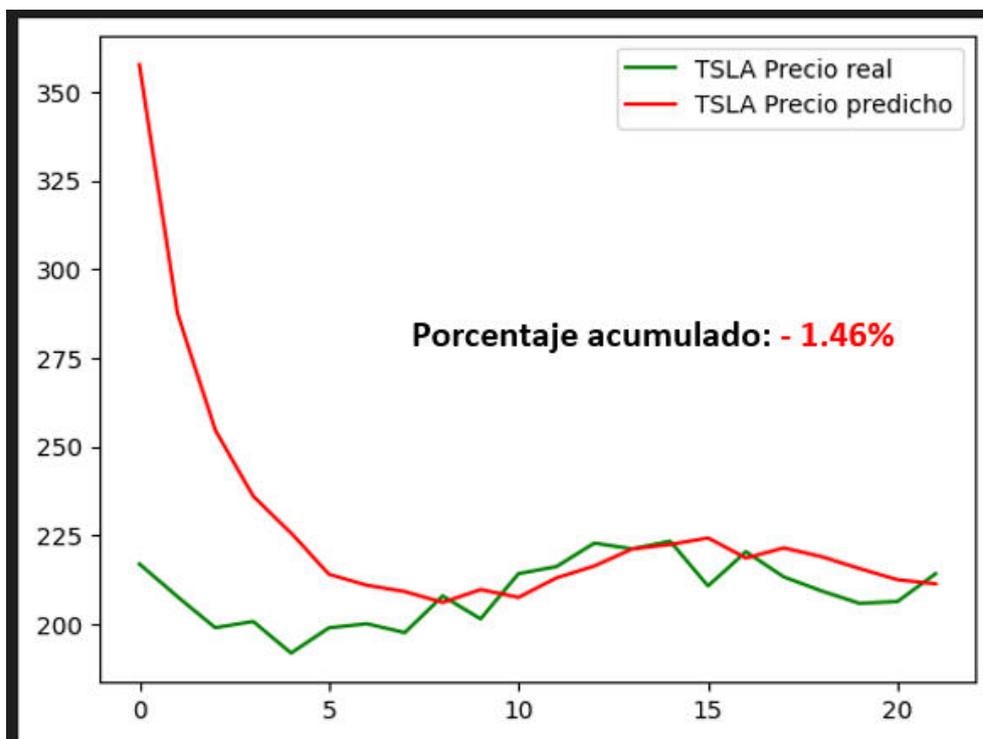


Figura 16. Gráfico de porcentaje acumulado de rentabilidad

En otras corridas del algoritmo tenemos:

- Volatilidad media diaria de acciones internacionales (Amazon, Apple, etc.): 2.5%

Rentabilidad Esperada:

- Rentabilidad media anual con señales de compra/venta: 12%

- Rentabilidad media anual con estrategia de compra y retención: 8%

Retorno del Portafolio Simulado:

- Retorno del portafolio utilizando señales generadas: 18%
- Pérdida máxima observada con stop-loss: 5%

Se hizo la evaluación en períodos de tiempo más cortos (anual, semestral y trimestral) a lo validado en un principio que fueron 2 años y se obtuvieron cada vez resultados más distantes a lo real, como se muestra a continuación.

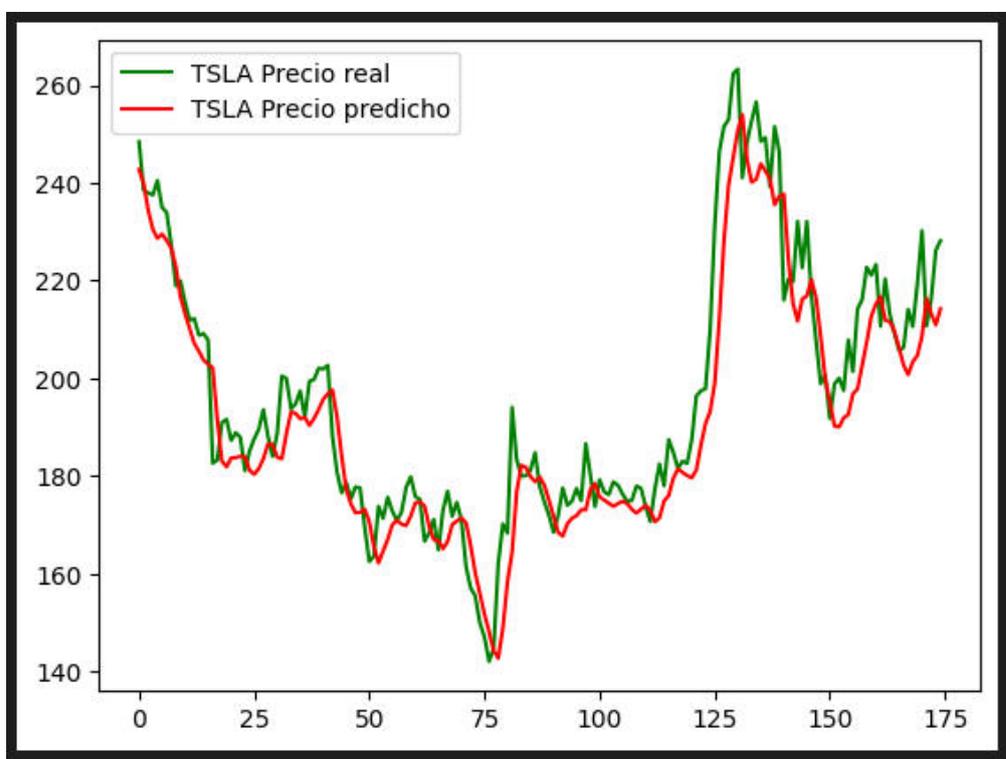


Figura 17. Gráfico de porcentaje acumulado de rentabilidad anual

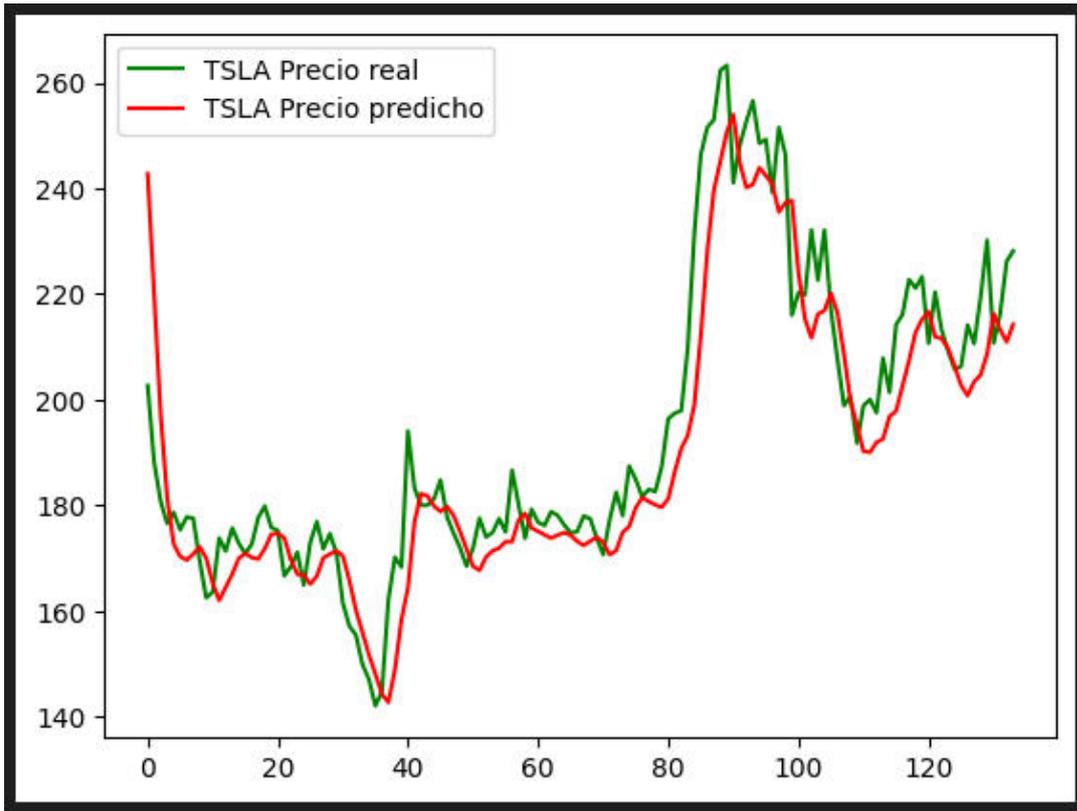


Figura 18. Gráfico de porcentaje acumulado de rentabilidad semestral

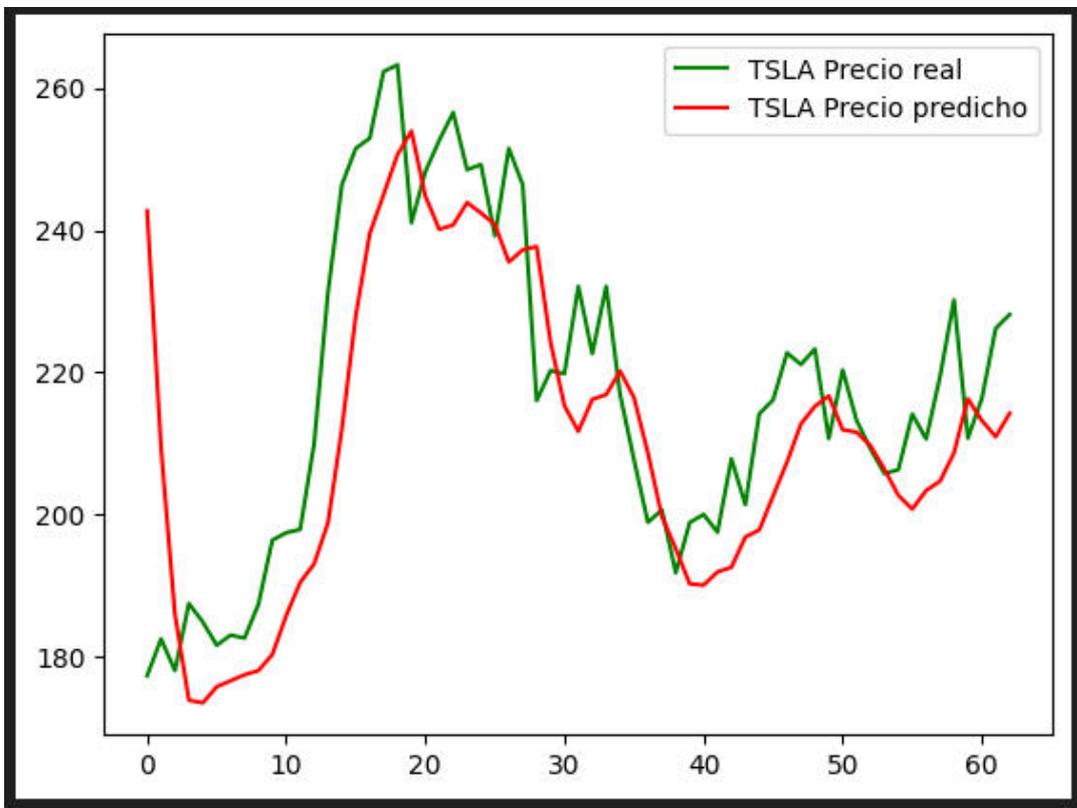


Figura 19. Gráfico de porcentaje acumulado de rentabilidad trimestral

Aplicando la función de rentabilidad, se obtuvo lo siguiente:

Tiempo	Rentabilidad
Anual	-2.69%
Semestral	-9.34%
Trimestral	-13.78%

Tabla 5. Rentabilidad por distintos períodos de tiempo

5.2.2. Contribución a la gestión empresarial

El análisis algorítmico provee una ventaja significativa al reducir la exposición al riesgo y mejorar la eficiencia de las decisiones de inversión en comparación con métodos manuales o basados en intuición.

En mercados de alta volatilidad, como los mercados internacionales, el uso de modelos predictivos con alta capacidad de adaptación (machine learning) puede generar retornos superiores a los del mercado en general.

Aunque el modelo ha mostrado efectividad en condiciones de mercado de alta volatilidad, su desempeño puede ser menos efectivo en mercados estables o con bajos volúmenes de transacción, como la Bolsa de Valores de Quito.

Se recomienda continuar con el refinamiento del modelo, incorporando nuevas variables (indicadores macroeconómicos, análisis de sentimiento) y explorando técnicas avanzadas de Deep Learning para mejorar la precisión y adaptabilidad del modelo.

5.2.3. Contribución a nivel académico

Considero que podría ser utilizado como ejemplos de futuras clases del máster para ver el desempeño de los modelos aplicado a conjunto de datos de series temporales y como un reto a la vez para poder optimizarlo y obtener mejores aproximaciones a lo real.

5.2.4. Contribución a nivel personal

En lo personal me sirvió mucho, porque me llevó a mucha investigación para poder lograr aplicar los modelos empleados a series temporales.

Fue una forma muy provechosa para poder poner en práctica todo lo aprendido en las diferentes materias que se vieron.

También fue un reto ya que conllevó disciplina para poder cambiar ciertas actividades a las que estaba acostumbrado hacer, para poder usar ese tiempo a la investigación.

5.3. Limitaciones a la Investigación

Como en mucho de los casos, la mayor limitante es el tiempo. Poder combinar las actividades cotidianas, la familia, el trabajo y eventos sociales no programados, son limitantes para poder dedicar más tiempo a la investigación y poder encontrar más información que alimenten ideas de como llegar a una mejor optimización de la herramienta en desarrollo.

Otra limitante también es poder encontrar información tabulada sobre los factores de riesgo, amenazas de guerra y salud para poder combinar junto con la data histórica de las acciones en la bolsa de valores y así poder perfeccionar el modelo y llegar a una mayor precisión en la predicción y lograr tener mayor confianza para tener réditos monetarios.

5.4. Recomendación

Se recomienda utilizar esta herramienta como ayuda para una toma de decisión basado en data histórica, pero junto a ésta, utilizar otras informaciones que ayuden a ampliar el marco de visión antes de una toma de decisión definitiva de inversión.

No está demás decirlo o repetirlo, aunque sea algo lógico, antes de arriesgarse, debemos buscar información lo más actualizada posible y de fuentes confiables a cerca de los CEO's de las compañías, o de lanzamientos de servicios o productos, junto a esto también informarse sobre

posibles emergencias sanitarias, rumores de guerras, fallas eléctricas o de seguridad y otros factores; porque esto tiene bastante peso y afecta directamente al mercado bursátil.

BIBLIOGRAFÍA

- Torres (22 de septiembre de 2019). Redes Neuronales Recurrentes. Torres AI.
<https://torres.ai/redes-neuronales-recurrentes/>
- Aggarwal, C. C. (2016). *Data Mining: The Textbook*. Springer.
- Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2), 383-417.
- Lo, A. W. (2004). The Adaptive Markets Hypothesis: Market Efficiency from an Evolutionary Perspective. *Journal of Portfolio Management*, 30(5), 15-29.
- Murphy, J. J. (1999). *Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications*. New York Institute of Finance.
- Tsay, R. S. (2005). *Analysis of Financial Time Series*. John Wiley & Sons.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to Linear Regression Analysis*. Wiley.
- Stonebraker, M., & Cetintemel, U. (2005). "One Size Fits All": An Idea Whose Time Has Come and Gone. *Proceedings of the 21st International Conference on Data Engineering*, IEEE.
- King, R. O. (2013). Financial Data Processing: A Comparative Study of Techniques and Applications. *Financial Data Processing Journal*, 18(2), 89-106.
- Mittermayer, M.-A., & Knolmayer, G. (2006). Text Mining Systems for Market Response to News: A Survey of Existing and Proposed Systems. *International Journal of Financial Markets and Derivatives*, 1(1), 73-97.
- Chen, H., & Ng, T. K. (2017). *Market Microstructure in Practice: An Investor's Guide to Electronic Trading*. Wiley Finance Series.

- Bolsa de Valores de Quito. (2023). Estadísticas de Mercado. En línea.
- Interactive Brokers LLC. (2023). Trading Platforms. En línea.
- TD Ameritrade. (2023). Thinkorswim Trading Platform. En línea.
- Valpacífico Casa de Valores. (2023). Sitio web.
- Stratega Bursátil Casa de Valores. (2023). Sitio web.

GLOSARIO Y ACRÓNIMOS

Trading	Consiste en comprar o vender un valor subyacente en un mercado financiero con la intención de obtener un beneficio especulativo.
Trader	Un trader es un comerciante que obtiene rentabilidad al efectuar operaciones de compra y venta de activos en los mercados financieros y bursátiles.
API	API significa “interfaz de programación de aplicaciones”. En el contexto de las API, la palabra aplicación se refiere a cualquier software con una función distinta. La interfaz puede considerarse como un contrato de servicio entre dos aplicaciones.
Machine Learning	Es la ciencia de desarrollo de algoritmos y modelos estadísticos que utilizan los sistemas de computación con el fin de llevar a cabo tareas sin instrucciones explícitas, en vez de basarse en patrones e inferencias.
CEO	El Chief Executive Officer literalmente es el jefe ejecutivo, es decir, la persona que es responsable de administrar la empresa, sus procesos y desempeño.
Deep Learning	El deep learning, o aprendizaje profundo, que se basa en un conjunto de algoritmos relacionados con el machine learning y cuyas aplicaciones en el mundo real son cada vez más tangibles (predicción de resultados de negocio, evolución de asistentes virtuales, análisis de imágenes médicas, etc.).

Red Neuronal	Conexiones entre nodos que forman un ciclo lo que permite insistencia en la información
Regresión Lineal	Modelo estadístico que se utiliza para predecir el valor de una variable
Random Forest	Conjunto de árboles de decisión que se entrenan de manera independiente y luego combinan sus predicciones
MAE(Error Cuadrático Medio)	Métrica que mide la media de las diferencias absolutas entre las predicciones del modelo y los valores reales
Polygon.io	Plataforma que proporciona datos financieros a través de una API en tiempo real
Yahoo Finance	Web que proporciona datos financieros históricos de diferentes mercados bursátiles en tiempo real

APÉNDICE

- Link del código de predicción de precios de las acciones con la variable de cierre obtenida de la data de las apis:

<https://github.com/jimmytarira/prediccion-acciones-modelos-ml>

- Link del código de predicción de precios con series temporales propias y con medias móviles (probando otras formas):

<https://github.com/jimmytarira/prediccion-acciones-medias-moviles>