



ESCUELA DE NEGOCIOS

**MAESTRÍA EN CIENCIA EN INTELIGENCIA DE NEGOCIOS Y CIENCIAS
DE DATOS**

TÍTULO DE LA INVESTIGACIÓN

**PRONÓSTICOS DE SERIES TEMPORALES UTILIZANDO REDES
NEURONALES ARTIFICIALES: UNA APLICACIÓN AL ESTUDIO DE LAS
SERIES DE TIEMPO EN EL MERCADO DE ACCIONES**

Profesor

Mario Salvador Gonzales Rodríguez Ph.D

Autor

Henry Fabian Bautista Vega

2023

RESUMEN

Este documento presenta una exploración exhaustiva de las técnicas avanzadas de redes neuronales profundas (DNN), incluida la integración de arquitecturas de memoria a corto plazo (LSTM) y redes neuronales convolucionales (CNN). La investigación se enfoca en analizar conjuntos de datos en el mercado de valores, empleando Python 3 junto con Tensor Flow y Keras. La importancia de la metodología radica en su potencial para revelar relaciones ocultas entre las variables observables del mercado de valores y las variables de estado. Además, permite el escrutinio en tiempo real de las tendencias del mercado y potencia la generación de pronósticos para los próximos períodos. La utilización de LSTM y redes convolucionales se aprovecha para lograr estos objetivos de manera efectiva. La evaluación de los resultados abarca una comparación mediante la métrica del error cuadrático medio (MSE) y error absoluto medio (MAE). A través de este análisis, el estudio muestra que las predicciones derivadas capturan fielmente los comportamientos intrincados inherentes a los datos de series temporales del mercado de valores.

Palabras Clave: Recurrent neural network, convolutional neural network , Long short-term memory (LSTM), Stock time series.

ABSTRACT

This paper presents a comprehensive exploration of advanced deep neural network (DNN) techniques, including the integration of short-term memory (LSTM) and convolutional neural networks (CNN) architectures. The research focuses on analyzing data sets in the market using Python 3 together with Tensor Flow and Keras. The importance of the methodology lies in its potential to reveal hidden relationships between observable stock market variables and state variables. In addition, it allows real-time scrutiny of the market trends and power the generation of forecasts for the next periods. The use of LSTM and convolutional networks is leveraged to achieve these objectives effectively. The evaluation of the results includes a comparison using the metric of the root mean square error (MSE) and Mean absolute error (MAE). Through this analysis, the study shows that the derived predictions faithfully capture the intricate behaviors inherent in stock market time series data.

Keywords: Recurrent neural network, convolutional neural network , Long short-term memory, Stock time series.

ÍNDICE DEL CONTENIDO

Table of Contents

| | |
|--|----|
| 1. RESUMEN | 2 |
| 2. ABSTRACT | 3 |
| 3. INTRODUCCIÓN..... | 1 |
| 4. REVISIÓN DE LITERATURA..... | 3 |
| 5. IDENTIFICACIÓN DEL OBJETO DE ESTUDIO | 6 |
| 6. PLANTEAMIENTO DEL PROBLEMA | 7 |
| 7. OBJETIVO GENERAL..... | 6 |
| 8. OBJETIVOS ESPECÍFICOS..... | 6 |
| 9. JUSTIFICACIÓN Y APLICACIÓN DE LA METODOLOGÍA | 9 |
| 10. RESULTADOS..... | 14 |
| 11. DISCUSIÓN DE LOS RESULTADOS Y PROPUESTA DE SOLUCIÓN..... | 19 |
| 12. CONCLUSIONES Y RECOMENDACIONES | 21 |
| 13. Referencias | 23 |

ÍNDICE DE TABLAS

| | |
|--|----|
| Tabla 1. Descripción estadística de series temporales no procesadas..... | 11 |
| Tabla 2. Comparación de los dos modelos usando MSE y MAE..... | 17 |

ÍNDICE DE FIGURAS

| | |
|--|----|
| Figura 1. Metodología para la previsión de series temporales de acciones.... | 9 |
| Figura 2. Precio de las acciones de la corporación de Nvidia..... | 10 |
| Figura 3. Diagrama de caja de la serie temporal de Nvidia..... | 10 |
| Figura 4. Ejemplo de una red neuronal | 12 |
| Figura 5. Arquitectura de la red de memoria a corto plazo (LSTM). | 12 |
| Figura 6. Predicción usando un modelo RNN en el conjunto de datos de entrenamiento..... | 15 |
| Figura 7. Predicción usando un modelo RNN en el conjunto de datos de testeo..... | 15 |
| Figura 8. Predicción usando un modelo CNN en el conjunto de datos de entrenamiento..... | 16 |
| Figura 9. Predicción usando un modelo RNN en el conjunto de datos de testeo..... | 17 |
| Figura 10. Error en las presiones usando los métodos LSTM y CNN..... | 18 |
| Figura 11. Scatter plot de los métodos LSTM y CNN.,..... | 18 |

INTRODUCCIÓN

La predicción de los índices bursátiles se presenta como un desafío primordial pero formidable dentro del ámbito de las finanzas, especialmente en el ámbito de la gestión de inversiones. La importancia de los índices bursátiles trasciende los meros indicadores de mercado; ejercen una influencia considerable sobre las decisiones de política económica y afectan las opciones que abarcan inversiones tales como acciones y bienes raíces. Sin embargo, el intrincado esfuerzo de pronosticar con precisión las tendencias del mercado de valores se enfrenta a complejidades multifacéticas arraigadas en la naturaleza misma de los mercados financieros.

Estos mercados exhiben un conjunto distintivo de atributos, una mezcla de volatilidad, ruido, no linealidad y complejidad intrincada. Su naturaleza no paramétrica desafía los modelos analíticos simples, mientras que la no estacionariedad elude hábilmente las predicciones estáticas. Tales características, sumadas a su propensión al comportamiento caótico, imbuén a los modelos de predicción con una propensión sustancial a errores de considerable magnitud.

Sin embargo, la complejidad se extiende más allá de la dinámica del mercado. Los flujos en los precios del mercado de valores no son únicamente productos de métricas comerciales históricas; están igualmente moldeados por un laberinto de influencias no lineales. Estos abarcan una variedad de elementos, desde las ondas de los eventos políticos hasta los flujos y reflujos de los sentimientos de los inversores, hasta los sucesos imprevistos que insertan sin problemas la imprevisibilidad en la ecuación del mercado.

En este entorno, la fusión de técnicas de red neuronal profunda (DNN) de vanguardia surge como un faro prometedor. Al integrar a la perfección la

destreza de las arquitecturas de memoria a corto plazo (LSTM) y red neuronal convolucional (CNN), esta investigación se esfuerza por navegar por las complejidades laberínticas de las tendencias del mercado de valores. A través de un innovador enfoque de ventana deslizante, nuestro objetivo es decodificar las dependencias temporales y los patrones latentes, impulsándonos hacia una comprensión más lúcida del enigmático ámbito del mercado de valores.

REVISIÓN DE LITERATURA

Existe un gran interés en la industria y la academia por mejorar la precisión de las predicciones de series de tiempo. Estos son importantes en una variedad de campos como: medicina, finanzas, ciencia debido al hecho de que los datos que se recopilan de estos campos muestran patrones que podemos cuantificar (estacionalidad, tendencia, variabilidad) (Tiao, 2001). La capacidad de predecir valores de series temporales requiere seleccionar el mejor modelo de predicción para realizar el pronóstico, lo cual es un problema desafiante debido a la complejidad de los datos y la cantidad de modelos disponibles (Vadlamani, Dadabada, & Deb, 2017).

El propósito fundamental del análisis de series temporales consiste en examinar las trayectorias que se observan en estas secuencias y desarrollar un modelo que caracterice la estructura de los datos, con el objetivo de anticipar los valores futuros de dichas secuencias temporales (Namini et al., 2018). Un ejemplo concreto es la capacidad para predecir los costos en proyectos, lo cual contribuye a optimizar los precios y a planificar las actividades de manera más eficiente. A lo largo del tiempo, se han propuesto diversos enfoques convencionales para llevar a cabo pronósticos en series temporales. Entre ellos se encuentran enfoques estocásticos, la aplicación de Máquinas de Vectores de Soporte (SVM), la utilización del modelo Autorregresivo de Media Móvil Integrada (ARIMA) y técnicas de suavizado exponencial, entre otros métodos (Liu & Deng, 2019). A pesar de que estos métodos han demostrado su eficacia al abordar distintos problemas de pronóstico en series temporales, presentan limitaciones en términos de escalabilidad cuando se aplican a conjuntos de datos extensos (Sen et al., 2019).

En tiempos recientes, las técnicas basadas en Redes Neuronales Profundas (DNN) han ganado creciente atención en el análisis de series temporales. A diferencia de los modelos de pronóstico convencionales, las DNN cuentan con una considerable habilidad para discernir patrones complejos tanto dentro de

una secuencia como entre relaciones no lineales. Estas redes pueden extraer características de mayor orden a partir de conjuntos de datos crudos de series temporales, y esto se logra con una inversión de esfuerzo humano significativamente menor (Rangapuram et al., 2018).

(Wu, Green, Ben, & O'Banion, 2020)

Los autores proponen un enfoque novedoso que utiliza modelos de aprendizaje automático basados en transformadores para pronosticar datos de series temporales. Al aprovechar los mecanismos de autoatención, este marco genérico puede manejar datos de series temporales tanto univariantes como multivariantes, así como incrustaciones de series temporales. Los autores demuestran la efectividad de su enfoque a través de un estudio de caso sobre el pronóstico de enfermedades similares a la influenza (ILI), mostrando resultados comparables a los métodos más avanzados.

(Shi, Jain, & Narasimhan, 2022)

El artículo se enfoca en la investigación del desempeño de diferentes modelos predictivos en el pronóstico de series de tiempo, considerando específicamente el impacto de los diferentes tamaños de ventana retrospectiva y la cantidad de puntos de tiempo para predecir el futuro. Los autores comparan cuatro métodos de aprendizaje profundo, a saber, RNN, LSTM, GRU y Transformer, junto con un método de referencia. Sus experimentos en el conjunto de datos de calidad del aire de Beijing destacan el rendimiento superior de los modelos de transformadores, con los errores promedio promedio (MAE) y los errores cuadráticos medios (RSME) más bajos para predicciones de un solo paso y de varios pasos.

(Yıldırım, Toroslu, & Fiore, 2021)

El ámbito del comercio de divisas (Forex) presenta una arena distintiva caracterizada por un alto potencial de ganancias y riesgos sustanciales. A diferencia de los mercados financieros convencionales, Forex permite a los comerciantes predecir las direcciones del tipo de cambio de divisas para obtener

ganancias, lo que, si se juzga mal, puede generar pérdidas significativas. Para abordar este desafío, una investigación reciente emplea la "memoria a corto plazo" (LSTM), una poderosa herramienta de aprendizaje profundo, para predecir las direcciones del tipo de cambio [5]. Al utilizar dos conjuntos de datos distintos que representan datos de indicadores macroeconómicos y técnicos, el estudio reconoce las técnicas analíticas duales empleadas en el mundo financiero. Este enfoque subraya la capacidad de LSTM para capturar patrones complejos a lo largo del tiempo y destaca su potencial para mejorar la precisión predictiva en Forex, un ámbito donde las predicciones efectivas son fundamentales para navegar por el delicado equilibrio entre ganancias y pérdidas.

Este trabajo estudia diferentes métodos y estrategias de pronóstico para la selección de modelos con el objetivo de mejorar la precisión de la predicción. Adicionalmente, investigaciones que emplean DNNs en el contexto de series temporales han arrojado resultados notables, ofreciendo una perspectiva amplia sobre cómo diversas arquitecturas de DNN pueden ser empleadas en el análisis de estas secuencias.

IDENTIFICACIÓN DEL OBJETO DE ESTUDIO

El objeto de estudio de esta investigación es el conjunto de datos de series temporales que representan los valores históricos de las bolsas de valores de empresas destacadas del sector tecnológico. El conjunto de datos se recopiló de fuentes financieras acreditadas y abarca los precios de cierre cada media hora de las acciones durante un período de ocho meses, desde enero de 2023 hasta agosto de 2023. La empresa de interés es NVIDIA, conocida por desarrollar circuitos integrados, que se utilizan en todo, desde consolas de juegos electrónicos hasta computadoras personales (PC).

La importancia de estudiar este conjunto de datos radica en su relevancia para el panorama financiero contemporáneo. El sector de la tecnología tiene una influencia sustancial sobre las economías globales y las estrategias de inversión, lo que hace que las predicciones precisas del precio de las acciones sean invaluable para la toma de decisiones informadas. Al seleccionar este conjunto de datos, el estudio tiene como objetivo contribuir al avance de las metodologías de pronóstico financiero al tiempo que aborda los desafíos únicos que plantea la dinámica volátil e interconectada del mercado de valores.

Comprender las tendencias, los patrones y los posibles movimientos futuros de los precios de las acciones de estas empresas de tecnología puede ayudar a los inversores, analistas financieros y legisladores a formular estrategias de inversión, técnicas de gestión de riesgos y decisiones políticas más eficaces. El enfoque del conjunto de datos en el sector de la tecnología se presta a una exploración de los impactos de las innovaciones tecnológicas, los cambios en el sentimiento del mercado y factores económicos más amplios en los movimientos del precio de las acciones. Como tal, este estudio busca brindar información que se extienda más allá del ámbito de las acciones individuales y contribuya a una comprensión más amplia de las intrincadas relaciones entre los desarrollos tecnológicos, los mercados financieros y el comportamiento de los inversores.

PLANTEAMIENTO DEL PROBLEMA

El papel fundamental de Nvidia Corporation en la configuración de los avances tecnológicos recientes ha llamado la atención de inversores, analistas y observadores de la industria. A medida que los precios de las acciones de la empresa fluctúan, la capacidad de predecir con precisión sus movimientos futuros se vuelve indispensable. Sin embargo, los métodos tradicionales de predicción del precio de las acciones a menudo no logran capturar los matices de tendencias de mercado tan dinámicas e interconectadas. El desafío radica en diseñar un modelo predictivo efectivo que aproveche el poder de las técnicas avanzadas de aprendizaje profundo para mejorar la precisión y dar cuenta de las complejidades del comportamiento del mercado de valores de Nvidia. Por lo tanto, el principal problema abordado en este estudio es desarrollar un modelo de predicción de precios de acciones robusto y preciso para Nvidia Corporation mediante la integración de metodologías de redes neuronales profundas de vanguardia.

\

OBJETIVO GENERAL

El objetivo general de esta investigación es aprovechar las técnicas avanzadas de aprendizaje profundo para mejorar la precisión de la predicción del precio de las acciones de Nvidia Corporation. Al combinar arquitecturas de memoria a largo plazo a corto plazo y red neuronal convolucional, el proyecto tiene como objetivo desarrollar un modelo integral que capture los patrones intrincados y las dependencias temporales dentro de los datos del mercado de valores de Nvidia. Este objetivo general busca brindar a las partes interesadas información más confiable sobre las trayectorias potenciales de los precios de las acciones de Nvidia, lo que permite decisiones de inversión informadas y planificación estratégica.

OBJETIVOS ESPECÍFICOS

- Recopilación y preprocesamiento de datos:

Recopilar datos históricos del precio de las acciones de Nvidia. Preprocesar los datos para garantizar la coherencia y la idoneidad para el análisis.

- Implementación del modelo y capacitación:

Implementar modelos LSTM y CNN adaptados al análisis de series temporales de acciones. Entrene a cada modelo utilizando los datos recopilados para capturar relaciones temporales, identificar tendencias y predecir los movimientos del mercado.

- Comparar el rendimiento predictivo de los modelos LSTM y CNN, teniendo en cuenta métricas como el error absoluto medio (MAE) y el error cuadrático medio (MSE).

JUSTIFICACIÓN Y APLICACIÓN DE LA METODOLOGÍA

La elección de emplear una combinación de metodologías de memoria a largo plazo a corto plazo (LSTM) y red neuronal convolucional (CNN) para la predicción del precio de las acciones, centrándose específicamente en los datos de mercado de Nvidia Corporation, se deriva de las complejidades intrínsecas de las series temporales financieras. Los modelos lineales tradicionales a menudo tienen dificultades para capturar las relaciones no lineales y las dependencias temporales presentes en dichos datos. Aquí, desarrollamos la justificación de esta metodología y su relevancia para el problema en cuestión.

La estrategia para aplicar y afinar los procedimientos de pronóstico de las series de tiempo se presenta de manera concisa en la Figura 2. Esta estrategia se fundamenta en el enfoque metodológico Crisp-DM y abarca las fases que se describen a continuación: comprensión del contexto empresarial, comprensión de los datos, preparación de los datos, construcción del modelo, evaluación y puesta en práctica.

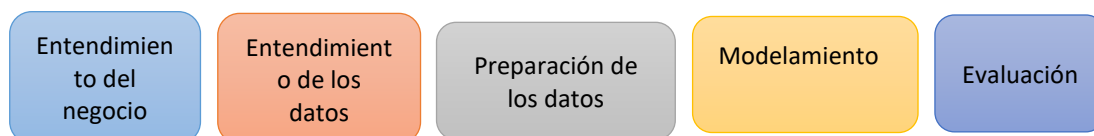


Figura 1. Metodología para la previsión de series temporales de acciones.

1. Entendimiento de los datos

El conjunto de datos se recopiló de fuentes financieras acreditadas en específico se utilizó una interfaz de programación de aplicaciones (API) desarrollada por Alpha Vantage la cual proporciona datos históricos y en tiempo real del mercado financiero a través de un conjunto de hojas de cálculo. Los precios de cierre fueron escogidos cada media hora desde enero del 2023 hasta junio del 2023. Recolectando un total de 1984 datos. Se escogió en específico esta serie de

datos ya que contiene una gran volatilidad de sus datos la Figura 2. Muestra el precio de sus acciones en el periodo mencionado.

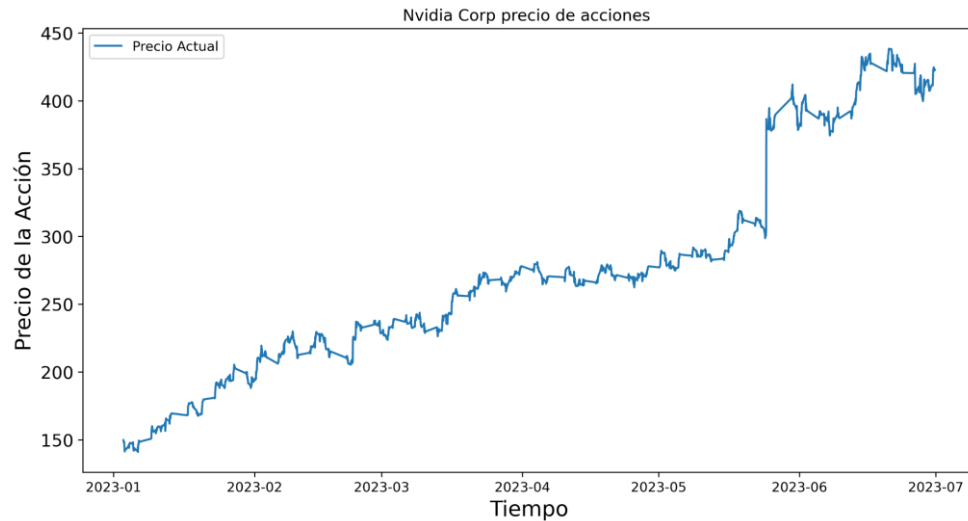


Figura 2. Precio de las acciones de la corporación de Nvidia.

Además, en la Fig. 3. se presenta un diagrama de caja que muestra los valores atípicos, la morfología y la simetría de los datos no procesados y una descripción completa en la Tabla 1.

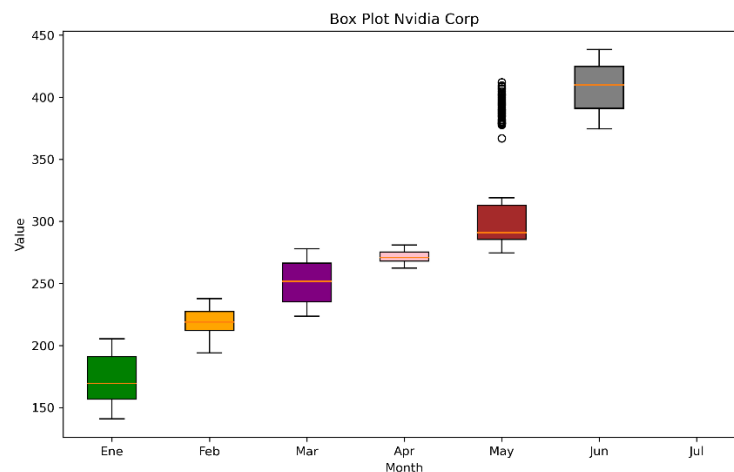


Figura 3. Diagrama de caja de la serie temporal de Nvidia

| Descripción | |
|-------------------|--------|
| Numero demuestras | 1984 |
| Media | 273.70 |
| Std | 77.60 |
| Min | 140.90 |
| 25% | 222.69 |
| 50% | 267.92 |
| 75% | 295.27 |
| Max | 438.39 |

Tabla 1. Descripción estadística de series temporales no procesadas.

2. Preparación y Modelamiento de los Datos

La preparación de los datos se realizó mediante la librería pandas de Python aquellos elementos faltantes fueron completados utilizando la media de todos los datos faltantes del mismo día. Para el modelamiento de los datos se utilizaron LSTM y CNN, los métodos son descritos a continuación.

3.1. Redes Neuronales Artificiales (RNA)

Las redes neuronales artificiales representan un sistema de procesamiento de información que imita de manera aproximada el funcionamiento del cerebro humano al simular las operaciones y conexiones presentes en las neuronas biológicas (Palani et al., 2008). Las ANN se utilizan típicamente cuando los datos incluyen componentes no lineales, debido a su eficacia en la resolución de este tipo de desafíos. De acuerdo a (Hyndman & Athanasopoulos, 2018), una red neuronal puede ser concebida como una estructura compuesta por "neuronas" dispuestas en capas. La capa inferior está formada por predictores (o entradas), mientras que la capa superior está compuesta por pronósticos (o salidas). También puede haber capas intermedias que contengan "neuronas ocultas", como se muestra en la Figura 4.

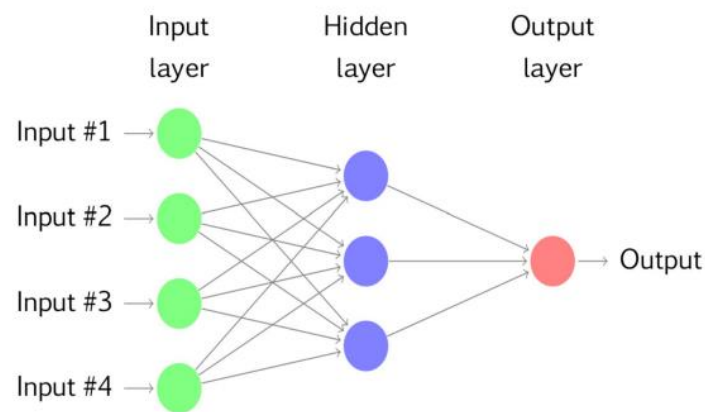


Figura 4. Ejemplo de una red neuronal

3.2. Red de memoria a largo plazo (LSTM)

La red de memoria a corto plazo es un tipo de red neuronal recurrente utilizada en el aprendizaje profundo. Según Gers, Eck, and Schmidhuber (2002), las unidades de entrada están completamente conectadas a una capa oculta que consta de bloques de memoria con 1 celda cada uno. Las salidas de la celda están completamente conectadas a las entradas de la celda, a todas las puertas y a las unidades de salida. Además, las puertas, la celda y las salidas están polarizadas. En pocas palabras, LSTM está organizado en celdas que incluyen varias operaciones, tiene una variable de estado interna, que se pasa de una celda a otra y se modifica mediante puertas de operación (Statsbot, 2021). La Figura 5 se presenta para ilustrar la arquitectura de LSTM.

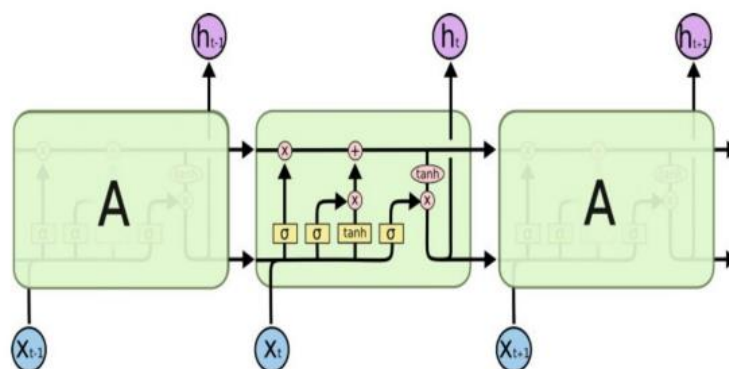


Figura 5. Arquitectura de la red de memoria a corto plazo (LSTM).

3. 3 Red neuronal convolucional

Las redes neuronales convolucionales (CNN) se asemejan a las redes neuronales artificiales (ANN) convencionales en su composición, ya que están formadas por neuronas dispuestas en tres dimensiones: las dimensiones espaciales de la entrada (alto y ancho) y la profundidad. Estas neuronas se ajustan automáticamente a través del proceso de aprendizaje (Shea y Nash, 2015). Además, estas neuronas se encuentran en tres tipos de capas diferentes (capas convolucionales, capas de agrupación y capas totalmente conectadas), que también desempeñan un papel en las CNN. Estas últimas son un tipo de redes neuronales profundas que han ganado recientemente popularidad debido a su eficacia en la resolución de problemas relacionados con series temporales (Wang et al., 2017). A diferencia de las ANN tradicionales, la salida de cada capa en una CNN está conectada únicamente a regiones locales en la entrada. Esto se logra mediante el desplazamiento de un filtro o una matriz de pesos a lo largo de la entrada, calculando un producto escalar en cada posición, es decir, realizando una convolución entre la entrada y el filtro. Esta estructura ha posibilitado que el modelo identifique patrones específicos en los datos de entrada al aprender a través de estos filtros (Borovykh et al., 2017).

Optimización de Adán

Adám se deriva de la abreviatura "Estimación de Momento Adaptativo". Es un algoritmo de optimización empleado en lugar del método convencional de Descenso de Gradiente Estocástico para actualizar de manera iterativa los pesos de una red neuronal basándose en los datos de entrenamiento (Brownlee, 2016). La técnica de Adam fusiona las concepciones de dos enfoques de optimización: el impulso y la propagación cuadrática media (RMSProp).

3. Testeo

Para construir un modelo predictivo robusto y bien validado, se tomó un enfoque meticuloso en el manejo del conjunto de datos. Específicamente, se empleó una división de datos de 80-20, con el 80 por ciento del conjunto de datos asignado al conjunto de entrenamiento y el 20 por ciento restante reservado para el conjunto de prueba. Este esquema de partición garantiza que una parte sustancial de los datos se dedique a entrenar el modelo, lo que le permite

aprender patrones y relaciones a partir de los datos históricos de precios de acciones.

Además, en la fase de entrenamiento del modelo se emplearon un total de 60 épocas. Una época representa un paso completo a través de todo el conjunto de datos de entrenamiento. La elección de 60 épocas se hizo para lograr un equilibrio entre la convergencia del modelo y la eficiencia computacional. Este número se determinó mediante experimentación y validación para garantizar que el modelo alcanzara un estado estable de aprendizaje sin sobre ajustarse a los datos de entrenamiento. También se alinea con las prácticas establecidas en el desarrollo de modelos de aprendizaje profundo.

Antes del análisis, el conjunto de datos se sometió a procedimientos de preprocesamiento para garantizar su idoneidad para la previsión de series temporales. Esto implicó abordar los puntos de datos faltantes, normalizar los precios de las acciones y eliminar posibles valores atípicos que podrían afectar negativamente la precisión de las predicciones. Además, los precios de cierre diarios se convirtieron en rendimientos logarítmicos para tener en cuenta la no estacionariedad inherente a las series temporales financieras.

RESULTADOS

En esta sección, presentamos los resultados de los dos sistemas meticulosamente estudiados a través de una combinación de representaciones visuales y datos tabulares, con el objetivo de proporcionar una comprensión integral de nuestros hallazgos.

1. Experimento 1 - Modelo LSTM

El modelo LSTM se ejecutó con un entrenamiento de 60 épocas. Estos parámetros se seleccionaron cuidadosamente en función de los conocimientos derivados de la Figura 4, que ilustra su importancia en la configuración del comportamiento del modelo.

Con el modelo ya entrenado se procedió a predecir el conjunto de datos con los cuales el sistema fue entrenado. Los resultados se muestran en la Figura 6.



Figura 6. Predicción usando un modelo RNN en el conjunto de datos de entrenamiento.

El siguiente paso en este experimento consistió en utilizar el conjunto de datos de testeo.

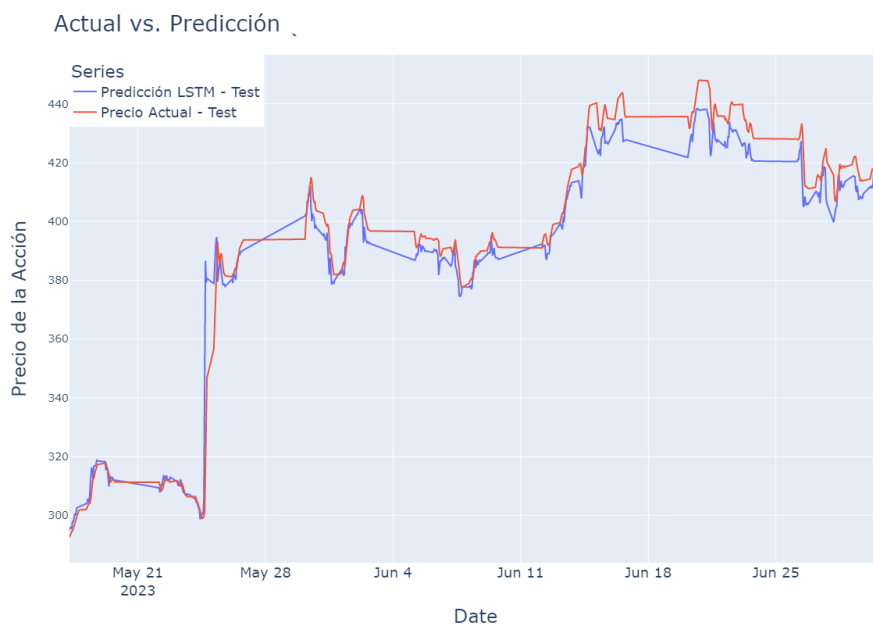


Figura 7. Predicción usando un modelo RNN en el conjunto de datos de testeo.

2. Experimento 2 - Modelo CNN:

De manera similar, el modelo CNN experimentó con parámetros determinados antes del inicio de la implementación. La lógica detrás de la selección de parámetros se basó en los conocimientos revelados por Figure 5, que arrojó luz sobre su papel fundamental en el rendimiento del modelo CNN.

Al ofrecer esta presentación estructurada, nuestro objetivo es proporcionar una visión más completa e intuitiva de los resultados obtenidos de ambos experimentos, facilitando una comprensión más profunda de sus implicaciones y significado en nuestro estudio.

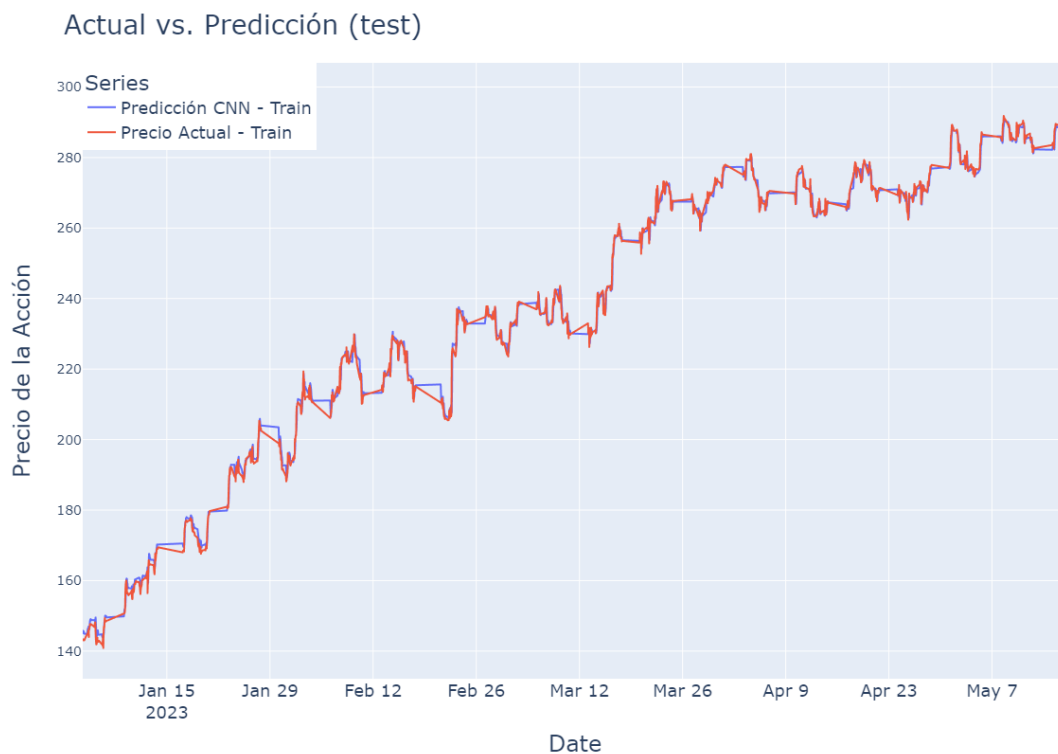


Figura 8. Predicción usando un modelo CNN en el conjunto de datos de entrenamiento.

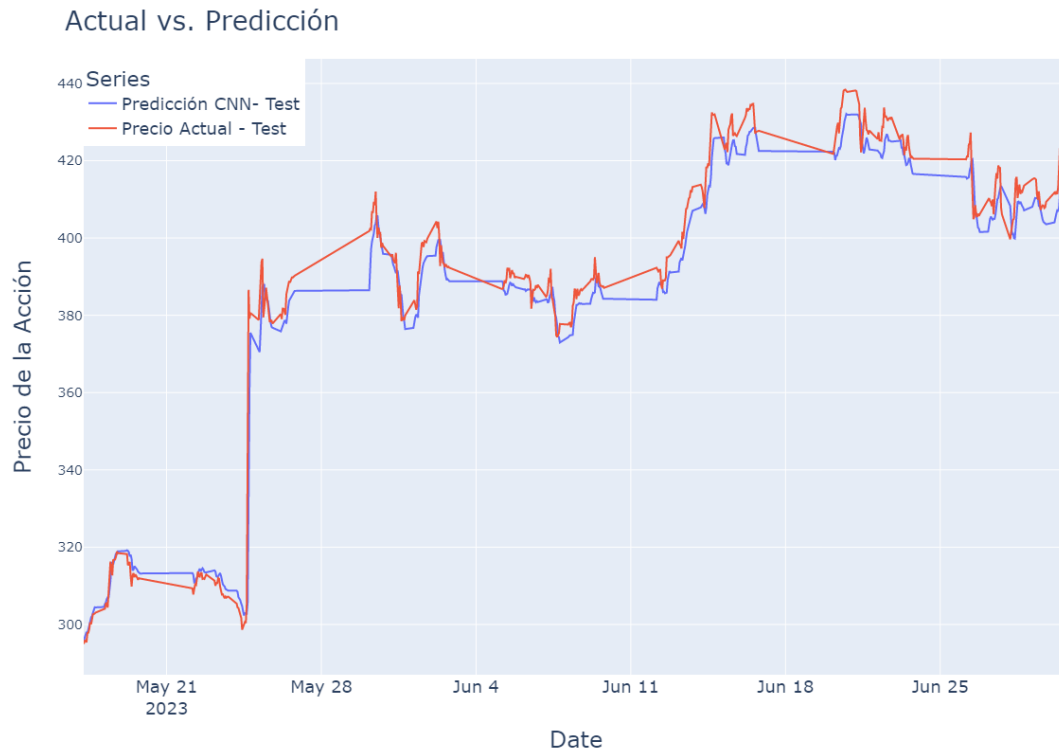


Figura 9. Predicción usando un modelo CNN en el conjunto de datos de testeo.

Los resultados aquí presentados se obtuvieron usando Python en conjunto con la librería TensorFlow.

3. Prueba de resultados MSE y MAE

La Tabla 2 proporciona una evaluación integral del rendimiento de los modelos de redes neuronales. Después de la fase de entrenamiento de la red, se realizaron pruebas rigurosas utilizando los conjuntos de datos de entrenamiento y prueba. Los resultados revelan una tendencia convincente, afirmando la superioridad del modelo CNN como una opción más competente para la predicción.

| | MSE | | MAE | |
|--------|-------|-------|-------|------|
| Modelo | Train | Test | Train | Test |
| LSTM | 22.85 | 88.75 | 3.63 | 7.16 |
| CNN | 19.85 | 69.40 | 3.25 | 6.47 |

Tabla 2. Comparación de los dos modelos usando MSE y MAE.

En la Figura 10. Podemos apreciar que los errores de comportan de manera similar durante el tiempo sin embargo el modelo LSTM se aleja más de la línea cero durante el tiempo.

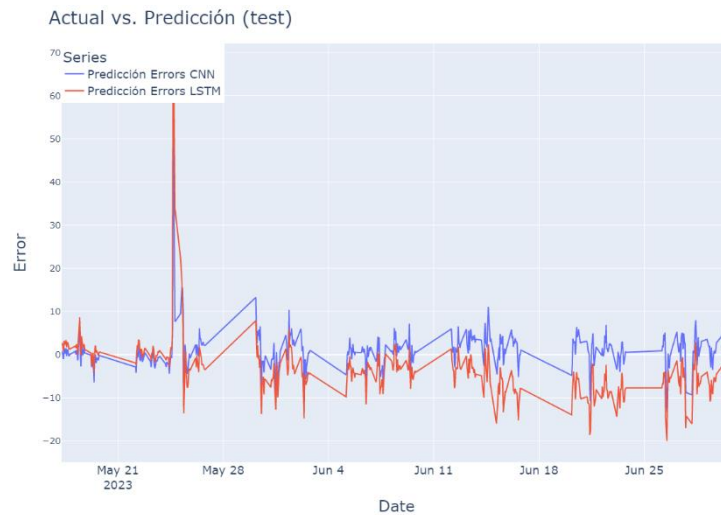


Figura 10. Error en las presiones usando los métodos LSTM y CNN.

4. Comparación mediante Scatter plot

El siguiente scatter plot muestra la relación entre dos variables entre los valores verdaderos y las predicciones. Donde el escenario ideal sería que los puntos formen una línea inclinada a 45 grados en la gráfica. Ponemos aprecio que los resultados obtenidos se acercan a las mismas.

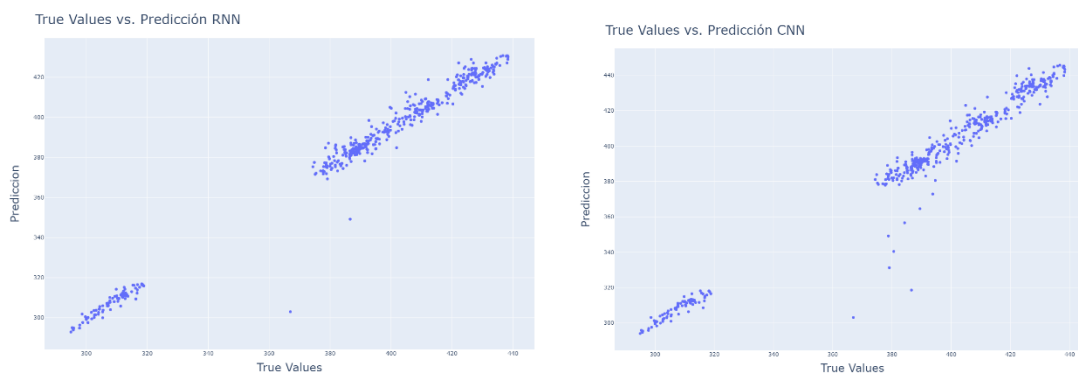


Figura 11. Scatter plot de los métodos.

DISCUSIÓN DE LOS RESULTADOS Y PROPUESTA DE SOLUCIÓN

La metodología elegida para este estudio, que combina técnicas de memoria a corto plazo (LSTM) y red neuronal convolucional (CNN), está respaldada por varias justificaciones clave y una aplicación clara al problema en cuestión.

Los datos de series temporales financieras, en particular la evolución de los precios de las acciones, son inherentemente complejos. Se caracteriza por la no linealidad, las dependencias temporales y la influencia de varios factores externos. En esta serie de tiempo en particular se logra apreciar en el mayo un salto considerable en el precio de las acciones. Desacuerdo con la teoría las redes LSTM son adecuadas para capturar dependencias de largo alcance, mientras que las CNN son efectivas para identificar patrones locales. Sin embargo, en este ejemplo en particular se pudo apreciar en la Figura 10. Que ambos modelos comparten un mismo patrón en la predicción de errores, no obstante, en LSTM en error en el tiempo aparenta aumentar. La combinación de estas arquitecturas permite una comprensión más holística de las complejidades de los datos.

Soluciones:

Los datos financieros pueden ser ruidosos, influenciados por irregularidades del mercado y eventos repentinos. La capacidad de LSTM para aprender de secuencias mientras se consideran estados anteriores puede ayudar a mitigar el ruido, mientras que los filtros espaciales de CNN pueden mejorar la solidez del modelo frente a valores atípicos y puntos de datos ruidosos.

Los precios de las acciones están influenciados por una miríada de factores, incluido el sentimiento del mercado, los indicadores económicos y las noticias específicas de la empresa. La fusión LSTM-CNN permite que el modelo capture estas relaciones interconectadas, proporcionando un análisis más completo de cómo varios factores afectan los movimientos de precios de las acciones.

CONCLUSIONES Y RECOMENDACIONES

La metodología elegida para este estudio, una combinación armoniosa de técnicas de memoria a corto plazo (LSTM) y redes neuronales convolucionales (CNN), se ve reforzada por una justificación sólida y una aplicación precisa a los desafíos actuales.

Los datos de series de tiempo financieras, con su intrincada dinámica, en particular las trayectorias de los precios de las acciones, encarnan inherentemente complejidad. Estas complejidades abarcan la no linealidad, las dependencias temporales y la influencia de diversas variables externas. En particular, en mayo se observó un notable aumento de los precios de las acciones, desviándose algo de las expectativas teóricas tradicionales. Las redes LSTM se destacan en capturar dependencias de corto alcance, mientras que las CNN disciernen hábilmente patrones a largo plazo. Curiosamente, la Figura 10 revela que ambos modelos comparten patrones de predicción de errores similares. Sin embargo, LSTM muestra una tendencia perceptible a aumentar los errores con el tiempo. La fusión de estos elementos arquitectónicos se desarrolla así como una estrategia primordial, que permite una comprensión más holística de las complejidades de los datos.

Ideas y recomendaciones clave:

Mitigación del ruido: Como se ha demostrado, los datos financieros pueden ser susceptibles al ruido, anomalías del mercado y acontecimientos imprevistos. La competencia de LSTM en el aprendizaje secuencial y la consideración de estados históricos puede amortiguar significativamente el impacto del ruido. Al mismo tiempo, los filtros espaciales de CNN refuerzan la solidez del modelo, minimizando efectivamente la influencia disruptiva de valores atípicos y puntos de datos erráticos.

Factores interconectados: Los precios de las acciones, al ser el resultado de una intrincada interacción entre el sentimiento del mercado, los indicadores económicos y las noticias específicas de la empresa, requieren un enfoque matizado para la modelización. La fusión LSTM-CNN surge como un aliado estratégico para desentrañar estas intrincadas relaciones. Proporciona un amplio panorama para evaluar cómo se entrelazan diversos factores multifacéticos para influir en la dinámica de los precios de las acciones.

En conclusión, este estudio defiende la integración de las arquitecturas LSTM y CNN, iluminando su sinergia armoniosa para decodificar las complejidades inherentes a los datos de series de tiempo financieras. Como resultado, este método no sólo mejora la precisión de las predicciones sino que también proporciona una comprensión más profunda de las fuerzas multifacéticas que dan forma a las tendencias del mercado de valores. Recomendamos una mayor exploración de este enfoque híbrido y su adaptación a una gama más amplia de desafíos de previsión financiera.

REFERENCIAS

- Gers, F. A., Eck, D., & Schmidhuber, J. (2002). Applying LSTM to time series predictable through time-window approaches. In *Neural Nets WIRN Vietri-01* (pp. 193-200). Springer.
- Palani, S., Liong, S.-Y., & Tkalich, P. (2008). An ANN application for water quality forecasting. *Marine Pollution Bulletin*, 56(9), 1586-1597.
- Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice*. OTexts.
- Shea, K., & Nash, R. (2015). An introduction to convolutional neural networks. arXiv preprint arXiv:1511.08458.
- Namini, S., Tavakoli, N., & Namin, A. S. (2018). A comparison of ARIMA and LSTM in forecasting time series. In *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)* (pp. 1394-1401). IEEE.
- Liu, F., & Deng, Y. (2019). A fast algorithm for network forecasting time series. *IEEE Access*, 7, 102554-102560.
- Sen, R., Yu, H.-F., & Dhillon, I. (2019). Think globally, act locally: A deep neural network approach to high-dimensional time series forecasting. arXiv preprint arXiv:1905.03806.
- Rangapuram, S. S., Seeger, M., Gasthaus, J., Stella, L., Wang, Y., & Januschowski, T. (2018). Deep state space models for time series forecasting. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems* (pp. 7796-7805).
- Brownlee, J. (2016). *Machine learning mastery with python*. Machine Learning Mastery Pty Ltd, 527, 100-120.
- Tiao, G. (2001). Time series: ARIMA methods. In N. J. Smelser & P. B. Baltes (Eds.), *International Encyclopedia of the Social Behavioral Sciences* (pp. 15,704–15,709). Oxford: Pergamon. Retrieved from <http://www.sciencedirect.com/science/article/pii/B0080430767005209>
- Vadlamani, R., Dadabada, P., & Deb, K. (2017). Financial time series prediction using hybrids of chaos theory, multi-layer perceptron and multi-objective evolutionary algorithms. *Swarm and Evolutionary Computation*, 36, May.

Wang, Z., Yan, W., & Oates, T. (2017). Time series classification from scratch with deep neural networks: A strong baseline. In 2017 International Joint Conference on Neural Networks (IJCNN) (pp. 1578-1585). IEEE.

ANEXOS

Anexo 1