

08

ANÁLISIS

**DE REDES NEURONALES Y SERIES DE TIEMPO EN ACTIVOS
FINANCIEROS**

ANÁLISIS

DE REDES NEURONALES Y SERIES DE TIEMPO EN ACTIVOS FINANCIEROS

ANALYSIS OF NEURAL NETWORKS AND TIME SERIES IN FINANCIAL ASSETS

Ximena Minshely Guillín-Llanos¹

E-mail: xguillin@uteq.edu.ec

ORCID: <https://orcid.org/0000-0002-7317-5990>

Adolfo Elizondo-Saltos¹

E-mail: aelizondos@uteq.edu.ec

ORCID: <https://orcid.org/0000-0001-7949-8697>

Miriam Patricia Cárdenas-Zea¹

E-mail: mcardenas@uteq.edu.ec

ORCID: <https://orcid.org/0000-0002-8687-5136>

Kevyn Adrián Alcívar-Méndez²

E-mail: adrianalcivar1@hotmail.com

ORCID: <https://orcid.org/0009-0005-0912-3204>

¹ Universidad Técnica Estatal de Quevedo. Ecuador.

² Investigador Independiente. Quevedo. Ecuador.

Cita sugerida (APA, séptima edición)

Guillín-Llanos, X. M., Elizondo-Saltos, A., Cárdenas-Zea, M. P., & Alcívar-Méndez, K. A. (2024). Análisis de redes neuronales y series de tiempo en activos financieros. *Revista Metropolitana de Ciencias Aplicadas*, 7(S1), 85-92.

RESUMEN

Una red neuronal artificial es un modelo matemático que utiliza un sistema de capas internas y externas conectadas a través de estructuras llamadas neuronas, que en conjunto simulan la arquitectura de las conexiones entre neuronas del cerebro humano. Sin embargo, estas redes neuronales han pasado por el proceso de aprendizaje sobre un conjunto de datos conocidos, se convierten en algoritmos capaces de predecir, dentro de un rango previamente establecido, el comportamiento de un conjunto del mismo tipo de datos del que sólo se conocen las etapas previas y no los resultados del comportamiento. Este trabajo tiene como objetivo obtener predicciones del precio de las acciones considerando diferentes tipos de modelos mediante el uso de herramientas computacionales especializadas. Se evidenció que la comparación de diferentes arquitecturas de aprendizaje profundo utilizando el método Naive reveló que los resultados obtenidos por este último fueron mejores que los resultados obtenidos por las arquitecturas de redes neuronales, así como la función FeedForward resultó ser el mejor de los modelos neuronales analizados.

Palabras clave:

Inteligencia artificial, aprendizaje supervisado, redes neuronales, deep-learning.

ABSTRACT

An artificial neural network is a mathematical model that uses a system of internal and external layers connected through structures called neurons, which together simulate the architecture of the connections between neurons in the human brain. However, these neural networks have gone through the process of learning on a set of known data, they become algorithms capable of predicting, within a previously established range, the behavior of a set of the same type of data for which they are only known. the previous stages and not the results of the behavior. This work aims to obtain stock price predictions considering different types of models through the use of specialized computational tools. It was evident that the comparison of different deep learning architectures using the Naive method revealed that the results obtained by the latter were better than the results obtained by neural network architectures, as well as the FeedForward function turned out to be the best of the neural models analyzed.

Keywords:

Artificial intelligence, supervised learning, neural networks, deep learning.

INTRODUCCIÓN

A mediados del siglo XX, el desarrollo de las computadoras se encontraba en auge con el advenimiento de las primeras máquinas electro-mecánicas orientadas a la resolución de problemas. Los distintos aportes de estos artefactos han permitido que, a día de hoy, los grupos de personas, pueden acceder a tecnologías cada vez más útiles para el diario vivir de todos los seres humanos (Arrieta Bechara et al., 2009; Marwala, 2013; Kommula, et al., 2013; LeCun et al., 2015; Bosse, 2020).

Algunas de las industrias en el último siglo han dependido de forma importante del desarrollo y avance de las computadoras, así como de sus distintas tecnologías y aplicaciones; por otro lado, otras compañías se han erguido tomando el desarrollo de las tecnologías como pivote en la oferta de un determinado servicio (Michalski et al., 2013; Cockburn et al., 2017; Ortiz Arango, 2017; Mota-Hernández et al., 2019; Navia-Rodríguez et al., 2020). La realidad es que hoy en día no puede concebirse un modelo de negocios que no tome en consideración a las computadoras y sus herramientas como elementos fundamentales para el desarrollo de productos y servicios, efecto que ha provocado una tremenda subida en el intercambio de bienes y servicios que día a día se lleva a cabo en todas las economías a nivel mundial (Escobar, 2002; Castel & Górriz, 2004; Stankevičienė & Akelaitis, 2014; Wallace, 2017; Vásquez, 2018; Vidal-Castillo et al., 2021).

Una de las herramientas de computación más utilizadas desde el punto de vista académico y empresarial ha sido la inteligencia artificial y sus múltiples ramas por la cantidad de aplicaciones que esta puede tener, que van desde el reconocimiento de voz hasta la predicción de fenómenos naturales (Cockburn et al., 2017; Henrique et al., 2018); en particular el aprovechamiento de las herramientas provistas por la inteligencia artificial en el estudio del comportamiento del mercado financiero es uno de los campos que cuenta con más adeptos, motivo por el cual se han abierto puertas para que muchos académicos y empresarios inviertan tiempo y recursos en la preparación de modelos cada vez más complejos que mejoren los resultados a nivel de comunidad mejorando la calidad de vida de las personas.

Las herramientas de inteligencia artificial para profesionales de los campos de la economía y la administración resultan de vital importancia para países en desarrollo que no han podido automatizar la mayoría de tareas financieras pues supone un desafío para las futuras generaciones desarrollar estas herramientas y proveerlas al sistema financiero tomando como propósito la mejora de la vida de las personas (Johnson & Padilla, 2005; Marwala, 2013).

En este trabajo se estudia, analiza y resume el contenido teórico y algunas técnicas relacionadas con el aprendizaje de máquinas y el deep-learning paso a paso, con

el propósito final de resumir y explicar los fundamentos dentro de la materia que permitan de forma accesible describir y desarrollar un conjunto de herramientas de redes neuronales que en el capítulo final serán usadas para hacer predicciones sobre un conjunto de datos tomados como series de tiempo.

A partir del conjunto de técnicas y métodos estudiados, se busca hacer una aplicación en predicción de precios de acciones sobre conjuntos de datos de libre acceso a través de la utilización de herramientas computacionales relacionadas con el ámbito del Deep-learning que permitan una comparación entre los diferentes métodos desarrollados en la parte teórica.

METODOLOGÍA

Los datos a utilizarse para los experimentos son valores históricos relacionados con el mercado accionario de la Empresa Google entre las que incluyen el valor accionario de apertura, cierre, ajuste y volumen diarios comprendidos entre las fechas del 3 enero del año 2007 y el 29 de junio de 2018. Esta base de datos fue obtenida a partir de la biblioteca “Quantmod”, disponible directamente en R. Dentro de la base de datos obtenida, la variable de interés principal es la del precio de ajuste de la acción puesto que la misma valora de forma adecuada el comportamiento del precio durante el día y a través de procesos directos de la bolsa de valores contiene con mayor precisión el funcionamiento de la empresa. El comportamiento de esta variable de cierre se muestra en la figura 1.



Figura 1. Precio diario ajustado de acciones de Google.

Las variables se tomarán como parte de los datos de entrada en los modelos, las mismas son:

- Precio de Apertura Precio más alto.
- Precio más bajo Precio de cierre.
- Volumen de acciones en el mercado.
- Rezago de un período del precio ajustado Variable binaria de tendencia.

Se considera que la tendencia en los datos Google, por ser una empresa considerada de alto precio de cotización en la bolsa tendrá una mayor volatilidad y que por tanto resultará en precios más variables (Stankevicien & Akelaitis, 2014; Ledezma-Yanez, et al., 2017). Es por ello que para este análisis del funcionamiento de las arquitecturas de redes neuronales como modelos de optimización se tomará en cuenta la base de datos segmentada de forma tal que pueda capturarse las regiones más apropiadas para el funcionamiento de las arquitecturas planteadas.

Como se ha explicado en la sección teórica del presente documento, en el desarrollo de los experimentos a continuación se busca reducir la distancia existente entre el resultado real y la predicción obtenida a través de una loss función que mejor se adapta a las necesidades de este problema en específico; a partir de Henrique, se entiende que la mayoría de autores dedicados a la predicción de valores financieros utilizan (entre otras) como loss function la Mean Absolute Error y la Mean Absolute Percentage Error puesto que tratando de predecir valores reales como lo son los precios y volúmenes, puede directamente medirse el error en la métrica usual (Henrique, Sobreiro, & Kimura, 2018). En este documento, todos los experimentos se llevaron a cabo con las dos métricas antes mencionadas, ponderando los resultados obtenidos a lo largo de las épocas de entrenamiento para obtener un promedio del error obtenido.

DESARROLLO

Como se ha mencionado, lo que se busca es predecir el precio ajustado de una acción, sin embargo, dado que variable como el volumen, tendencia y precios tienen unidades completamente diferentes se requiere hacer una normalización de la base de datos antes de proceder con los cálculos. Esta normalización se hace para cada una de las variables, incluyendo la del precio ajustado de la acción, lo que provoca que el resultado en términos de loss function no pueda traducirse de forma exacta a lenguaje convencional por lo que sería necesaria una conversión sobre los resultados obtenidos.

Esta conversión consiste en multiplicar el valor obtenido como promedio en la loss function por la desviación estándar del conjunto de datos, a partir de lo cual los resultados pueden entenderse en el lenguaje convencional como la cantidad de dólares con la cual estamos fallando en nuestras predicciones:

Dólares de falla = Loss std (P. Ajustado)

Regiones de similar crecimiento

Una primera estrategia de optimización, consiste en tomar las regiones de la base de datos cuyos comportamientos sean similares y separar estos datos como los conjuntos de entrenamiento, validación y prueba respectivamente. Por la tendencia creciente que tiene la base de datos, se disponen de los siguientes períodos con las especificaciones del conjunto de datos seleccionados respectivamente:

1. Conjunto de Entrenamiento. - Los datos comprendidos entre el 6 de noviembre de 2015 y el 1 de mayo de 2017.
2. Conjunto de Validación. - Los datos comprendidos entre el 11 de mayo de 2012 y el 12 de diciembre de 2013.
3. Conjunto de Prueba. - Los datos comprendidos entre el 25 de noviembre de 2008 y el 22 de diciembre del 2009.

Cuyo comportamiento se puede visualizar en la figura 2

Región Estable

Por la complejidad de la predicción del modelo, una selección más adecuada del conjunto de datos necesarios para el entrenamiento y validación, resulta de tomar la región más estable dentro del total de la base de datos. Por esta razón se toman los conjuntos de datos a partir de un subconjunto de la base original, considerando la región más estable, comprendida entre el 22 de diciembre del año 2009 y el 15 de diciembre del año 2012. Este subconjunto de datos puede visualizarse en la figura 3.

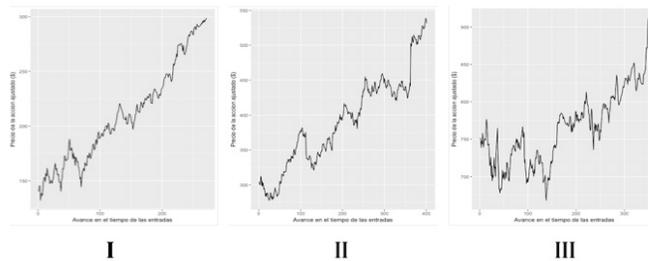


Figura 2. Donde el primer grupo de datos (I) representa el conjunto de prueba; el segundo grupo de datos (II) representa el conjunto de validación y finalmente el tercer grupo de datos (III) representa el conjunto de entrenamiento.

Tomando una línea base: Método Naive

Previo al análisis de los resultados con las arquitecturas de redes neuronales, debe pensarse que un modelo tiene que resultar con mejores características de predicción que aquello que consideremos trivial; es por ello que incorporar un método de predicción trivial como aquel que vaticina al día siguiente el mismo resultado del valor

ajustado de la acción para el día anterior puede ser un buen inicio al momento de medir y comparar el modelo.

Es por ello que, considerando el conjunto de datos de la región estable, se genera una función ca-paz de medir el error de un método de predicción cuya salida sea exactamente igual al valor obtenido el día anterior; este método será incorporado como Naive en la tabla de resultados y permitirá tener una idea más clara de lo que está ocurriendo.

Tendencia del precio diario ajustado de GOOGLE

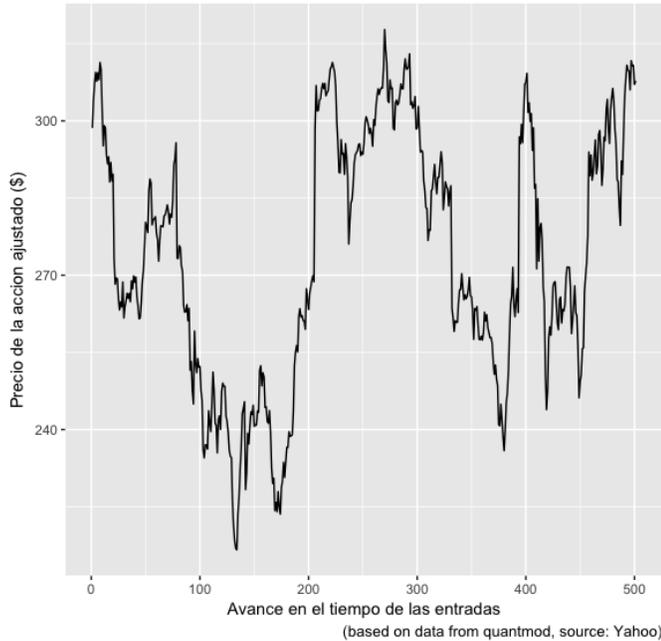


Figure 3. Precio diario ajustado de acciones de Google en la región estable.

Es menester señalar que los precios de las acciones no fluctúan tanto entre un día y otro, sino que sus cambios se evidencian con el paso del tiempo (que es lo que se considera para el entrenamiento) por lo que un modelo que solo repite la predicción del día anterior puede ser mejor en términos de loss function puede ser mejor, pero no resulta práctico para una aplicación en la práctica.

Feed Forward Network

Considerando la primera arquitectura descrita en el presente documento, el primer modelo de predicción es una red completamente conectada (FeedForwardNetwork). En este caso se escoge una capa, con un total de 28 neuronas además de la capa de salida con una sola neurona. Esta arquitectura se escoge considerando otros modelos de algunas áreas en predicción de tiempo y a su vez optimizando la cantidad de neuronas en la capa. Por su naturaleza, en esta arquitectura todos los datos de entrada resultarán como insumos para la arquitectura, es por ello que la dimensión de entrada depende directamente de la cantidad de variables y el rezago en el tiempo.

A través de una repetición constante del modelo, modificando los parámetros externos y propios del modelo, se encuentra que el modelo con los mejores resultados para la arquitectura aquí descrito son los siguientes:

Rezago = 5

Paso = 2

Resultados obtenidos en las regiones de similar crecimiento

Con los parámetros escogidos, se obtiene la siguiente figura 4 de los valores de loss function, a lo largo de las diferentes épocas:

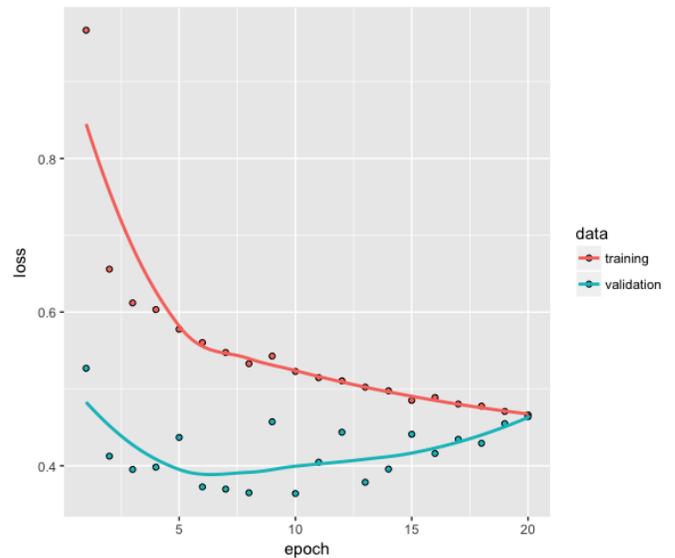


Figure 4. Loss functions de la primera arquitectura en regiones de similar comportamiento.

De este modelo se obtienen valores para la loss de 0.5507, validation loss de 0.4179 y en la evaluación con los datos de prueba un valor de 1.46.

Resultados obtenidos en la región estable

Sobre este conjunto de datos, dado que los valores son continuos se toma como conjunto de entrenamiento y de validación las primeras 400 entradas, en dos conjuntos parejos, cada uno con un total de 200 entradas. De este entrenamiento y de su correspondiente validación se obtienen los resultados que se muestran a continuación:

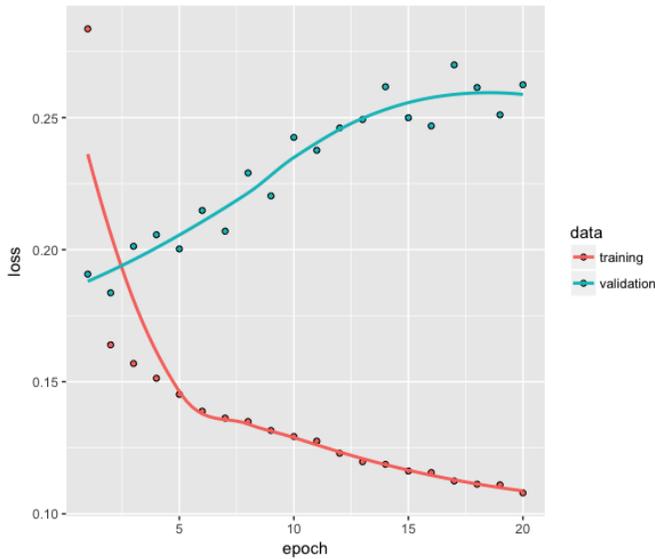


Figure 5. Loss functions de la primera arquitectura en región estable.

Se encuentra entonces que los valores de loss function a lo largo de las etapas se han reducido de forma significativa. Esto se debe a que se encuentra un patrón más repetitivo en el flujo de los datos dentro de esta región estable. No obstante, los valores tienen una tendencia creciente, lo cual es causado por problemas de sobre-entrenamiento en el modelo.

LSTM network

El segundo modelo que se toma en cuenta en el presente documento es una red LSTM como se describió en el capítulo anterior. Se recuerda que este modelo puede ofrecer mejores resultados pues toma en cuenta información del pasado como remanente que se introduce en la predicción del presente. Esta arquitectura contiene una sola capa LSTM interna con un total de 32 neuronas, así como la capa exterior de salida que arroja un out-put de un solo valor.

Se procede de la misma forma que con la arquitectura anterior, modificando uno por uno los parámetros externos, encontrando que los mejores resultados se obtienen a partir de los siguientes parámetros:

Rezago = 4 Paso = 2

Por conveniencia, se usa la base de datos dentro de la región estable como se describió en la arquitectura anterior, tomando los mismos conjuntos de datos para obtener los resultados que se muestran en la figura 6.

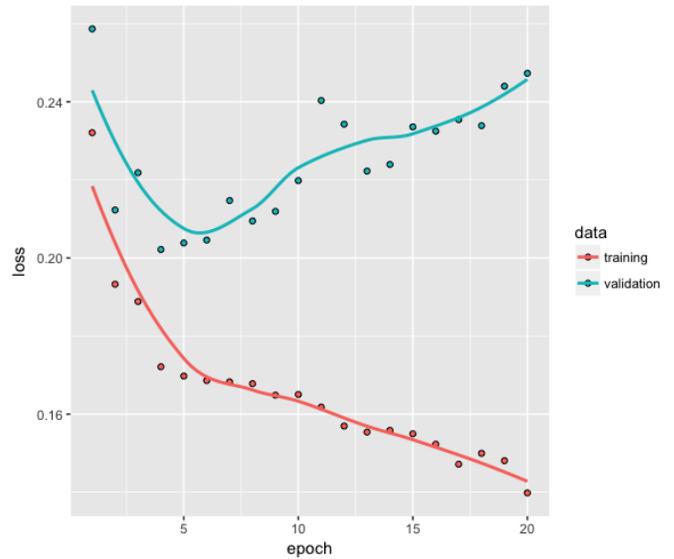


Figure 6. Loss functions de la segunda arquitectura en región estable.

A través del gráfico se puede identificar que los valores obtenidos para las loss functions mantienen un rango de valores cercanos (aunque en promedio más bajos) que los obtenidos por la arquitectura FFN, sin embargo, vale rescatar de ahora los valores no tuvieron un crecimiento exagerado como ocurrió con la arquitectura anterior.

Ligera modificación del modelo anterior

Una de las variantes actuales de la red LSTM es la red GRU, de la cual no se ha hecho detalle dentro del presente documento, pero la cual funciona de forma muy similar a la LSTM. Esta red incluye ciertos componentes de activación o no activación de las memorias de los cálculos de las épocas pasadas mejorando ligeramente la red.

Los parámetros que se utilizan dentro del contexto externo del problema, son similares a los utilizados en la red LSTM en esta misma sección al igual que el conjunto de datos, tomados desde la región estable. Los resultados obtenidos se muestran en la figura 7.

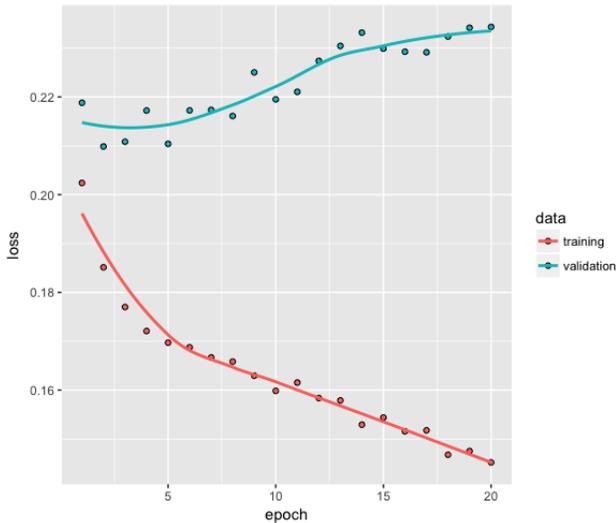


Figure 7. Loss functions de la segunda arquitectura (red GRU) en región estable.

Doble LSTM

En esta arquitectura se introduce el recurrent dropout. Además, con el objetivo de mejorar el procesamiento de los datos se hace uso de dos capas internas del tipo LSTM: Una de ellas con 64 neuronas internas y la segunda con un total de 32 neuronas. Al final de la arquitectura se utiliza la acostumbrada capa de salida.

Una vez más, el proceso de elección de los parámetros externos se hizo probando uno por uno los resultados, con lo que se toma en para este caso las valores como sigue:

$$\text{Rezago} = 4 \text{ Paso} = 2$$

A partir de los cuales se obtiene los resultados que se muestran en la figura 8.

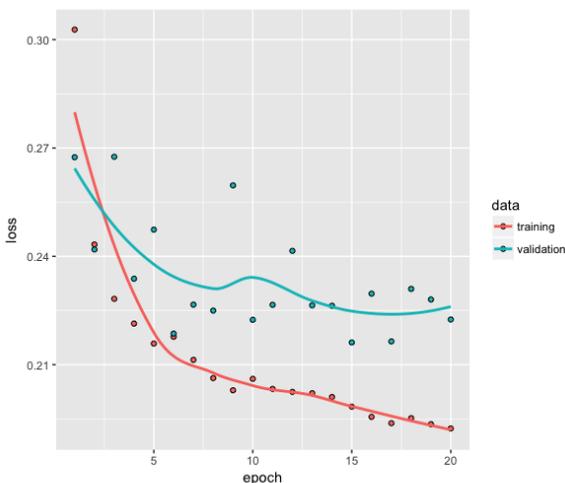


Figura 8. Loss functions de la tercera arquitectura (doble LSTM) en región estable.

Comparación entre modelos de la región estable

A partir de los 4 modelos programados y desarrollados en el presente capítulo que toman en cuenta el conjunto de datos obtenido de la región estable y tomando como referencia el promedio como medida, se obtiene en la tabla 1 de resultados del valor de loss function (entrenamiento y validación) para cada uno de ellos:

Tabla 1. Loss functions para los modelos.

RESULTADOS DE ENTRENAMIENTO Y VALIDACIÓN		
Loss and Validation loss para cada uno de los modelos		
Arquitectura	Loss Function (MAE)	Validation Loss (MAE)
Feed Forward Network	0.1367	0.2315
LSTM	0.1656	0.2253
GRU	0.1629	0.2231
Doble LSTM	0.2116	0.2337

El modelo más adecuado va a ser el que provea mejores resultados en términos de mayor precisión en el vaticinio del precio ajustado de la acción. Sin embargo, se tiene que hacer uso del conjunto de datos prueba para poder medir la precisión de los modelos pues no se puede concluir a partir de la loss function de validación si un modelo es mejor que otro. Como los cuatro modelos de la región estable presentados en este documento ocupan el mismo conjunto de datos de prueba, se puede comparar que tan precisos resultaron ser los modelos, como se detalla en la tabla 2.

Las obtenida a través de predict_generator

Tabla 2. Resultados para todos los modelos.

Arquitectura	Loss Function (MAE) para conjunto de prueba	Valor en dólares (\$)
Feed Forward Network	0.2785	6,78
LSTM	0.3248	7,91
GRU	0.3203	7,80
Doble LSTM	0.3055	7,44
Naive	0.2208	5,37

CONCLUSIONES

La creación de métodos matemáticos y computacionales permitieron predecir el valor de un activo financiero comparando diferentes arquitecturas utilizadas en aprendizaje profundo, así como considerando las operaciones involucradas.

La comparación de diferentes arquitecturas de aprendizaje profundo utilizando el método Naive reveló que los resultados obtenidos por este último fueron mejores que los resultados obtenidos por las arquitecturas de redes neuronales.

Las predicciones realizadas por el método ingenuo se deben a dos hechos básicos que son, los precios entre acciones y las fluctuaciones de precios.

La función FeedForward resultó ser el mejor de los modelos neuronales analizados, porque se utilizaron como variables predictivas los precios de apertura, cierre, máximo y mínimo, permitiendo una arquitectura más robusta.

REFERENCIAS BIBLIOGRÁFICAS

- Arrieta Bechara, J. E., Torres Cruz, J. C., & Velásquez Ceballos, H. (2009). Predicciones de modelos económicos y redes neuronales: el caso de la acción de SU-RAMINV. *Semestre Económico*, 12(25), 95-109.
- Bosse, N. (2020). An Introduction to Deep Learning and the Concept of Regularization. *Reading Processing Applying*, 23.
- Castel, A. G., & Górriz, C. G. (2004). Importancia de los recursos humanos en la productividad de las tecnologías de la información. En J. A. Fraiz Brea y M. Vila Alonso, *La empresa y su entorno: best papers proceedings 2004*. (pp. 969-977). Universidade de Vigo.
- Cockburn, I. M., Henderson, R., & Stern, S. (2018). The impact of artificial intelligence on innovation: An exploratory analysis. En A. Agrawal, J. Gans Y A. Goldfarb, *The economics of artificial intelligence: An agenda*. (pp. 115-146). University of Chicago Press.
- Escobar, V. (2002). Análisis de datos con redes neuronales aplicadas al diagnóstico de la solvencia empresarial. *Cuestiones Económicas* (3), 79-132.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.
- Górriz, C. G., & Castel, A. G. (2004). *Impacto de las Tecnologías de la Información en la Productividad de las Empresas Españolas*. Facultad de Ciencias Económicas y Empresariales.
- Griffin, J., Nardari, F., & Stulz, R. M. (2004). Stock market trading and market conditions. (Working Paper No. 10719). National Bureau of Economic Research. <http://www.nber.org/papers/w10719> doi: 10.3386/w10719
- Henrique, B. M., Sobreiro, V. A., & Kimura, H. (2018). Stock price prediction using support vector regression on daily and up to the minute prices. *The Journal of finance and data science*, 4(3), 183-201.
- Johnson, C. A., & Padilla, M. A. (2005). Regularidades no lineales en índices accionarios. Una aproximación con redes neuronales. *El trimestre económico*, 72(288), 765-821.
- Kommula, VP, Reddy, KO, Shukla, M., Marwala, T. y Rajulu, AV (2013). Caracterización fisicoquímica, de tracción y térmica de hebras de fibras de pasto Napier (nativo africano). *Revista Internacional de Análisis y Caracterización de Polímeros*, 18 (4), 303-314.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436-444.
- Ledezma-Yanez, I., Wallace, W. D. Z., Sebastián-Pascual, P., Climent, V., Feliu, J. M., & Koper, M. (2017). Interfacial water reorganization as a pH-dependent descriptor of the hydrogen evolution rate on platinum electrodes. *Nature Energy*, 2(4), 1-7.
- Michalski, R. S., Carbonell, J. G., & Mitchell, T. M. (Eds.). (2013). *Machine learning: An artificial intelligence approach*. Springer Science & Business Media.
- Mota-Hernández, C. I., Alvarado-Corona, R., & Jiménez García, B. M. (2019). RNA's en el Análisis a Subyacentes de Divisas y Deuda Cotizadas en el MexDer. *Acta Nova*, 9(2), 190-203.
- Navia-Rodríguez, J. R., Cobos-Lozada, C. A., & Mendoza-Becerra, M. E. (2020). Trading Algorítmico para la Predicción de Series de Tiempo Financieras: Una Revisión Sistemática. *Revista Ibérica de Sistemas e Tecnologías de Informação*, (E38), 337-357.
- Ortiz Arango, F. (2017). Pronóstico de precios de petróleo: una comparación entre modelos GARCH y redes neuronales diferenciales. *Investigación económica*, 76(300), 105-126.
- Stankevičienė, J., & Akelaitis, S. (2014). Impact of public announcements on stock prices: Relation between values of stock prices and the price changes in Lithuanian stock market. *Procedia-Social and Behavioral Sciences*, 156, 538-542.
- Vasquez, M. (2018). Predicción de series de tiempo usando un modelo híbrido basado en la descomposición wavelet. *Comunicaciones en Estadística*, 11(2), 257-283.
- Vidal-Castillo, A. A., Venegas-Martínez, F., & Ortiz-Arango, F. (2021). Análisis comparativo de pronósticos del IPC obtenidos mediante modelos GARCH y redes neuronales. *Revista de investigación en ciencias contables y administrativas*, 7(1), 134-166.