



GRADO EN INGENIERÍA EN TECNOLOGÍAS DE TELECOMUNICACIÓN

TRABAJO FIN DE GRADO

Uso de técnicas de machine learning para la
predicción del valor de tipos de cambio en los siete
pares principales y su posterior aplicación al Bitcoin

Autor: Pedro González Schleissner

Director: Atilano Fernández-Pacheco Sánchez-Migallón

Madrid

Declaro, bajo mi responsabilidad, que el Proyecto presentado con el título
Uso de técnicas de machine learning para la predicción del valor de tipos de cambio en los
siete pares principales y su posterior aplicación al Bitcoin
en la ETS de Ingeniería - ICAI de la Universidad Pontificia Comillas en el
curso académico 2022/23 es de mi autoría, original e inédito y
no ha sido presentado con anterioridad a otros efectos.
El Proyecto no es plagio de otro, ni total ni parcialmente y la información que ha sido
tomada de otros documentos está debidamente referenciada.



Fdo.: Pedro González Schleissner Fecha: 4 / 7 / 23

Autorizada la entrega del proyecto

EL DIRECTOR DEL PROYECTO

Fdo.: Atilano Fernández-Pacheco Sánchez-Migallón Fecha: 4 / 7 / 23



COMILLAS

UNIVERSIDAD PONTIFICIA

ICAI

GRADO EN INGENIERÍA EN TECNOLOGÍAS DE TELECOMUNICACIÓN

TRABAJO FIN DE GRADO

Uso de técnicas de machine learning para la
predicción del valor de tipos de cambio en los siete
pares principales y su posterior aplicación al Bitcoin

Autor: Pedro González Schleissner

Director: Atilano Fernández-Pacheco Sánchez-Migallón

Madrid

USO DE TÉCNICAS DE MACHINE LEARNING PARA LA PREDICCIÓN DEL VALOR DE TIPOS DE CAMBIO EN LOS SIETE PARES PRINCIPALES Y SU POSTERIOR APLICACIÓN AL BITCOIN.

Autor: González Schleissner, Pedro.

Director: Fernández-Pacheco Sánchez-Migallón, Atilano.

Entidad Colaboradora: ICAI – Universidad Pontificia Comillas

RESUMEN DEL PROYECTO

El proyecto busca desarrollar modelos de predicción del valor de apertura y de cierre de tipos de cambio. Para ello se emplean en una primera instancia un modelo de regresión lineal, dos modelos autorregresivos y una red LSTM, siendo entrenados por datos procedentes de la serie a predecir. A la vista de los resultados, se decide incluir variables macroeconómicas y financieras que mejoren la predicción. Finalmente, se aplican un modelo de regresión lineal y un modelo LSTM a la serie de Bitcoin, con el objetivo de encontrar similitudes y diferencias entre las divisas tradicionales y las criptomonedas.

Palabras clave: FOREX, predicción, machine learning, Bitcoin

1. Introducción

El proyecto se centra en el desarrollo de modelos de predicción de tipos de cambio, que relacionan las principales divisas del mundo. Se emplearán diferentes técnicas de machine learning para comparar y analizar la capacidad de adaptación de los modelos al problema. Además, se explorará la viabilidad de aplicar los modelos a la predicción del valor de las criptomonedas, como el Bitcoin, estableciendo un paralelismo entre las monedas tradicionales y las alternativas descentralizadas.

El mercado de divisas, conocido como Forex, es el más grande y líquido del mundo, con un volumen diario de transacciones de billones de dólares. La liquidez y la oferta-demanda de las principales divisas son aspectos clave en este mercado. Las ocho divisas más intercambiadas son el dólar estadounidense, el euro, el yen japonés, la libra inglesa, el franco suizo, el dólar canadiense, el dólar australiano y el dólar neozelandés. Además, existen las divisas exóticas que pertenecen a economías emergentes.

Por otro lado, las criptomonedas son un mercado joven y en constante innovación, con una naturaleza descentralizada y seguridad/anonimidad, lo que las posiciona como alternativas al dinero fiduciario controlado por los bancos centrales.

El desarrollo de modelos de predicción de precios en estos mercados es de gran importancia, ya que brinda una herramienta excepcional para anticipar los cambios en los precios y tomar decisiones informadas en el ámbito financiero.

2. Definición del proyecto

El proyecto tiene como objetivo desarrollar modelos para predecir el tipo de cambio de diferentes pares de divisas. Se busca predecir el valor de apertura y cierre para cada intervalo temporal, así como identificar tendencias y puntos de inflexión en las series. También se plantea comparar la predicción de divisas tradicionales y criptomonedas.

La metodología en cascada o waterfall se utilizará para el desarrollo del proyecto, dividiéndolo en seis etapas: requisitos, análisis, diseño, codificación, prueba y comparación de resultados. En caso de que los resultados no sean satisfactorios, se realizará una nueva iteración de la metodología.

Tras la selección de los datos, se procederá a su depuración para que éstos puedan ser utilizados por los modelos. Concretamente, se entrenará un modelo de regresión lineal, dos modelos autorregresivos y una red LSTM.

Se llevará a cabo, posteriormente, una segunda iteración donde un modelo lineal y una red LSTM serán entrenados con variables financieras y macroeconómicas, con el objetivo de mejorar el poder predictivo de los modelos.

3. Descripción del modelo/sistema/herramienta

La *Figura 1* muestra la arquitectura del sistema que se busca desarrollar. A partir de unas bases de datos, de las cuales se extrae la información, se alimenta la fase de procesamiento de los datos. Una vez éstos estén limpios, se emplearán para el entrenamiento de un modelo que predecirá el valor de apertura del intervalo y otro que predecirá el valor de cierre.

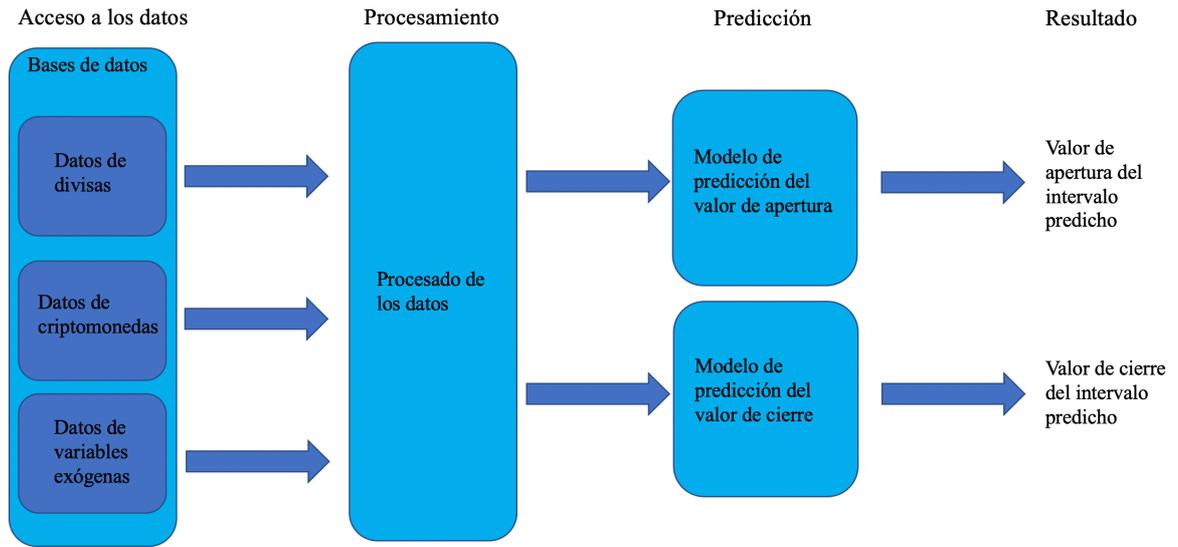


Figura 1: Diagrama de la arquitectura desarrollada

(Fuente: Elaboración propia)

El primer modelo es un modelo de regresión lineal en el cual se emplean todas las variables de la serie de divisas, salvo la variable a predecir, en el entrenamiento del modelo. A continuación, se desarrolla un modelo autorregresivo empleando una regresión lineal, en la cual la única variable predictiva es la variable que predecir, pero del instante temporal previo. El tercer modelo, también autorregresivo, se desarrolla empleando una librería especializada que permite incluir integración y media móvil en el modelo. Concretamente, este tipo de modelo es conocido como ARIMA y tendrá como variable independiente distintos retrasos de la variable dependiente. Finalmente, se desarrollará una red LSTM, un tipo de red neuronal recurrente, la cual empleará las 14 últimas observaciones de la variable dependiente como predictores.

Con el objetivo de mejorar los resultados, se incluyen variables de tipos de interés, inflación, índices de los principales mercados y del precio del oro. Éstas se emplean en dos modelos de regresión lineal, el primero empleando solo las nuevas variables y el segundo usando todas las variables disponibles. El procedimiento se repite con la red LSTM.

Por último, los modelos que mejor desempeño hayan demostrado serán empleados para la predicción de los valores de apertura y cierre de la serie de Bitcoin.

4. Resultados

Los resultados de los ocho modelos desarrollados a lo largo del proyecto aparecen resumidos en la *Tabla 1*.

Iteración	Modelo	Resultados
Iteración 1	Regresión lineal	Resultados erróneos en la predicción del valor de apertura. El modelo de cierre obtiene R2 altos, pero las predicciones son réplicas de la observación anterior.
	Autorregresivo I	Buenos R2, pero las predicciones son réplicas de la observación anterior.
	Autorregresivo II	Buenos R2, pero las predicciones son réplicas de la observación anterior.
	LSTM	Buenos R2, pero las predicciones son réplicas de la observación anterior.
Iteración 2	Regresión lineal (Todas las variables)	Buenos R2, pero las predicciones son réplicas de la observación anterior.
	Regresión lineal (Solo variables exógenas)	El modelo no es capaz de ajustar la predicción a la serie real empleando únicamente variables macroeconómicas y financieras.
	LSTM (Todas las variables)	Buenos R2, pero las predicciones son réplicas de la observación anterior.
	LSTM (Solo variables exógenas)	El ajuste del modelo es notablemente peor que en el modelo con todas las variables, pero es capaz de predecir correctamente la tendencia en determinados tramos de la serie.

Tabla 1: Resumen de resultados

(Fuente: Elaboración propia)

Los resultados han quedado marcados por un factor. Cualquier modelo que incluyese datos de la propia serie a predecir ha conseguido ajustes altos, pero los resultados del modelo no eran predicciones útiles, sino el valor de la observación previa. Es decir, estos modelos no aportaban ninguna información predictiva.

Sin embargo, los modelos que emplearon variables exógenas únicamente consiguieron modelar tendencias en ciertos puntos de las series, siendo este el mejor resultado obtenido.

Para la predicción de las series de Bitcoin, se decidió emplear dos modelos de la segunda iteración. El primer modelo era la regresión lineal con todas las variables, ya que éste es similar al resto de modelos que emplean variables de la serie. El segundo modelo fue la red LSTM con variables exógenas, debido a que este modelo ha sido el único que ha conseguido obtener predicciones reales.

Los resultados obtenidos son similares a los resultados de las monedas tradicionales. El modelo que empleaba variables de la serie ha replicado los valores de la observación previa en la predicción y el modelo LSTM ha tenido un peor ajuste, pero ha sido capaz de modelar correctamente algunas tendencias.

5. Conclusiones

En resumen, el estudio ha analizado diferentes modelos de predicción de activos financieros, centrándose en las criptomonedas y las monedas tradicionales. Se ha observado que los modelos que mejor ajuste han mostrado se comportaban de forma similar a modelos de Random Walk, que se basan en sumar una perturbación aleatoria al último valor. Por otro lado, los modelos LSTM, que utilizan variables exógenas en lugar de los datos de la propia divisa, han logrado modelar tendencias y puntos de inflexión en las series, aunque con una fiabilidad baja debido a las predicciones erróneas.

Se ha encontrado que las criptomonedas se comportan de manera similar a las series tradicionales, pero con una mayor volatilidad. Esto ha empeorado los resultados de los modelos, tanto en las criptomonedas como en las monedas tradicionales con alta volatilidad.

Se concluye que la creación de modelos precisos y capaces de anticipar eventos requiere el uso de variables financieras y macroeconómicas. Se sugiere que futuros trabajos en este campo deben enfocarse en la selección adecuada de variables predictivas, teniendo en cuenta la situación macroeconómica, política y social que afecta a las dos divisas del par. Además, se sugiere considerar el uso de noticias y el sentimiento que se encuentra en ellas, especialmente en el caso de las criptomonedas, cuya volatilidad a menudo está relacionada con noticias o mensajes en redes sociales de personas influyentes.

6. Referencias

Chitenderu, T., Maredza, A., & Sibanda, K. (2014). *The Random Walk Theory And Stock Prices: Evidence From Johannesburg Stock Exchange*.

Forex Market Trading Online. (2023). Obtenido de Forex Early Warning:
<https://www.forexearlywarning.com/forex-trading-course/forex-market-trading>

Hochreiter, S., & Schmidhuber, J. (1997). *LONG SHORT-TERM MEMORY*.

MESSE, R., & ROGOFF, K. (1983). *EMPIRICAL EXCHANGE RATE MODELS OF THE SEVENTIES*.

Tang, Y., Song, Z., Zhu, Y., Yuan, H., Hou, M., Ji, J., . . . Li, J. (2022). *A survey on machine learning models for financial time series forecasting*.

YEŞİN, P. (2016). *Exchange Rate Predictability and State-of-the-Art Models*.

USE OF MACHINE LEARNING TECHNIQUES FOR PREDICTING THE VALUE OF EXCHANGE RATES IN THE SEVEN MAJOR PAIRS AND THEIR SUBSEQUENT APPLICATION TO BITCOIN.

Author: González Schleissner, Pedro.

Supervisor: Fernández-Pacheco Sánchez-Migallón, Atilano.

Collaborating Entity: ICAI – Universidad Pontificia Comillas

ABSTRACT

The project seeks to develop models for the prediction of the opening and closing values of exchange rates. For this purpose, a linear regression model, two autoregressive models and an LSTM network are used in the first instance, being trained by data from the series to be predicted. In view of the results, it is decided to include macroeconomic and financial variables to improve the prediction. Finally, a linear regression model and a LSTM model are applied to the Bitcoin series, with the aim of finding similarities and differences between traditional currencies and cryptocurrencies.

Keywords: FOREX, prediction, machine learning, Bitcoin

1. Introduction

The project focuses on the development of exchange rate prediction models, which relate the main currencies of the world.

Different machine learning techniques will be used to compare and analyze the adaptability of the models to the problem. In addition, the feasibility of applying the models to the prediction of the value of cryptocurrencies, such as Bitcoin, will be explored, drawing a parallel between traditional currencies and decentralized alternatives.

The foreign exchange market, known as Forex, is the largest and most liquid market in the world, with a daily trading volume of trillions of dollars. Liquidity and supply-demand of major currencies are key aspects of this market. The eight most traded currencies are the US dollar, the euro, the Japanese yen, the British pound, the Swiss franc, the Canadian dollar, the Australian dollar, and the New Zealand dollar. In addition, there are exotic currencies belonging to emerging economies.

On the other hand, cryptocurrencies are a young and constantly innovating market, with a decentralized nature and security/anonymity, which positions them as alternatives to fiat money controlled by central banks.

The development of price prediction models in these markets is of great importance, as it provides an exceptional tool to anticipate price changes and make informed financial decisions.

2. Definition of the project

The project aims to develop models to predict the exchange rate of different currency pairs. It seeks to predict the opening and closing value for each time interval, as well as to identify trends and inflection points in the series. It is also proposed to compare the prediction of traditional currencies and cryptocurrencies.

The waterfall methodology will be used for the development of the project, dividing it into six stages: requirements, analysis, design, coding, testing and comparison of results. In case the results are not satisfactory, a new iteration of the methodology will be performed.

After the selection of the data, it will be debugged so that it can be used by the models. Specifically, a linear regression model, two autoregressive models and an LSTM network will be trained.

Subsequently, a second iteration will be carried out where a linear model and an LSTM network will be trained with financial and macroeconomic variables, with the aim of improving the predictive power of the models.

3. Description of the model/system/tool

Figure 2 shows the architecture of the system to be developed. The data processing phase is fed from databases, from which the information is extracted. Once the data are clean, they will be used to train a model that will predict the opening value of the interval and another one that will predict the closing value.

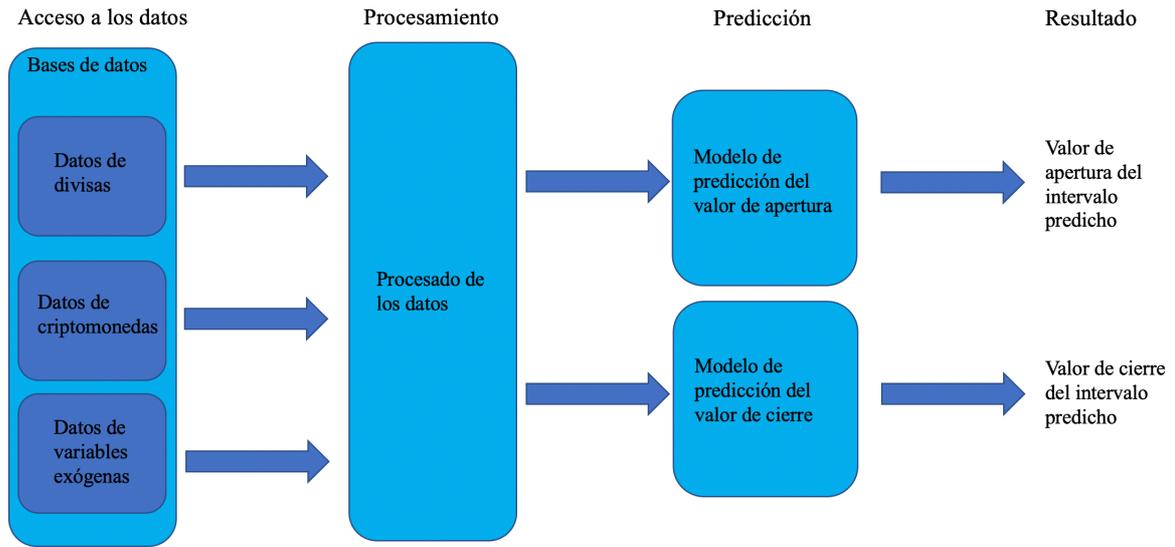


Figura 2: Diagram of the developed architecture

(Source: Own elaboration)

The first model is a linear regression model in which all the variables of the currency series, except the variable to be predicted, are used in the training of the model. Next, an autoregressive model is developed using linear regression, in which the only predictive variable is the variable to be predicted, but from the previous time instant. The third model, also autoregressive, is developed using a specialized library that allows integration and moving average to be included in the model. Specifically, this type of model is known as ARIMA and will have as independent variable different lags of the dependent variable. Finally, a LSTM network, a type of recurrent neural network, will be developed, which will use the last 14 observations of the dependent variable as predictors.

4. Results

The results of the eight models developed throughout the project are summarized in *Tabla 2*.

Iteración	Modelo	Resultados
Iteration 1	Lineal regression	Erroneous results in the prediction of the opening value. The closure model obtains high R2, but the predictions are replicas of the previous observation.

	Autoregressive I	Good R2, but predictions are replicas of the previous observation.
	Autoregressive II	Good R2, but predictions are replicas of the previous observation.
	LSTM	Good R2, but predictions are replicas of the previous observation.
Iteration 2	Lineal regression (All variables)	Good R2, but predictions are replicas of the previous observation.
	Lineal regression (Exogenous variables only)	The model is not able to fit the prediction to the real series using only macroeconomic and financial variables.
	LSTM (All variables)	Good R2, but predictions are replicas of the previous observation.
	LSTM (Exogenous variables only)	The model fit is noticeably worse than in the model with all variables, but it is able to correctly predict the trend in certain parts of the series.

Tabla 2: Summary of results

(Source: Own elaboration)

The results have been marked by one factor. Any model that included data from the series itself to be predicted achieved high fits, but the model results were not useful predictions, but the value of the previous observation. In other words, these models did not provide any predictive information.

However, the models that used exogenous variables only managed to model trends at certain points in the series, this being the best result obtained.

For the prediction of the Bitcoin series, it was decided to use two models from the second iteration. The first model was linear regression with all variables, as this is similar to the other models using series variables. The second model was the LSTM network with

exogenous variables, since this model was the only one that managed to obtain real predictions.

The results obtained are similar to the results for traditional currencies. The model that used variables from the series replicated the values of the previous observation in the prediction and the LSTM model had a worse fit but was able to correctly model some trends.

5. Conclusions

In summary, the study has analyzed different prediction models for financial assets, focusing on cryptocurrencies and traditional currencies. It has been observed that the models that showed the best fit behaved similarly to Random Walk models, which are based on adding a random perturbation to the last value. On the other hand, LSTM models, which use exogenous variables instead of the data of the currency itself, have managed to model trends and inflection points in the series, albeit with low reliability due to erroneous predictions.

Cryptocurrencies have been found to behave similarly to traditional series, but with higher volatility. This has worsened model results for both cryptocurrencies and traditional currencies with high volatility.

It is concluded that the creation of accurate models capable of anticipating events requires the use of financial and macroeconomic variables. It is suggested that future work in this field should focus on the proper selection of predictive variables, taking into account the macroeconomic, political and social situation affecting the two currencies in the pair. In addition, it is suggested to consider the use of news and the sentiment found in them, especially in the case of cryptocurrencies, whose volatility is often related to news or messages in social networks of influential people.

6. References

Chitenderu, T., Maredza, A., & Sibanda, K. (2014). *The Random Walk Theory And Stock Prices: Evidence From Johannesburg Stock Exchange*.

Forex Market Trading Online. (2023). Obtenido de Forex Early Warning: <https://www.forexearlywarning.com/forex-trading-course/forex-market-trading>

Hochreiter, S., & Schmidhuber, J. (1997). *LONG SHORT-TERM MEMORY*.

MESSE, R., & ROGOFF, K. (1983). *EMPIRICAL EXCHANGE RATE MODELS OF THE SEVENTIES*.

Tang, Y., Song, Z., Zhu, Y., Yuan, H., Hou, M., Ji, J., . . . Li, J. (2022). *A survey on machine learning models for financial time series forecasting*.

YEŞİN, P. (2016). *Exchange Rate Predictability and State-of-the-Art Models*.

ÍNDICE DE LA MEMORIA

ÍNDICE DE FIGURAS	III
ÍNDICE DE TABLAS	VI
CAPÍTULO 1. INTRODUCCIÓN	8
1.1 MOTIVACIÓN DEL PROYECTO.....	8
CAPÍTULO 2. DESCRIPCIÓN DE LAS TECNOLOGÍAS	12
2.1 PYTHON.....	12
2.2 VISUAL STUDIO	13
2.3 KERAS	14
2.4 SCIKIT-LEARN	15
2.5 STATSMODELS.....	16
CAPÍTULO 3. ESTADO DE LA CUESTIÓN	18
3.1 MODELOS ACTUALES	18
3.2 SERIES TEMPORALES	20
3.3 MODELOS A EMPLEAR	21
3.3.1 Regresión lineal	22
3.3.2 Modelos autorregresivos.....	22
3.3.3 Redes neuronales	23
CAPÍTULO 4. DEFINICIÓN DEL TRABAJO	30
4.1 JUSTIFICACIÓN	30
4.2 OBJETIVOS	31
4.3 METODOLOGÍA	32
4.4 PLANIFICACIÓN Y ESTIMACIÓN ECONÓMICA	34
CAPÍTULO 5. SISTEMA/MODELO DESARROLLADO	37
5.1 REQUISITOS	38
5.2 ANÁLISIS.....	43
5.3 DISEÑO.....	52

5.3.1	Regresión lineal	52
5.3.2	Autorregresivo I.....	53
5.3.3	Autorregresivo II.....	54
5.3.4	LSTM.....	54
5.4	PRUEBA	55
5.4.1	Regresión lineal	55
5.4.2	Autorregresivo I.....	58
5.4.3	Autorregresivo II.....	61
5.4.4	LSTM.....	63
5.5	COMPARACIÓN DE RESULTADOS	67
5.6	REQUISITOS II	68
5.7	ANÁLISIS II	74
5.8	DISEÑO II.....	80
5.8.1	Regresión lineal	80
5.8.2	LSTM.....	81
5.9	PRUEBA II	81
5.9.1	Regresión lineal	81
5.9.2	LSTM.....	83
5.10	COMPARACIÓN DE RESULTADOS II.....	85
5.11	PRUEBA DE LOS MEJORES MODELOS EN LA SERIE DE BITCOIN	86
CAPÍTULO 6.	ANÁLISIS DE RESULTADOS	91
CAPÍTULO 7.	CONCLUSIONES Y TRABAJOS FUTUROS.....	93
CAPÍTULO 8.	BIBLIOGRAFÍA	95
ANEXO I:	ALINEACIÓN DEL PROYECTO CON LOS ODS.....	96

Índice de figuras

Figura 1: Diagrama de la arquitectura desarrollada.....	9
Figura 2: Diagram of the developed architecture	15
Figura 3: Volumen de los principales mercados financieros.....	9
Figura 4: Logo de Python	12
Figura 5: Logo de Visual Studio.....	13
Figura 6: Logo de Keras	14
Figura 7: Logo de SciKit Learn.....	15
Figura 8: Logo de Statsmodels	16
Figura 9: Métodos de predicción de series temporales financieras	21
Figura 10: Estructura de un perceptrón simple.....	23
Figura 11: Estructura del perceptrón multicapa	26
Figura 12: Estructura de una red neuronal recurrente	27
Figura 13: Estructura de una unidad LSTM	28
Figura 14: Diagrama de la arquitectura desarrollada (Fuente: Elaboración propia)	38
Figura 15: Serie diaria de EURUSD (Fuente: Elaboración Propia)	44
Figura 16: Serie diaria de USDJPY (Fuente: Elaboración Propia)	45
Figura 17: Serie diaria de GBPUSD (Fuente: Elaboración Propia)	45
Figura 18: Serie diaria de USDCHF (Fuente: Elaboración Propia)	46
Figura 19: Serie diaria de AUDUSD (Fuente: Elaboración Propia)	46
Figura 20: Serie diaria de USDCAD (Fuente: Elaboración Propia).....	47
Figura 21: Serie diaria de NZDUSD (Fuente: Elaboración Propia).....	47
Figura 22: Serie semanal de EURUSD (Fuente: Elaboración Propia)	48
Figura 23: Serie semanal de USDJPY (Fuente: Elaboración Propia)	48
Figura 24: Serie semanal de GBPUSD (Fuente: Elaboración Propia)	49
Figura 25: Serie semanal de USDCHF (Fuente: Elaboración Propia)	49
Figura 26: Serie semanal de AUDUSD (Fuente: Elaboración Propia)	50
Figura 27: Serie semanal de USDCAD (Fuente: Elaboración Propia).....	50
Figura 28: Serie semanal de NZDUSD (Fuente: Elaboración Propia).....	51

Figura 29: Serie diaria de Bitcoin (Fuente: Elaboración Propia)	51
Figura 30: Serie semanal de Bitcoin (Fuente: Elaboración Propia)	52
Figura 31: Gráfico de autocorrelación de la serie euro/dólar diario (Fuente: Elaboración propia).....	53
Figura 32: Serie real vs serie predicha para el valor de cierre del par dólar/franco suizo semanal (Fuente: Elaboración propia).....	58
Figura 33: Serie real vs serie predicha para el valor de apertura del par euro/dólar estadounidense semanal (Fuente: Elaboración propia)	60
Figura 34: Zoom de la serie real vs serie predicha para el valor de apertura del par dólar neozelandés/dólar estadounidense semanal (Fuente: Elaboración propia).....	63
Figura 35: Serie real vs serie predicha para el par dólar estadounidense /dólar canadiense semanal (Fuente: Elaboración propia).....	65
Figura 36: Evolución del tipo de interés del Banco Central Europeo (Fuente: Elaboración propia).....	76
Figura 37: Evolución del tipo de interés de la Reserva Federal (Fuente: Elaboración propia)	76
Figura 38: Evolución del IPC en Estados Unidos (Fuente: Elaboración propia)	77
Figura 39: Evolución del IPC en Europa (Fuente: Elaboración propia).....	77
Figura 40: Evolución del índice del NASDAQ (Fuente: Elaboración propia).....	78
Figura 41: Evolución del índice del NYSE (Fuente: Elaboración propia)	78
Figura 42: Evolución del índice del SSE (Fuente: Elaboración propia).....	79
Figura 43: Evolución del precio del oro (Fuente: Elaboración propia).....	79
Figura 44: Serie real vs serie predicha con regresión lineal para el valor de cierre del par franco suizo/dólar estadounidense (Fuente: Elaboración propia).....	82
Figura 45: Zoom de la serie real vs serie predicha con regresión lineal para el valor de cierre del libra esterlina/dólar estadounidense (Fuente: Elaboración propia)	83
Figura 46: Serie real vs serie predicha con LSTM para el valor de cierre del par franco suizo/dólar estadounidense (Fuente: Elaboración propia).....	84
Figura 47: Serie real vs serie predicha con LSTM para el valor de cierre del par dólar australiano/dólar estadounidense (Fuente: Elaboración propia).....	85

Figura 48: Serie real vs serie predicha con modelo de regresión lineal para el valor de cierre de la serie semanal de Bitcoin (Fuente: Elaboración propia)	87
Figura 49: Serie real vs serie predicha con modelo de regresión lineal para el valor de cierre de la serie diaria de Bitcoin (Fuente: Elaboración propia)	88
Figura 50: Serie diaria de Bitcoin (Fuente: Elaboración Propia)	89
Figura 51: Serie real vs serie predicha con LSTM para el valor de cierre de la serie diaria de Bitcoin (Fuente: Elaboración propia)	89
Figura 52: Serie real vs serie predicha con LSTM para el valor de apertura de la serie diaria de Bitcoin (Fuente: Elaboración propia)	90

Índice de tablas

Tabla 1: Resumen de resultados	10
Tabla 2: Summary of results.....	16
Tabla 3: Los siete principales pares de divisas y los pares exóticos más comunes (Fuente: Elaboración propia)	10
Tabla 4: Descripción de los conjuntos de datos (Fuente: Elaboración propia).....	40
Tabla 5: Descripción de las variables contenidas en las series (Fuente: Elaboración propia)	42
Tabla 6: Equivalencias entre símbolos de unidades y su correspondiente orden de magnitud (Fuente: Elaboración propia).....	43
Tabla 7: Resultados del modelo de regresión lineal para el valor de apertura (Fuente: Elaboración propia)	56
Tabla 8: Resultados del modelo de regresión lineal para el valor de cierre (Fuente: Elaboración propia)	57
Tabla 9: Resultados del modelo autorregresivo I para el valor de apertura (Fuente: Elaboración propia)	59
Tabla 10: Resultados del modelo autorregresivo I para el valor de cierre (Fuente: Elaboración propia).....	60
Tabla 11: Resultados del modelo autorregresivo II para el valor de apertura (Fuente: Elaboración propia)	61
Tabla 12: Resultados del modelo autorregresivo II para el valor de cierre (Fuente: Elaboración propia)	62
Tabla 13: Resultados del modelo LSTM para el valor de apertura (Fuente: Elaboración propia).....	64
Tabla 14: Resultados del modelo LSTM para el valor de cierre (Fuente: Elaboración propia)	65
Tabla 15: Descripción de los conjuntos de datos de variables exógenas (Fuente: Elaboración propia).....	70

Tabla 16: Explicación de las variables de los conjuntos de datos exógenos (Fuente: Elaboración Propia) 74

Capítulo 1. INTRODUCCIÓN

El proyecto plantea el desarrollo de modelos capaces de predecir el valor de los tipos de cambio que relacionan a las divisas más importantes del mundo. Se emplearán y compararán diferentes tipos de modelos, empleando diferentes técnicas de machine learning. Posteriormente, se comparará su capacidad de adaptarse al problema y se analizarán posibles formas de mejorar el rendimiento de los modelos en caso de que sea necesario. Finalmente, se estudiará la viabilidad de aplicar los modelos obtenidos a la predicción del valor de las criptomonedas, concretamente el Bitcoin, tratando de trazar un paralelismo entre las monedas tradicionales y aquellas que se posicionan como alternativa a estas.

1.1 MOTIVACIÓN DEL PROYECTO

El mundo financiero está compuesto de numerosos mercados, en los cuales se intercambian todo tipo de bienes y activos financieros, pero ninguno de ellos se acerca al mercado de divisas en términos de volumen.

El mercado de divisas, también conocido como Forex, del término inglés *Foreign Exchange*, es el mercado financiero más grande y líquido del mundo. Permite a individuos, empresas y gobiernos convertir una moneda en otra. A diferencia de otros mercados financieros, el mercado de divisas opera las 24 horas del día de lunes a viernes. Esta accesibilidad sumada a los bajos costes de operar en este mercado hace que de media en 2019 se realizasen transacciones por valor de 5 billones de dólares estadounidenses al día, en comparación con los 20 mil millones de dólares diarios que alcanzaban los mayores mercados de acciones del mundo.

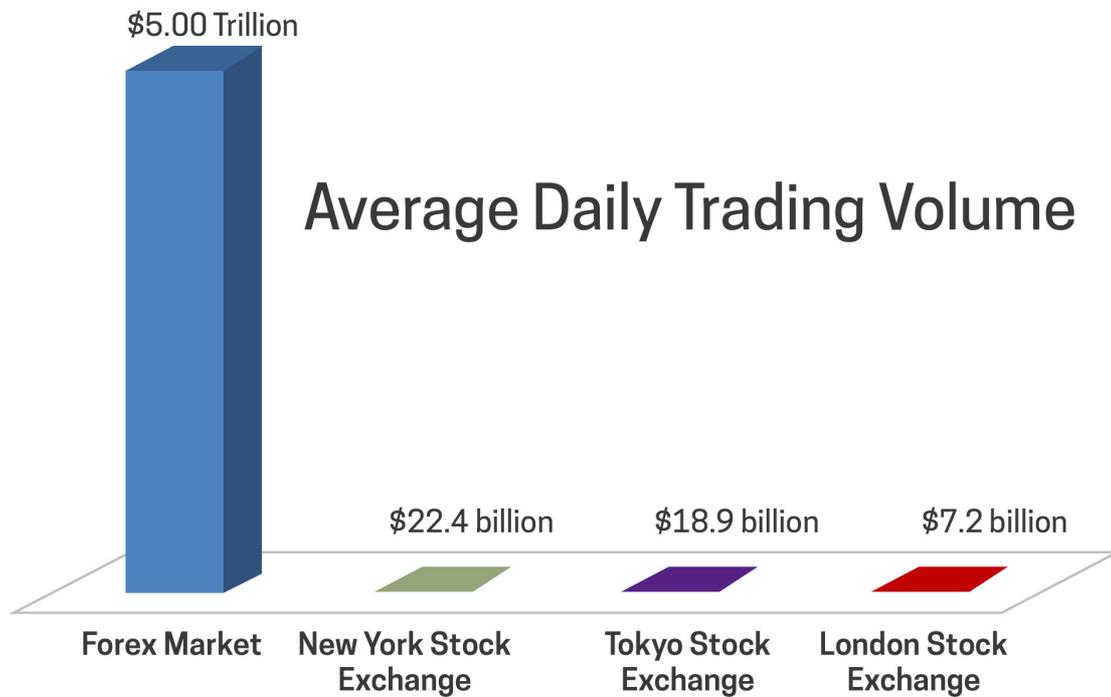


Figura 3: Volumen de los principales mercados financieros

(Fuente: (Forex Market Trading Online, 2023))

Para comprender mejor el mercado de divisas es necesario entender los activos con los que se opera. Estos activos son las monedas de los distintos países y son clasificadas en función de la liquidez que tienen en los mercados, es decir el volumen de transacciones que se realizan con ellas. Este factor es importante ya que una alta liquidez garantiza una alta oferta y demanda de las principales divisas, garantizando que las operaciones de cambio van a poder llevarse a cabo en cualquier momento. Las ocho divisas más intercambiadas, conocidas como *major currencies*, son: el dólar estadounidense (USD), el euro (EUR), el yen japonés (JPY), la libra inglesa (GBP), el franco suizo (CHF), el dólar canadiense (CAD), el dólar australiano (AUD) y el dólar neozelandés (NZD). El USD es la divisa más intercambiada de todas, siendo utilizada en un 91% de todas las operaciones [2], por lo que

los 7 principales pares de divisas son las resultantes de las combinaciones del USD con las 7 divisas restantes.

Otro grupo de divisas es el conocido como divisas exóticas. Estas son aquellas divisas no incluidas en las principales, pertenecientes a economías emergentes. La liquidez de este grupo varía dependiendo de la moneda en cuestión, desde monedas muy líquidas hasta poco líquidas. La pareja formada por una divisa principal y una exótica se conoce como pares de divisas exóticos.

Los pares principales y los pares exóticos más usados son los mostrados en la *Tabla 3*.

Pares principales	Pares exóticos
EUR/USD	EUR/TRY
USD/JPY	USD/ZAR
GBP/USD	USD/MXN
USD/CHF	USD/CZK
AUD/USD	EUR/HKD
USD/CAD	
NZD/USD	

Tabla 3: Los siete principales pares de divisas y los pares exóticos más comunes (Fuente: Elaboración propia)

Dejando de lado un mercado tradicional, como es el de divisas, pasamos a otro relativamente joven y en constante innovación: el de las criptomonedas. Fenómeno que hace apenas una

década era desconocido para la mayoría de las personas ahora forma parte de la economía mundial y se prevé que su importancia solo vaya en aumento conforme pasen los años. Su naturaleza descentralizada, es decir, que no están sujetas a regulaciones ni al control de bancos centrales, además de su seguridad y anonimidad, las ha posicionado como una firme alternativa al dinero FIAT, aquel cuyo valor se encuentra respaldado por el banco central que lo emite en vez de por una commodity o materia prima como el oro.

Observamos pues la importancia que ambos mercados tienen en los mercados y en la economía global, por lo que un modelo que sea capaz de indicar como van a cambiar los precios del mercado es una herramienta excepcional.

Capítulo 2. DESCRIPCIÓN DE LAS TECNOLOGÍAS

2.1 PYTHON

El desarrollo de los modelos se realizará con Python, un lenguaje de programación orientado a objetos, siendo en la actualidad uno de los lenguajes más populares y extendidos debido a su facilidad, gran documentación y a su amplia comunidad de desarrolladores que contribuyen de forma constante con nuevas librerías y herramientas de gran utilidad. Además, el estado del arte en materia de modelos se ha realizado en gran parte usando este lenguaje.



Figura 4: Logo de Python

(Fuente: <https://www.python.org/community/logos/>)

Para el desarrollo, la compilación y la depuración del código en Python, se empleará Visual Studio. Dentro de Python se usarán dos librerías para la construcción de los modelos: Keras y Scikit-learn, la primera siendo usada para realizar complejos modelos como redes

neuronales, mientras que la segunda se empleará para la creación de regresiones lineales, modelos que en comparación son más simples.

2.2 VISUAL STUDIO

Visual Studio es un software o programa conocido como un entorno de desarrollo integrado o IDE, el cual proporciona las herramientas necesarias para la creación y el desarrollo completo de programas. Las principales funcionalidades que cumple son la incorporación de un editor de texto, de un compilador y que permite depurar y probar el código. En este caso de uso las dos últimas funciones no son imprescindibles, pero Visual Studio nos permitía tener todas las funcionalidades en un único entorno y agilizaba de forma significativa el proceso de creación de código mediante su funcionalidad de autocompletar el código.

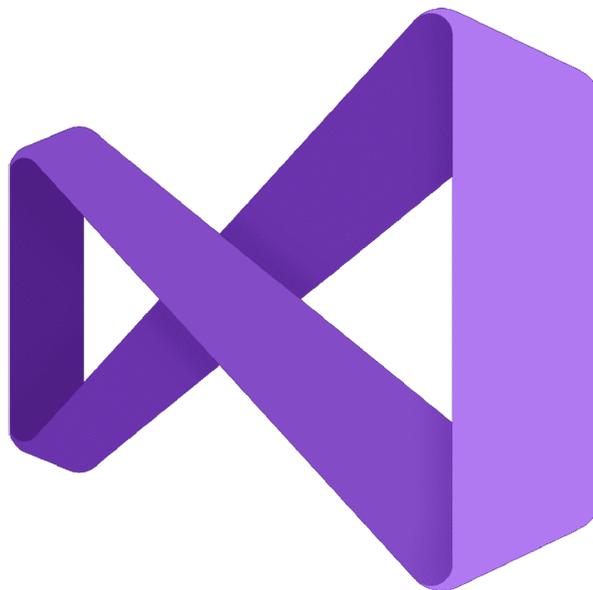


Figura 5: Logo de Visual Studio

(Fuente: <https://1000marcas.net/visual-studio-logo/>)

Para este proyecto, se ha decidido que el código se desarrollará mediante Notebooks de Python, tipo de archivo el cual difiere ligeramente de un programa tradicional de Python. En un Notebook, el código se divide en celdas cuya compilación es individual al resto del código, pudiendo llevar a cabo un enfoque mucho más modular.

2.3 KERAS

Keras es una biblioteca escrita en Python y creada para el diseño e implementación de redes neuronales, la cual basa su funcionamiento en librerías de nivel más bajo como TensorFlow, que emplea tensores para llevar a cabo los cálculos, obteniendo mejoras en el rendimiento.

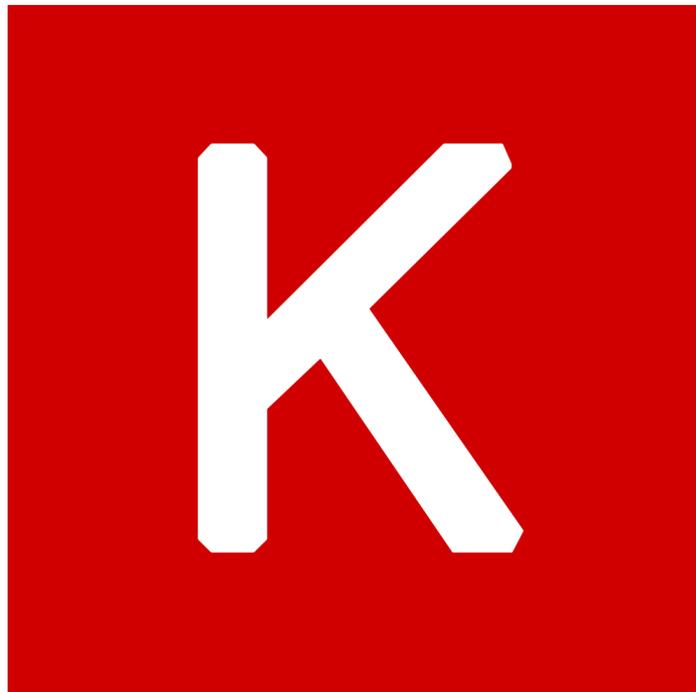


Figura 6: Logo de Keras

(Fuente: https://commons.wikimedia.org/wiki/File:Keras_logo.svg)

Se ha elegido esta librería debido a su interfaz de alto nivel que permite crear y entrenar modelos de aprendizaje profundo de manera sencilla, así como su extensiva documentación por parte de los creadores como por parte de la comunidad. Su extensivo uso ha permitido que los modelos creados con esta librería se encuentren entre los modelos del estado del arte en numerosas áreas del aprendizaje profundo.

En este proyecto, la librería de Keras será usada para desarrollar los diferentes tipos de redes neuronales que se emplearán en la predicción de los tipos de cambio.

2.4 SCIKIT-LEARN

Scikit-learn, también conocido como Sklearn, es una biblioteca que proporciona herramientas para la minería de datos y el análisis supervisado y no supervisado de los mismos.



Figura 7: Logo de SciKit Learn

(Fuente: https://commons.wikimedia.org/wiki/File:Scikit_learn_logo_small.svg)

Se empleará debido a su sencillez e integración con otras librerías de uso frecuente como NumPy y Matplotlib, lo cual facilitará enormemente el proceso de análisis y visualización de los resultados.

Sklearn será la librería elegida para la creación de los modelos de regresión lineal.

2.5 STATSMODELS

Statsmodels es una librería de Python especializada en la creación de modelos estadísticos. Permite desarrollar modelos estadísticos, regresiones, modelos autorregresivos, realizar la exploración de los datos y el diagnóstico de dichos modelos.



Figura 8: Logo de Statsmodels

(Fuente: https://www.statsmodels.org/stable/_images/statsmodels-logo-v2-horizontal.svg)

La librería se empleará para el desarrollo de modelos autorregresivos, concretamente del modelo ARIMA.

Capítulo 3. ESTADO DE LA CUESTIÓN

El tercer capítulo del proyecto se centrará en el estudio y análisis del marco teórico y del estado de la cuestión de los conceptos relacionados con la predicción de los movimientos de los diferentes pares de cambio de divisa.

Se comenzará realizando un breve estudio sobre los modelos que se emplean en la actualidad, posteriormente, se estudiará el concepto de serie temporal, entrando en detalle en sus principales características. A continuación, se explicará el funcionamiento de los diferentes modelos que se han elegido para desarrollar en este proyecto.

3.1 MODELOS ACTUALES

Desde hace décadas hay un gran interés por ser capaz de predecir comportamientos o valores futuros en prácticamente cualquier ámbito, ya sea para adelantarnos a problema o para maximizar el beneficio o rendimiento de una actividad. El mundo de las divisas no es ninguna excepción, la búsqueda de un modelo que de ventaja sobre los competidores podría suponer enormes beneficios para aquella persona o empresa que lo tuviese.

Cabe destacar que, aunque aparentemente parezca una tarea trivial la predicción de acciones o divisas ha resultado ser extremadamente compleja, especialmente cuando se busca realizar predicciones en el corto plazo. En este caso han sido pocos los modelos capaces de obtener resultados mejores que un simple modelo de Random Walk, tal y como se menciona en (MESSE & ROGOFF, 1983). Un modelo de Random Walk se basa en la idea de que, en una determinada serie de valores, como puede ser una serie histórica de una acción o de un par de divisas, los valores futuros son impredecibles y que únicamente dependen del valor anterior más una perturbación aleatoria. Por otra parte, la teoría de análisis fundamental o

intrínseco expone que una acción tiene un valor de equilibrio que se puede determinar mediante un análisis de la acción. Según se expone en (Chitenderu, Maredza, & Sibanda, 2014) las perturbaciones ocasionadas por la intervención humana causarán que el valor se mueva de forma aleatoria en torno a su valor intrínseco, es decir, las acciones se comportan de forma natural como un Random Walk.

Dejando de lado la mentalidad cortoplacista, dominada por la aparente aleatoriedad de las series, nos encontramos con modelos que tratan de adelantarse a los movimientos o tendencias de las divisas a medio o largo plazo. En estos casos, el comportamiento de las divisas puede ser explicado en cierta medida por el uso de variables macroeconómicas (YEŞIN, 2016).

Un ejemplo de modelos a medio plazo es el del Fondo Monetario Internacional (FMI), una organización compuesta por 190 países cuyos objetivos, entre otros, es promover la estabilidad económica y el comercio internacional. Para ello, es de vital importancia adelantarse a posibles cambios en las tendencias de las divisas. Para ello se desarrollaron entre 2006 y 2011 tres modelos con la finalidad de predecir a medio plazo si una moneda se encontraba sobrevalorada o infravalorada (YEŞIN, 2016). Los modelos se basaban en el uso de variables macroeconómicas para predecir un valor de equilibrio conocido como *equilibrium real effective exchange rate (REER)*, el cual se comparaba con el REER real para determinar si la divisa se encontraba infravalorada o sobrevalorada. Dicha diferencia entre el valor de equilibrio y el valor real determina si la moneda se apreciará, en el caso de que se encuentre infravalorada, o si se depreciará, si la divisa se encuentra sobrevalorada.

3.2 SERIES TEMPORALES

Los valores del tipo de cambio de los diferentes pares de divisas vienen expresados en forma de serie temporal. Una serie temporal es un conjunto de datos que se organiza en base a una temporalidad, es decir, los primeros datos son aquellos que sucedieron primero en el tiempo. Esto es importante, ya que el orden en el que se encuentran los datos es importante, ya que arroja información.

Debido a su estructura temporal, podemos identificar tres características principales que deben ser analizadas en cada una de las series a estudiar.

- **Tendencia:** hace referencia a los cambios sistemáticos que se dan en las series temporales. Estos cambios sistemáticos ocurren en largos periodos de tiempo y pueden tomar formas lineales, como incrementar o disminuir siguiendo una recta de diferente pendiente, o pueden tomar distribuciones no lineales como una tendencia exponencial.
- **Estacionalidad:** son patrones que se repiten de forma regular a lo largo de un periodo de tiempo fijo. Un ejemplo puede ser un pico en la demanda de helados durante el verano o un aumento en las reservas de restaurantes durante los fines de semana.
- **Ciclo:** concepto similar a la estacionalidad, hace referencia a patrones repetitivos en las series temporales. Las principales diferencias entre la estacionalidad y los ciclos son su frecuencia y su duración. Los ciclos no tienen una frecuencia de ocurrencia fija, aunque se suele dar cada determinados años, esta medida puede variar, mientras que la estacionalidad se da en intervalos fijos, por ejemplo, todos los veranos. Por otro lado, los ciclos suelen darse en periodos de tiempo superiores a un año, mientras que la estacionalidad hace referencia a repeticiones que se dan dentro de un año.

3.3 MODELOS A EMPLEAR

En esta sección del tercer capítulo, se realizará una revisión de los modelos que van a desarrollar a lo largo de las próximas secciones del proyecto. En concreto, se van a emplear modelos tanto estadísticos, como modelos basados en machine learning. Dentro de los modelos estadísticos se encuentran los modelos de regresión lineal y los modelos autorregresivos. Los modelos basados en machine learning serán las redes neuronales recurrentes (RNN) y los modelos de memoria a corto y largo plazo, o Long Short-Term Memory (LSTM).

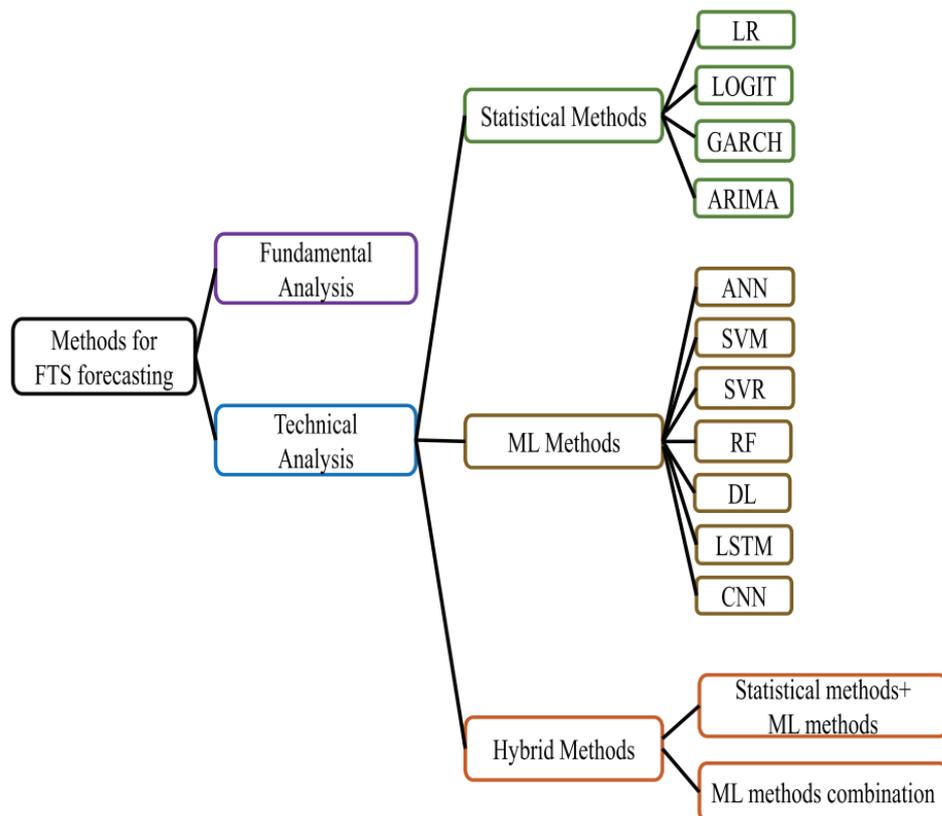


Figura 9: Métodos de predicción de series temporales financieras

(Fuente: (Tang, y otros, 2022))

Es común dentro de los modelos de predicción de series temporales, que se generen nuevas variables con el fin de mejorar el poder predictivo de dichas variables. Estas nuevas variables son atrasos temporales de la variable a predecir, ya que, en numerosas ocasiones, el valor en el momento t depende de los valores en momentos previos. Este concepto se desarrollará más adelante, cuando se incorpore en los diferentes modelos.

3.3.1 REGRESIÓN LINEAL

El primer modelo que se va a emplear en la predicción de las series del precio de los pares de divisas será el modelo de regresión lineal. Como se ha mencionado previamente, este modelo se encuentra dentro de la categoría de modelos estadísticos, siendo uno de los modelos más sencillos a nivel teórico y a nivel de implementación.

Su funcionamiento se basa en la llevar a cabo una combinación lineal de las variables independientes o predictivas con el objetivo de predecir la variable dependiente. La expresión de una regresión lineal es la siguiente:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon$$

Donde y representa la variable dependiente, x_i representa el conjunto de k variables independientes que son usadas para predecir y , donde $i = 1, 2, \dots, k$. β_j representa los coeficientes de cada una de las variables independientes, para los cuales $j = 0, 1, 2, \dots, k$. El término β_0 no acompaña a ninguna variable independiente y es conocido como la ordenada en el origen. ε hace referencia al error, es decir es la diferencia entre la predicción realizada por el modelo y el valor real de la variable dependiente. La expresión del valor predicho de y , representado como \hat{y} , es la siguiente:

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k$$

3.3.2 MODELOS AUTORREGRESIVOS

Desarrollando el concepto planteado previamente, en el cual se mencionaba que los valores presentes de una serie temporal dependían de los valores previos de dicha serie, introducimos

los modelos autorregresivos. Estos se basan en la autocorrelación de la serie, la cual es una medida estadística que analiza la relación de valores de la serie de diferentes puntos temporales.

Los modelos autorregresivos se rigen por una serie de parámetros fijos, entre los cuales destaca el parámetro que controla el número de observaciones previas que se emplearán como variables predictivas. Posteriormente, se incluyen más parámetros en función del tipo de modelo autorregresivo que se esté empleando.

3.3.3 REDES NEURONALES

Uno de los tipos de modelos con los que se van a realizar predicciones sobre el valor de los tipos de cambio son las redes neuronales, también conocidas como redes neuronales artificiales o ANN por sus siglas en inglés. A finales de los años 50 se propuso por primera vez el concepto de perceptrón, dando nacimiento al perceptrón simple o SLP (*single layer perceptron*). El SLP es el tipo de red neuronal más simple y la arquitectura sobre la cual se construyen las redes más complejas.

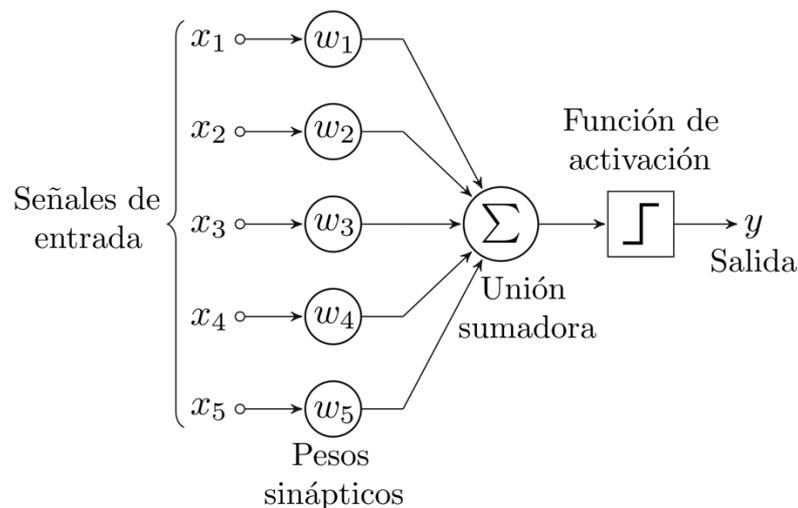


Figura 10: Estructura de un perceptrón simple

(Fuente: https://es.wikipedia.org/wiki/Archivo:Perceptr%C3%B3n_5_unidades.svg)

El perceptrón simple consta de 4 partes principales:

- Una **capa de entrada** mediante la cual se introducen los datos en el sistema.
- Los **pesos**. De manera similar a una regresión lineal, los valores provenientes de la capa de entrada se multiplican por su peso correspondiente. Al comienzo del entrenamiento de la red, los pesos son inicializados de forma aleatoria.
- La **neurona**. Cada uno de los productos de los distintos pesos con su correspondiente peso se introducen en la neurona, donde se suman. Además, cada neurona cuenta con un peso o bias, el cual también es añadido a la suma.

$$w_1x_1 + w_2x_2 + \dots + w_nx_n + \text{bias}$$

- La **función de activación**. La combinación lineal de las entradas se pasa por una función, denominada función de activación, obteniendo un valor que será la salida de la neurona. El objetivo de las funciones de activación es otorgarles a las redes la no-linealidad necesaria para poder resolver problemas más complejos. Las principales funciones de activación son:

•

○ Función escalón:

- $f(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 & \text{si } x \geq 0 \end{cases}$

○ Función sigmoide:

- $f(x) = \frac{1}{1 + e^{-x}}$

○ Función ReLU:

- $f(x) = \max(x, 0)$

○ Función tangente hiperbólica:

- $f(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}}$

El proceso de entrenamiento se basa en el ajuste de los pesos. Dicho ajuste dependerá de la diferencia entre el valor obtenido a la salida y el valor esperado. Para ser más precisos, los pesos se ajustan en base a la siguiente fórmula.

$$\Delta w = \eta \cdot d \cdot x$$

- η es la tasa de aprendizaje, factor que ajusta la cantidad que cambian los pesos en cada iteración.
- d es el valor predicho a la salida menos el valor esperado.
- x es el valor del dato de entrada.

3.3.3.1 Perceptrón multicapa (MLP)

El perceptrón multicapa consiste en una arquitectura basada en el perceptrón simple, donde varias neuronas se agrupan en diferentes capas. El MLP recibe su nombre debido a los tres grupos de capas que posee. La primera es la capa de entrada, mediante la cual se introducen los valores a la red. El segundo grupo es conocido como capa oculta, ya que no tiene conexión directa con la entrada o la salida. Dentro de este grupo se puede encontrar desde una capa de neurona en adelante. Por último, la capa de salida, encargada de obtener la predicción en función de los valores recibidos del último conjunto de neuronas de la capa oculta.

El funcionamiento del perceptrón multicapa se basa en una red de perceptrones simples conectados entre sí. Cada una de estas conexiones tiene un peso asignado, además de que cada neurona cuenta con su bias.

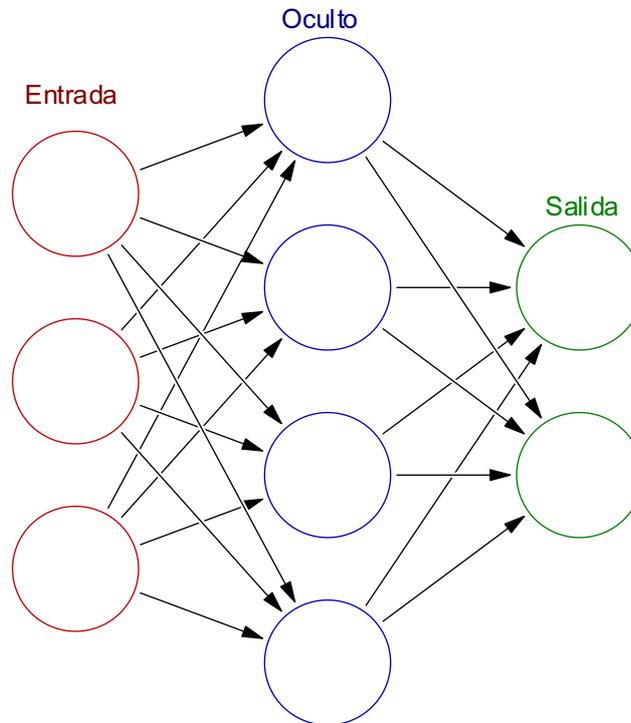


Figura 11: Estructura del perceptrón multicapa

(Fuente:

https://upload.wikimedia.org/wikipedia/commons/thumb/1/11/Colored_neural_network_es.svg/1920px-Colored_neural_network_es.svg.png)

De la misma forma que en el perceptrón simple, los pesos y bias de la red han de actualizarse tras cada iteración, sin embargo, al ser una red con múltiples capas, el error calculado a la salida solo afecta directamente a los parámetros de la última capa. Para poder actualizar el resto de las capas se emplea un algoritmo que propaga los errores desde la capa de salida a la capa de entrada. El algoritmo actualizará los pesos y bias con el objetivo de minimizar el valor de la función de coste, función que depende de la diferencia entre el valor predicho y el valor esperado.

3.3.3.2 Redes neuronales recurrentes (RNN)

Una vez entendidas las redes neuronales, analizaremos aquellos tipos de redes empleadas para la predicción de datos secuenciales o series temporales. Las redes neuronales recurrentes, RNN por sus siglas en inglés, son aquellas redes con capacidad de recordar datos

previos para aplicarlos a la predicción de valores futuros. Esta memoria se consigue de forma muy sencilla, la salida de la red para un tiempo t se emplea como entrada en la predicción del tiempo $t+1$.

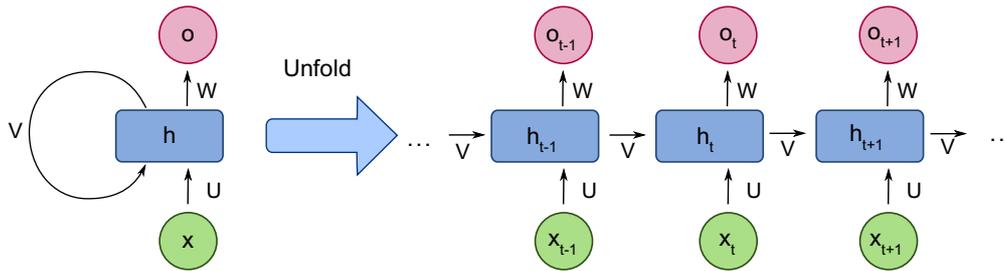


Figura 12: Estructura de una red neuronal recurrente

(Fuente:

https://en.wikipedia.org/wiki/Recurrent_neural_network#/media/File:Recurrent_neural_network_unfold.svg)

Observamos en el diagrama anterior que, además de las salidas o_t de la red a cada entrada x_t , la salida de la red de cada instante temporal es la entrada del siguiente paso temporal.

Sin embargo, las redes neuronales recurrentes no son capaces de detectar dependencias o relaciones que se extienden en el tiempo. Este problema, conocido como desvanecimiento del gradiente, se debe a que la actualización de los pesos de la red se realiza mediante el algoritmo de Back-Propagation Through Time (BPTT), que propaga los errores desde la salida de la red hacia la entrada. Esta propagación hace que los errores tiendan a cero o diverjan según van avanzando hacia etapas pasadas (Hochreiter & Schmidhuber, 1997). Para solucionar este problema se desarrollaron las redes Long Short-Term Memory (LSTM).

3.3.3.3 Redes Long Short-Term Memory (LSTM)

Creadas con el objetivo de solucionar el problema del desvanecimiento del gradiente, las redes Long Short-Term Memory reciben su nombre debido a la capacidad que tienen de trazar dependencias tanto a corto como a largo plazo.

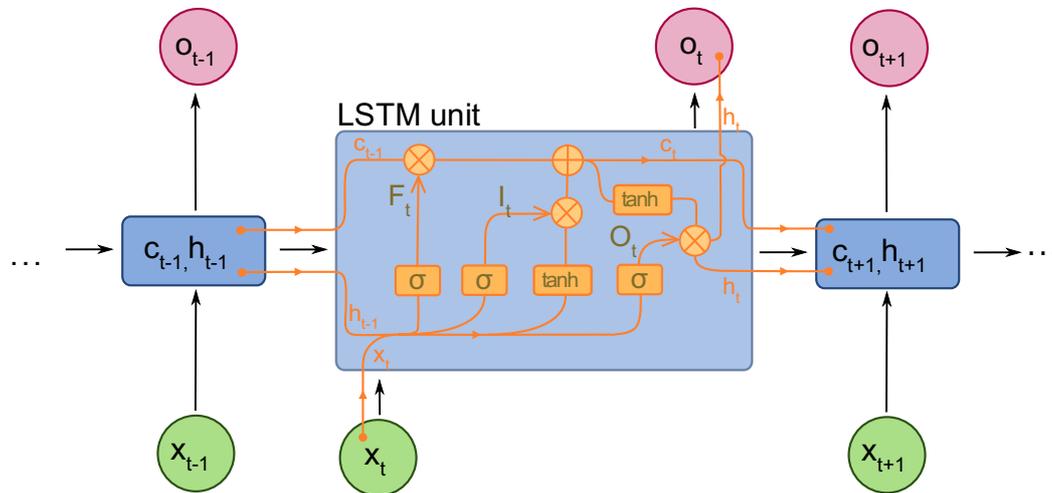


Figura 13: Estructura de una unidad LSTM

(Fuente: https://en.wikipedia.org/wiki/Recurrent_neural_network#/media/File:Long_Short-Term_Memory.svg)

Para ser capaz de establecer dependencias a largo plazo, una red LSTM está compuesta de tres puertas, las cuales controlan los flujos de información; un estado conocido como celda de memoria o cell state (c_t), encargado de transportar la información a largo plazo debido a que sufre solo ligeras modificaciones a lo largo de todos los estados; y el estado h_t , el cual será manipulado por las puertas y determinará la salida de la unidad en t y, por lo tanto, la entrada de $t+1$. Los tres tipos de puerta son los siguientes:

- **Puerta de olvido o forget gate (F_t):** Es la puerta encargada de determinar qué información previa, que se encuentra almacenada en el cell state, se debe de olvidar. Para ello, recibe como entrada la salida previa h_{t-1} , y la entrada actual x_t y se pasan a través de una función de activación sigmoide.
- **Puerta de entrada o input gate (I_t):** Determina a partir de la salida previa h_{t-1} y la entrada actual x_t qué información se añadirá a la celda de memoria mediante el

uso de una función de activación sigmoide y otra función de activación de tangente hiperbólica.

- **Puerta de salida u output gate (O_t):** Esta puerta recibe como entradas la salida previa h_{t-1} , la entrada actual x_t y el valor actualizado por la puerta de entrada de la celda de memoria c_t . Al último valor lo pasa por una función de tangente hiperbólica, mientras que, a las dos primeras entradas, se las introduce en una función sigmoide, para finalmente multiplicar ambas. El resultado de esta puerta será la salida de la red en el instante t y la entrada de la siguiente etapa.

Capítulo 4. DEFINICIÓN DEL TRABAJO

4.1 JUSTIFICACIÓN

La predicción eficiente del valor de los tipos de cambio podría tener un profundo impacto en el mundo que conocemos a día de hoy. Este mercado, el más líquido y que más volumen de transacciones experimenta diariamente, es empleado por países, empresas y particulares, es decir, puede afectar a todas las partes que componen nuestra sociedad. Su impacto puede ser positivo en diferentes áreas, como, por ejemplo:

- La toma de decisiones financieras por parte de empresas y particulares. En numerosas ocasiones, es necesario realizar cambios de divisa para poder desarrollar la actividad diaria de las empresas o de los individuales. El conocimiento de los posibles movimientos futuros de un tipo de cambio puede resultar en que se escoja el mejor momento o la mejor estrategia de conversión de las divisas, obteniendo la posición financiera más favorable. Ligado a esto, un modelo eficiente puede resultar extremadamente beneficioso en el control de riesgos que está asumiendo una empresa a la hora de operar con capital extranjero.
- Desarrollo de políticas monetarias y económicas. Los bancos centrales de los países o de las uniones monetarias son los que se encargan de mantener el valor de la moneda en rangos que consideren adecuados. En situaciones normales, esto se consigue de forma eficiente con el cambio de los tipos de interés y de la cantidad de moneda que se encuentra en circulación. Sin embargo, el uso de modelos que den información relevante con mayor margen de tiempo puede hacer que los bancos centrales tengan mayor margen de maniobra y puedan aplicar las políticas oportunas de forma más eficiente.
- La realización de inversiones financieras. Un modelo eficiente puede ser capaz de indicarnos como se va a comportar el mercado antes de que esto suceda. Una posición

aventajada como esa es extremadamente valiosa para aquellas personas o empresas que quieran obtener beneficios del intercambio de divisas.

La utilidad de un modelo con buena capacidad predictiva es enorme, tanto para el inversor particular que opera con pequeñas cantidades, como para el banco central que busca mantener la estabilidad económica y monetaria de un país.

4.2 OBJETIVOS

Este proyecto persigue como objetivo el desarrollo de modelos para la predicción del tipo de cambio de diferentes pares de divisas. Se persigue que los modelos cumplan los siguientes objetivos:

- Predicción del valor de apertura y cierre para cada intervalo temporal
- Obtener modelos capaces de predecir tendencias y puntos de inflexión en las series
- Trazar similitudes y diferencias entre la predicción de divisas tradicionales y criptomonedas

Como mencionado previamente, los modelos empleados hoy en día por bancos y organizaciones internacionales, como el Fondo Monetario Internacional (FMI), persiguen como objetivo estudiar a medio plazo si una moneda se encuentra por debajo o por encima de un valor equilibrio. Se trata de estimar si la tendencia que tomará la moneda en un futuro será alcista o bajista, pudiendo así tomar las medidas que consideren oportunas.

Por el contrario, el objetivo de este proyecto reside en la predicción de un valor concreto para el par de divisas. Concretamente, al ser datos de intervalo, es decir, al no tratarse de un valor concreto, sino de un intervalo acotado por dos valores concretos. Por ello, se buscará predecir el valor de apertura y el valor de cierre del intervalo. Para ello se llevarán a cabo cuatro modelos diferentes: uno basado en una regresión lineal, otro en métodos autorregresivos, otro basado en redes neuronales recurrentes y, finalmente, un modelo basado en redes Long Short-Term Memory (LSTM). Se comenzará tratando de predecir los valores futuros únicamente con los valores pasados de la serie. En caso de que no se obtengan

modelos capaces de ajustarse con precisión a la serie, se emplearán variables externas como variables macroeconómicas y financieras.

Por otro lado, a diferencia del FMI, que trata de predecir tendencias a medio plazo (YEŞİN, 2016), los modelos a desarrollar en este trabajo buscan predecir en el corto plazo, centrándose en intervalos de días o semanas.

4.3 METODOLOGÍA

El desarrollo del proyecto se realizará siguiendo la metodología en cascada o waterfall. Este enfoque o modelo es uno de los pioneros en el mundo del desarrollo software, el cual aplica un enfoque secuencial, donde el desarrollo se divide en diferentes fases. Las diferentes fases se encuentran conectadas entre sí, de tal forma que la salida de una fase es la entrada de la siguiente. Además, en cada etapa se especifican un conjunto de objetivos a cumplir antes de poder pasar a la siguiente fase.

Se ha elegido el modelo en cascada debido a que su enfoque granular y secuencial permite afrontar un proyecto extenso, como el que planteamos en este trabajo, de forma más organizada y asequible, basándose en la consecución de pequeños objetivos.

El modelo consta de seis diferentes etapas:

- **Requisitos:** En esta etapa se documentan los requisitos que debe tener el programa o sistema, incluyendo sus características y funcionalidades. También, se definirán los conjuntos de datos que van a ser empleados para el entrenamiento de los modelos.
- **Análisis:** Una vez definidos los objetivos procedemos a realizar un análisis de los datos con los que se va a llevar a cabo el proyecto. Se llevará a cabo la limpieza y preparación de los datos. Las series pueden provenir de fuentes distintas, siendo posible que no haya consistencia en los tipos de datos, como por ejemplo en el formato de fecha. Por ello es necesario realizar un preprocesado.

Por último, en modelos de forecasting de series temporales, la variable dependiente o variable a predecir es la siguiente observación a la última tomada para el entrenamiento. Se necesita entonces obtener el vector de la variable objetivo para poder entrenar el modelo.

- **Diseño:** La tercera fase consiste en una descripción detallada del código y de la arquitectura que se va a emplear para alcanzar los objetivos planteados en la primera etapa.
- **Codificación:** Basándonos en las especificaciones expuestas en la anterior etapa se procede al desarrollo del código.

Al ser un proyecto donde la complejidad del código no es alta, las etapas de diseño y codificación se llevarán a cabo dentro de una misma fase, a la cual nos referiremos a partir de ahora como la fase de diseño.

- **Prueba:** Los modelos creados en la etapa de codificación deben ser evaluados, estudiando así su rendimiento y definiendo posibles cambios en la arquitectura con el objetivo de mejorar los resultados.
- La etapa final del desarrollo consistirá en la comparación de los resultados de los diferentes modelos y pares de divisas, así como la obtención de conclusiones.

La metodología en cascada permite realizar nuevas iteraciones de las etapas con el fin de mejorar los modelos mediante cambios en los datos, en las arquitecturas o en otros componentes.

Por último, se probarán los modelos con mejor comportamiento sobre la serie de Bitcoin, con el objetivo de estudiar las posibles similitudes entre las divisas tradicionales y las criptomonedas.

4.4 PLANIFICACIÓN Y ESTIMACIÓN ECONÓMICA

La búsqueda de modelos que sean capaces de predecir el valor de activos financieros es uno de los grandes objetivos de empresas e inversores particulares. Conseguir una ventaja sobre el mercado supondría tener una fuente de ingresos prácticamente ilimitada. Este proyecto busca alcanzar dicho objetivo y tratar de comercializar los modelos.

Más allá de la complejidad de la tarea, hay una serie de factores limitantes o que pueden acarrear gastos en caso de comercialización de los modelos. El primer factor que debemos tener en cuenta es el acceso a los datos que se pretenden emplear en el entrenamiento de los modelos. Es necesario tener fuentes de datos fiables y con datos disponibles prácticamente en tiempo real. A pesar de que hay numerosas fuentes de información pública y gratuita, como las que se han empleado en este proyecto para el desarrollo de los modelos, en una fase de comercialización es posible que se requiera de fuentes de pago que soporten un alto volumen de acceso en tiempo real. El coste de acceso a los datos dependerá de la fuente que se quiera acceder y del volumen de datos a emplear, por ello, para una estimación más concreta del coste sería necesario conocer los datos a emplear.

El segundo factor limitante es la capacidad de computación necesaria. En el caso de comercializar la herramienta, ya sea ofreciendo un servicio a terceros o vendiendo los modelos directamente, uno de los factores más importantes es el acceso rápido a predicciones sobre el mercado. El problema es que la información se va actualizando constantemente, requiriendo que los modelos se entrenen en tiempo real, lo cual puede resultar en altos tiempos de computación para obtener la predicción. Como solución, se pueden implantar sistemas computacionales distribuidos que repartan las altas cargas computacionales mediante la ejecución de los procesos en paralelo y conseguir así tiempos de ejecución menores. A pesar de ser una buena solución, esto supone un gasto a tener en cuenta en caso de comercializar los modelos. De nuevo, conocer el valor de los costes asociados a la capacidad de computación dependerá del tamaño de la operación y de la cantidad de recursos que se necesiten.

Existen diversas formas de monetizar el proyecto en caso de que su comercialización sea viable. La primera sería ofreciendo servicios de consultoría financiera basados en las predicciones obtenidas por el modelo o directamente establecer un fondo de inversión basado en el modelo, donde se invierta el dinero de terceros. Otra opción es la concesión temporal del uso del modelo a empresas, a cambio de unas tasas mensuales y un porcentaje de los ingresos obtenidos con el modelo.

Se van a considerar tres posibles escenarios en la comercialización de los modelos: un escenario pesimista, uno realista y uno optimista.

- El escenario **pesimista** recoge aquellas posibles situaciones en las que se comercialicen los modelos y los ingresos no sean mayores que los gastos fijos que proceden de los dos factores mencionados previamente. Un caso posible sería que los modelos comenzasen dando buenos resultados y se llevasen a cabo las inversiones iniciales necesarias para su uso, pero por cambios ajenos a los modelos, como la guerra entre Rusia y Ucrania o el COVID-19, la situación financiera cambie y las predicciones dejen de ajustarse a la realidad.
- En el escenario **realista** los modelos son capaces de predecir correctamente algunas características de las series, como posibles cambios de tendencia o cambios de valor, pero sin tener una fiabilidad que sea atractiva para el grueso de clientes. En este caso las expectativas de resultados coinciden con la realidad y las inversiones iniciales son acordes a los ingresos esperados.
- El escenario **optimista** trata el caso en el que los modelos tengan un excelente poder predictivo y que, por lo tanto, garanticen un incremento de las rentabilidades obtenidas por los clientes. En esta situación el valor de nuestro proyecto es difícil de calcular, pues las mayores empresas del mundo e incluso los bancos centrales de los países resultarían muy beneficiadas de su uso.

Capítulo 5. SISTEMA/MODELO DESARROLLADO

En este capítulo del proyecto se desarrollarán los diferentes modelos, siguiendo la metodología de cascada previamente introducida. Los pasos serán los siguientes: requisitos, análisis, diseño, prueba y comparación de los resultados.

El flujo de información se rige mediante la arquitectura expuesta en la *Figura 14*. Se accede a las diferentes fuentes de datos, las cuales podemos separar en tres: la fuente de datos que contiene las series históricas de los siete pares de divisas principales, la fuente de datos que contiene la información de las criptomonedas, en concreto del Bitcoin; y, finalmente, la fuente de información que contiene las variables macroeconómicas y financieras que se emplean en el desarrollo de determinados modelos. Tras haber accedido a la información, se pasa a la fase de procesamiento, en la cual los datos se depuran de tal forma que puedan ser usados por los modelos en la siguiente etapa. La tercera etapa, donde tiene lugar la predicción, recibe como entrada los datos procesados y devuelve los valores de apertura y cierre. Para los valores de apertura se emplea un modelo, mientras que para los de cierre otro.

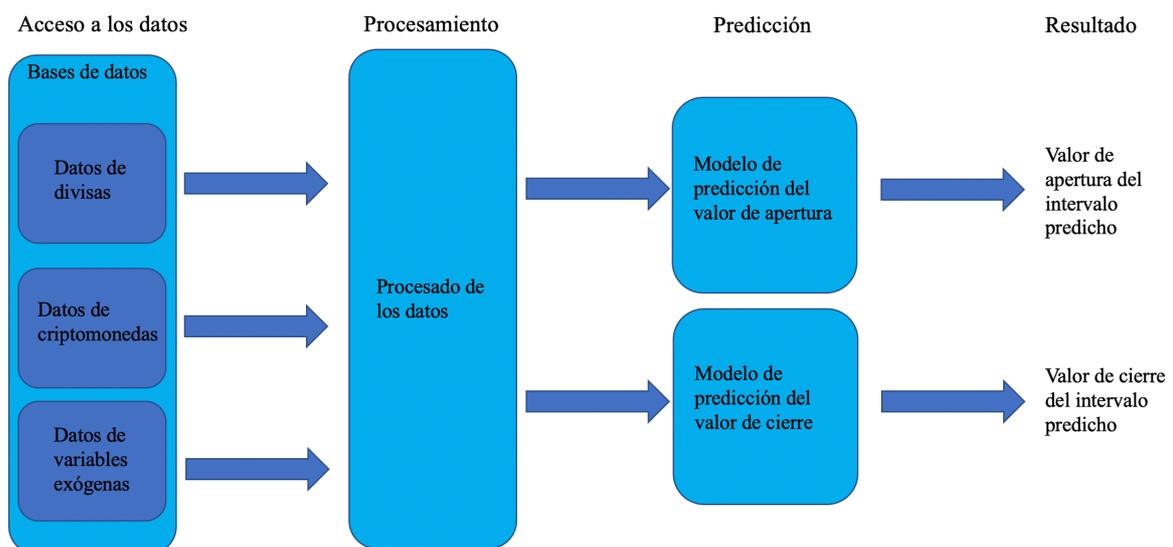


Figura 14: Diagrama de la arquitectura desarrollada (Fuente: Elaboración propia)

5.1 REQUISITOS

El proyecto tiene como objetivo desarrollar y entrenar cuatro modelos cuya finalidad sea la predicción del tipo de cambio entre los siete principales pares de divisas. Posteriormente, se aplicarán dichos modelos al precio del Bitcoin, tratando de estudiar si hay similitudes entre las monedas tradicionales y las criptomonedas.

El sistema o programa recibirá un conjunto de datos, al cual aplicará una serie de transformaciones, que serán descritas en el siguiente apartado, con el objetivo de que los datos tengan la forma adecuada para ser utilizados por las redes. Tras aplicar el tratamiento pertinente a los conjuntos de datos, estos serán utilizados para el entrenamiento y valoración de los modelos.

Los datos de los siete principales pares fueron obtenidos del portal web del banco online suizo Dukascopy Bank, mientras que el conjunto histórico de datos de Bitcoin se obtuvo del portal web *Investing.com*. En total se han descargado 16 series de datos, siendo ocho de ellas de frecuencia diaria y las otras ocho de frecuencia semanal. A continuación, la *Tabla 4* muestra el nombre de los conjuntos de datos, indicando su frecuencia y su fecha de inicio y final.

Nombre del archivo	Frecuencia	Fecha de inicio	Fecha de finalización
EURUSD_d.csv	Diaria	31.12.2012 23:00:00.000	12.03.2023 22:00:00.000
USDJPY_d.csv	Diaria	31.12.2012 23:00:00.000	11.03.2023 23:00:00.000

GBPUSD_d.csv	Diaria	31.12.2012 23:00:00.000 GMT+0100	12.03.2023 22:00:00.000 GMT+0100
USDCHF_d.csv	Diaria	31.12.2012 23:00:00.000 GMT+0100	12.03.2023 22:00:00.000 GMT+0100
AUDUSD_d.csv	Diaria	31.12.2012 23:00:00.000 GMT+0100	12.03.2023 22:00:00.000 GMT+0100
USDCAD_d.csv	Diaria	31.12.2012 23:00:00.000 GMT+0100	12.03.2023 22:00:00.000 GMT+0100
NZDUSD_d.csv	Diaria	31.12.2012 23:00:00.000 GMT+0100	12.03.2023 22:00:00.000 GMT+0100
EURUSD_w.csv	Semanal	26.12.2004 23:00:00.000 GMT+0100	12.03.2023 22:00:00.000 GMT+0100
USDJPY_w.csv	Semanal	26.12.2004 23:00:00.000 GMT+0100	12.03.2023 22:00:00.000 GMT+0100
GBPUSD_w.csv	Semanal	26.12.2004 23:00:00.000 GMT+0100	12.03.2023 22:00:00.000 GMT+0100

USDCHF_w.csv	Semanal	26.12.2004 23:00:00.000 GMT+0100	12.03.2023 22:00:00.000 GMT+0100
AUDUSD_w.csv	Semanal	26.12.2004 23:00:00.000 GMT+0100	12.03.2023 22:00:00.000 GMT+0100
USDCAD_w.csv	Semanal	26.12.2004 23:00:00.000 GMT+0100	12.03.2023 22:00:00.000 GMT+0100
NZDUSD_w.csv	Semanal	26.12.2004 23:00:00.000 GMT+0100	12.03.2023 22:00:00.000 GMT+0100
Bitcoin_d.csv	Diaria	Jan 01, 2013	Mar 15, 2023
Bicoin_w.csv	Semanal	Jan 06, 2013	Mar 12, 2023

Tabla 4: Descripción de los conjuntos de datos (Fuente: Elaboración propia)

Las columnas correspondientes a las fechas de inicio y fin aparecen con el formato de fecha original de los conjuntos de datos, siendo necesario su cambio en futuros pasos con el objetivo de tener homogeneidad en el tipo y en el formato de las variables de todos los conjuntos de datos.

En la *Tabla 5* que se expone a continuación, aparecen las variables que incluye cada conjunto de datos. Podemos separar los datos en aquellos provenientes de cada fuente, ya que estos incluyen las mismas variables.

Fuente de datos	Variable	Tipo de dato	Descripción
Dukascopy Bank (siete principales pares)	Local time	Fecha	Fecha final del intervalo representado
	Open	Float	Valor de apertura del intervalo temporal
	High	Float	Valor máximo del intervalo temporal
	Low	Float	Valor de mínimo del intervalo temporal
	Close	Float	Valor de cierre del intervalo temporal
	Volume	Float	Cantidad intercambiada en el intervalo
Investing.com (Bitcoin)	Date	Fecha	Fecha final del intervalo representado

	Price	Float	Valor de cierre del intervalo temporal
	Open	Float	Valor de apertura del intervalo temporal
	High	Float	Valor máximo del intervalo temporal
	Low	Float	Valor mínimo del intervalo temporal
	Vol.	Float	Cantidad intercambiada en el intervalo
	Change %	Float	Porcentaje de variación en el precio entre el comienzo y el final del intervalo

Tabla 5: Descripción de las variables contenidas en las series (Fuente: Elaboración propia)

5.2 ANÁLISIS

El primer paso que tomar tras la presentación inicial de los datos y el establecimiento de los objetivos del sistema es la depuración de los conjuntos de datos. Se comenzará convirtiendo el formato de la fecha de los 14 conjuntos de los principales pares. Actualmente, el formato de los siete principales pares es dd.mm.yyyy hh:mm:ss.fff GMT+0100. Para las dos series de Bitcoin el formato de la fecha es mmm dd, yyyy . El formato objetivo para las variables de fecha de todas las series es yyyy-mm-dd.

El siguiente paso en el formateo de los datos es la conversión de sufijos de unidades a números con decimal en la columna de volumen de las series de Bitcoin. Para ello, se han definido las siguientes equivalencias:

Símbolo	Orden de magnitud
K	10^3
M	10^6
B	10^9

Tabla 6: Equivalencias entre símbolos de unidades y su correspondiente orden de magnitud (Fuente: Elaboración propia)

A continuación, para homogeneizar todas las series de datos, se cambiarán los nombres de las columnas que representen la misma variable, pero no coincidan. Las variables de fecha pasarán a llamarse *Fecha*, las variables de volumen serán *Volume* y la variable *Change %* se renombrará a *Change*, con la finalidad de evitar potenciales errores al incluir caracteres especiales en el nombre de la variable. Finalmente, la variable *price* de las series de Bitcoin pasará a ser *Close*, manteniendo así consistencia con el resto de las series y con la nomenclatura empleada en las series de activos financieros.

Por último, las series de datos correspondientes al Bitcoin organizaban los datos de la fecha más reciente a la fecha más antigua. Esto resulta en una serie temporal cuyo tiempo se encuentra invertido. Es por ello por lo que hemos reorganizado las filas para que los datos vayan de más antiguo a más reciente.

Las series finales, representadas gráficamente, se muestran a continuación.

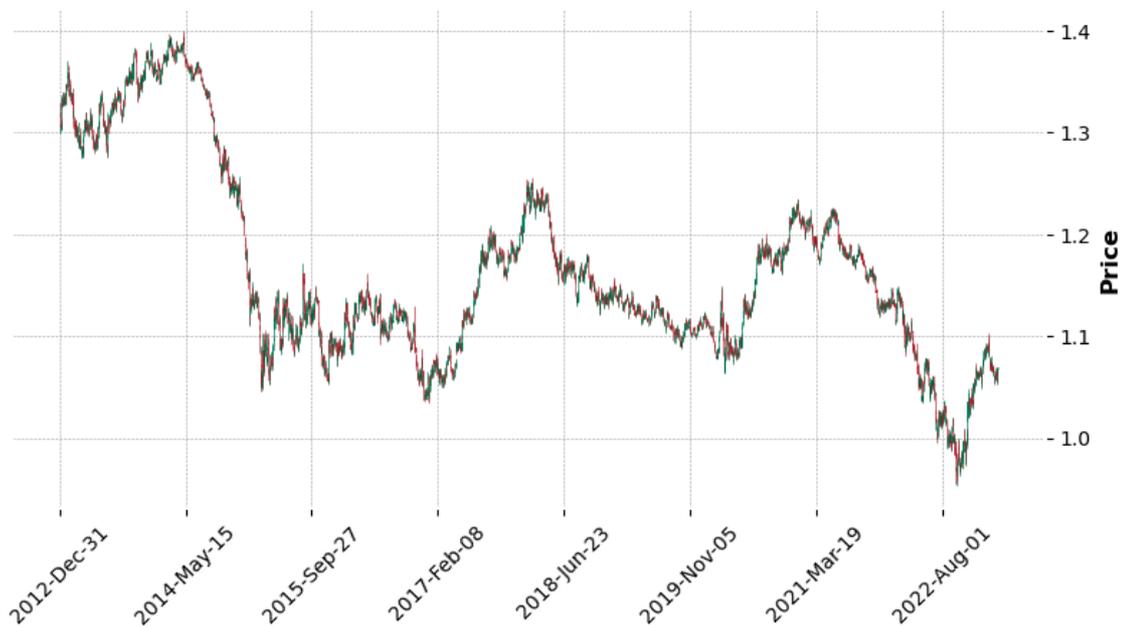


Figura 15: Serie diaria de EURUSD (Fuente: Elaboración Propia)

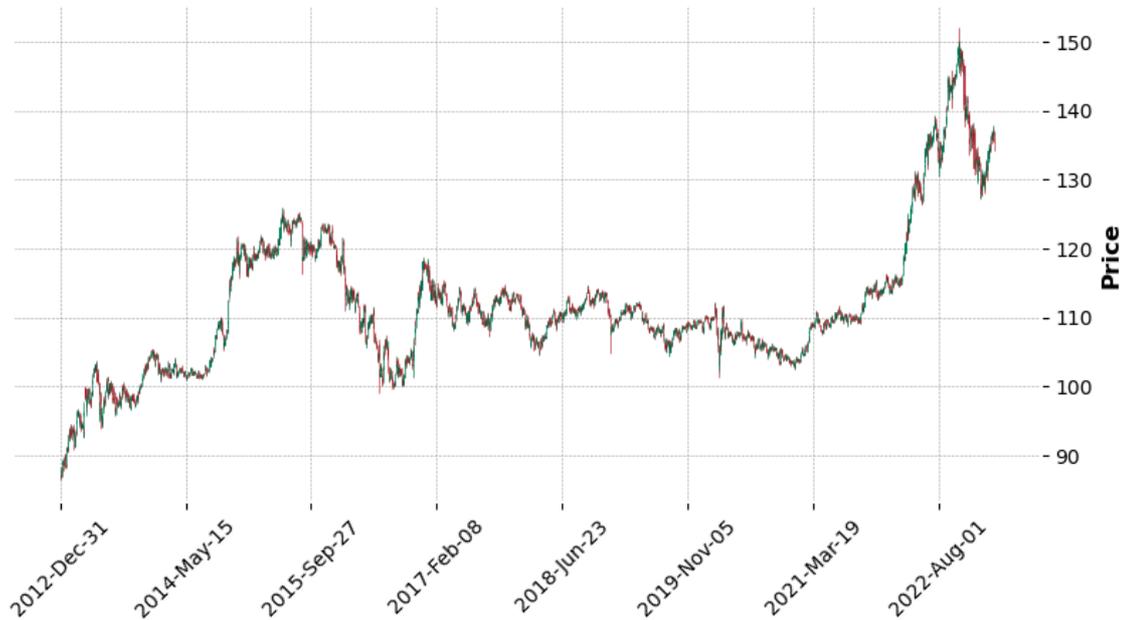


Figura 16: Serie diaria de USDJPY (Fuente: Elaboración Propia)

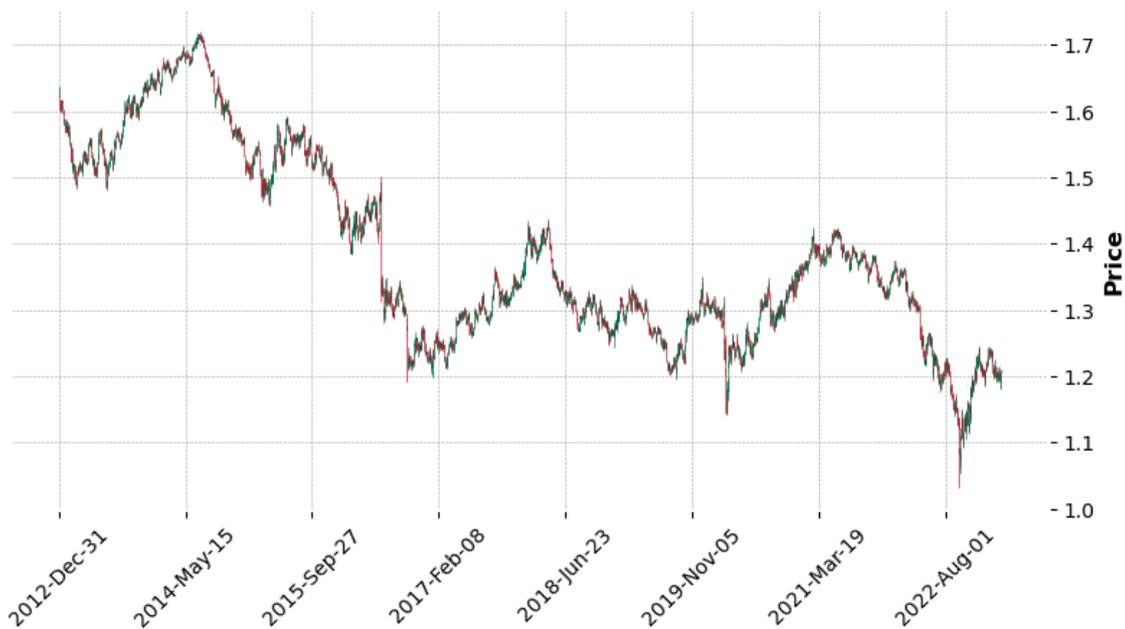


Figura 17: Serie diaria de GBPUSD (Fuente: Elaboración Propia)



Figura 18: Serie diaria de USDCHF (Fuente: Elaboración Propia)

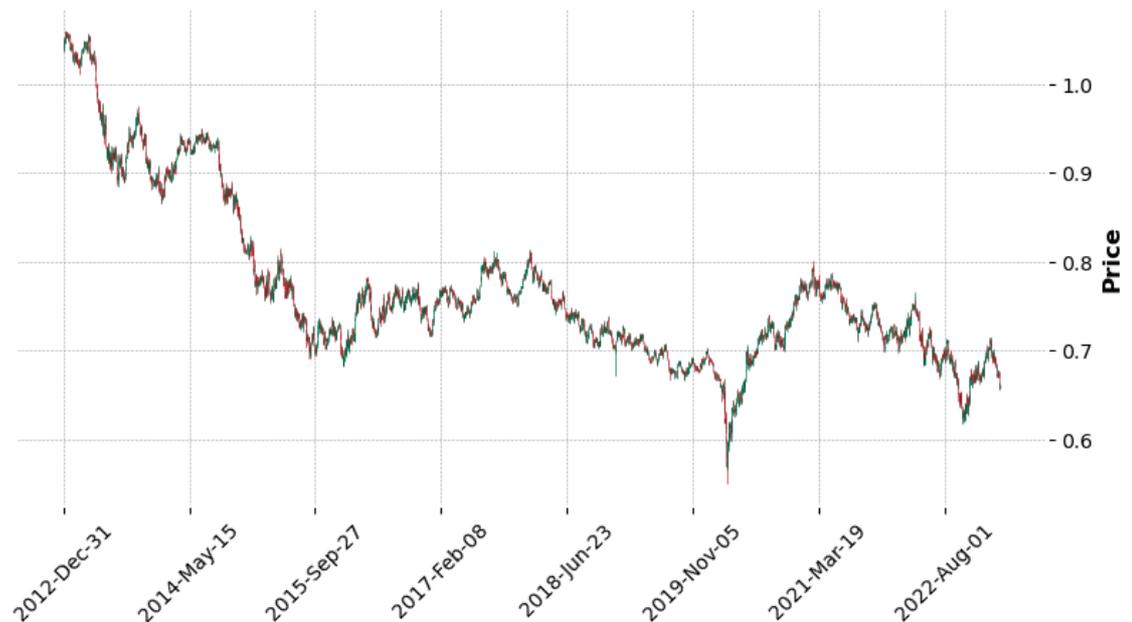


Figura 19: Serie diaria de AUDUSD (Fuente: Elaboración Propia)

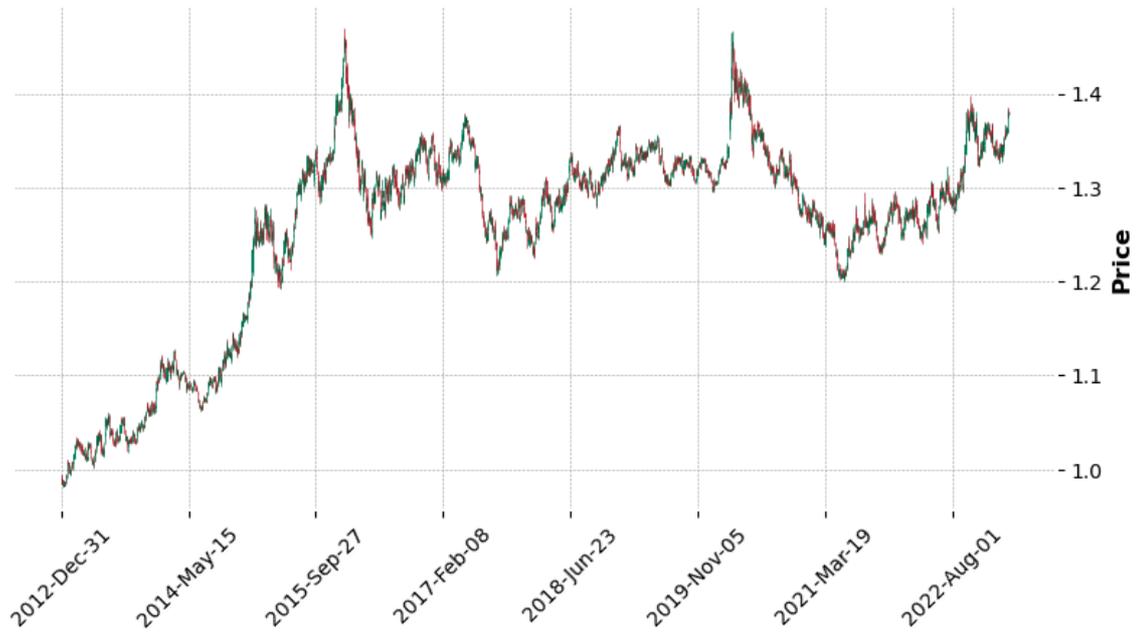


Figura 20: Serie diaria de USDCAD (Fuente: Elaboración Propia)

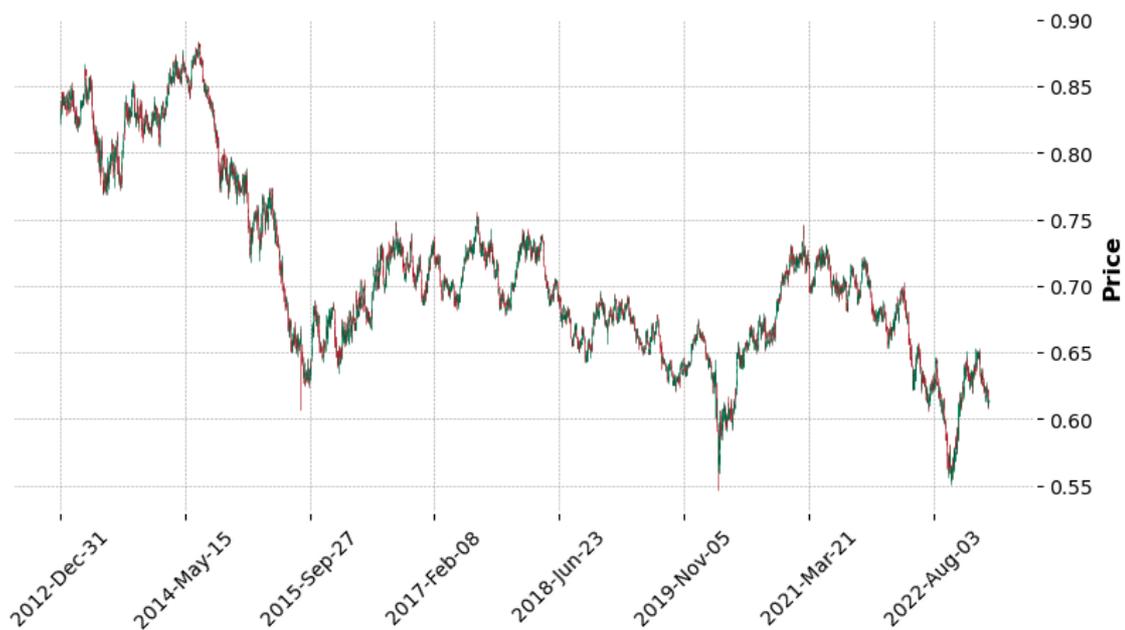


Figura 21: Serie diaria de NZDUSD (Fuente: Elaboración Propia)

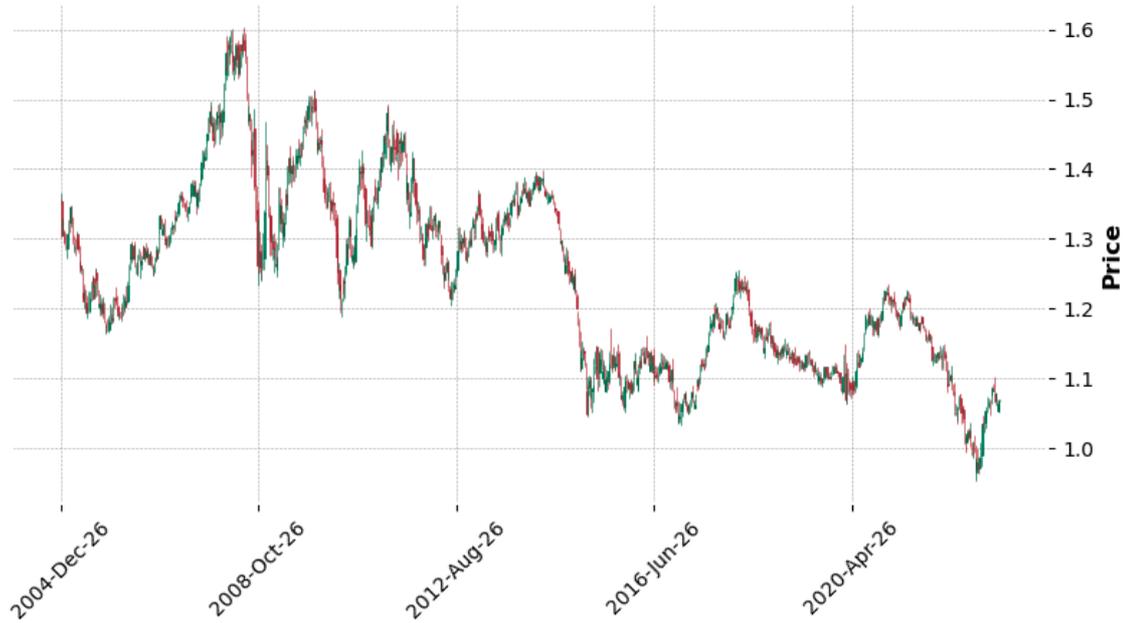


Figura 22: Serie semanal de EURUSD (Fuente: Elaboración Propia)

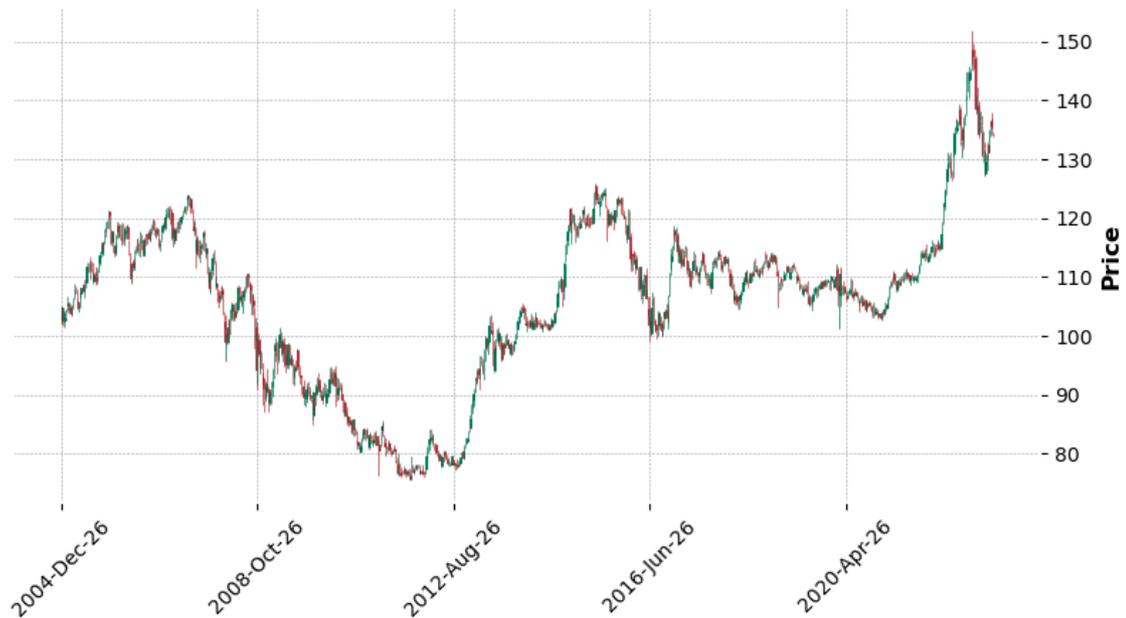


Figura 23: Serie semanal de USDJPY (Fuente: Elaboración Propia)

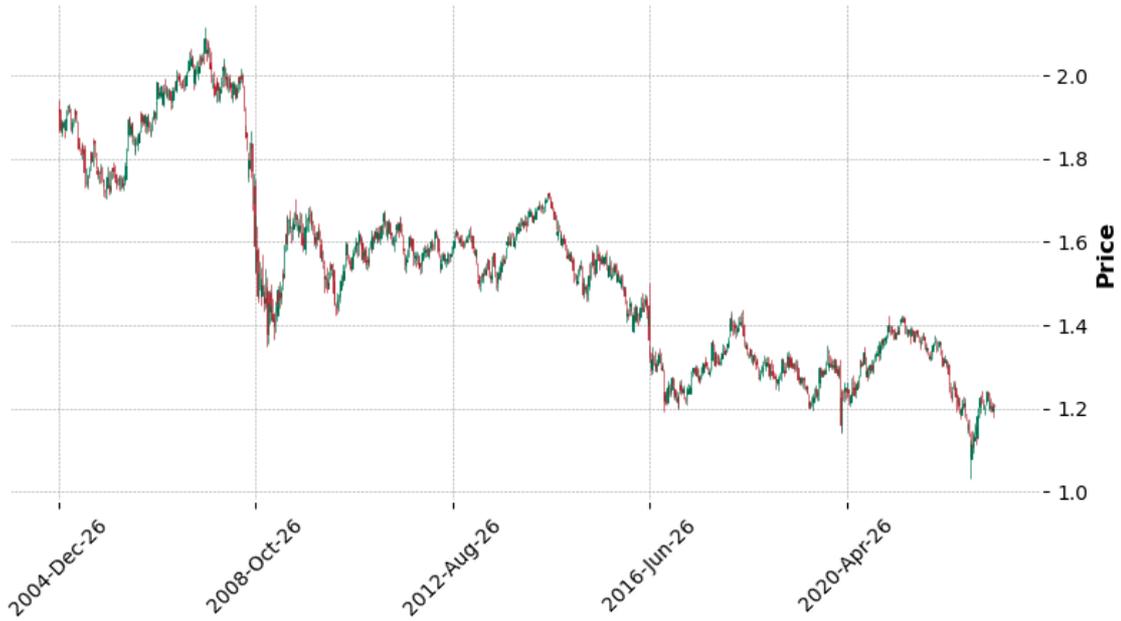


Figura 24: Serie semanal de GBPUSD (Fuente: Elaboración Propia)

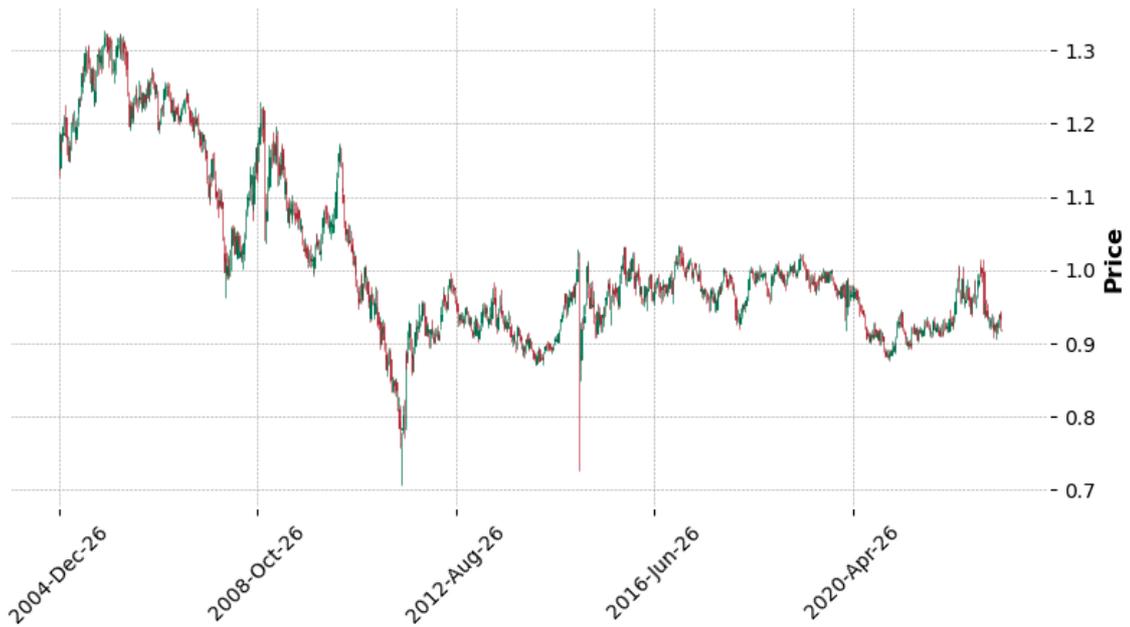


Figura 25: Serie semanal de USDCHF (Fuente: Elaboración Propia)

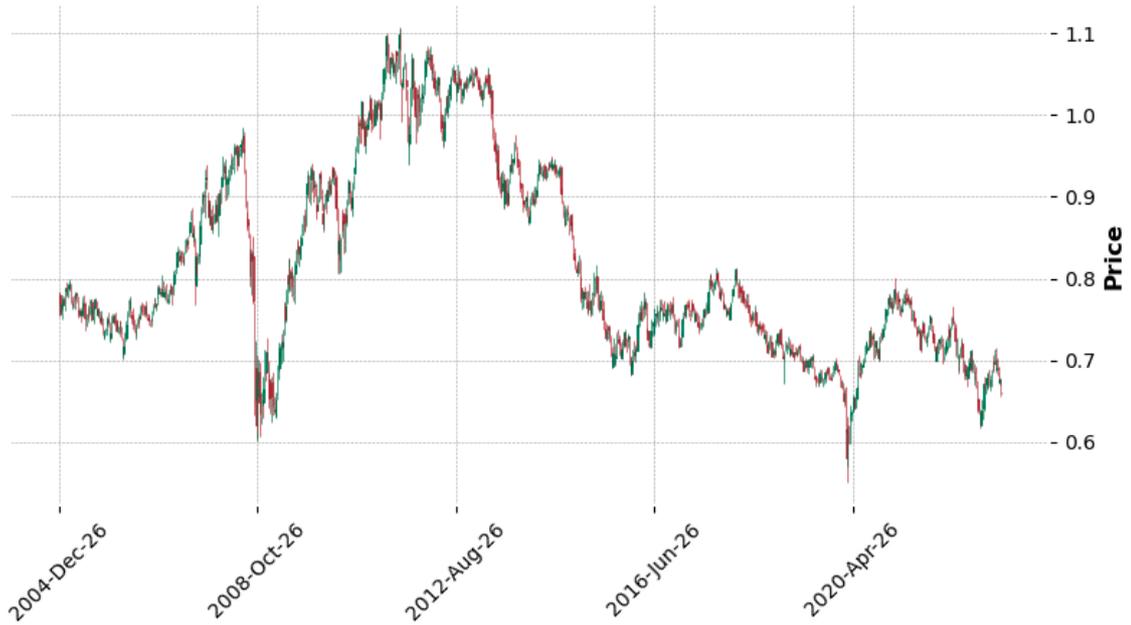


Figura 26: Serie semanal de AUDUSD (Fuente: Elaboración Propia)

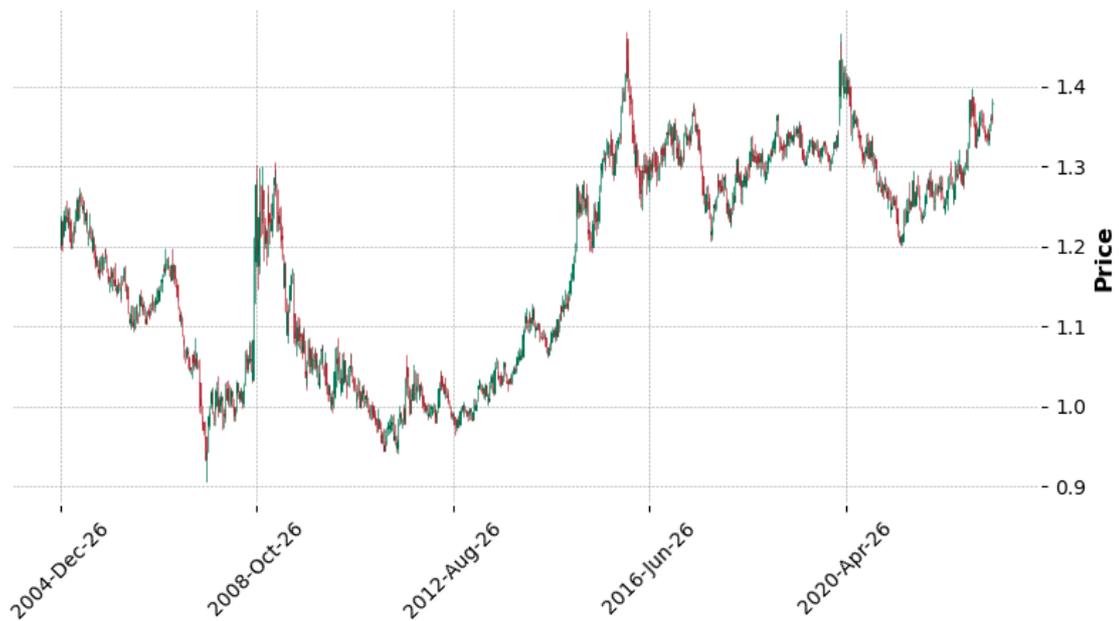


Figura 27: Serie semanal de USDCAD (Fuente: Elaboración Propia)

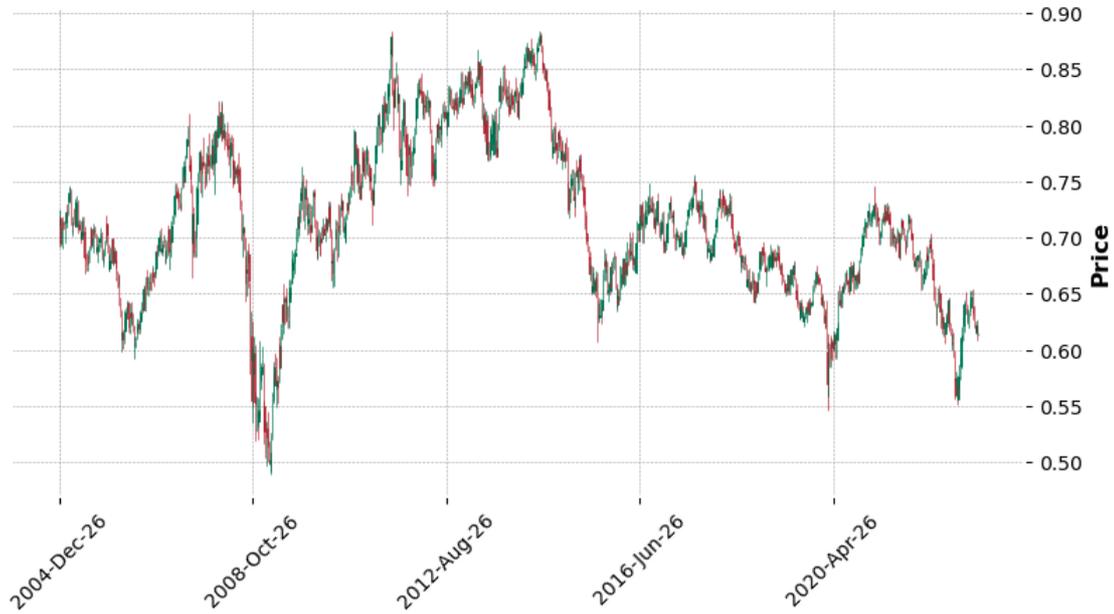


Figura 28: Serie semanal de NZDUSD (Fuente: Elaboración Propia)

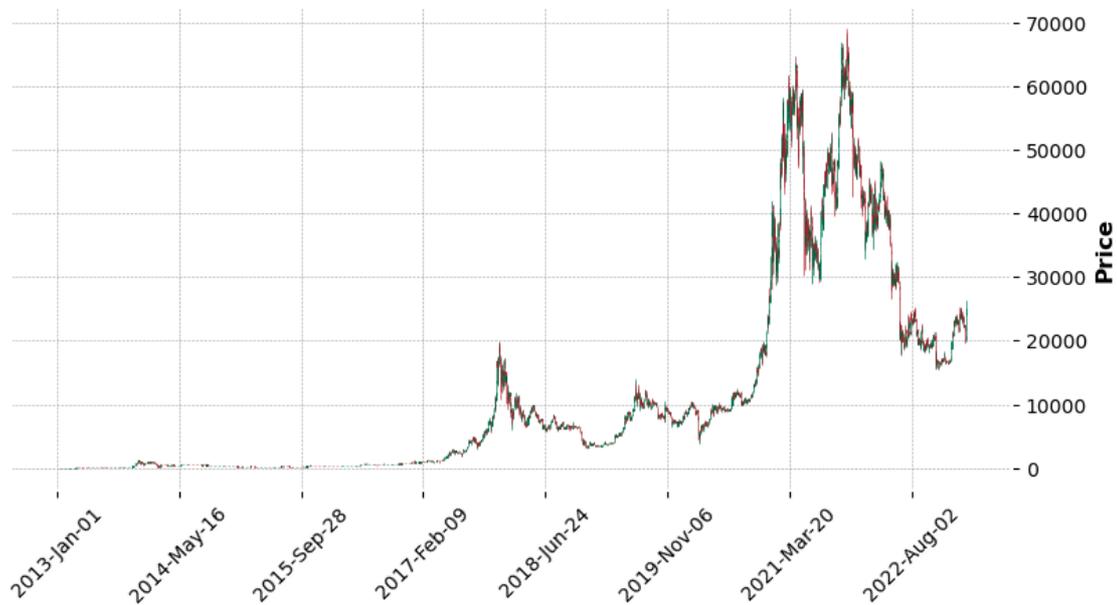


Figura 29: Serie diaria de Bitcoin (Fuente: Elaboración Propia)

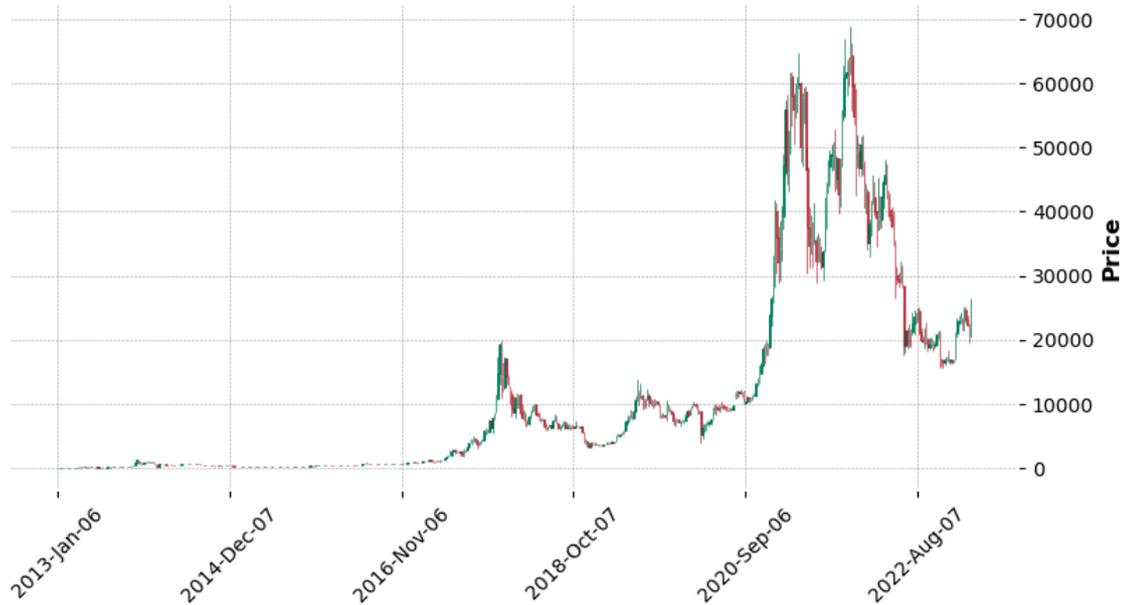


Figura 30: Serie semanal de Bitcoin (Fuente: Elaboración Propia)

5.3 DISEÑO

5.3.1 REGRESIÓN LINEAL

El modelo de regresión lineal planteado empleará los atributos de una observación para predecir la siguiente. Por ejemplo, en la predicción del valor de cierre en el tiempo t , emplearemos los atributos de valor de apertura, máximo, mínimo y volumen de la observación de $t-1$. Es decir, descartaremos el uso de la variable a predecir, ya que este enfoque será revisado en los modelos autorregresivos.

Se procede a seleccionar las variables independientes y, posteriormente, dividir el conjunto de datos en cuatro subconjuntos diferentes. Dos subconjuntos serán los que contengan las variables independientes, empleándose uno para el entrenamiento y el otro para el test. Los otros dos conjuntos son los que contienen los valores reales de la variable dependiente, es decir, de la variable que se va a predecir. El conjunto de entrenamiento engloba los datos del 1 de enero de 2013 hasta el 1 de enero de 2022, mientras que el conjunto de datos de test va del 2 de enero de 2022 hasta el 12 de marzo del 2023.

5.3.2 AUTORREGRESIVO I

Para el desarrollo del modelo autorregresivo se va a seguir empleando una arquitectura de regresión lineal, con la particularidad de que las variables predictivas que se van a emplear en este caso serán la variable dependiente con diferentes números de retrasos. En otras palabras, trataremos de predecir la variable dependiente del instante t , con los valores de la variable dependiente en los instantes $t-i$, para diferentes valores de i .

El primer paso que se debe llevar a cabo es determinar el tamaño de la ventana de valores pasados que queremos incluir. Para ello, emplearemos la prueba de la autocorrelación parcial, la cual nos indica la correlación que existe entre las observaciones pasadas y la observación que se busca predecir. Los resultados son claros, la correlación es alta únicamente con la observación previa. Esto sucede para las 14 series de pares de divisas de las que disponemos. A continuación, se observa la gráfica que indica la autocorrelación entre los distintos instantes temporales en la serie del euro/dólar diario, aunque la gráfica es virtualmente idéntica en el resto de las series. Seleccionaremos para el modelo con autorregresivos únicamente la observación previa.

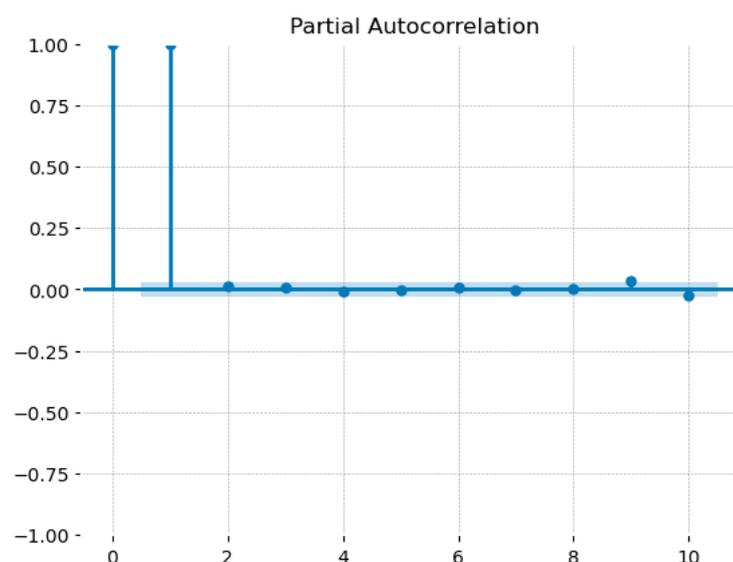


Figura 31: Gráfico de autocorrelación de la serie euro/dólar diario (Fuente: Elaboración propia)

5.3.3 AUTORREGRESIVO II

Se empleará un nuevo modelo autorregresivo, el cual se basará en la arquitectura ARIMA. Las siglas indican que se trata de un modelo autorregresivo (AR), integrado (I) y de media móvil.

Que el modelo sea integrado hace referencia a que el modelo calcula la diferencia entre una observación y su observación previa, convirtiendo la serie en estacionaria. La media móvil es un indicador que calcula el valor medio de las n últimas observaciones. La finalidad de dicho indicador es obtener un precio medio que no sea sensible a fluctuaciones rápidas en el precio.

Se entrenará un modelo que incluya las cinco observaciones previas, que calcule la diferencia solo con el valor anterior y que calcule la media móvil de una observación. Se planteó inicialmente calcular la media móvil de más valores, pero los tiempos de computación eran demasiado elevados, por lo que se decidió emplear un valor pequeño y realizar cambios, si es que se considera oportuno, en futuros modelos.

5.3.4 LSTM

El modelo final que se va a emplear es un tipo de red neuronal recurrente conocido como LSTM. La arquitectura elegida para este modelo es sencilla, primero se añade una capa de 100 unidades LSTM, seguida de una capa de dropout al 15%, es decir, en cada iteración del entrenamiento se desactiva el 15% de las unidades con el objetivo de combatir el sobreajuste de la red. Después de la capa de dropout, se añade una nueva capa LSTM con 50 unidades y finalmente otra capa de dropout al 15%. La capa de salida consta de una única neurona.

Como función de optimización se emplea *Adam*, un optimizador cuyo uso en redes neuronales está muy extendido. Éste incluye una serie de valores predefinidos para los hiperparámetros que controlan el aprendizaje, consiguiendo generalmente los mejores resultados de cualquier función de optimización predefinida. Para la función de pérdida se ha optado por el error cuadrático medio.

Finalmente, se requiere hacer un tratamiento de los datos para el entrenamiento de la red. Se ha elegido una ventana de 14, es decir, se emplearán las 14 previas observaciones en la predicción del valor. Por ello, la red recibe un tensor que contiene vectores de 14 observaciones. A pesar de haber comprobado en modelos previos que la autocorrelación solo es significativa con un retraso, queremos tratar de sacar partido a la memoria a largo plazo que tiene el modelo LSTM incluyendo muestras más antiguas.

5.4 PRUEBA

La evaluación de los resultados de los diferentes modelos se realizará empleando el coeficiente de determinación, comúnmente conocido como R cuadrado. Ésta es una medida, con valores comprendidos entre 0 y 1, la cual determina qué porcentaje de la variabilidad de la serie real es capaz de ser explicada por el modelo. Valores cercanos a 1 muestran que el modelo es capaz de capturar la mayor parte de la variabilidad y, por tanto, podremos decir que el modelo es bueno prediciendo la serie.

El cálculo del coeficiente de determinación se obtiene de dos medidas de diferencia: la suma de los cuadrados de los residuos, es decir, los valores reales menos los predichos; y la suma de los cuadrados de las diferencias entre los valores reales y la media. La fórmula exacta se muestra a continuación, donde \hat{y}_t representa el valor predicho para la observación t , \bar{y} la media e y_t el valor real de la observación t .

$$R^2 = 1 - \frac{\sum_t (y_t - \hat{y}_t)^2}{\sum_t (y_t - \bar{y})^2}$$

5.4.1 REGRESIÓN LINEAL

Los resultados obtenidos mediante el modelo de regresión lineal se muestran a continuación. Se ha elegido emplear la métrica de R2 para medir el ajuste de la serie predicha a la serie real.

Valor de apertura

Divisas	R2
EURUSD_d	99,96%
USDJPY_d	99,98%
GBPUSD_d	99,98%
USDCHF_d	99,94%
AUDUSD_d	99,95%
USDCAD_d	99,98%
NZDUSD_d	99,96%
EURUSD_w	99,74%
USDJPY_w	99,91%
GBPUSD_w	99,85%
USDCHF_w	99,65%
AUDUSD_w	99,66%
USDCAD_w	99,86%
NZDUSD_w	99,76%

Tabla 7: Resultados del modelo de regresión lineal para el valor de apertura (Fuente: Elaboración propia)

Valor de cierre	
Divisas	R2
EURUSD_d	98,36%
USDJPY_d	99,02%
GBPUSD_d	98,61%
USDCHF_d	95,87%

AUDUSD_d	96,77%
USDCAD_d	97,70%
NZDUSD_d	98,15%
EURUSD_w	87,94%
USDJPY_w	93,17%
GBPUSD_w	90,56%
USDCHF_w	68,15%
AUDUSD_w	77,39%
USDCAD_w	84,25%
NZDUSD_w	84,77%

Tabla 8: Resultados del modelo de regresión lineal para el valor de cierre (Fuente: Elaboración propia)

Lo primero que llama la atención de los resultados es la gran diferencia que hay entre los resultados del modelo de apertura y del de cierre. Se ha tratado de buscar una explicación, pero no se ha averiguado porque los resultados obtenidos en el modelo de apertura son casi perfectos. Al tratarse de coeficientes de determinación tan cercanos al 100% y debido a la gran diferencia de resultados con el modelo de cierre, vamos a descartar los resultados de la serie de apertura y clasificarlos como erróneos.

En cuanto a los resultados del modelo de cierre, las series predichas se ajustan de forma correcta a las series reales, tal y como se puede comprobar de los altos valores del R2. Sin embargo, estos resultados son engañosos. Al emplear los datos de la observación previa a la que se busca predecir, el modelo ha aprendido únicamente a calcular un valor cercano al del instante previo al que se quiere predecir. En aquellas series en las cuales la volatilidad es baja, el ajuste de la predicción, marcado por el R2, es bueno, pues los valores no se separan en gran medida de una observación a la siguiente. En casos con mayor volatilidad, como en USDCHF_w, las predicciones no son buenas, pues al haber tanta volatilidad la mayoría de los picos en las predicciones se encuentran invertidos. En la gráfica que se muestra a continuación se puede observar este fenómeno.

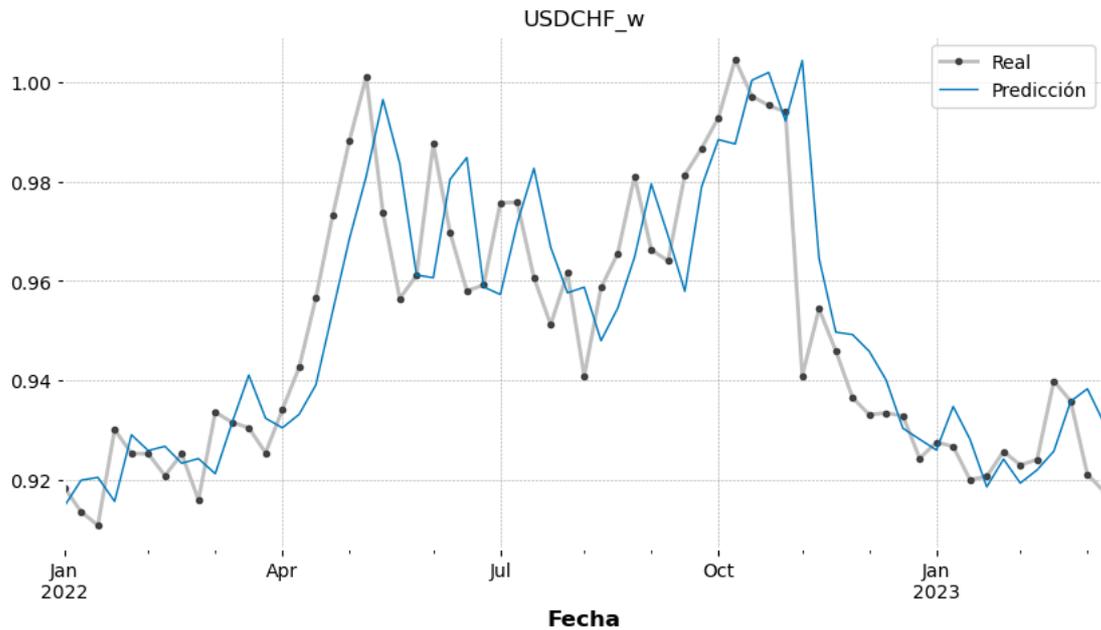


Figura 32: Serie real vs serie predicha para el valor de cierre del par dólar/franco suizo semanal
(Fuente: Elaboración propia)

El problema observado con este modelo es que no se está consiguiendo predecir movimientos futuros, solo se consigue emular valores pasados. Con el objetivo de mejorar los resultados, se realizará una nueva regresión incluyendo variables externas a la serie.

5.4.2 AUTORREGRESIVO I

Los resultados obtenidos mediante el modelo con autorregresivos de orden 1 se muestran a continuación.

Valor de apertura	
Divisas	R2
EURUSD_d	98,61%
USDJPY_d	99,20%
GBPUSD_d	98,82%
USDCHF_d	96,69%

AUDUSD_d	97,18%
USDCAD_d	97,95%
NZDUSD_d	98,37%
EURUSD_w	91,25%
USDJPY_w	94,41%
GBPUSD_w	93,19%
USDCHF_w	74,54%
AUDUSD_w	84,81%
USDCAD_w	90,74%
NZDUSD_w	90,12%

Tabla 9: Resultados del modelo autorregresivo I para el valor de apertura (Fuente: Elaboración propia)

Valor de cierre	
Divisas	R2
EURUSD_d	98,63%
USDJPY_d	99,20%
GBPUSD_d	98,83%
USDCHF_d	96,75%
AUDUSD_d	97,22%
USDCAD_d	97,98%
NZDUSD_d	98,45%
EURUSD_w	91,26%
USDJPY_w	94,27%
GBPUSD_w	93,49%
USDCHF_w	76,57%

AUDUSD_w	85,48%
USDCAD_w	90,63%
NZDUSD_w	89,89%

Tabla 10: Resultados del modelo autorregresivo I para el valor de cierre (Fuente: Elaboración propia)

Igual que sucedió en el primer modelo de regresión lineal, los ajustes de las series predichas son buenos debido a que únicamente se reproduce el valor previo al que se quiere predecir. En este caso es muy evidente, el modelo solo desplaza la serie original una posición para obtener la serie predicha. La serie representada a continuación es la predicción frente al valor real de la serie euro/dólar semanal, donde vemos que la predicción son valores casi idénticos a los previos.

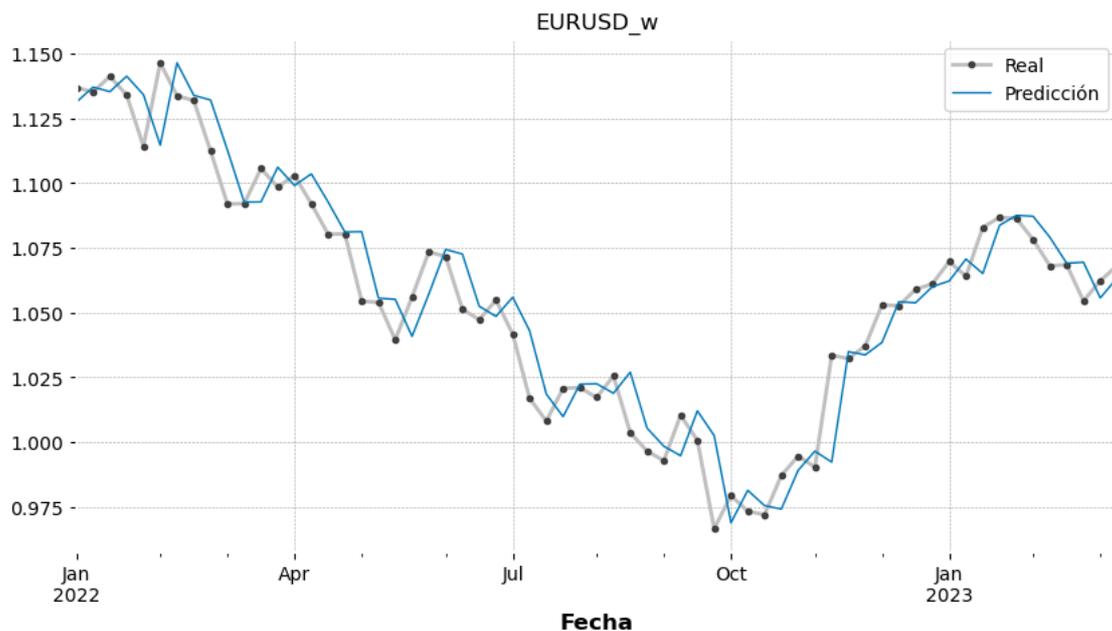


Figura 33: Serie real vs serie predicha para el valor de apertura del par euro/dólar estadounidense semanal (Fuente: Elaboración propia)

5.4.3 AUTORREGRESIVO II

Tanto los resultados obtenidos con el modelo de regresión lineal como en el modelo autorregresivo previo han sido únicamente desplazamientos de la serie original. Se busca con el segundo modelo de autorregresivos inferir algún comportamiento de las series.

Valor de apertura	
Divisas	R2
EURUSD_d	99,49%
USDJPY_d	99,77%
GBPUSD_d	99,15%
USDCHF_d	98,87%
AUDUSD_d	98,90%
USDCAD_d	98,76%
NZDUSD_d	99,07%
EURUSD_w	95,61%
USDJPY_w	97,57%
GBPUSD_w	93,92%
USDCHF_w	90,58%
AUDUSD_w	93,35%
USDCAD_w	90,30%
NZDUSD_w	93,22%

Tabla 11: Resultados del modelo autorregresivo II para el valor de apertura (Fuente: Elaboración propia)

Valor de cierre

Divisas	R²
EURUSD_d	99,50%
USDJPY_d	99,77%
GBPUSD_d	99,16%
USDCHF_d	98,88%
AUDUSD_d	98,93%
USDCAD_d	98,74%
NZDUSD_d	99,11%
EURUSD_w	95,61%
USDJPY_w	97,60%
GBPUSD_w	93,99%
USDCHF_w	90,46%
AUDUSD_w	93,34%
USDCAD_w	90,25%
NZDUSD_w	93,22%

Tabla 12: Resultados del modelo autorregresivo II para el valor de cierre (Fuente: Elaboración propia)

En línea con los resultados de los modelos previos, los altos coeficientes de determinación resultantes no representan una buena predicción, puesto que se sigue empleando el valor previo como la predicción en el modelo. Este resultado era de esperar, debido a que, en el análisis de los coeficientes de autocorrelación, comprobamos que de las observaciones previas la más significativa era la anterior a la que se quería predecir. Por ello, el aumentar el número de observaciones previas que incluimos en el modelo no ha tenido gran efecto. A continuación, en la *Figura 34*, podemos observar el efecto de retardo que tiene el modelo.

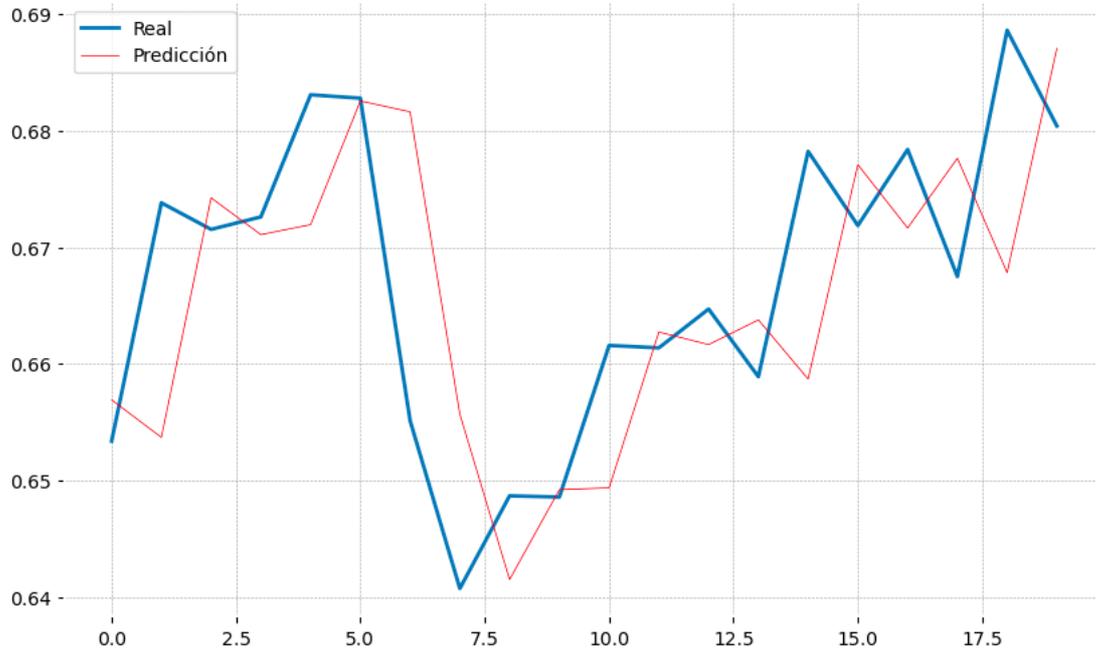


Figura 34: Zoom de la serie real vs serie predicha para el valor de apertura del par dólar neozelandés/dólar estadounidense semanal (Fuente: Elaboración propia)

5.4.4 LSTM

Los resultados de los modelos LSTM de apertura y cierre aparecen en la *Tabla 13* y *Tabla 14*, respectivamente.

Valor de apertura	
Divisas	R2
EURUSD_d	93,07%
USDJPY_d	99,17%
GBPUSD_d	94,45%
USDCHF_d	68,63%
AUDUSD_d	65,55%
USDCAD_d	87,98%

NZDUSD_d	93,23%
EURUSD_w	88,87%
USDJPY_w	95,25%
GBPUSD_w	91,63%
USDCHF_w	72,77%
AUDUSD_w	81,87%
USDCAD_w	79,36%
NZDUSD_w	91,30%

Tabla 13: Resultados del modelo LSTM para el valor de apertura (Fuente: Elaboración propia)

Valor de cierre	
Divisas	R2
EURUSD_d	92,55%
USDJPY_d	98,60%
GBPUSD_d	98,26%
USDCHF_d	88,82%
AUDUSD_d	97,70%
USDCAD_d	96,24%
NZDUSD_d	97,38%
EURUSD_w	83,47%
USDJPY_w	95,79%
GBPUSD_w	87,71%
USDCHF_w	55,68%
AUDUSD_w	66,11%

USDCAD_w	55,85%
NZDUSD_w	48,68%

Tabla 14: Resultados del modelo LSTM para el valor de cierre (Fuente: Elaboración propia)

Los resultados obtenidos con los modelos LSTM son variados. Hemos podido observar que aquellas series con mayor volatilidad han tenido peores ajustes. Un ejemplo de ello es el par entre el dólar estadounidense y el dólar canadiense, mostrado en la *Figura 35*, donde podemos observar que el modelo no ha sido capaz de ajustarse a los picos más pronunciados. Esto se puede deber a la inclusión de un mayor número de observaciones pasadas, haciendo que la contribución conjunta de éstas suavice algún valor extremo que se pueda dar.

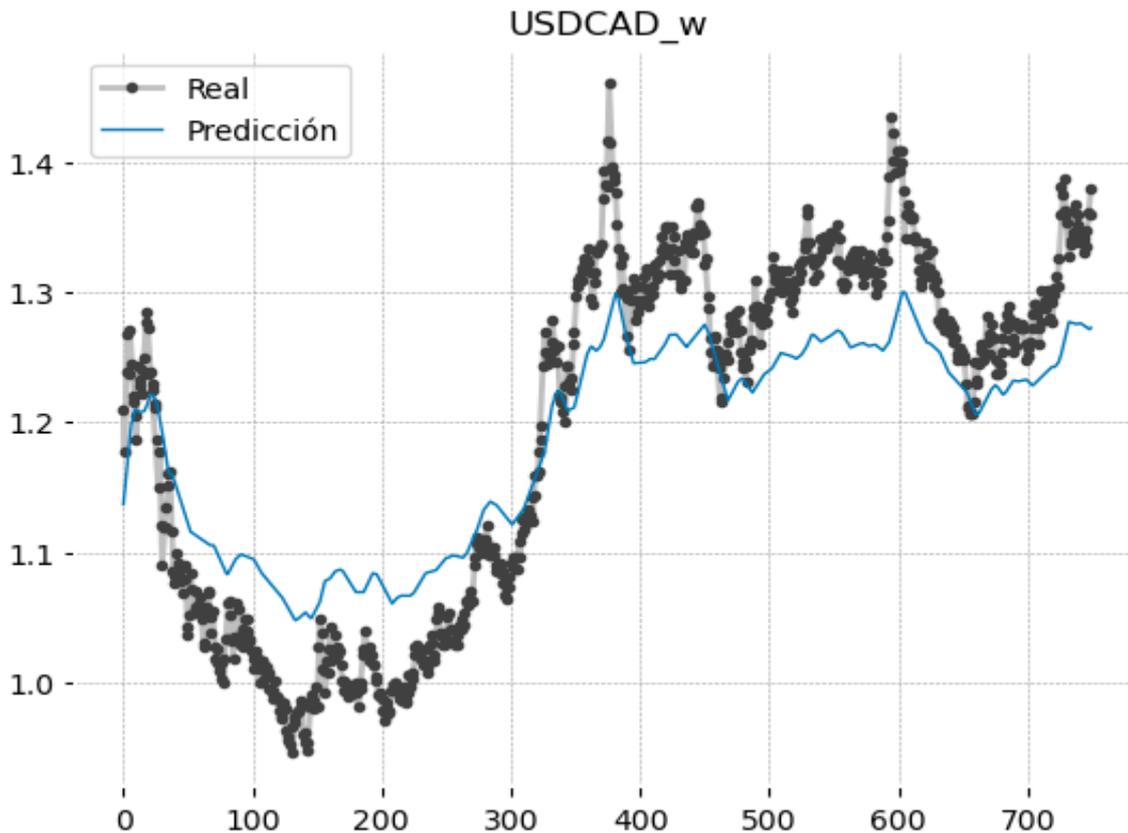


Figura 35: Serie real vs serie predicha para el par dólar estadounidense /dólar canadiense semanal (Fuente: Elaboración propia)

Además, el modelo no es capaz de realizar predicciones. Necesita observar un cambio en la tendencia para que la predicción también cambie la tendencia.

5.5 COMPARACIÓN DE RESULTADOS

Los resultados obtenidos en los cuatro modelos realizados no han sido satisfactorios. A pesar de que los modelos hayan sido capaces de ajustarse a las series, no han sido capaces de modelar el comportamiento de éstas. Simplemente se replicaba un valor cercano a la última observación y, si no se producía un cambio brusco en el valor, el valor predicho era cercano al real. El problema de esto es que, cuando se produce una subida pronunciada seguida de una bajada, cuando se produzca la bajada el modelo predecirá que el valor sigue subiendo, generando un desajuste importante. Si este caso lo traducimos a una situación real de inversión, la operación se saldaría con pérdidas significativas. Por este motivo no vamos a considerar los modelos como válidos para su uso.

El objetivo de los modelos es que éstos sean capaces de adelantarse a comportamientos que puedan tener los activos financieros. Concluimos que únicamente con los datos de la serie este objetivo no se puede alcanzar. Para solucionar esta situación, se va a desarrollar una nueva tanda de modelos que incluyan variables macroeconómicas y financieras. El motivo es que los mercados se mueven como reacción a sucesos previos, algunos de índole excepcional, como una guerra, pero otros, como los tipos de interés o la inflación, se pueden cuantificar e introducir en los modelos. Se espera así que los modelos generen relaciones de causalidad entre sucesos ajenos a la serie y la propia serie.

Se emplearán los modelos de regresión lineal y de LSTM. Los modelos autorregresivos se han descartado, ya que hemos podido comprobar en los apartados previos que eventos previos la serie no arrojan información útil para la predicción.

5.6 REQUISITOS II

Para la segunda iteración en el desarrollo de los modelos, se van a incluir una serie de variables exógenas, compuestas por variables macroeconómicas y financieras. Además de las variables externas, se seguirán empleando las variables usadas en la primera iteración.

Se han buscado variables macroeconómicas y financieras que tengan relación con movimientos en los tipos de cambio. La primera variable macroeconómica que se relaciona con el valor de las divisas son los tipos de interés. Esto se debe a que cuando los tipos de interés suben en un país se genera un efecto llamada de capital externo como consecuencia del aumento de las rentabilidades. Consecuentemente, el aumento en la demanda de esa moneda hará que se aprecie con respecto a otras divisas. Se han seleccionado dos tipos de interés, el marcado por el Banco Central Europeo y el marcado por la Reserva Federal de Estados Unidos, ya que más del 80% de los intercambios de moneda son entre el euro y el dólar estadounidense.

De acuerdo con la importancia de la zona euro y de Estados Unidos, se han incluido indicadores de la inflación de ambos. Para la zona euro y para Estados Unidos se ha elegido el índice de precios del consumo (IPC) como indicador, el cual mide la inflación comparando la evolución en los precios de una cesta de los productos más consumidos por los ciudadanos de un país. Hemos decidido incluir métricas de inflación ya que éstas representan el valor de una moneda, por lo que una pérdida de valor de una divisa podría reducir su atractivo y generar una depreciación en su valor.

A continuación, se han seleccionado los índices de los tres mercados de acciones más grandes del mundo: el New York Stock Exchange (NYSE), el NASDAQ y el Shanghái Stock Exchange (SSE). Con la selección de estos indicadores financieros se busca encontrar relación, si es que hay, entre mercados en situación de expansión o de recesión y el valor de las divisas.

Finalmente, incluiremos una materia prima en la predicción de los modelos: el oro. Esto es debido a que, en épocas de incertidumbre y volatilidad, se suele buscar refugio en el oro debido a su característica estabilidad.

Los datos han sido obtenidos de las diferentes fuentes:

- SSE: Investing.com.
- NASDAQ y NYSE: Nasdaq.com.
- Tipo de interés de la Reserva Federal: Investing.com.
- IPC de Estados Unidos: Bureau of Labor Statistics.
- Oro: World Gold Council.
- Tipo de interés del Banco Central Europeo: EPData, citando al BCE como fuente.
- IPC de Europa: ECB Statistical Data Warehouse.

En la *Tabla 15* se encuentra un resumen de los datos descargados, mostrando el nombre del archivo, la frecuencia de las observaciones y las fechas iniciales y finales.

Nombre del archivo	Frecuencia	Fecha de inicio	Fecha de finalización
BCE Interest Rate.csv	Solo se marcan las fechas en las que se produjeron cambios del tipo de interés	14 de octubre de 2008	15 de junio de 2023
Federal Funds Composite Interest Rate.csv	Diaria	25.02.2014	10.03.2023
inflacion_eu.csv	Mensual	1999-01-30 24:00:00	2023-05-30 24:00:00

US IPC.csv	Mensual	Diciembre de 2012	Mayo de 2023
NASDAQ.csv	Diaria	07/01/2013	06/28/2023
NYSE.csv	Diaria	07/01/2013	06/28/2023
SSE.csv	Diaria	01/01/2014	06/29/2026
Gold.xlsx	Diaria	1978-12-29	2023-06-23

Tabla 15: Descripción de los conjuntos de datos de variables exógenas (Fuente: Elaboración propia)

El formato de las fechas representado en la tabla es el que aparece en los conjuntos de datos. Como se puede observar, una de las primeras tareas que se necesitará realizar en el apartado de análisis será homogeneizar el formato de las fechas.

Finalmente, la *Tabla 16* muestra las variables de cada uno de los conjuntos de datos, junto a su tipo de dato y a una breve explicación.

Fuente de datos	Variable	Tipo de dato	Descripción
Tipo de interés BCE	Año	Fecha	Año en el que se produce el cambio de tipo
	Periodo	Fecha (String)	Día y mes en el que se produce el cambio de tipo (Día XX de mes)
	Tipo de interés	Float	Valor del tipo de interés

	Tasa de depósito	Float	Valor de la tasa de depósito
	Tasa de facilidad de préstamo	Float	Valor de la tasa de facilidad de préstamo
Tipo de interés Reserva Federal	Fecha	Fecha	Fecha del intervalo medido
	Último	Float	Valor de cierre del intervalo
	Apertura	Float	Valor de apertura del intervalo
	Máximo	Float	Valor máximo del intervalo
	Mínimo	Float	Valor de mínimo del intervalo
	Vol.	Columna vacía	-
	% var.	Porcentaje	Variación entre observaciones
IPC Estados Unidos	Series Id	String	ID de la serie
	Year	Integer	Año de la observación

	Period	String	Mes de la oobservación (MXX)
	Value	Float	Valor del IPC
IPC Unión Europea	date	Fecha	
	conf	Columna vacía	-
	comment	String	Comentarios del valor registrado
	status 1	String	“Normal value” como único valor
	s1	Float	Valor del IPC
NASDAQ	Date	Fecha	Fecha del intervalo
	Close/Last	Float	Valor de cierre del intervalo
	Volume	Columna vacía	-
	Open	Float	Valor de apertura del intervalo
	High	Float	Valor máximo del intervalo
	Low	Float	Valor de mínimo del intervalo

NYSE	Date	Fecha	Fecha del intervalo
	Close/Last	Float	Valor de cierre del intervalo
	Volume	Columna vacía	-
	Open	Float	Valor de apertura del intervalo
	High	Float	Valor máximo del intervalo
	Low	Float	Valor de mínimo del intervalo
SSE	Date	Fecha	Fecha del intervalo
	Price	Float	Valor de cierre del intervalo
	Open	Float	Valor de apertura del intervalo
	High	Float	Valor máximo del intervalo
	Low	Float	Valor de mínimo del intervalo

	Vol.	Integer	Volumen del intervalo
	Change %	Float	Variación entre observaciones
Gold	Unnamed 0:	Columna vacía	-
	Unnamed 1:	Columna vacía	-
	Unnamed 2:	Fecha	Fecha del intervalo
	USD	Float	Valor del intervalo

Tabla 16: Explicación de las variables de los conjuntos de datos exógenos (Fuente: Elaboración Propia)

5.7 ANÁLISIS II

Para el procesado de los datos, el primer paso que se realizará será retirar las variables de los conjuntos de datos que no vayamos a emplear en el entrenamiento de los modelos. De todos los conjuntos nos quedaremos únicamente con la variable o variables que indiquen la fecha y con la variable del valor de interés.

La siguiente fase en el procesado de los datos es la selección de la ventana temporal en la que se trabajará. La fecha de inicio más restrictiva es la del conjunto de datos de los tipos de interés de la Reserva Federal, siendo ésta el 25 de febrero de 2014. En cuanto a la fecha de finalización también es el conjunto de datos de la reserva federal el que acota el intervalo, teniendo información hasta el 10 de marzo de 2023. Por ello, los datos se seleccionarán desde el 25 de febrero de 2014 hasta el 10 de marzo de 2023.

A continuación, ajustaremos el formato de las fechas para que coincida con el formato de las series originales de los pares de divisas, concretamente yyyy-mm-dd. En el caso de las variables cuya frecuencia no sea diaria, se extenderán los valores existentes. En el caso de las variables mensuales, todos los días del mes tendrán el valor especificado. En el caso de los tipos de interés del BCE, al solo informar cuando se producen cambios en el tipo, se extenderán los valores hasta que se produzca un cambio en el tipo de interés.

El siguiente paso es el formato de los números. Al descargar datos de diferentes fuentes, los números aparecen representados empleando tanto la coma como el punto como separador decimal. Los modelos que se emplearán en un futuro solo aceptan el punto como separador, por lo tanto, se procede a la conversión del formato de aquellos valores que emplean la coma como separador.

El último paso en el procesado de los datos va a ser otorgar a las variables nombres únicos, ya que posteriormente todas serán unidas en un mismo conjunto de datos con cada uno de los pares de divisas. La lógica que se va a seguir es renombrar la columna a “valor_” seguido del nombre del conjunto de datos. Además, todas las columnas de fecha se han renombrado a “Fecha”.

Completado el procesado de los datos procedemos a representar gráficamente las series macroeconómicas y financieras seleccionadas.

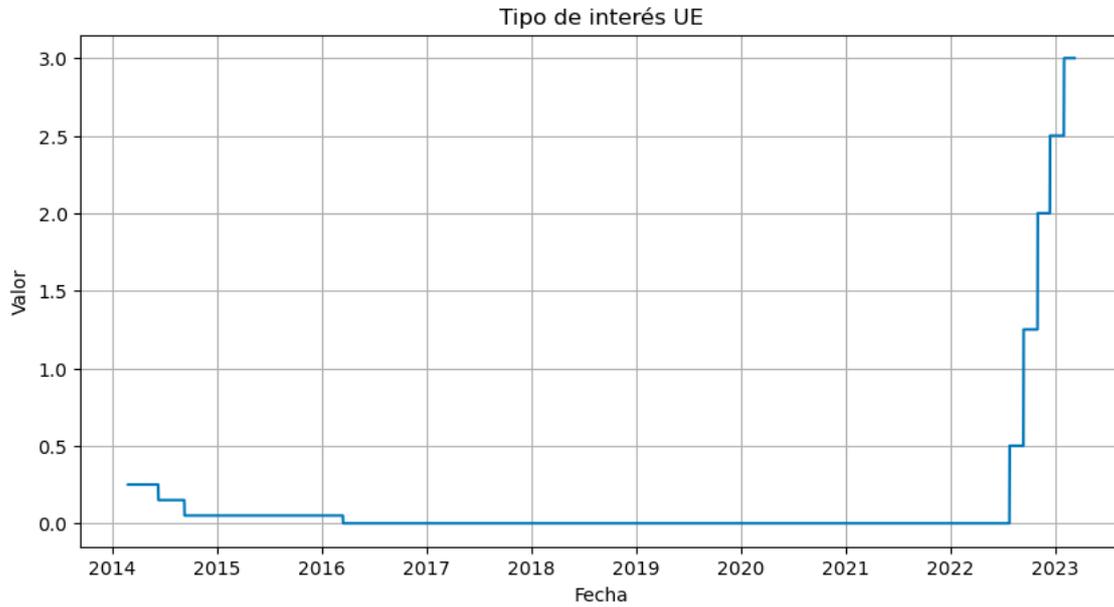


Figura 36: Evolución del tipo de interés del Banco Central Europeo (Fuente: Elaboración propia)

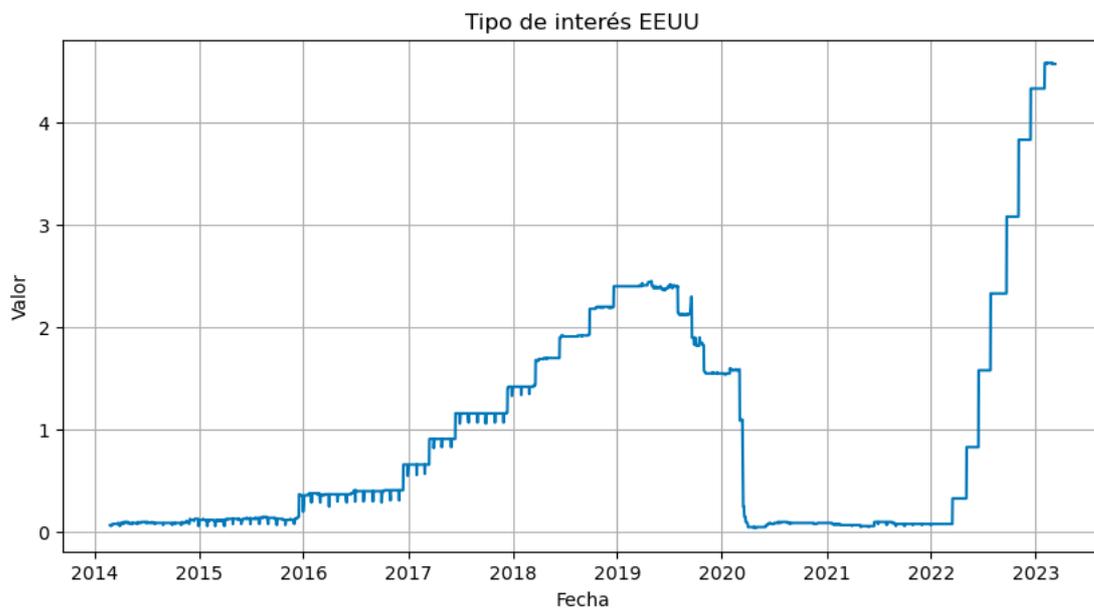


Figura 37: Evolución del tipo de interés de la Reserva Federal (Fuente: Elaboración propia)

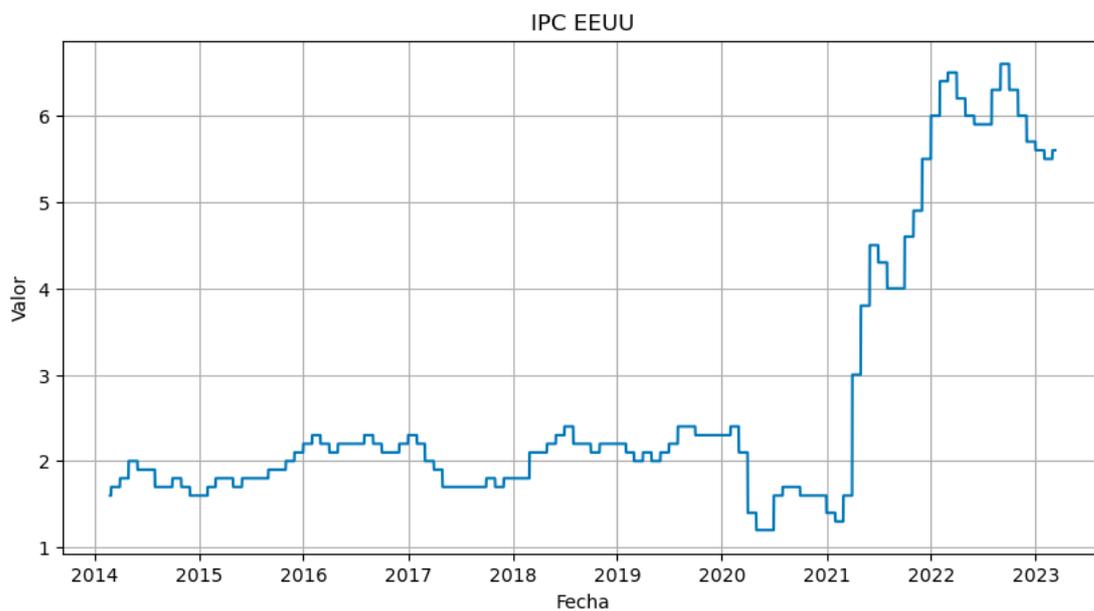


Figura 38: Evolución del IPC en Estados Unidos (Fuente: Elaboración propia)

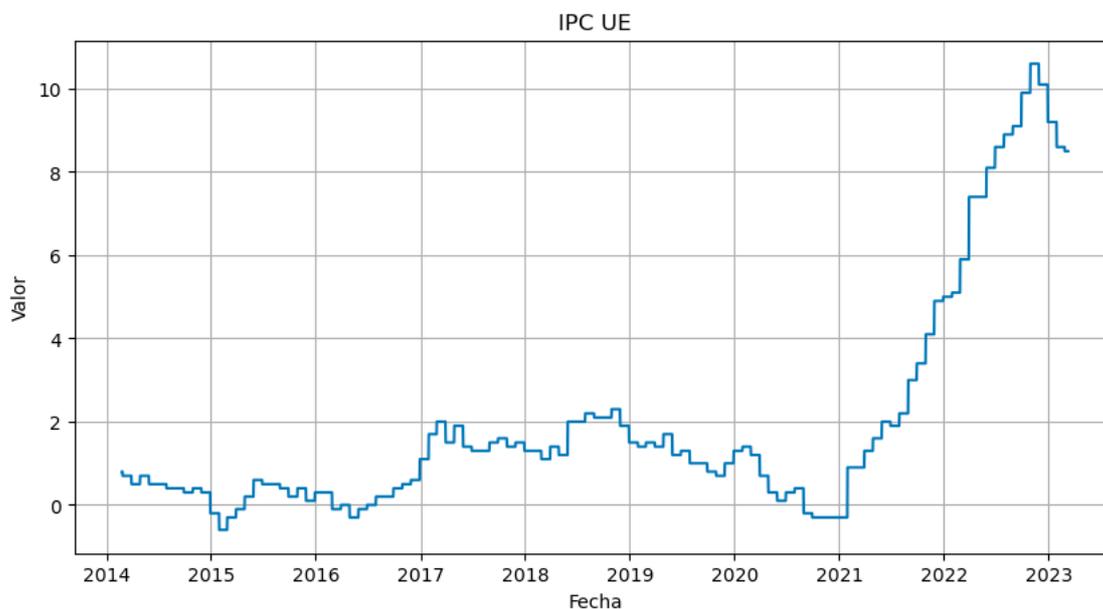


Figura 39: Evolución del IPC en Europa (Fuente: Elaboración propia)



Figura 40: Evolución del índice del NASDAQ (Fuente: Elaboración propia)



Figura 41: Evolución del índice del NYSE (Fuente: Elaboración propia)

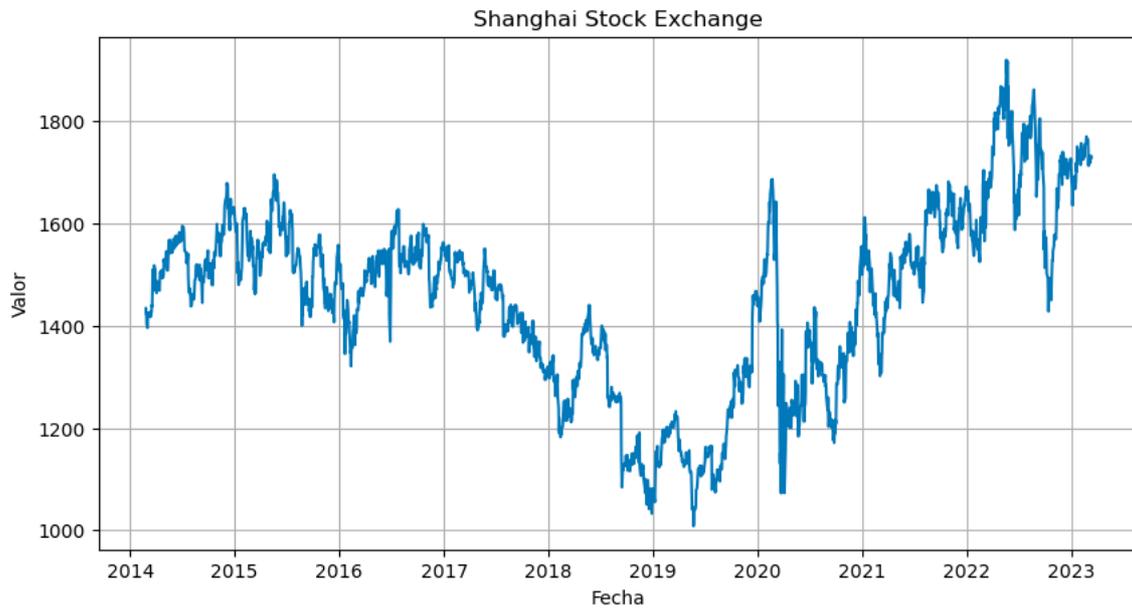


Figura 42: Evolución del índice del SSE (Fuente: Elaboración propia)

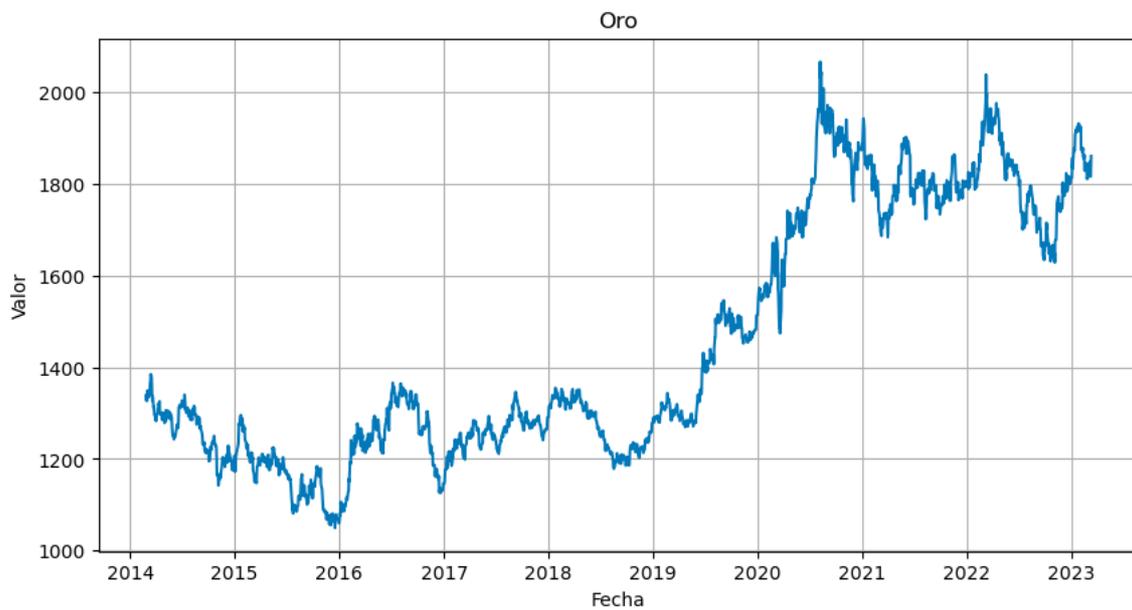


Figura 43: Evolución del precio del oro (Fuente: Elaboración propia)

El paso final, antes de poder emplear los datos en los diferentes modelos, requiere juntar las series de los pares de divisas con las variables financieras y macroeconómicas. Para ello emplearemos un left join, con el objetivo que todas las observaciones tengan datos reales de la serie de divisas. Los valores NAN que queden en las nuevas variables se sustituirán por el primer valor superior que esté informado. El razonamiento es que los valores que no se encuentren informados será debido a que los mercados se encontraban cerrados dicho día, por lo que su valor será el del día anterior. Además, acotamos el espacio temporal de las series de divisas al especificado previamente, del 25 de febrero de 2014 hasta el 10 de marzo de 2023.

Para esta segunda tanda de modelos, se van a emplear únicamente las series diarias de los datos de las divisas. Esto se debe a dos motivos principales, el primero es que los resultados de los primeros modelos fueron mejores para las series diarias y el segundo es debido que el ajuste de fechas entre las nuevas series y las series originales no era posible con datos semanales, al darse muchas fechas no coincidentes.

5.8 DISEÑO II

5.8.1 REGRESIÓN LINEAL

Para el desarrollo del modelo de regresión lineal emplearemos como variables independientes únicamente las variables macroeconómicas y financieras. El objetivo es comprobar si existe alguna relación entre la serie y los nuevos datos, sin nublar los resultados con el resto de las variables de la propia serie. A continuación, se entrenarán las regresiones lineales con todos los datos disponibles, tanto de las series macroeconómicas y financieras como el resto de las variables de las series de las divisas.

Será necesario llevar a cabo una normalización del conjunto de los datos, pues las variables están representadas en escalas diferentes, habiendo desde porcentajes a valores en las decenas de miles.

Para la predicción, se emplearán los datos de la serie del instante temporal previo al que se quiere predecir. Para ello, se desplazan las variables independientes una posición hacia el futuro.

5.8.2 LSTM

La arquitectura propuesta inicialmente para la red LSTM consistía en dos capas LSTM, con 50 unidades cada una. Sin embargo, tras un primer entrenamiento de la red se observó que la arquitectura no era suficientemente profunda como para modelar el comportamiento de la serie. Por ello, se añadió una tercera capa LSTM de 50 unidades. A las capas LSTM les seguía una capa de salida compuesta por una única neurona, encargada de generar el valor predicho a raíz de las entradas recibidas de la capa previa.

Además, fue necesario aumentar el número de épocas que se entrenaba el modelo, es decir, el número de veces que se utiliza el conjunto de datos completo para entrenar el modelo, concretamente, se incrementó dicho valor de 10 a 30, puesto que el modelo seguía reduciendo el error hasta aproximadamente la época 25.

De la misma forma que en la regresión lineal, los primeros modelos incluían únicamente las variables macroeconómicas y financieras y, a continuación, se realizó una segunda tanda de redes LSTM con todas las variables. También, fue necesario normalizar las variables que se empleaban, al haber variables expresadas en distintos órdenes de magnitud.

Finalmente, se emplearán los datos de la serie de divisas y de las series macroeconómicas y financieras del instante temporal previo al que se quiere predecir. Para ello, se desplazan las variables independientes una posición hacia el futuro.

5.9 PRUEBA II

5.9.1 REGRESIÓN LINEAL

Los dos primeros modelos de regresión lineal predicen los valores de apertura y cierre con las variables macroeconómicas y financieras. Estos modelos no han sido capaces de modelar

las series. En todos ellos, la primera mitad de los valores predichos son similares a los reales, pero a partir de ese punto la serie predicha diverge de la real. Además, en todos los casos lo hace de una forma muy característica, siguiendo un patrón escalonado.

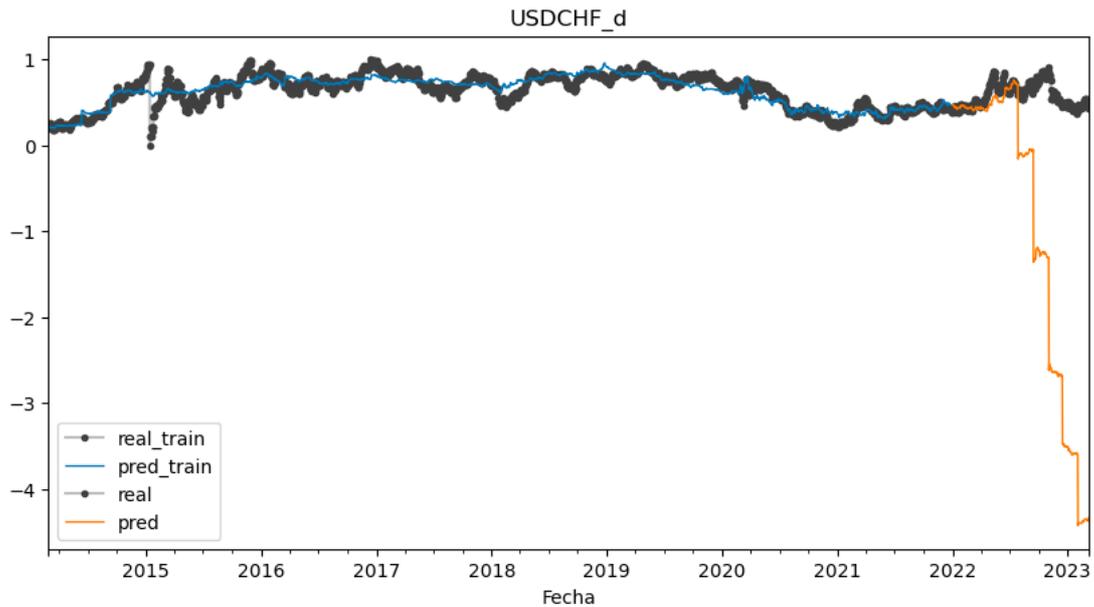


Figura 44: Serie real vs serie predicha con regresión lineal para el valor de cierre del par franco suizo/dólar estadounidense (Fuente: Elaboración propia)

Los dos siguientes modelos, que incluyen todas las variables, son muy similares a los modelos que no usaban las nuevas variables macroeconómicas y financieras. El peso de la predicción lo llevan las variables de la propia serie, por lo que el modelo replica la serie original, pero con un retardo. La *Figura 45* refleja esta situación.

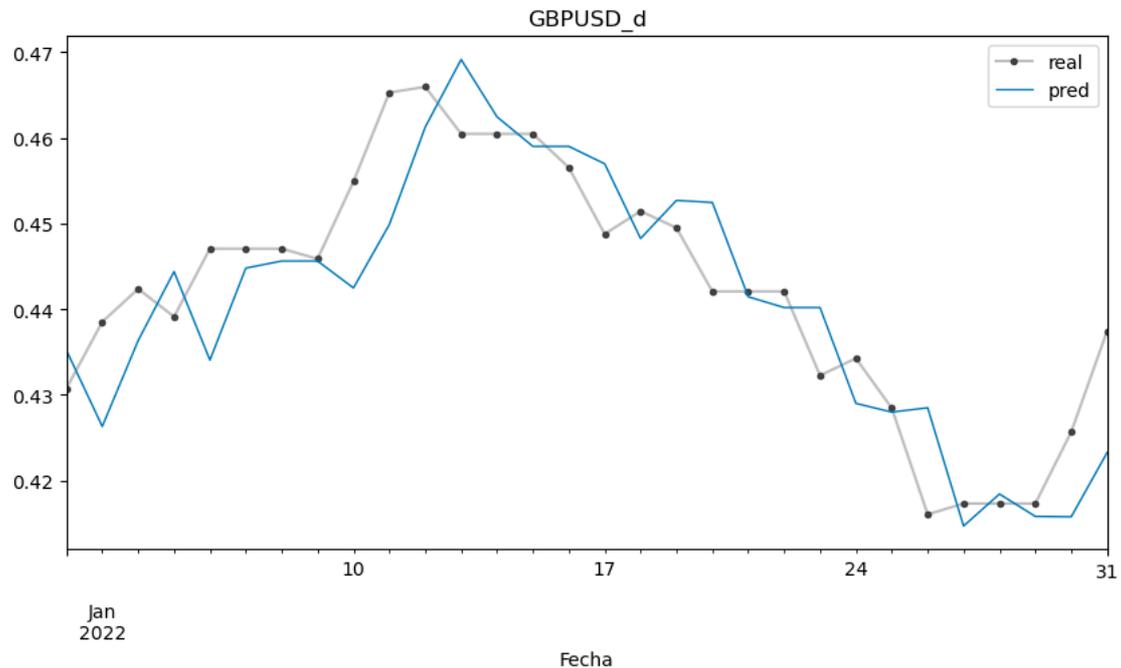


Figura 45: Zoom de la serie real vs serie predicha con regresión lineal para el valor de cierre del libra esterlina/dólar estadounidense (Fuente: Elaboración propia)

5.9.2 LSTM

Los resultados obtenidos en las redes entrenadas únicamente con datos financieros y macroeconómicos no han sido satisfactorios. La serie predicha no se conseguía ajustar a los movimientos de la serie real. De hecho, los resultados del coeficiente de determinación no superan el 15% para ninguna de las series, ni en la predicción del valor de apertura ni el de cierre.

A pesar de ello, en determinadas zonas de las series, el modelo es capaz de captar la tendencia real del tipo de cambio. En la *Figura 46*, se puede observar que, en el tramo inicial, hasta junio del 2022, el modelo es capaz de predecir las tendencias de la serie. Capta el primer tramo, donde la serie es más estable, siendo posteriormente capaz de determinar la subida, la posterior bajada y de nuevo la subida. A pesar de ello, desde ese punto hasta noviembre, la serie predicha no se ajusta lo más mínimo a la real. Finalmente, el modelo determina correctamente que se produce una tendencia bajista en el par.

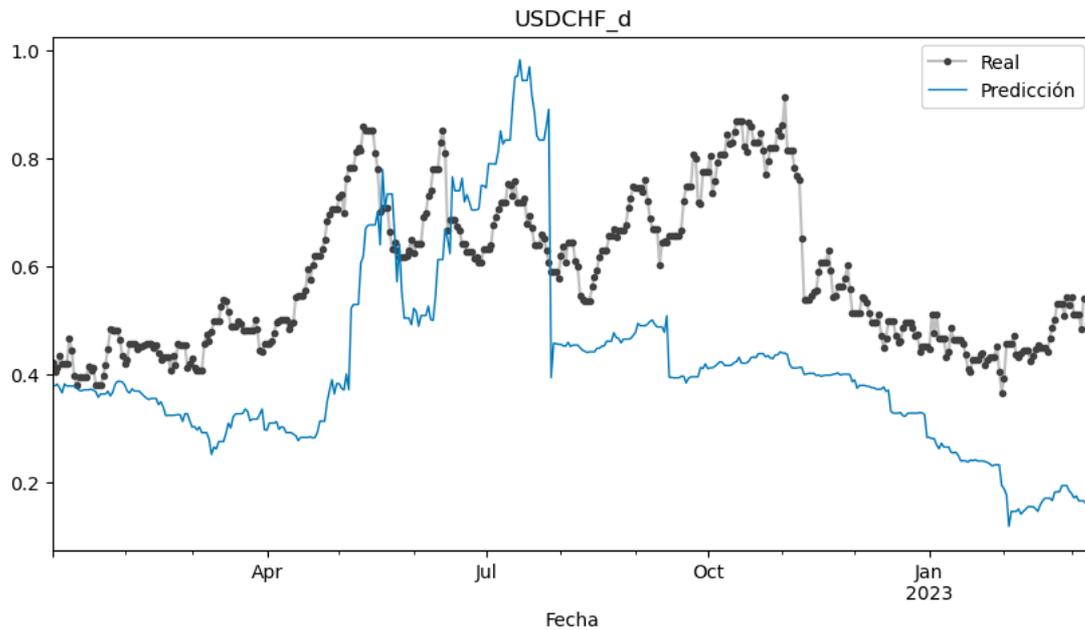


Figura 46: Serie real vs serie predicha con LSTM para el valor de cierre del par franco suizo/dólar estadounidense (Fuente: Elaboración propia)

En cuanto a los modelos entrenados con todas las variables, los resultados estuvieron condicionados por los valores previos de la serie. Como ocurría en los primeros modelos desarrollados, las predicciones son réplicas de los valores previos de la serie. Además, el uso de variables exógenas no ha mejorado los modelos previos, sino que ha provocado que el ajuste sea peor. La *Figura 47* muestra la predicción de la serie del par dólar australiano y dólar estadounidense, donde se aprecia claramente que los movimientos de la serie predicha son una réplica de los movimientos previos de la serie.

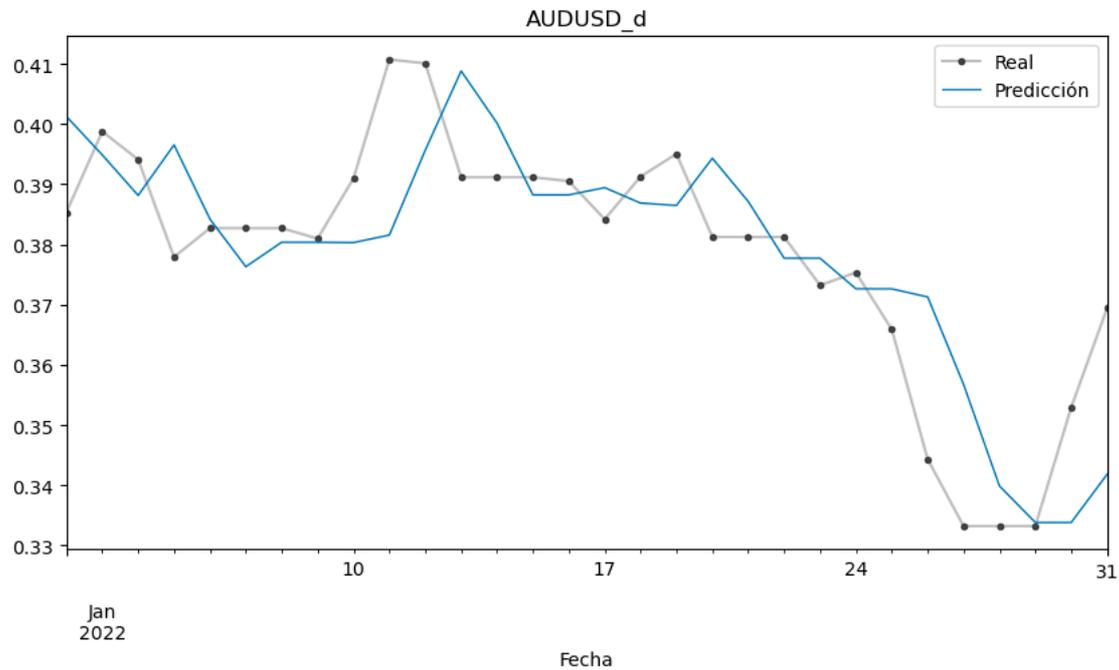


Figura 47: Serie real vs serie predicha con LSTM para el valor de cierre del par dólar australiano/dólar estadounidense (Fuente: Elaboración propia)

5.10 COMPARACIÓN DE RESULTADOS II

La inclusión de datos financieros y macroeconómicos no ha conseguido mejorar los modelos. Además, no ha sido capaz de conseguir que los modelos no basaran sus predicciones en replicar los datos previos.

Los modelos lineales en los que solo se emplearon las variables exógenas han demostrado el peor comportamiento hasta ahora. La simplicidad del modelo no ha sido capaz de replicar las relaciones que pueden encontrarse ocultas en los datos. Por lo tanto, se descarta el uso de estos modelos.

Al añadir además las variables de la serie, el desempeño ha sido similar al de los modelos de la primera iteración. El ajuste era bueno, con coeficientes de determinación altos, pero el modelo no era capaz de predecir, su función era replicar. Se descartará también el uso de

estos modelos, ya que se obtienen los mismos resultados que en los modelos que no empleaban las variables exógenas.

En cuanto a los modelos LSTM, comenzaremos repasando el que únicamente empleaba para el entrenamiento el conjunto de variables macroeconómicas y financieras. En términos de ajuste de la serie predicha el modelo no consiguió modelar la serie real con precisión, obteniendo coeficientes de determinación bajos. Sin embargo, en determinadas zonas de las series, el modelo fue capaz de predecir las tendencias que tomaba la serie real, así como algunos picos en los que se pasaba de una tendencia alcista a bajista, y viceversa.

El último modelo que se ha entrenado en la segunda iteración ha sido el modelo LSTM con todas las variables. De la misma forma que en el resto de los modelos que empleaban el mismo conjunto de variables, las variables de la serie de la divisa son las que han llevado la mayor parte del peso en la predicción. Es por este motivo que los valores predichos por el modelo han resultado ser valores muy similares a los de la serie en la muestra anterior, es decir, no arrojaban información nueva sobre el posible comportamiento del par de divisas.

5.11 PRUEBA DE LOS MEJORES MODELOS EN LA SERIE DE BITCOIN

A la vista de los resultados obtenidos, se han seleccionado dos modelos para estudiar su rendimiento en las series de Bitcoin, con el objetivo de encontrar similitudes y diferencias entre el comportamiento de los modelos con monedas tradicionales y criptomonedas.

En cuestión del tratamiento de los datos, el procedimiento que se ha seguido es análogo al explicado en secciones anteriores para las divisas tradicionales. Fue necesario cambiar el formato de la fecha, cambiar el nombre de las variables, cambiar el separador decimal de coma a punto y convertir las variables finales a tipo numérico, a excepción de la fecha.

En los modelos previos se ha observado que aquellos modelos que se entrenaban con las propias variables de la serie, éstas predominaban en las predicciones, las cuales eran valores muy similares a de la observación previa. Por este motivo, el primer modelo que se va a

emplear pertenece a la primera iteración y es una regresión logística con solo las variables de la serie de Bitcoin.

El modelo de regresión lineal produce resultados parecidos a los que se habían obtenido previamente con el resto de las divisas. Los coeficientes de determinación, es decir, el ajuste de la predicción, son altos, especialmente en el modelo que predice el valor de apertura. A pesar de ello, el buen ajuste se debe a que se están replicando los valores pasados, como podemos observar en la *Figura 48*. Los cambios de tendencias en la serie predicha se dan un instante temporal después de que haya ocurrido en la serie real.

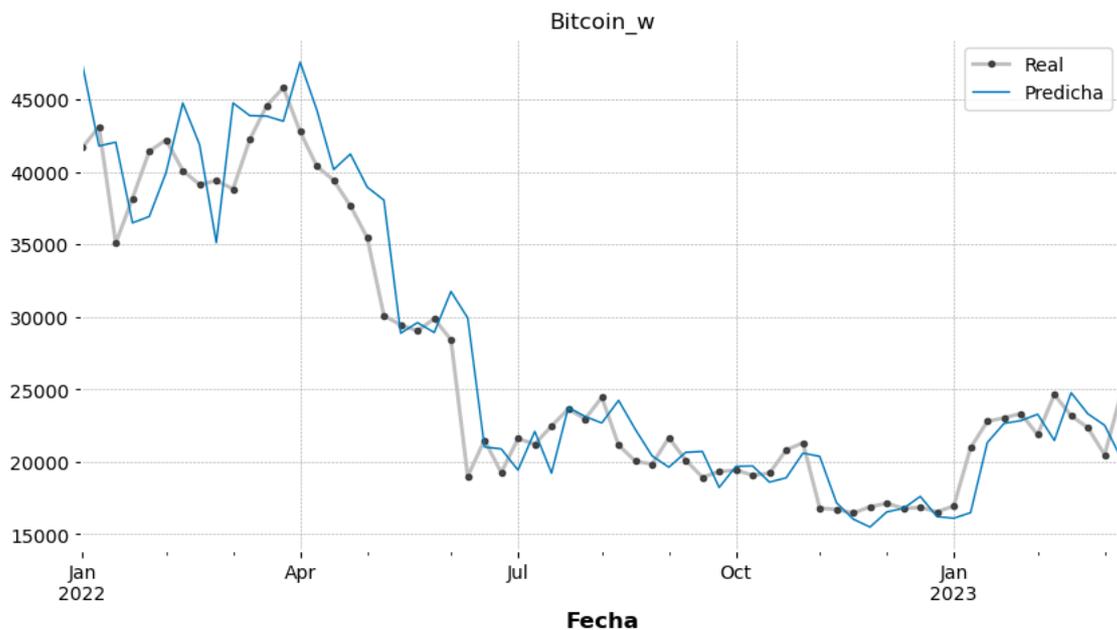


Figura 48: Serie real vs serie predicha con modelo de regresión lineal para el valor de cierre de la serie semanal de Bitcoin (Fuente: Elaboración propia)

A pesar del buen ajuste, en la predicción del valor de cierre para la serie diaria se produce una predicción muy dispar a la serie real. Probablemente, se pueda deber a la volatilidad que caracteriza a la serie total, en la cual se producen subidas o bajadas del precio muy significativas en muy poco tiempo. Esto ha podido resultar en que se repitiera un patrón que en entrenamiento iba seguido de una fuerte bajada, pero que en el caso del test no. La *Figura*

49 muestra que en marzo del 2022 el modelo esperaba una fuerte bajada, similar a la que tuvo lugar en las mismas fechas del año anterior (*Figura 50*).

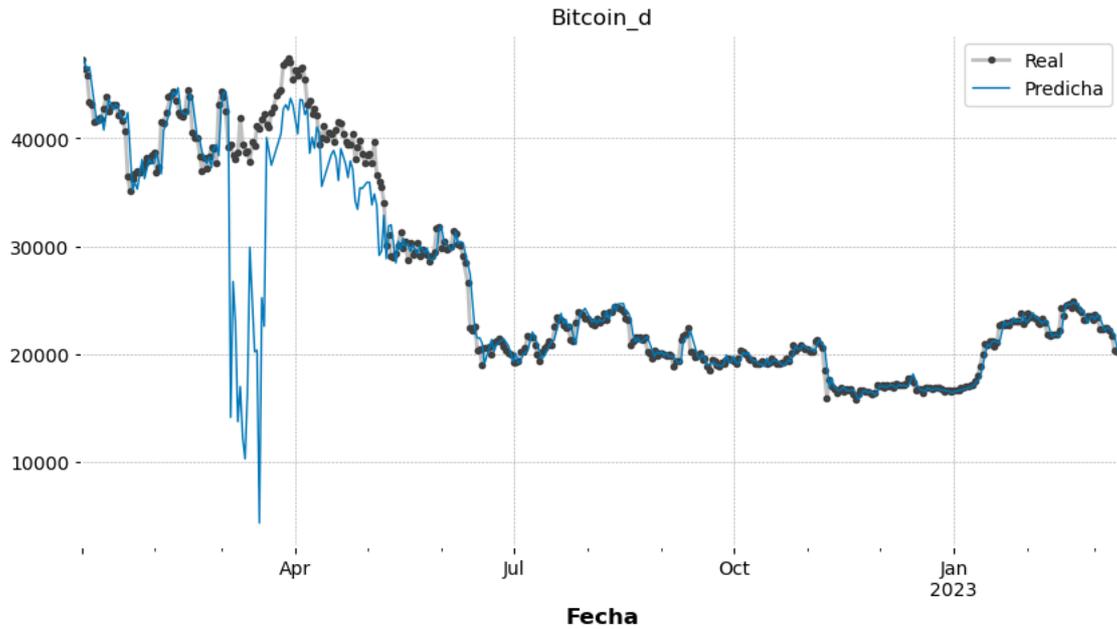


Figura 49: Serie real vs serie predicha con modelo de regresión lineal para el valor de cierre de la serie diaria de Bitcoin (Fuente: Elaboración propia)

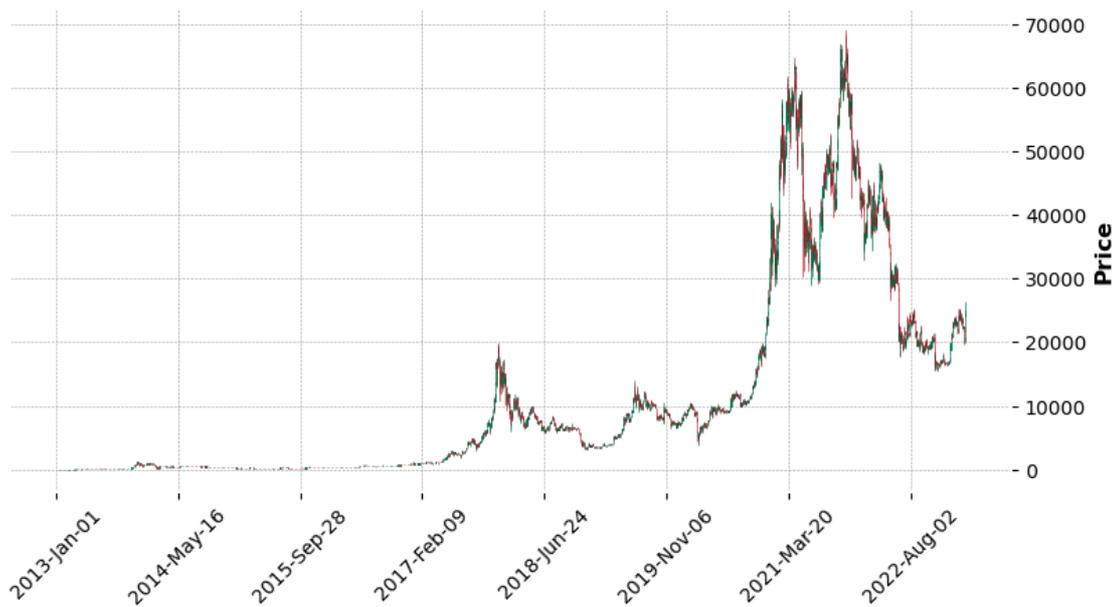


Figura 50: Serie diaria de Bitcoin (Fuente: Elaboración Propia)

El segundo modelo que se va a emplear en la predicción de la serie de Bitcoin va a ser una red LSTM en la que emplearemos únicamente las variables macroeconómicas y financieras para su entrenamiento. Previamente, este modelo no ha dado buenos resultados en cuanto a ajuste, pero ha sido capaz de predecir la tendencia que tomaba la serie real en determinadas ocasiones.

El comportamiento ha sido similar al esperado, con malos ajustes a la serie real, pero zonas en las que el modelo ha conseguido predecir la tendencia ligeramente. Tanto en el modelo del valor de apertura como el del valor de cierre las tendencias han sido modeladas bien en zonas de baja volatilidad, sin embargo, en las zonas de alta volatilidad la predicción no arroja información útil.

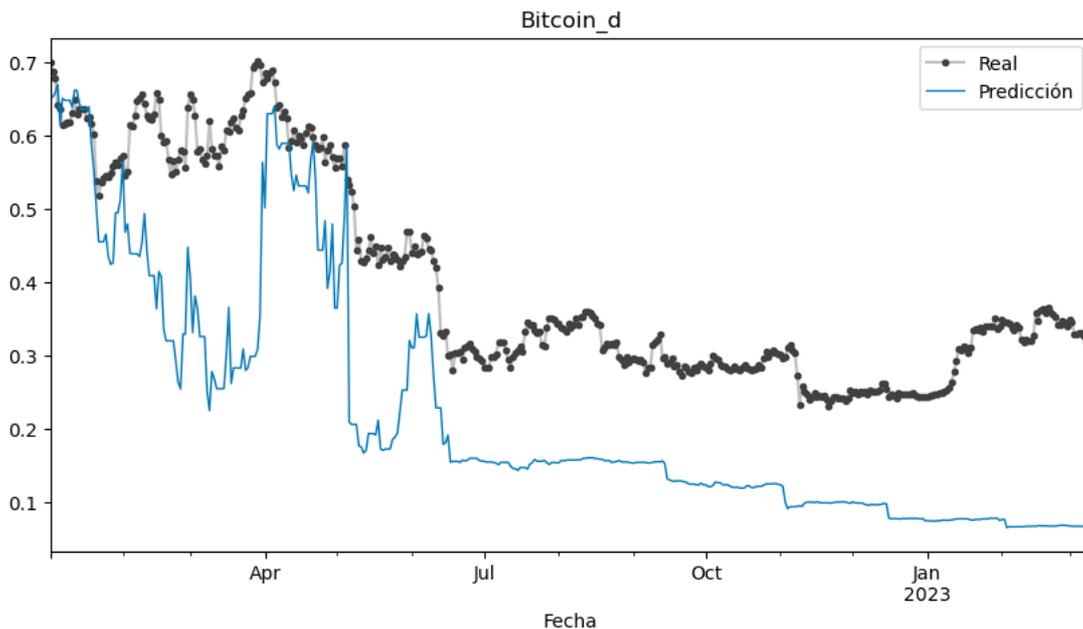
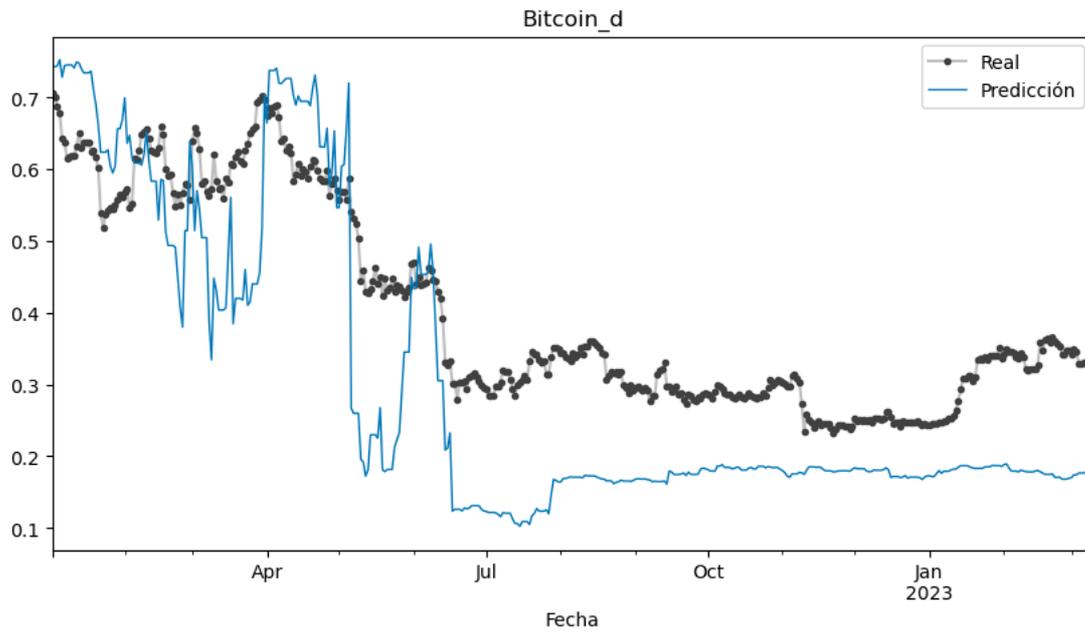


Figura 51: Serie real vs serie predicha con LSTM para el valor de cierre de la serie diaria de Bitcoin (Fuente: Elaboración propia)



*Figura 52: Serie real vs serie predicha con LSTM para el valor de apertura de la serie diaria de Bitcoin
(Fuente: Elaboración propia)*

Capítulo 6. ANÁLISIS DE RESULTADOS

Los resultados de los modelos han venido determinados principalmente por las variables que se han empleado en el entrenamiento del modelo. El tipo de modelo y su arquitectura han caído en un segundo plano. En aquellos modelos donde hubiese datos de la propia serie éstos han predominado en el cálculo de la predicción.

Los modelos buscan reducir el error total que se da en sus predicciones, siendo la forma más sencilla de hacerlo la réplica de los valores previos de la propia serie. La volatilidad en las series, en líneas generales, era baja, lo cual se traduce en poca variación de una observación a la siguiente. Esto ha conseguido que los modelos redujesen el error de la forma más eficiente al generar un valor de predicción similar al valor previo. Si el objetivo de estos modelos fuera alcanzar un ajuste alto se estaría realizando un trabajo adecuado. Sin embargo, el entrenamiento de los modelos buscaba que éstos fueran capaces de predecir valores o tendencias antes de que se produjesen y con seguir así una ventaja sobre el mercado.

Aquellos modelos en los que se han empleado las variables financieras y macroeconómicas y no los datos de la serie, han conseguido arrojar algo de información acerca de los comportamientos de las series. En determinados rangos de tiempo, los modelos han sido capaces de identificar las tendencias que tomaban las series, es decir, han conseguido generar información que dé una ventaja sobre el mercado. El problema es que el ajuste obtenido no es bueno, por lo que su uso es más de carácter cualitativo que cuantitativo. Además, otro problema que presenta el modelo es que hay zonas en las que no se acierta la tendencia y la serie predicha no se ajusta a la real, por lo que en una situación real no se podrá saber cuándo la predicción va a ser correcta y cuándo no.

El resultado obtenido en las series de Bitcoin ha sido similar al de las divisas tradicionales. Hay un claro predominio de las variables de la propia serie a la hora de que el modelo genere la predicción y los coeficientes de determinación son bajos. También, el modelo LSTM ha sido capaz de inferir determinados comportamientos de la serie cuando se entrenaba con

variables exógenas, pero, de la misma forma que con las divisas tradicionales, no es posible saber en qué momento la predicción de la tendencia es correcta o no.

Por otro lado, un aspecto en el que se ha diferenciado el comportamiento de las divisas tradicionales y del Bitcoin ha sido en el ajuste. El Bitcoin, como la mayoría de las criptomonedas, se caracteriza por su volatilidad. Esto ha resultado en situaciones donde la predicción mostraba una volatilidad superior a la de la serie real.

Capítulo 7. CONCLUSIONES Y TRABAJOS FUTUROS

En capítulos previos se revisó bibliografía que afirmaba que, en la predicción de activos financieros, los mejores resultados se obtenían generalmente con modelos de Random Walk, es decir, modelos cuya predicción se basa en sumar una perturbación aleatoria al último valor. Se ha podido comprobar que esta teoría es cierta, ya que los mejores ajustes se basaban en proyecciones que replicaban el valor previo más una perturbación aleatoria.

Por el contrario, en aquellos modelos en los que los datos de la propia divisa no se empleaban para su predicción, se han observado signos de que los modelos LSTM han conseguido modelar tendencias en base a predictores externos. A pesar de que su fiabilidad sea baja, al incluir también predicciones erróneas, se ha alcanzado parcialmente uno de los objetivos del trabajo, la obtención de modelos capaces de predecir tendencias y puntos de inflexión en las series.

El comportamiento de las criptomonedas ha sido más parecido al de las series tradicionales del que se esperaba en un principio. Las predicciones han venido determinadas por los valores previos de la serie, mientras que los modelos que solo incluían variables exógenas, se ha conseguido modelar correctamente ciertas tendencias. Sin embargo, la principal diferencia que se ha observado ha sido que la mayor volatilidad de la serie ha empeorado los resultados de los modelos. Esto también ha ocurrido, aunque en menor medida, en las series de las monedas tradicionales que presentaban alta volatilidad. Concluimos, por lo tanto, que las monedas tradicionales y las criptomonedas son muy similares en su comportamiento, salvo por la mayor volatilidad que caracteriza a las criptomonedas.

Se ha llegado a la conclusión de que la creación de modelos precisos y capaces de adelantarse a los sucesos pasa por el uso de variables financieras y macroeconómicas. Consideramos que el motivo de que haya zonas temporales donde la predicción no se ajusta a los movimientos reales es que no se ha incluido en los datos la variable exógena que tiene el efecto de causa sobre el movimiento de la serie. Por ejemplo, los movimientos provocados

por la crisis del 2008 o por el COVID-19 puede ser que estuvieran mejor representados con otras variables que no hemos incluido y, por tanto, el modelo no ha encontrado la relación de causalidad.

Por este motivo, se considera que para futuros trabajos de debe aplicar un enfoque diferente a la creación de los modelos. Se debe centrar en la selección adecuada de las variables predictivas en función de la situación macroeconómica, política y social que rodee a las dos divisas que constituyen el par. Incluso, el uso de noticias, o el sentimiento que se halla en ellas, podría mejorar los resultados de las predicciones, especialmente en las criptomonedas, ya que su volatilidad viene marcada por en muchas ocasiones por noticias relacionadas o incluso mensajes en redes sociales de personas influyentes.

Capítulo 8. BIBLIOGRAFÍA

Chitenderu, T., Maredza, A., & Sibanda, K. (2014). *The Random Walk Theory And Stock Prices: Evidence From Johannesburg Stock Exchange*.

Forex Market Trading Online. (2023). Obtenido de Forex Early Warning:
<https://www.forexearlywarning.com/forex-trading-course/forex-market-trading>

Hochreiter, S., & Schmidhuber, J. (1997). *LONG SHORT-TERM MEMORY*.

MESSE, R., & ROGOFF, K. (1983). *EMPIRICAL EXCHANGE RATE MODELS OF THE SEVENTIES*.

Tang, Y., Song, Z., Zhu, Y., Yuan, H., Hou, M., Ji, J., . . . Li, J. (2022). *A survey on machine learning models for financial time series forecasting*.

YEŞİN, P. (2016). *Exchange Rate Predictability and State-of-the-Art Models*.

ANEXO I: ALINEACIÓN DEL PROYECTO CON LOS ODS

Los Objetivos de Desarrollo sostenible son 17 objetivos desarrollados por las Naciones Unidas cuya motivación es abordar los principales desafíos a los que la humanidad tiene que hacer frente.

1. Fin de la pobreza
2. Hambre cero
3. Salud y bienestar
4. Educación de calidad
5. Igualdad de género
6. Agua limpia y saneamiento
7. Energía asequible y no contaminante
8. Trabajo decente y crecimiento económico
9. Industria, innovación e infraestructura
10. Reducción de las desigualdades
11. Ciudades y comunidades sostenibles
12. Producción y consumo responsables
13. Acción por el clima
14. Vida submarina
15. Vida de ecosistemas terrestres
16. Paz, justicia e instituciones sólidas
17. Alianza para lograr los objetivos

El proyecto desarrollado ha consistido en crear modelos que sean capaces de predecir los tipos de cambio. Este trabajo puede alinearse con dos de los Objetivos de desarrollo sostenible:

- El primero es el número ocho, trabajo decente y crecimiento económico. El uso de estos modelos puede contribuir al crecimiento económico, al mejorar las herramientas actuales que se emplean en la estimación de los tipos de cambio. Además, se podría controlar de forma más efectiva el riesgo que implica el cambio de divisas, fomentando la estabilidad económica y financiera de empresas y países y, consecuentemente, ayudando a la aparición de trabajo decente y estable.
- El siguiente objetivo es el nueve, industria, innovación e infraestructura. La aparición de modelos con capacidad de predecir los tipos de cambio supondría una innovación en la industria financiera.

Además, se podrían alcanzar otros ODS como consecuencia del alineamiento con los objetivos ocho y nueve. Estos dos objetivos se traducirían en una economía más sana y próspera, por lo que se podría mejorar la situación de forma indirecta de los objetivos uno (Fin de la pobreza) y 10 (Reducción de las desigualdades).