

Foreign Exchange Forecasting Models: LSTM and BiLSTM Comparison [†]

Fernando García ¹, Francisco Guijarro ¹, Javier Oliver ^{1,*} and Rima Tamošiūnienė ²

¹ Department of Economics and Social Sciences, Universitat Politècnica de València, 46022 Valencia, Spain; fergarga@esp.upv.es (F.G.); fraguima@esp.upv.es (F.G.)

² Department of Financial Engineering, Vilnius Gediminas Technical University, 10223 Vilnius, Lithuania; rima.tamosiuniene@vilniustech.lt

* Correspondence: jaolmun@ade.upv.es

[†] Presented at the 10th International Conference on Time Series and Forecasting, Gran Canaria, Spain, 15–17 July 2024.

Abstract: Knowledge of foreign exchange rates and their evolution is fundamental to firms and investors, both for hedging exchange rate risk and for investment and trading. The ARIMA model has been one of the most widely used methodologies for time series forecasting. Nowadays, neural networks have surpassed this methodology in many aspects. For short-term stock price prediction, neural networks in general and recurrent neural networks such as the long short-term memory (LSTM) network in particular perform better than classical econometric models. This study presents a comparative analysis between the LSTM model and BiLSTM models. There is evidence for an improvement in the bidirectional model for predicting foreign exchange rates. In this case, we analyse whether this efficiency is consistent in predicting different currencies as well as the bitcoin futures contract.

Keywords: LSTM; BiLSTM; foreign exchange prediction; bitcoin



Citation: García, F.; Guijarro, F.; Oliver, J.; Tamošiūnienė, R. Foreign Exchange Forecasting Models: LSTM and BiLSTM Comparison. *Eng. Proc.* **2024**, *68*, 19. <https://doi.org/10.3390/engproc2024068019>

Academic Editors: Olga Valenzuela, Fernando Rojas, Luis Javier Herrera, Hector Pomares and Ignacio Rojas

Published: 4 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The foreign exchange market is worth more than USD 1.3 billion a year. It is considered the largest market in the world. It is a very important market for transactions of international goods and services. Also, it is important for reducing foreign currency risk for companies with receipts and payments in currencies other than their local currency. This implies the existence of predictive models for the prices of different currency pairs, as well as their development. Thus, economic agents and corporations could establish strategies for hedging and risk mitigation. Moreover, investors and speculators need to optimise their positions in the foreign exchange market. Thus, there is an important objective of reducing forecasting errors with models that are able to capture movements beyond a simple linear relationship between past and current prices. To do this, it is necessary to acquire knowledge about how prices move and to determine which types of models are most appropriate. There are many currencies traded against the dollar, as it is considered the reference currency. The most traded currencies are referred to as major currencies, which generally coincide with economic powerhouses. Less traded currencies may be considered exotic, even if they are traded against the US dollar (USD).

There are different models and methodologies for time series forecasting and, in this case, for forecasting currency prices. For example, the authors of [1] apply the Purchasing Power Parity Theory to forecast the USD/EUR exchange rate. This is based on the assumption that there are no arbitrage processes. The price differential between the currencies of the two countries needs to be known. The Goal Programming methodology, which has been used in other studies for the construction of rankings [2,3], could be applied to time series forecasting [4]. However, the classical models most commonly used with time series

are ARIMA models. For example, ref. [5] establishes an AR (1) model using the Box–Jenkins methodology for forecasting between the Naira and Dollar currencies over the period 1982–2011. Not only has this methodology been applied for the prediction of currencies qualified as “majors”, but also for the so-called “exotics”. Ref. [6] analyses the US dollar versus the Pakistani rupee currency pair. They analyse in detail the process of determining the stationarity of the series in these types of currencies by applying differences in the time series of prices. In this way, the observed results show a consistent model that has a forecasting error of 1%. Nevertheless, other publications have shown no clear benefits using ARIMA models to forecast other types of time series. Thus, the authors of [7] apply this methodology not for the prediction of prices but for their volatilities, measured by the variance of yields. In this case, they do not seem to obtain satisfactory results.

On the other hand, with the advent of neural networks, there has been an advancement in time series forecasting models. With them, it is possible to capture non-linear relationships between prices (inputs) and forecasts (output), significantly improving the results compared to ARIMA models. Ref. [8] analyses a comparison between classical ARIMA models and a neural network such as backpropagation to predict daily quotes of different exotic currencies against the dollar. A significant reduction in the prediction error measures applied in the comparison of both models is observed. For example, ref. [9] additionally incorporates more variables into the neural network. In this case, these are different moving averages. However, there is no consensus in the financial literature when it comes to establishing the number of variables to incorporate as inputs or the number of neurons or nodes in the construction of the network. This leads to variability in the results depending on modifications in the structure and parameters of the network [10].

A significant improvement in the use of neural networks in time series prediction is provided by recurrent neural networks. The long short-term memory recurrent network stands out. This network obtains significant improvements using only the prices of the asset to be predicted without including additional variables, as is the case with other types of networks. Thus, there are comparative studies between different neural networks and the LSTM network. For example, the authors of ref. [11] compare an Elman network with an LSTM network, obtaining excellent results in short-term predictions. Other authors compare the LSTM network with other methodologies such as VaR and Support Vector Machines, obtaining very high accuracy in the prediction of the USD/INR exchange rate, close to 98% [12].

It is concluded that the application of the LSTM recurrent neural network in currency forecasting shows superior results to other types of neural networks and short-term econometric models [13]. However, in recent years, it has been identified that the construction of hybrid models can improve the results compared to a single model. Such hybrid models can combine more than one methodology. One of the advantages that is evident is the reduction in risk obtained through the use of a single model, which improves the accuracy of predictions [14]. Combining models with at least one neural network significantly improves the results [15].

There are many references that analyse the advantages of these hybrid models. There is evidence of a growing interest in recent decades, highlighting, among others, hybrid models formed by ARIMA models and neural networks. The hybridisation process can be established sequentially or simultaneously depending on the combination of models and methodologies selected [16]. For example, the authors of [17] build a hybrid model sequentially. First, they estimate an ARIMA model for the monthly prediction of the USD/ALL exchange rate. The residuals of this first model form part of the inputs to the neural network. The results suggest that the combination of linear and non-linear models can provide favourable results for different measures of prediction error, such as the RMSE, MAE, and MAPE. Following this idea, ref. [18] multiplies the predictions obtained by the ARIMA model with the predictions obtained by the neural network. Evidence indicates that this type of modelling can work well for long-term predictions but may not be efficient in some cases for short-term predictions.

As demonstrated in the financial literature, it is evident that the LSTM recurrent neural network outperforms the ARIMA model [19]. The evolution of this type of network, which is unidirectional, into a bidirectional network such as the bidirectional LSTM (BiLSTM) shows significant improvements in time series forecasting. For example, ref. [20] improves the accuracy of forecasting by 37.78% on average in the prediction of different stock market indices. Following this line of research, this article compares the performance of the LSTM neural network with a bidirectional recurrent network (BiLSTM) in the daily price prediction of the EUR/USD, GBP/USD, EUR/NZD, EUR/JPY, EUR/GBP, and BTC/USD (bitcoin) currency exchange rates.

The methodology section (Methods) describes the long short-term memory neural network and its bidirectional variant that will be compared in the prediction of the different foreign exchange rates. Finally, the main conclusions are presented based on the prediction errors obtained with each model, as well as the main limitations of this work.

2. Methods

This section proposes methodologies for forecasting the daily closing price of different currencies and bitcoin. Two models are proposed. First, the main features of the LSTM recurrent neural network are presented. This network has been shown in multiple studies to improve time series prediction results compared to classical econometric models. Next, a variant of this network, the bidirectional LSTM, is presented. This network is mainly based on the directionality of the information flow within the neurons.

2.1. Long Short-Term Memory Neural Network

Neural networks, in general, are able to determine more complex (non-linear) relationships between variables in a time series. The long short-term recurrent neural network (LSTM) was proposed in [21]. This network allows for the detection of possible dependencies in a time series or sequential data in the long term.

The long short-term memory (LSTM) neural network is based on deep learning, with promising results in the prediction of time series. This network has the ability to activate different neurons that process different time sequences of data, thereby capturing long-term dependency relationships [22]. Usually, neural networks such as backpropagation present a problem during the learning process called the vanishing problem. Recurrent neural networks, such as LSTM, solve this problem by grouping neurons into different blocks with cells and gates. The flow of information that passes through the block is controlled by different doors. They process information differently. The LSTM network generally has a structure composed of a memory cell controlled by three doors (Figure 1). The information passes through the cells of each block by modifying their state (m_t). As indicated, the gates control this flow of information. These doors are activated depending on what the previous state of the cell was (m_t), how the information was output from the neuron in the previous process (h_{t-1}), and the new information that is incorporated into the process (x_t). In this way, the so-called forget door controls the information that must be rejected and therefore forgotten in the internal process of learning by the cell. On the contrary, the relevant gate or gateway determines how much information about past temporary states is incorporated as input into the neuron. That is, it is a process of memorization. Thus, each neuron stores different information. Finally, the cell obtains an output as a result of the process that has taken into account what its previous state was and what new information has been incorporated.

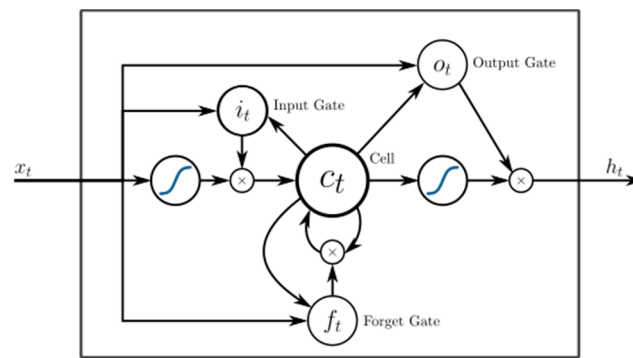


Figure 1. Cell structure of an LSTM network. Source: “CC” by E. A. Santos. Licenced under BY CC-SA 4.0.

2.2. BiLSTM Model

The bidirectional long short-term memory (BiLSTM) neural network has two hidden layers composed of the same type of network [23]. The first LSTM network performs a learning process by processing information in one direction (forward). The second LSTM network uses information in the opposite direction (backward) [24]. Respectively, the networks maintains its own structure with the gates already mentioned in Section 2.1. Figure 2 shows the structure of a BiLSTM network. As can be seen, in the first step (forward), the information is processed step by step starting with the initial sample data. At every stage, the bidirectional long short-term memory network computes and updates the state of every cell. In the backward process, the information is treated starting from the last time step of the time series data. The combination of both obtains more complete information in the learning process by capturing the intertemporal interactions of the time series.

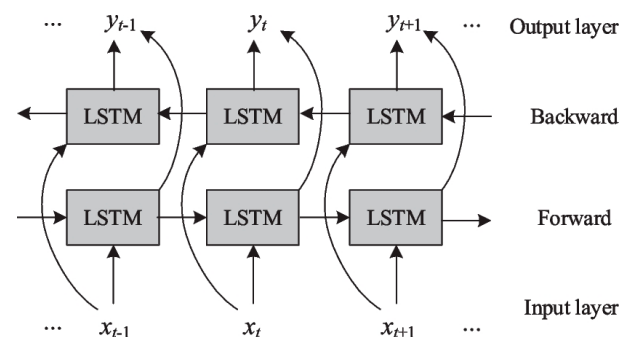


Figure 2. BiLSTM structure. Source: [25]. Licenced under BY CC BY 4.0.

3. Results

This section presents the different results obtained in the prediction of different currencies and bitcoin using the two models already discussed. In order to perform a comparative analysis, the prediction errors of both models (MAE, MAPE, RMSE) were calculated.

3.1. Database

The selected sample includes the following currency exchange rates: EUR/GBP, EUR/JPY, EUR/NZD, EUR/USD, GBP/USD, and BTC/USD (bitcoin). Daily closing prices were collected from 18 December 2017 until 16 January 2024 (1535 observations).

3.2. Data Analysis and Processing

Foreign coins behave in a similar way to any other time series and therefore often present kurtosis and asymmetry. Table 1 shows the main descriptive statistics for each series. It can be seen that GBP/USD is the only currency exchange rate in the sample with positive kurtosis.

Table 1. Descriptive statistics (test period).

	Mean	Median	Sd	Skew.	Kurt.
BTC/USD	29,677	27,915	9330.58	0.528	−0.761
EUR/GBP	0.861	0.861	0.016	−0.116	−0.750
EUR/JPY	144.621	143.404	9.894	0.123	−1.027
EUR/NZD	1.709	1.702	0.066	0.035	−0.961
EUR/USD	1.070	1.075	0.041	−0.556	−0.135
GBP/USD	1.244	1.242	0.058	−0.048	0.041

3.3. Model Estimation and Results

For the training of the two neural networks (LSTM and BiLSTM), 32 cells and a batch size of 32 and 10 epochs were used in both cases. The sample was divided into two parts: from 18 December 2017 to 01 December 2021 (1000 observations) for the training process and from 2 December 2021 to 16 January 2024 (535 observations) for the prediction (test). Three measures of prediction error (the MAPE, MAE, and RMSE) were used for the comparative analysis of the two models. Table 2 exhibits the main results that were acquired.

Table 2. Error prediction measures.

Model	GBP/USD	EUR/USD	EUR/NZD	EUR/JPY	EUR/GBP	BTC
MAPE						
LSTM	0.0153	0.0376	0.0161	0.0405	0.0110	0.1232
BiLSTM	0.0160	0.0219	0.0137	0.0375	0.0094	0.0707
MAE						
LSTM	0.0132	0.0331	0.0145	0.0351	0.0102	0.0638
BiLSTM	0.0139	0.0189	0.0122	0.0326	0.0087	0.0339
RMSE						
LSTM	0.0172	0.0366	0.0178	0.0395	0.0123	0.0798
BiLSTM	0.0185	0.0224	0.0151	0.0372	0.0105	0.0465

As can be seen in Table 2, for the case of the MAPE, the BiLSTM network has better results compared to the LSTM network for all the currencies analysed and the BTC, except for the GBP/USD rate. In the latter case, the error increased by 4.8%. Noteworthy are the reductions in the prediction error by more than 40% for EUR/USD and BTC/USD. For EUR/GBP and EUR/NZD, the error reduction obtained was around 15%. Finally, for EUR/JPY, the error reduction, as the most moderate, was 7.4%. The second measure of error analysed, the MAE, again showed a reduction in error for all currencies and BTC, with the exception of GBP/USD. As in the previous case, the EUR/USD and BTC/USD exchange rates showed reductions of more than 40%. For EUR/GBP and EUR/NZD, similar reductions of 15% were obtained, and a reduction of 7% was found for EUR/JPY. Finally, for the RSME, similar results to the previous ones were obtained.

It can be concluded that the bidirectional neural network (BiLSTM) improves the predictions in 83.3% of the assets analysed, with error reductions ranging from 6% to 47%. It should be noted that for the GBP/USD currency pair, the BiLSTM neural network did not improve the predictions obtained by the LSTM network. In view of the descriptive statistics shown in Table 1, it can be seen that this currency pair shows a positive kurtosis compared to the rest of the sample. However, this conclusion should be confirmed with further analysis by extending the sample to other currencies. The improvement in prediction using the BiLSTM network obtained in this study is consistent with other studies, even in other areas. For example, [26] improves the prediction of construction costs.

4. Conclusions

This paper compares the long short-term memory (LSTM) recurrent neural network with the bidirectional network (BiLSTM). To achieve this objective, the daily closing prices of various foreign currencies were predicted, as well as the bitcoin cryptocurrency. Different currencies were selected, some considered “Majors” and others “Exotic”, the former having a higher transaction volume compared to the latter. On the other hand, bitcoin was selected for the study because it is the most widely used and best known cryptocurrency at an international level. On the one hand, this comparative study tries to verify the accuracy of the sample’s closing price predictions by using the bidirectional network (BiLSTM). On the other hand, the diversity of the selected sample allows us to find out whether or not there is any kind of improvement that might allow some kind of distinction between “Majors” and “Exotic” currencies. In any case, the analysis is based on the premise that this type of network is very efficient in the prediction of short-term prices, as suggested by the literature review. The results suggest that the BiLSTM network significantly reduced the prediction errors that were analysed using different measures of error. The efficiency of this network is in line with other comparative studies, not only with financial data. It was observed that the error reductions obtained ranged from approximately 7% to 47%. However, this disparity was not found to be attributable to either of the two currency typologies or to their descriptive statistics.

One of the limitations of this paper concerns the selected sample. The analysis should be deepened by expanding the number of currency pairs. In addition, it would be interesting to see whether there is also a reduction in error with different sample timeframes. In other words, an intraday analysis should be carried out. The splitting of the sample into rolling windows should also be studied, both for price prediction for the next time as well as for different time predictions.

Author Contributions: Conceptualization, F.G. (Francisco Guijarro) and F.G. (Fernando García); literature review, J.O. and R.T.; writing, original draft, J.O., F.G. (Fernando García) and R.T.; writing, review, F.G. (Francisco Guijarro). All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analysed in this study. This data can be found here <https://www.visualchart.es>.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Ghalayini, L. Modeling and Forecasting the US Dollar/Euro Exchange Rate. *Int. J. Econ. Financ.* **2014**, *6*, 194–207. [[CrossRef](#)]
2. García, F.; Guijarro, F.; Oliver, J. A Multicriteria Goal Programming Model for Ranking Universities. *Mathematics* **2021**, *9*, 459. [[CrossRef](#)]
3. Burmann, C.; García, F.; Guijarro, F.; Oliver, J. Ranking the Performance of Universities: The Role of Sustainability. *Sustainability* **2021**, *13*, 13286. [[CrossRef](#)]
4. Repetto, M.; La Torre, D.; Tariq, M. Goal Programming in Federated Learning: An Application to Time Series Forecasting. In Proceedings of the 2022 International Conference on Decision Aid Sciences and Applications (DASA), Chiangrai, Thailand, 23–25 March 2022; pp. 1672–1677. [[CrossRef](#)]
5. Nwankwo, S.C. Autoregressive Integrated Moving Average (ARIMA) Model for Exchange Rate (Naira to Dollar). *Acad. J. Interdiscip. Stud.* **2014**, *3*, 429–433. [[CrossRef](#)]
6. Asadullah, M. Forecast Foreign Exchange Rate: The Case Study of PKR/USD. *Mediterr. J. Soc. Sci.* **2020**, *11*, 129–137. [[CrossRef](#)]
7. Dunis, C.L.; Huang, X. Forecasting and trading currency volatility: An application of recurrent neural regression and model combination. *J. Forecast.* **2002**, *21*, 317–354. [[CrossRef](#)]
8. Mbaga, Y.V.; Olubusoye, O.E. Foreign Exchange Prediction: A Comparative Analysis of Foreign Exchange Neural Network (FOREXNN) and ARIMA Models. 2014. Available online: <https://www.researchgate.net/publication/280040546> (accessed on 3 July 2024).

9. Kamruzzaman, J.; Sarker, R.A. Comparing ANN Based Models with ARIMA for Prediction of Forex Rates. *ASOR Bull.* **2003**, *22*, 2–11.
10. Huang, W.; Lai, K.K.; Wang, S. Forecasting Foreign Exchange Rates With Artificial Neural Networks: A Review. *Int. J. Inf. Technol. Decis. Mak.* **2004**, *3*, 145–165. [[CrossRef](#)]
11. Escudero, P.; Alcocer, W.; Paredes, J. Recurrent Neural Networks and ARIMA Models for Euro/Dollar Exchange Rate Forecasting. *Appl. Sci.* **2021**, *11*, 5658. [[CrossRef](#)]
12. Kaushik, M.; Giri, A.K. Forecasting Foreign Exchange Rate: A Multivariate Comparative Analysis between Traditional Econometric Contemporary Machine Learning & Deep Learning Techniques. *arXiv* **2002**, arXiv:2002.10247. [[CrossRef](#)]
13. Islam, M.S.; Hossain, E.; Rahman, A.; Shahadat, M.; Andersson, K. A Review on Recent Advancements in Forex Currency Prediction. *Algorithms* **2020**, *13*, 186. [[CrossRef](#)]
14. Hajirahimi, Z.; Khashei, M. Hybrid structures in time series modeling and forecasting: A review. *Eng. Appl. Artif. Intell.* **2019**, *86*, 83–106. [[CrossRef](#)]
15. Khashei, M.; Bijari, M. A new class of hybrid models for time series forecasting. *Expert Syst. Appl.* **2012**, *39*, 4344–4357. [[CrossRef](#)]
16. Zougagh, N.; Charkaoui, A.; Echchatbi, A. Artificial intelligence hybrid models for improving forecasting accuracy. *Procedia Comput. Sci.* **2021**, *184*, 817–822. [[CrossRef](#)]
17. Mucaj, R.; Sinaj, V. Exchange Rate Forecasting using ARIMA, NAR and ARIMA-ANN Hybrid Model. *J. Multidiscip. Eng. Sci. Technol.* **2017**, *4*, 8581–8586.
18. Wang, L.; Zou, H.; Li, L.; Chaudhry, S. An ARIMA-ANN Hybrid Model for Time Series Forecasting. *Syst. Res. Behav. Sci.* **2013**, *30*, 244–259. [[CrossRef](#)]
19. García, F.; Guijarro, F.; Oliver, J.; Tamošiūnienė, R. Foreign Exchange Forecasting Models: ARIMA and LSTM Comparison. *Eng. Proc.* **2023**, *39*, 81. [[CrossRef](#)]
20. Siami-Namini, S.; Tavakoli, N.; Namin, A.S. The Performance of LSTM and BiLSTM in Forecasting Time Series. In Proceedings of the 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 24 February 2020; pp. 3285–3292. [[CrossRef](#)]
21. Box, G.E.P.; Jenkins, G.M. *Time Series Analysis: Forecasting and Control*; Holden-Day: San Francisco, CA, USA, 1976.
22. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)] [[PubMed](#)]
23. Huang, C.-G.; Yin, X.; Huang, H.-Z.; Li, Y.-F. An Enhanced Deep Learning-Based Fusion Prognostic Method for RUL Prediction. *IEEE Trans. Reliab.* **2020**, *69*, 1097–1109. [[CrossRef](#)]
24. Schuster, M.; Paliwal, K.K. Bidirectional recurrent neural networks. *IEEE Trans. Signal Process.* **1997**, *45*, 2673–2681. [[CrossRef](#)]
25. Altché, F.; de La Fortelle, A. An LSTM Network for Highway Trajectory Prediction. In Proceedings of the IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), Yokohama, Japan, 15 March 2018; 2017; pp. 353–359. [[CrossRef](#)]
26. Wang, C.; Qiao, J. Construction Project Prediction Method Based on Improved BiLSTM. *Appl. Sci.* **2024**, *14*, 978. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.