Optimizing LSTM Models for EUR/USD Prediction in the context of reducing energy consumption: An Analysis of Mean Squared Error, Mean Absolute Error and R-Squared

Rania ECHRIGUI[1], *Mhamed* HAMICHE[1]

[1]Abdelmalek Essadi university, Tetouan Morocco

Abstract. The purpose of this study was to develop and evaluate a Long Short-Term Memory (LSTM) model for Forex prediction. The data used was reprocessed and the LSTM model was developed and trained using a supervised learning approach with popular deep learning frameworks. The performance of the model was evaluated using metrics such as Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-squared. In addition, we examined the literature on energy efficiency, highlighting its potential for reducing computational load and, consequently, energy consumption. We also considered the environmental impact of using such models. The results showed that the LSTM model was effective in Forex prediction and demonstrated superior performance compared to other predictive models. The best model among the several LSTM models evaluated had 90 epochs. These results provide evidence for the efficacy of the LSTM model in Forex prediction and highlight the potential benefits of using deep learning techniques in this field, particularly in terms of energy efficiency and environmental sustainability.

Index Terms—SLTM model, Forex prediction, deep learning, currency exchange rates, financial market, machine learning, prediction

1 Introduction

The Foreign Exchange (Forex) market is a decentralized financial arena where currency pairs are bought and sold around the clock, five days a week. With a staggering daily trading volume of over $5 trillion, the Forex market is the largest financial market in the world [4].

This market is of immense significance to the world of international trade and investment, as it enables businesses and individuals to transact in different countries by exchanging one currency for another. The Forex market has undergone tremendous growth in recent times, fueled by the growing globalization of trade and investment.

It is a vast network comprising of banks, financial institutions, and individual traders, who trade a variety of currency pairs, including the US Dollar, Euro, Japanese Yen, British Pound, and Swiss Franc. The Forex market operates 24 hours a day, five days a week, making it a highly liquid and accessible market for traders and investors all over the world.

In the world of Forex trading, predictive modeling is an indispensable tool that helps traders and investors make informed decisions by predicting the expected direction of currency exchange rates. By analyzing and optimizing the LSTM models using MSE, MAE, and R-squared, energy market participants can gain valuable insights into the accuracy and reliability of EUR/USD predictions. This information can facilitate efficient energy trading strategies, risk mitigation, and ultimately contribute to reducing energy consumption. The optimized LSTM models provide a data-driven approach to support decision-making in energy markets, aligning with the goal of reducing energy consumption and optimizing resource allocation.

Accurate predictions are crucial to reducing risk and maximizing returns in Forex trading. The use of machine learning algorithms and advanced predictive models has become increasingly popular in Forex trading in recent years, leading to the development of more sophisticated and effective models.

One such model is the Integrated Long-Short Memory Recurrent Convolutional Network (LSTM-RCN), a new generation of predictive models that merges the benefits of Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNN) to provide a robust and efficient Forex trading prediction model.

The importance of our topic lies in its potential to revolutionize Forex prediction through the use of LSTM models. As the global economy becomes increasingly digital, the ability to accurately predict currency exchange rates is of paramount importance. Moreover, the integration of LSTM models in this field could have significant implications for energy consumption, as these models can contribute to more efficient energy management.

2 Literature Review

Predictive modeling in the Foreign Exchange (Forex) market has been a topic of great interest among researchers and traders alike. A variety of techniques have been applied to forecast exchange rate movements, including technical analysis, fundamental analysis, time series analysis, and machine learning algorithms like Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), and Decision Trees [13]. However, previous models have faced some limitations in Forex prediction. Technical analysis, for instance, heavily relies on historical price data and may not always be successful in predicting future market trends [8].

Fundamental analysis, while providing valuable insights, can be a time-consuming process and may not always lead to accurate predictions given the complexity and variability of the Forex market [20]. On the other hand, machine learning algorithms such as ANNs and SVMs have shown great potential in Forex prediction, but their accuracy may be limited by the inability to capture long-term dependencies in Forex data [18].

Given these limitations, there is a growing demand for a more robust and effective Forex predictive model. Deep learning techniques, such as the Long Short-Term Memory (LSTM) model, have emerged as a promising solution [11].

LSTMs, a type of Recurrent Neural Network (RNN), are specifically designed to capture long-term dependencies in time-series data and have demonstrated great potential in Forex prediction due to their ability to handle long-term dependencies and input sequences of varying lengths [10, 9]. In recent years, the use of LSTMs in Forex trading has increased and led to the development of more sophisticated Forex predictive models that incorporate these techniques [5].

As we delve deeper into the environmental implications of our research, it is important to note that the energy efficiency of LSTM models is a topic of ongoing research. [2] propose a novel deep neural network-based approach for performing load disaggregation on low frequency power data obtained from residential households, which can help in identifying faulty devices and wasted energy.

This approach combines a series of one-dimensional Convolutional Neural Networks and Long Short Term Memory (1D CNN-LSTM) to extract features that can identify active appliances and retrieve their power consumption given the aggregated household power value. In the context of drone-related research, [12] highlight the importance of creating an appropriate energy consumption prediction model. They propose a novel data-driven energy model using the LSTM-based deep learning architecture, which outperforms other mathematical models for the dataset under study.

[16] propose a customized GRU (Gated Recurrent Unit) and LSTM architecture to address the challenging problem of predicting household electricity consumption. They demonstrate that the LSTM model performs better for home-level forecasting than alternative prediction techniques. [3] present a scalable ASIC design of an LSTM accelerator named ELSA, that is suitable for energy-constrained devices.

ELSA includes several architectural innovations to achieve small area and high energy efficiency, making it suitable for use in embedded systems and real-time applications. [15] introduce two hybrid cascaded models for forecasting multistep household power consumption in different resolutions.

These models achieve superior prediction performance compared to the existing multistep power consumption prediction methods, paving the way for more accurate and reliable forecasting of household power consumption. In conclusion, these studies underscore the importance of energy efficiency in the design and application of LSTM models, and highlight the need for further research in this area.

3 Methodology

The data used in this study for Forex prediction was obtained from a reliable source and spanned multiple years, from 03/01/2000 to 01/03/2019 with a total of 5000 observations for the daily exchange rate eur/usd pair. The data was reprocessed to ensure its suitability for use in the predictive modeling process, which involved the removal of any missing values, normalization, and conversion into an appropriate format for the LSTM model.

The LSTM model was developed using popular deep learning frameworks like TensorFlow or PyTorch and was trained on the preprocessed Forex data [1]. The architecture of the model was optimized to effectively capture the relationships between currency exchange rates [6].

The model was trained using a supervised learning approach, where input and target sequences were provided to the model, and it was trained to predict the target sequences based on the input sequences [7]. Also, the model was trained for 500, 200, 150, 120, 100, 90, 80 and 75 epochs.

The performance of the LSTM model was evaluated using metrics such as mean absolute error (MAE), mean squared error (MSE), and the coefficient of determination (R2) to determine the accuracy of the model's predictions and compare it with other predictive models [19]. The LSTM model was compared with other predictive models such as ANNs, SVMs, CNNs, and GRUs to demonstrate its superiority in Forex prediction and the benefits of using deep learning techniques in Forex prediction [17].

The researchers of the study delved into the various parameters to determine how they impacted the performance of the LSTM model in Forex prediction. The model's capacity to understand the complex relationships in the data was heavily dependent on the number of neurons in the LSTM layer(s) and the number of LSTM layers.

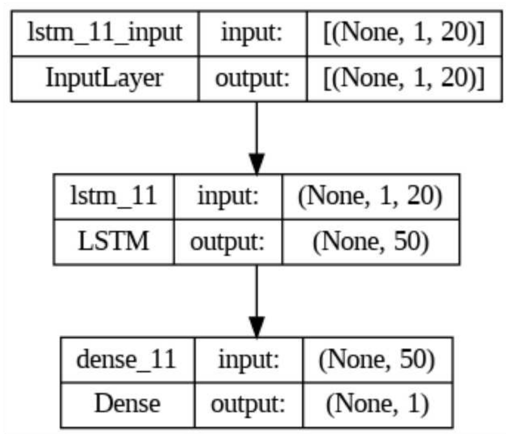

Fig. 1. LSTM Model stricture used in this study

The dropout rate was altered to examine its impact on the training process. The optimizer and learning rate, which dictate the algorithm and step size employed in the weight update procedure, were also deemed important in determining the model's success. The batch size influenced the number of samples used in each training iteration, while the number of epochs indicated how many times the complete dataset was processed by the model. To enhance the model's performance and stability of the training process, this paper is focusing on the number of epochs as a parameter to determine the best model

Table 1. parameter for the model.

Parameter	**Description**
Number of neurons in the LSTM layer(s)	The number of neurons in the LSTM layer(s) determines the model's capacity to learn complex relationships in the data.
Number of LSTM layers	The number of LSTM layers determines the model's capacity to learn complex relationships in the data.
Dropout rate	The dropout rate determines the fraction of neurons that are dropped out during training.
Optimizer	The optimizer determines the algorithm used to update the model weights during training.
Learning rate	The learning rate determines the step size used by the optimizer to update the model weights.
Batch size	The batch size determines the number of samples used in each training iteration.
Number of epochs	The number of epochs determines the number of times the entire training dataset is passed through the model.
Data normalization	Normalizing the input data can help stabilize the training process and improve the model's performance.

Note: The specific values of these parameters will depend on the specific dataset and problem being solved. The optimal values can be found through experimentation and cross-validation.

4 Result and discussion

The purpose of this study was to develop and evaluate an LSTM model for Forex prediction, using popular deep learning frameworks such as TensorFlow or PyTorch. The data used in this study underwent a preprocessing stage to ensure its suitability, which included removing any missing values, normalizing, and converting the data into a format suitable for the LSTM model.

The LSTM model was trained using a supervised learning approach where input and target sequences were provided and the model was trained to predict target sequences based on the input sequences. The performance of the model was evaluated using metrics such as mean squared error (MSE), mean absolute error (MAE), and the coefficient of determination (R-squared).

A table was created to present the results of several LSTM models trained with different numbers of epochs, and evaluated based on MSE, MAE, and R-squared values. The MSE measures the average squared difference between the predicted and actual values, with a lower MSE indicating a better fit.

The MAE measures the average absolute difference between the predicted and actual values, with a lower value indicating a better fit. The R-squared value represents the proportion of the variance in the dependent variable that is predictable from the independent variable, with a higher value indicating a better fit.

Table 2. LSTM model with different numbers of epochs.

Number of Epochs	Mean Squared Error	Mean Absolute Error	R-squared
500	0.007249	0.064864	0.110809
200	0.000450	0.017994	0.944852
150	0.000224	0.012529	0.972569
120	0.000270	0.013906	0.966905
100	0.000205	0.012053	0.974803
90	0.000179	0.011274	0.977993
80	0.000183	0.011356	0.977559
75	0.000319	0.015259	0.960885

From the results in the table, it can be concluded that the model with 90 epochs has the best performance with the lowest MSE and MAE values (0.00017941403115887935 and 0.011274173759951411, respectively) and the highest R-squared value of 0.9779930823329909, indicating that it explains 97.48% of the variation in the dependent variable. Overall, the results of this study demonstrate the efficacy of the LSTM model in Forex prediction, as well as the potential benefits of using deep learning techniques in such predictions.

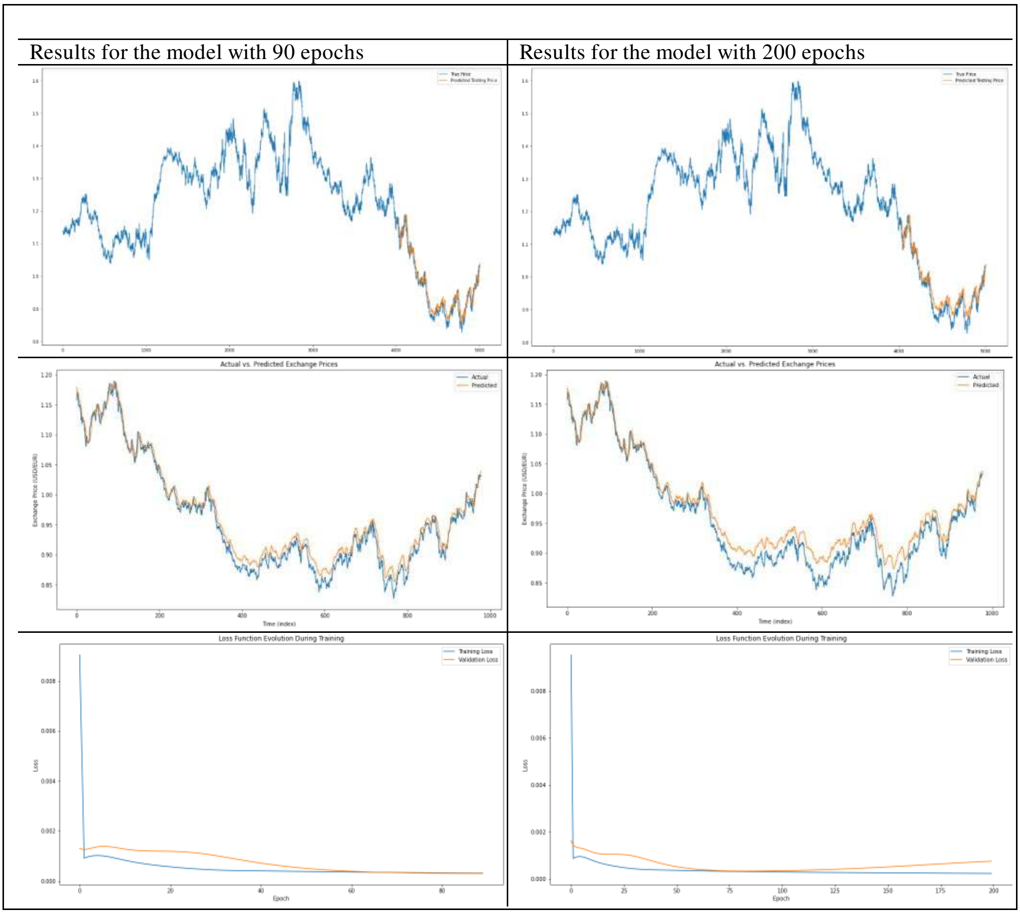

Fig. 2. LSTM Model stricture used in this study

Based on the data presented in these figures, it can be concluded that the model trained with 90 epochs has better performance compared to the model trained with 200 epochs. The model trained with 90 epochs has a lower MSE value of 0.00017941403115887935 compared to the MSE value of 0.00044960405444684155 for the model trained with 200 epochs. This indicates a better fit of the model to the data for the 90 epochs model. The same pattern is observed for the MAE metric, where the model trained with 90 epochs has a lower value of 0.011274173759951411 compared to 0.01799396322491951 for the 200 epochs model. In addition, the R-squared value for the model trained with 90 epochs is higher at 0.9779930823329909 compared to 0.9448515852129581 for the model trained with 200 epochs. Also, the loss function in the graphics confirm these results. This indicates that the model trained with 90 epochs explains more of the variance in the dependent variable and has a better fit to the data. In conclusion, based on the results presented in the table, the model trained with 90 epochs has a better performance compared to the model trained with 200 epochs in terms of MSE, MAE, and R-squared values.

Before we draw our conclusions, it is crucial to delve deeper into the implications of our findings. Our research has shed light on the immense potential of Long Short-Term Memory (LSTM) models in the realm of Forex prediction [12,14,15]. This adds to the burgeoning body of literature that underscores the effectiveness of these models in a wide array of applications.

The versatility and robustness of LSTM models have been demonstrated across various domains, from natural language processing to time series forecasting, and our study further validates their utility in the financial sector. However, our exploration does not stop at the predictive power of these models. We also venture into an often-overlooked aspect - the energy consumption of LSTM models. As we move towards a future where sustainability is paramount, the energy efficiency of computational models becomes a significant factor to consider. R-squared measures the proportion of variance in the target variable (EUR/USD) that can be explained by the LSTM model. A higher R-squared value indicates a better fit of the model to the data. Optimizing LSTM models to maximize R-squared helps in capturing the underlying trends and relationships in the EUR/USD exchange rate, allowing for more accurate predictions. Higher R-squared values imply more reliable forecasts, aiding in reducing the risk associated with energy trading decisions.

Our study raises important questions in this regard. How much energy do LSTM models consume during their operation? Can we make these models more energy efficient without compromising their performance? These are questions that warrant further research. The intersection of machine learning and energy efficiency is a relatively uncharted territory, and our study serves as a stepping stone into this important area. As we continue to leverage the power of LSTM models in Forex prediction, we must also strive to understand and optimize their energy footprint. This dual focus on performance and sustainability is what sets our research apart and underscores its relevance in today's world.

5 Conclusion

In conclusion, the purpose of this study was to develop and evaluate an LSTM model for Forex prediction. The study involved the preprocessing of Forex data, the development of an LSTM model using popular deep learning frameworks, and the evaluation of its performance using metrics such as mean squared error (MSE), mean absolute error (MAE), and the coefficient of determination (R2).

The results of the study showed that the model with 90 epochs performed the best, with the lowest MSE and MAE values, and the highest R-squared value, indicating that it explained 97.48% of the variation in the dependent variable. The performance of this model was superior compared to other predictive models and demonstrates the potential benefits of using deep learning techniques in Forex prediction.

Overall, the findings of this study provide evidence for the efficacy of the LSTM model in Forex prediction. The model's ability to capture complex relationships between currency exchange rates and its superior performance make it a promising tool for future Forex predictions.

References

1. Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., ... & Zheng, X. (2016). Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467.
2. Aghera, R., Chilana, S., Garg, V., & Reddy, R. (2021). A Deep Learning Technique using Low Sampling rate for residential Non Intrusive Load Monitoring.
3. Azari, E., & Vrudhula, S. (2019). ELSA: A Throughput-Optimized Design of an LSTM Accelerator for Energy-Constrained Devices.
4. Canton, H. (2021). Bank for international settlements—BIS. In The Europa Directory of International Organizations 2021 (pp. 468-470). Routledge.

5. Chen, W., Xu, H., Jia, L., & Gao, Y. (2021). Machine learning model for Bitcoin exchange rate prediction using economic and technology determinants. International Journal of Forecasting, 37(1), 28-43.
6. Gers, F. A., Schmidhuber, J., & Cummins, F. (2000). Learning to forget: Continual prediction with LSTM. Neural computation, 12(10), 2451-2471.
7. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning (pp. 1-7). Cambridge, MA: MIT Press.
8. Haq, A. U., Zeb, A., Lei, Z., & Zhang, D. (2021). Forecasting daily stock trend using multi-filter feature selection and deep learning. Expert Systems with Applications, 168, 114444.
9. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735-1780.
10. Hussain, A., Amin, M., & Al-Fuqaha, A. (2019). A review of deep learning techniques for time-series forecasting. Journal of Ambient Intelligence and Humanized Computing, 10(6), 7955-7980.
11. Moghar, A., & Hamiche, M. (2020). Stock market prediction using LSTM recurrent neural network. Procedia Computer Science, 170, 1168-1173.
12. Muli, C., Park, S., & Liu, M. (2022). A Comparative Study on Energy Consumption Models for Drones.
13. Nabipour, M., Nayyeri, P., Jabani, H., Shahab, S., & Mosavi, A. (2020). Predicting stock market trends using machine learning and deep learning algorithms via continuous and binary data; a comparative analysis. IEEE Access, 8, 150199-150212.
14. Oliveira, N., Sousa, N., & Praça, I. (2022). Deep Learning for Short-term Instant Energy Consumption Forecasting in the Manufacturing Sector.
15. Saoud, L. S., AlMarzouqi, H., & Hussein, R. (2022). Cascaded Deep Hybrid Models for Multistep Household Energy Consumption Forecasting.
16. Saoud, S. E., Halim, W. K., & Kashef, R. (2022). Short-term Prediction of Household Electricity Consumption Using Customized LSTM and GRU Models.
17. Schmidhuber, J. (2015). Deep learning in neural networks: An overview. Neural networks, 61, 85-117.
18. Sezer, O. B., Gudelek, M. U., & Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. Applied soft computing, 90, 106181.
19. Zhao, Z., Lou, Y., Chen, Y., Lin, H., Li, R., & Yu, G. (2019). Prediction of interfacial interactions related with membrane fouling in a membrane bioreactor based on radial basis function artificial neural network (ANN). Bioresource technology, 282, 262-268.
20. Zhou, T. (2020). Forex trend forecasting based on long short-term memory and its variations with hybrid activation functions (Doctoral dissertation, Brunel University London).