S&P/BMV IPC Forecasting Using Quantum Long Short-Term Memory

Jordi Fabián González-Contreras, Jesús Yaljá Montiel-Pérez, Erik Zamora-Gómez and Luis Enrique Andrade-Gorjoux

May 28, 2024

S&P/BMV IPC forecasting using Quantum Long Short-Term Memory

Jordi Fabián González-Contreras[1 [0009−0008−6289−0818]], Jesús Yaljá Montiel-Pérez[1 [0000−0002−0214−1080]], Erik Zamora-Gómez[1 [0000−0002−3682−8585]], and Luis Enrique Andrade-Gorjoux[1 [0000−0001−8924−8368]]

Instituto Politécnico Nacional, Centro de Investigación en Computación, Av. Juan de Dios Batiz S/N, Gustavo A. Madero, 07738, Ciudad de México, México
[jgonzalezc2023, jyalja, landradeg2022]@cic.ipn.mx, [ezamorag]@ipn.mx

Abstract. This paper presents results on time series forecasting using a quantum recurrent neural network called Quantum Long Short-Term Memory, known by its acronym QLSTM. In this study, we present experimental results about the forecasting on a financial closing price dataset (S&P/BMV IPC), where the data is analyzed to determine its correlation dimension to validate whether the time series exhibits non-linear behavior.
To assess the performance of the QLSTM, the mean squared error metric is calculated to underscore the advantages of this quantum variant compared to its classical counterpart, the LSTM. It is noteworthy that this work represents the first publication in non-linear time series prediction applied to a Mexican stock index using quantum computing.https://github.com/JordiFGonzalezC/QLSTM

Keywords: QLSTM · Quantum Finance · Quantum Time Series Forecasting · Non-Linear Systems.

1 Introduction

1.1 Background

Our study builds on the QLSTM [1], proposed by Samuel Yen-Chi Chen, Shinjae Yoo, and Yao-Lung L. Fang in 2020 in their article "Quantum Long Short-Term Memory". They introduced an architecture for the QLSTM which implements Variational Quantum Circuits (VQC) instead of using classical artificial neural networks to reinterpret the classical LSTM into a quantum computational neural model.

As outlined in [1], the proposed architecture enables the encoding of arrays of numbers into a quantum state, essentially mapping these values to an angular value within a qubit. The aim is to predict non-linear time series, experiments that could validate the decrease in the mean square error of quantum circuits in making predictions about real-world systems. In this case, estimating closing values in the stock market for S&P/BMV IPC index with a better simulation performance and with a faster decrease in the mean square error than the classical computing algorithms.

S&P/BMV IPC, also known as the Mexican Stock Exchange (Bolsa Mexicana de Valores or BMV), is one of the major stock exchanges in Mexico. IPC (Índice de Precios y Cotizaciones) is the main stock market index in Mexico, and it represents a selection of the most actively traded stocks on the BMV.

Therefore, knowing those future values is of great interest to the Mexican financial sector.

The paper is structured as follows: Introduction, this section reviews the QLSTM architecture and showcases its main components. It characterizes the data that will be used and concludes on why these data are considered non-linear. In the subsequent section, Experimental Development, we delve into the methodology employed to address the forecasting problem using QLSTM. Additionally, experimental results regarding error during training are presented. Finally, in the Results section, we analyze and compare the prediction outcomes of the QLSTM against the LSTM and we highlight the improvements achieved by employing a quantum algorithm for time series forecasting.

1.2 QLSTM architecture

The Quantum Long Short-Term Memory (QLSTM) architecture is an innovative variant of the classical Long Short-Term Memory (LSTM) recurrent neural network that incorporates principles of quantum computing. In QLSTM, quantum gates and quantum-inspired mechanisms are integrated into the architecture to enhance its computational power and memory capabilities.

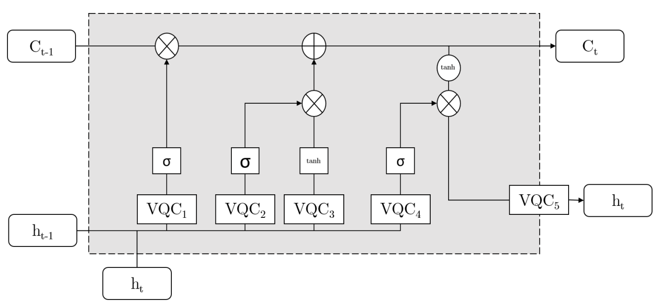


Fig. 1: QLSTM architecture [1]

Like traditional LSTM, its quantum version is designed to process sequential data, making it well-suited for tasks like natural language processing and time series analysis. However, QLSTM leverages quantum properties such as superposition and entanglement to perform certain operations more efficiently than classical counterparts, potentially offering advantages in tasks that involve complex temporal dependencies or large-scale data. Basically the QLSTM architecture is kind of same structure than its classical counterpart, with the difference that the classical neural networks (where usually synaptic weights are trained) are changed for variational quantum circuits as is shown in the architecture in the Fig.1.

1.3 Data categorization usaing Grassberger-Procaccia algorithm

Financial markets exhibit complex and often non-linear behavior [23], making it crucial for analysts and traders to employ sophisticated tools to grasp the inherent unpredictability. The use of the Grassberger-Procaccia algorithm to estimate the correlation for the analysis of time series and dynamical systems to quantify complexity and fractal structure is a common tool used in the literature [25].

The correlation dimension of the time series will be discussed as the result obtained after applying the Grassberger-Procaccia algorithm as D_2.

The correlation dimension, originally developed in the field of nonlinear dynamics and chaos theory [25], have found a unique and increasingly important role in the realm of finance. These mathematical constructs offer a quantitative framework for assessing the sensitivity of financial systems to initial conditions, shedding light on the intricate interplay of factors that drive market fluctuations. By examining the correlation dimension of financial time series data, analysts can gain valuable insights into the underlying dynamics of asset prices, market volatility, and risk.

To find the correlation dimension, it is necessary to apply the following mathematical expression [27] to the financial data source [2], to simplify the calculations for the Grassberger-Procaccia algorithm, a Github code implementation is used [26].

The Grassberger-Procaccia algorithm is a technique used in time series analysis to estimate the correlation dimension, a measure of the complexity and fractal structure of the underlying system. This algorithm is particularly useful for detecting the presence of chaos in a time series.

Phase Space Reconstruction First, the phase space is reconstructed from the one-dimensional time series using the method of time delay. Vectors of embedding dimension m and a time delay τ are created. For a time series $x(t)$, the vectors in the phase space are:

$$\mathbf{X}(t) = (x(t), x(t+\tau), x(t+2\tau), \ldots, x(t+(m-1)\tau))$$

Calculation of Distances in Phase Space Next, the Euclidean distance between all pairs of points in the reconstructed phase space is calculated:

$$d_{ij} = \|\mathbf{X}(i) - \mathbf{X}(j)\|$$

Correlation Function The correlation function $C(r)$ is then defined, which is the fraction of pairs of points whose distance is less than a threshold r:

$$C(r) = \frac{2}{N(N-1)} \sum_{i<j} H(r - d_{ij})$$

where H is the Heaviside step function, N is the total number of points in the series, and d_{ij} is the distance between points i and j.

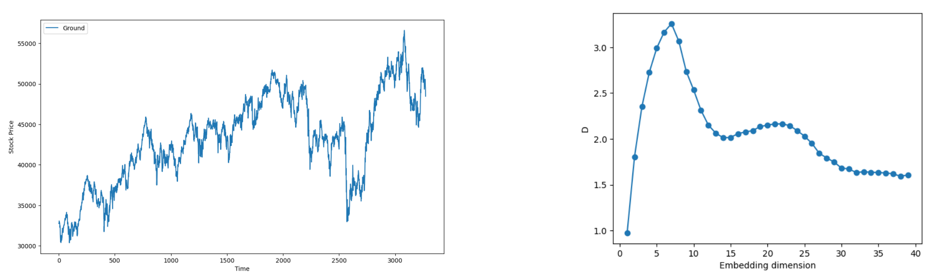
Estimation of the Correlation Dimension The correlation dimension D_2 is estimated by observing how $C(r)$ changes as r varies:

$$D_2 = \lim_{r \to 0} \frac{\log C(r)}{\log r}$$

In practice, this is done by performing a linear regression of $\log C(r)$ against $\log r$ over a range of r where this relationship is approximately linear.

This algorithm helps researchers better understand the nature of time series data and identify patterns that are not evident through traditional linear methods. [27].

Estimation of correlation dimension for S&P/BMV IPC time series. It is feasible to estimate certain metrics that indicate its non-linear nature. In 2021, a comprehensive analysis was conducted to characterize the non-linearity of the S&P/BMV IPC [28]. This involved examining the trends in behavior and memory of the prices of the shares comprising the S&P/BMV IPC index in the Mexican stock market, as well as the evolution of returns over time. To assess the normality of returns, the Jarque-Bera test was employed, revealing non-normality in the generated returns from the stations. Notably, the Hurst exponent was utilized as a measure of memory in prices. The findings from these analyses led to the conclusion that the index exhibits a persistent behavior [28].



(a) S&P/BMV IPC time series

(b) Correlation dimension

Fig. 2: Correlation dimension for S&P/BMV IPC time series

Using correlation function, the obtained value for D is shown in the graph (2b) for different candidate embedding dimensions (1 to 40), this measure provides information about the complexity and structure of the underlying time series. A correlation dimension decreasing suggests a more complex or non-linear structure in the dynamic system [27], but as it is inconclusive, the definition of persistent time series will be employed [28].

2 Experimental development

2.1 Methodology

As a result of our modifications in the classical LSTM [9], the trainable parameters are now primarily the rotational parameters embedded within the quantum circuits. This transformation has given birth to a hybrid quantum-classical neural network layer, with the optimization process relying on classical techniques. Our research, outlined in the QLSTM paper, provides compelling evidence that QLSTM outperforms traditional LSTM networks. QLSTM exhibits an exceptional ability to gather an abundance of information right from the start of the training process, and its loss decreases consistently and rapidly compared to its classical counterpart. Consequently, our primary objective in this study is to substantiate the superior capabilities of QLSTM in terms of the error and the speed of its decrease.

In our adaptation, we have made specific adjustments to align the model with the Variational Quantum Circuit (VQC) illustrated in the next section. Additionally, we've leveraged the Pennylane [6] simulator, which is integrated into our setup, to execute these VQCs. Pennylane seamlessly integrates quantum computing with classical machine learning frameworks like TensorFlow and PyTorch, enabling efficient development of hybrid algorithms. Its robust autodifferentiation capabilities and extensive documentation make it ideal for optimizing variational circuits, while its flexibility and active community support foster innovation and ease of use.

In the following section, our primary goal is to train a QLSTM model tailored for forecasting non-linear time series. This work serves as a proof for showcasing that QLSTM not only holds potential for time series prediction but also surpasses its classical counterpart in various aspects. Our chosen configuration includes 4 qubits, a single variational layer, and a learning rate set at 0.05. The choice of 4 qubits was driven by our desire to demonstrate QLSTM's effectiveness even with a limited number of qubits [1]. We determined the latter parameters through experimentation over multiple training epochs to identify the settings that produced the most favorable results.

2.2 Training the Variational Quantum Circuit (VQC)

The process of training a Variational Quantum Circuit (VQC) using Torch [10], a robust deep learning framework, is a sophisticated undertaking that combines the domains of quantum computing and machine learning. It commences by encoding the specific problem of interest into a quantum circuit, typically represented as a sequence of quantum gates. This parameterized VQC then becomes the focal point for Torch's involvement, which is employed to optimize these parameters utilizing classical machine learning techniques.

The VQC consists of three layers. The first layer, known as the encoding layer, is where a mapping from real numbers (classical) to angular rotations on

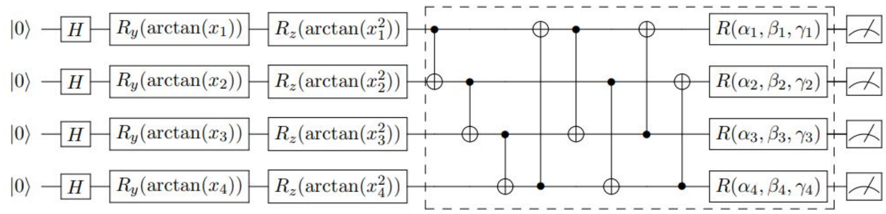


Fig. 3: Variational Quantum Circuit for QLSTM [1]

the qubit is performed. This is achieved using Hadamard gates (H) and rotational operators (R_y, R_z) along the y, z axes. In the Fig.3, it is shown that each element in the time series is mapped to each qubit using the *arctan* function. The subsequent layer is the variational part, where qubits are entangled and rotated by certain angles (α, β, γ) around each x, y, z axis, respectively. These angles represent hyperparameters that need to be tuned using established techniques, such as gradient descent, to optimize predictions. Finally, there is the measurement layer, which maps the qubit's value to a classical bit.

2.3 Training loss metrics for LSTM and QLSTM

To evaluate the training progress over epochs in the field of Machine Learning, it is common to use mean squared error (also known as loss metric) to determine whether the training is improving or deteriorating. This is done in order to know when to stop the training and avoid the overtraining zone. This is crucial because training further can lead to gradient explosion and an increase in error instead of a decrease.

As seen in Fig.4, in a) represents the mean squared error for the training of 100 epochs for the LSTM. As evident from epoch 100 to later, the error stabilizes and exhibits a trend that converges around zero. It is noteworthy that, even though they may appear similar, the training of the quantum model significantly accelerates the error reduction rate. According to the graphs, the error diminishes up to ten times faster in the quantum version compared to the classical one.

On the other hand, for its quantum counterpart, in Figure b) the QLSTM displays a similar behavior as the error approaches zero. However, it's noteworthy that this has been executed on a classical computer using the Pennylane simulator [6]. A quantum version of a QLSTM algorithm executed on a quantum computer could offer significant advantages in terms of efficiency, exploration capability, and modeling of quantum correlations, provided challenges associated with error correction and efficient implementation on quantum hardware are overcome.

Table 1 provides a comparison of the hyperparameters used in the code to ensure similar and as fair conditions as possible for determining which of them would achieve improved performance. Since both are recurrent neural networks

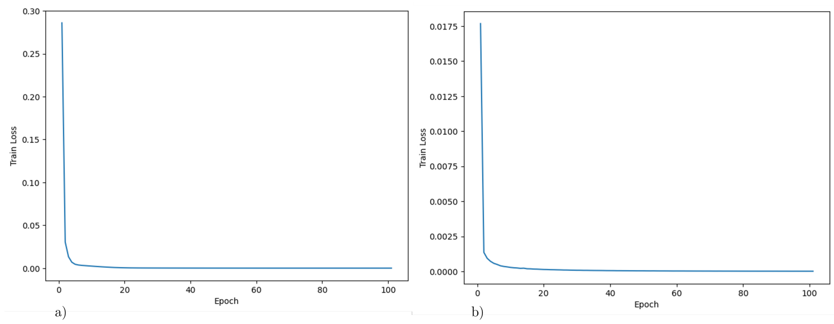


Fig. 4: Training: a) LSTM loss vs. epoch , b) QLSTM loss vs. epoch

with a single layer, the Adam optimizer from the Torch library was utilized to obtain optimal values for the necessary parameters in both neural networks. In both cases, the output of the neural network represents the next value in the series.

Table 1: Hyperparameters comparison for LSTM vs. QLSTM

Hyperparameter	Classical	Quantum
Learning rate	1×10^{-4}	5×10^{-2}
Number of sensors	16	16
Number of layers	1	1
Number of qubits	N/A	4

The two loss metric graphs illustrate a similar trend after epoch 10, with one representing the results of a classical LSTM algorithm and the other its quantum counterpart.

In the classical LSTM graph, there is a noticeable and swift decrease in loss after epoch , indicating effective learning and adaptation to the training data. This aligns with the expected behavior of a well-tuned LSTM, showcasing its ability to capture sequential dependencies and patterns.

Surprisingly, the quantum counterpart graph exhibits a comparable trend in the loss metric, despite the fundamentally different approach of quantum computing. The rapid decrease in loss suggests that the quantum model is also effectively learning and adapting to the training data, starting from epoch 8. This parallel behavior implies that the quantum model, with its unique computational principles, is achieving results similar to its classical counterpart.

The commonality in the loss reduction patterns after epoch 10 between the classical LSTM and its quantum counterpart underscores the effectiveness of the quantum model in utilizing its distinct computational capabilities, the error values for both neural networks in the training zone behaves in the same order

Table 2: Error values during training at epoch 100

	Classical	Quantum
Test error	1.2×10^{-4}	1.6×10^{-4}

error around 1e−4 (Table 2). This observation highlights the potential of quantum machine learning to provide effective solutions, even in situations where classical algorithms have traditionally excelled.

3 Results

The financial time series graphs of the S&P/BMV IPC span from 2010 to 2022, showcasing an overlay of the real data, the results of a classical LSTM, and a QLSTM. The dashed blue region starting from 67% of the data represents the testing zone.

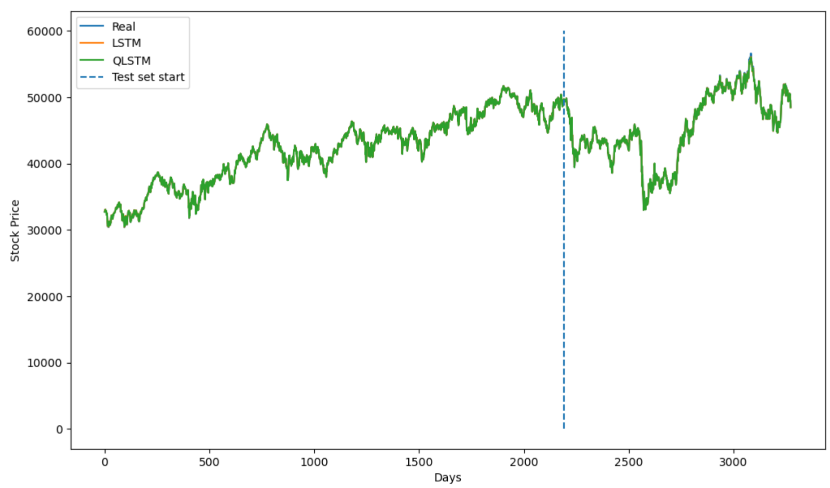


Fig. 5: S&P/BMV IPC prediction results LSTM vs. QLSTM. Both recurrent neural networks behave similarly, its highlighted that even if the training didn't considers the pick around 3000 day, both could predict this increment in the time series without issues.

In Figure 5, is possible to note the comparison between the LSTM vs. the QLSTM predictions, corresponding to the S&P/BMV IPC time series that represents the historical fluctuations of the index. The curve of the classical LSTM closely follows the real data in graph, as expected, the QLSTM curve behaves very closely to the real data. While the performance of the QLSTM is very similar to that of the LSTM, a slight superiority of the classical version may be noticeable, suggesting a potentially higher predictive capacity in this specific context.

It is crucial to note that the QLSTM was executed on a classical computer with a simulator, which could have influenced its performance. Emphasis is placed on the possibility that the quantum version could significantly improve if run on a real quantum computer, surpassing the classical version. However, it is underscored that overcoming challenges associated with running on a real quantum computer, such as noise and qubit topology, is essential to fully leverage the potential of quantum computing in finance.

To reinforce the notion that the quantum version of the LSTM may perform equally or even better, we can examine the prediction results in the testing zone. In the first graph a), there is a smoother error reduction as it approaches zero, followed by stabilization. In contrast in graph b), the error reduction in the quantum version is highly accelerated and abrupt, rapidly reaching zero and immediately stabilizing around it.

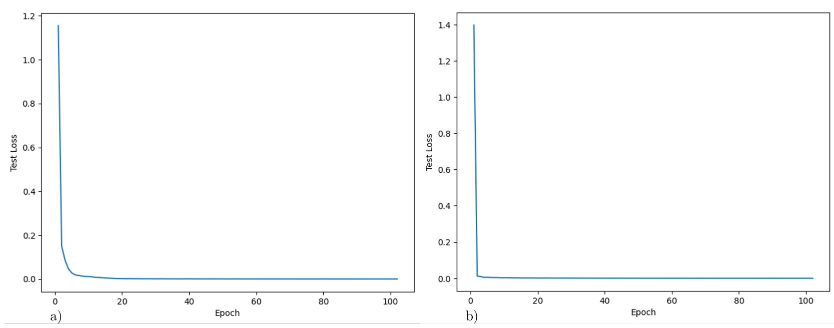


Fig. 6: Testing: a) LSTM loss vs. epoch , b) QLSTM loss vs. epoch

The error values for both neural networks in the testing zone were in the same order error around 1e−4 (Table 3), the values are close but it must be noted that the quantum algorithm was executed in a quantum computer simulator, so the results obtained from the QLSTM could be improved by making an adequate implementation in a real quantum computer. Although, due to current hardware limitations, perhaps it is most likely that it will have worse performance than its classic version.

Table 3: Error values during testing at epoch 100

	Classical	Quantum
Test error	1.2×10^{-4}	1.5×10^{-4}

4 Conclusion

This work represents a continuation of the exploration into the application of Quantum Long Short-Term Memory (QLSTM), a concept introduced by Samuel Yen-Chi Chen, Shinjae Yoo, and Yao-Lung L. Fang in their 2020 article "Quantum Long Short-Term Memory." The proposed architecture utilizes Variational Quantum Circuits (VQC) instead of classical artificial neural networks in a conventional LSTM, aiming to predict non-linear time series, particularly closing values in the S&P/BMV IPC stock market index.

An achievement of this work is demonstrating the possibility of matching the performance of the most effective classical algorithms with quantum computing. In the metrics presented, it was observed how the quantum version showed no distinction from the classical one. Evaluations were conducted under the most equitable conditions for both approaches. Furthermore, the application of these methods for time series prediction demonstrates their effectiveness even in the case of non-linear time series, which, in theory, pose a significant challenge for analysis and prediction. This serves as a clear example of the capabilities quantum machine learning has in predicting systems that, in principle, cannot be forecasted from a deterministic or stochastic perspective.

Our application of the QLSTM, focusing on rotational parameters embedded within quantum circuits, led to the development of a hybrid quantum-classical neural network layer. The trainable parameters became predominantly the rotational parameters, transforming the traditional LSTM into a QLSTM. Our research suggests that QLSTM outperforms traditional LSTM networks, showcasing an exceptional ability to rapidly gather information and decrease loss consistently from the outset of the training process.

The training process involves a Variational Quantum Circuit (VQC) using Torch, a robust deep learning framework. This intricate process combines the realms of quantum computing and machine learning, encoding the problem into a quantum circuit and optimizing parameters using classical machine learning techniques.

The three-layer variational quantum circuit includes an encoding layer, a variational part, and a measurement layer. The encoding layer maps real numbers to angular rotations on qubits, the variational part entangles and rotates qubits using hyperparameters, and the measurement layer converts qubit values to classical bits.

The comparison of financial data graphs from 2010 to 2022, featuring actual data, classical LSTM results, and QLSTM results, suggests that both models perform similarly. While the classical LSTM may exhibit a slight advantage, it's essential to consider that the QLSTM was executed on a classical computer with a simulator. The potential for significant improvement exists if the QLSTM runs on a real quantum computer, surpassing the classical version. As future work, we aim to implement the algorithm on a real quantum computer to further explore its capabilities and overcome the challenges associated with noise and qubit topology in quantum computing for finance.

References

1. Samuel Yen-Chi Chen, Shinjae Yoo, and Yao-Lung L. Fang: Quantum Long Short-Term Memory. 2020. https://www.osti.gov/servlets/purl/1842795.
2. S&P Dow Jones: S&P/BMV IPC Mexico. Obtained from https://www.spglobal.com/spdji/es/indices/equity/sp-bmv-ipc/overview.
3. Yahoo Finance, NASDAQ Composite USA. Obtained from https://finance.yahoo.com/quote/5EIXIC/history.
4. Linda Reichl, The Transition to Chaos. Springer, Third Edition. 2021. ISSN 2365-6425.
5. Yuji Cao, Xiyuan Zhou, Xiang Fei, Huan Zhao: Linear-layer-enhanced quantum long short-term memory for carbon price forecasting. Springer, Third Edition. 2023. https://link.springer.com/article/10.1007/s42484-023-00115-2.
6. Bergholm V., Izaac J., Schuld M.: Pennylane: Automatic differentiation of hybrid quantum-classical computations. ArXiv Prepr ArXiv181104968. 2018. https://arxiv.org/abs/1811.04968.
7. Ceschini A., Rosato A., Panella M.: Hybrid quantum-classical recurrent neural networks for time series prediction. Int. Jt. Conf. Neural Netw. IJCNN. IEEE, pp 1–8. 2022.
8. Cong I., Choi S., Lukin M.D.: Quantum convolutional neural networks. Nat Phys 15(12):1273–1278. 2019. https://doi.org/10.1038/s41567-019-0648-8.
9. Hochreiter S., Schmidhuber J.: Long short-term memory. Neural Comput 9(8):1735–1780. 1997. https://doi.org/10.1162/neco.1997.9.8.1735.
10. Paszke A., Gross S., Massa F.:Pytorch: An imperative style, high-performance deep learning library.Adv Neural Inf Process Syst 32. 2019.
11. H. P. Nautrup, N. Delfosse, V. Dunjko, H. J. Briegel, and N. Friis:Optimizing quantum error correction codes with reinforcement learning. 2018. https://arxiv.org/abs/1812.08451.
12. A. Cross:The ibm q experience and qiskit open-source quantum computing software. APS Meeting Abstracts. 2018.
13. K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii, Quantum circuit learning. Physical Review A, vol. 98, no. 3, p. 032309, 2018.
14. Y. Du, M.-H. Hsieh, T. Liu, and D. Tao, The expressive power of parameterized quantum circuits. arXiv preprint arXiv:1810.11922, 2018.
15. V. Havlíček, A. D. Córcoles, K. Temme, A. W. Harrow, A. Kandala, J. M. Chow, and J. M. Gambetta. Supervised learning with quantum-enhanced feature spaces. Nature, vol. 567, no. 7747, pp. 209–212, 2019.
16. M. Schuld and F. Petruccione, Supervised learning with quantum computers, vol. 17. Springer, 2018.
17. S. J. Whalen, A. L. Grimsmo, and H. J. Carmichael, Open quantum systems with delayed coherent feedback. Quantum Science and Technology, vol. 2, no. 4, p. 044008, 2017.
18. G. Lai, Z. Dai, Y. Yang, and S. Yoo, Re-examination of the role of latent variables in sequence modeling. Advances in Neural Information Processing Systems, pp. 7812–7822, 2019.
19. V. Bergholm, J. Izaac, M. Schuld, C. Gogolin, C. Blank, K. McKiernan, and N. Killoran, Pennylane: Automatic differentiation of hybrid quantum-classical computations. ArXiv preprint arXiv:1811.04968, 2018.
20. S. M. Girvin, M. H. Devoret, and R. J. Schoelkopf. Circuit qed and engineering charge-based superconducting qubits. Physica Scripta, vol. T137, p. 014012, 2009.

21. M. Schuld, V. Bergholm, C. Gogolin, J. Izaac, and N. Killoran, "Evaluating analytic gradients on quantum hardware" Physical Review A, vol. 99, no. 3, p. 032331, 2019.
22. S. Haris, G. Georg and L. Jacques. Chaos Detection and Predictability. Springer: Lecture Notes in Physics, vol. 915, no. 3, p. 032331, 2016.
23. C. Espinosa Méndez. Comportamiento caótico en los mercados bursátiles latinoamericanos utilizando Visual Recurrence Analysis, Análisis Económico, p. 159-183. Universidad Autónoma Metropolitana Unidad Azcapotzalco, 2008.
24. A. Mene Hevia. Cálculo de los exponentes característicos de Lyapunov. Universidade De Santiago De Compostela, 2019.
25. Le Barón, B. Chaos and nonlinear forecastability in Economics and Finance. Philosophical Transactions of Royal Society of London, Series A, 1994.
26. Sebastiano. Implementation of the Grassberger-Procaccia algorithm to estimate the Correlation Dimension of a set of points. Gihub. 2021.
27. W.A. Brock, W.D. Dechert, J.A. Scheinkman, and B. LeBaron. Economic Reviews. 15, 197 1996.
28. Morales Castro A., Lizola Margolis P.E., and Álvarez Tostado Ceballos H.E.. El exponente de Hurst y la memoria de los precios en las acciones del sector de bienes de consumo frecuente de Índice S&P/BMV IPC. 2021. http://doi.org/10.5281/zenodo.7415892