RECURRENT QUANTUM NEURAL NETWORKS: A REVIEW

Gleydson Fernandes de Jesus[1,a], Valéria Loureiro da Silva[1,b].

[1,a] HPC, SENAI CIMATEC, Brazil,
[1,b] Departamento de Eletronica, SENAI CIMATEC, Brazil.

Abstract: The analysis of historical data allows the execution of predictive tasks such as weather and stock price forecasting. To achieve these goals, Recurrent Neural Networks are implemented in classical computers and, in recent years, quantum methods have also emerged to perform prediction tasks based on the analysis of historical series, which have been called Quantum Recurrent Neural Network (QRNN). The objective of this work is to identify and review the main QRNNs discussed in the literature. A literature search in google scholar resulted in eight relevant papers that were reviewed. In general, the QRNNs show better training accuracy and stability compared to classical methods. It is not possible to speak of a training time advantage with the noisy and low-scale quantum computers currently available.

Keywords: Quantum Computing; RNN; LSTM.

REDES NEURAIS RECORRENTES QUÂNTICAS: UMA REVISÃO

Resumo: A análise de dados históricos permite a execução de tarefas preditivas como, por exemplo, a previsão do tempo e de preço de ações. Para alcançar esses objetivos, Redes Neurais Recorrentes são implementadas em computadores clássicos e, nos últimos anos, surgiram também métodos quânticos para realizar tarefas de previsão baseadas na análise de séries históricas, que vêm sendo denominados Quantum Recurrent Neural Network (QRNN). O objetivo deste trabalho é identificar e revisar as principais QRNNs encontradas na literatura. Uma busca da literatura no google scholar resultou em oito artigos relevantes que foram revisados. Em geral, as QRNNs mostram melhor acurácia e estabilidade de treino em comparação aos métodos clássicos. Não é possível falar em vantagem no tempo de treinamento com os computadores quânticos ruidosos e de baixa escala atualmente disponíveis.

Palavras-chave: Computação Quântica; RNN; LSTM.

1. INTRODUCTION

Recurrent Neural Networks (RNNs) [1] are machine learning models used in time series forecasting, such as weather forecasting and natural language processing. Training of these models is done by processing historical data on classical computers, is a costly process, requiring high computational capacity and processing time.

Quantum computing [2] has been gaining prominence in recent years, especially due to the interest of BigTechs such as Google, Microsoft and IBM, which have developed quantum computers and leveraging the market and the development of new applications. In the case of IBM, the computer usage is available to the scientific community.

The main advantage of quantum computing is in the processing speed which can be significantly accelerated (exponentially in the best case) with respect to classical computing.

Thus, quantum RNN models [3-10], as well as other quantum machine learning models [2], naturally emerge as an attempt to circumvent the problems of classical RNNs.

In this paper, we review the main quantum computing methods alternatives for recurrent neural networks found in the literature.

1.1. Recurrent Neural Networks

Recurrent Neural Networks (RNNs) [1] are machine learning models capable of maintaining memory over historical input data, adjusting their learning not only through the last information received but also through even older information. Because of this memory capacity, RNNs are used in predictive models based on time series processing, as well as in natural language processing. There are also more elaborate models of RNNs, such as Long-Short Term Memory (LSTM), which emerge as a classical solution to the problems faced by RNNs during the calculation of gradients in the training process.

1.2 Quantum Machine Learning

Quantum machine learning [2] is a recent research area that seeks to find more accurate and easier to train machine learning models, relying on the use of quantum computers instead of classical computers.

Being a recent area, the development of quantum machine learning models is still incipient and limited by the lack of large-scale and error-corrected quantum computers. For this reason, current models are developed to work on devices with low number of qubits and no error correction, also known as NISQ (Noise Intermediate Scale Quantum) devices, saving scarce computational resources. In this way, the learning process can be interleaved between classical and quantum computers, with classical computers taking on the role of optimizing the input parameters of quantum circuits.

These models are generally based on 3 steps [2]: (1) the encoding of the data in the quantum computers; (2) the application of a variational layer; and (3) the application of measurements on the computational basis. After these 3 steps, the obtained results are classically processed through an optimizer, which will lead to new parameters to be used by the quantum network.

1.2.1 Encoding

In order for data processing to be performed on quantum computers, it is necessary that they are first loaded into these devices. This rewriting process is fundamental so that possible quantum advantages can be extracted and depends on the algorithm used.

Two common approaches to data encoding in quantum computers are amplitude embedding and angle embedding. In the first case, information is stored in the range of probabilities. However, this method of encoding is costly in itself, which makes the use of encoding in rotation angles more attractive in machine learning tasks.

1.2.2 Variational Layer

After encoding the data in the quantum computer, the variational layer is applied to approximate the output to the target function. The variational layer is composed by parametrized quantum gates that should be optimized to achieve the minimization of the cost function.

1.2.3 Measurement

The reading of the information processed by quantum computers requires the measurement of the quantum states, resulting in the collapse of each qubit state into the |0> or |1> states, destroying the superposition and projecting the system state into a computational basis state. This step is indispensable for the utilization of the information processed in quantum devices.

2. METHODOLOGY

The algorithms in this review were found through a Google Scholar search with the terms "Quantum RNN", "Quantum LSTM". The papers that presented a description of the implementation and comparison with classical models were selected. The methods were analyzed in terms of circuit design, implementation and results obtained.

3. RESULTS AND DISCUSSION

3.1 Bausch 2020

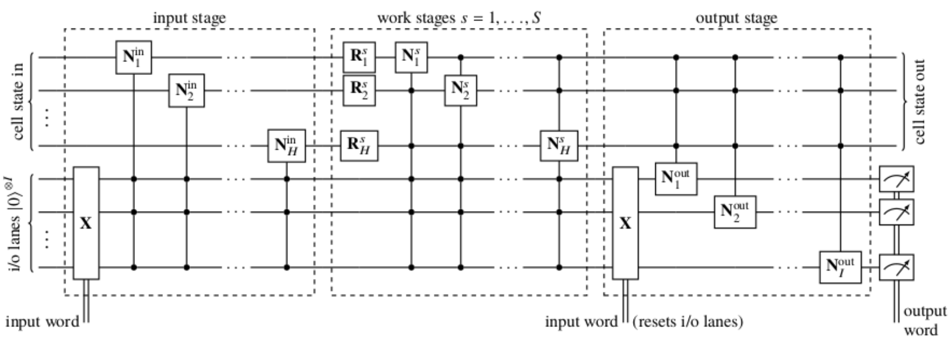
The first reference to a quantum RNN found in the literature was the work of Bausch, published at the 34th Conference on Neural Information Processing Systems (NeurIPS 2020), in Vancouver, Canada, and made available on Arxiv. The authors use the term Recurrent Quantum Neural Network, but choose the acronym "QRNN", which can be confused with the established term Quasi Recurrent Neural Network. We believe that the acronym RQNN would avoid possible confusion in the use of the term. In addition, we note that the acronym QRNN is repeated in the other references that appear about quantum recurrent neural networks that appear after this one.

The Bausch algorithm was structured using a parameterized quantum circuit, a common structure in quantum machine learning models, and was based on the work of Cao et al. [11]. The implementation was done in simulation using the deep learning library "pytorch" and the tests were done using 2-8 CPUs and 500MB and 35GB of memory per core. However, the authors did not report whether they used quantum computing SDKs in conjunction to accomplish the task.

The framework was used in the memorization tests of an integer sequence, for the classification of images using the MNIST dataset and for the generation of synthetic data using the same dataset. Regarding the analysis of the algorithm for the identification of long sequences, the authors tested the ability to predict the next base in a sequence of letters of a genetic code.

The main cell of Bausch's work is shown in Figure 1.

Figure 1. Recurrent block (Bausch 2020)



The algorithm uses amplitude amplification, which increases its depth and reduces the need to make several measurements to obtain the desired results.

3.2 Emmanoulopoulos 2022

In 2022, Emmanoulopoulos published a paper on task prediction using parameterized quantum circuits (PQCs) in prediction tasks. The data used was built specifically to perform the algorithm tests.

The implementation was done using the Keras libraries and Tensorflow. The PQCs were designed using Google Cirq, Google's quantum computing SKD, and tensors transformation using TensorFlow Quantum.

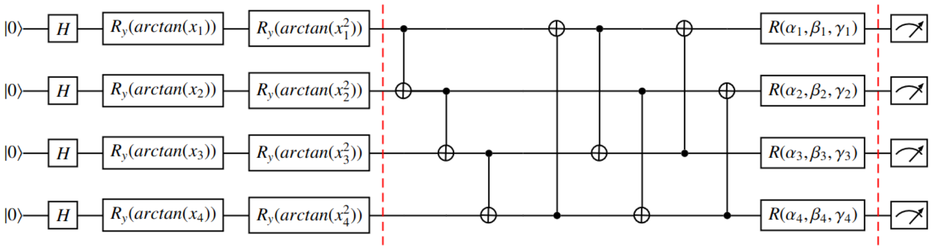
The results obtained were compared with classical results using bidirectional LSTMs and show that quantum methods achieve similar metrics to the classical and slightly outperform it on high noise signals.

3.3 Chen 2020

Chen 2020 presents the first model of quantum LSTM. The tests were based on periodic functions and population dynamics of the density matrix of an open quantum system.

The results obtained show that the model proposed by Chen learns faster than classical LSTMs, and converges more stably, without loss function increases, when using the same number of parameters.

Figure 2. Recurrent block (Chen 2020)



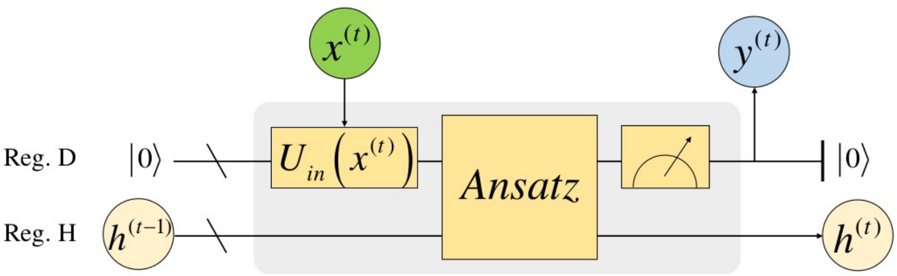
3.4 Takaki 2020

In 2020, Takaki et al. presented a recurrent quantum algorithm implemented through a variational quantum circuit and based on the recursive design of a classical RNN. The algorithm is fully quantum and uses two sets of qubits. The first set is measured periodically so that the network has the recurrent structure while the second set is not measured, so it can keep the encoded information. The tests were made in the prediction of a cosine function and a triangular wave, in addition to the prediction of the dynamics of the Lyndblad master equation for a system of 3 spins. No comparisons were made with classical algorithms. This model is especially difficult to implement in NISK devices, since the decoherence of the system makes it difficult to store information in quantum systems for long periods. The authors do not mention the use of quantum computing SDKs to perform the simulations. For the classical optimizer the authors used the BFGS algorithm using SciPy.

3.5 Li 2023

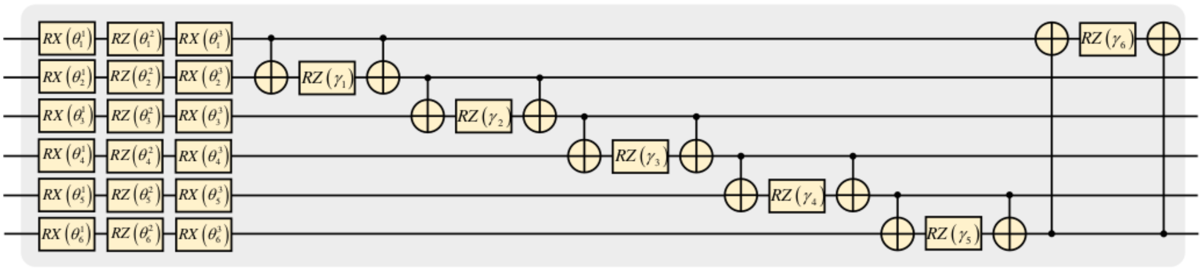
Li 2023 proposed characteristics that a good quantum RNN needs to have, namely: (1) be flexible to be implemented on various NISK platforms; (2) be fully quantum and not hybrid (classical-quantum); (3) be efficient in sequential learning of classical data. From these characteristics, the authors created a quantum recurrent neural network model. Li's recurrent block is presented in Figure 2.

Figure 3. Recurrent block (Li 2023)



The circuit consists of two registers. Data is encoded in the first register and ansatz is applied using qubits from the first and second registers. Only the first register is measured while the second one is kept intact to maintain the information of the original data. The structure of the ansatz used is shown in figure 3.

Figure 4. Ansatz (Li 2023)



From this structure, two recurrent models are presented, the "plain Quantum Recurrent Neural Network" (pQRNN) and the "staggered Quantum Recurrent Neural Network" (sQRNN). The structure of the two models is similar, since both models use the same recurrent block. However, in the pQRNN the application of the units that encode the information, as well as the qubit that is measured, are always applied to the same qubits, as shown in figure 4. On the other hand, in the sQRNN these qubits are changed at each step, as shown in figure 5.

Figure 5. pQRNN (Li 2023)

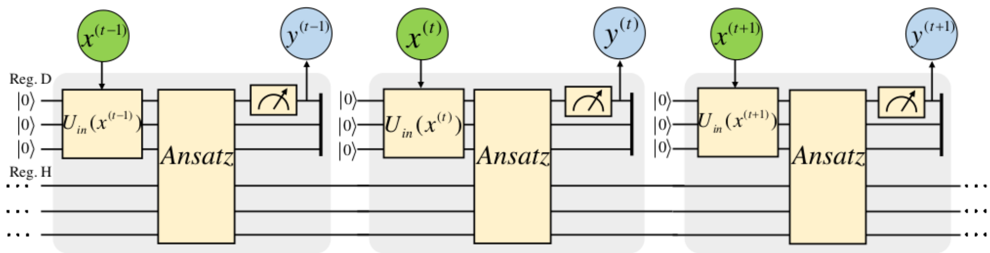
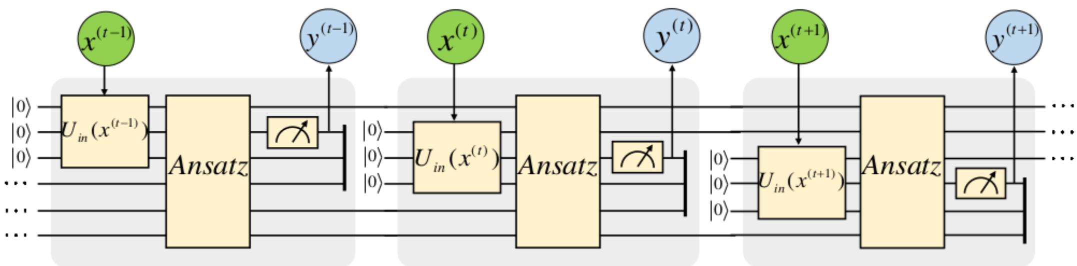


Figure 6. sQRNN (Li 2023)



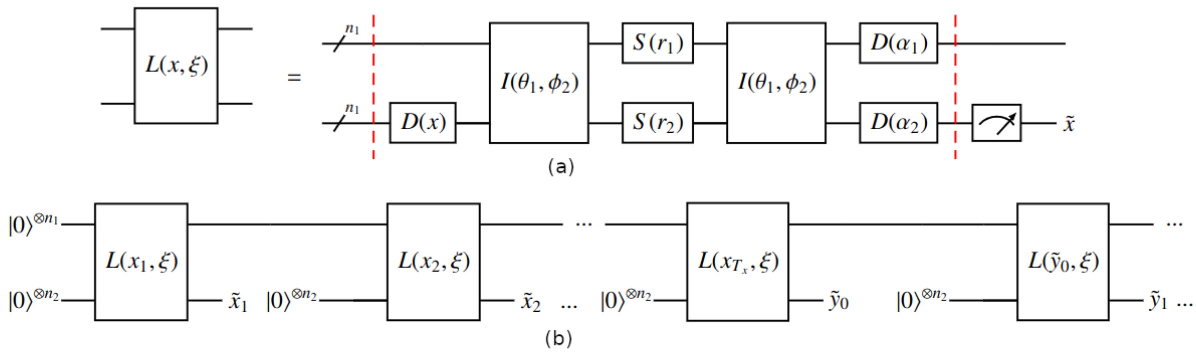
Tests were performed with weather indicators, stock prices and text categorization. Both methods show considerable advantages over a classical RNN. In particular, quantum methods were better at predicting the sharp fluctuations that occur in data variation.

3.6 Siemaszko 2023

Siemaszko et al. 2023 presents a quantum RNN model to be applied in continuous variable quantum computers, which can be implemented in already available photonic quantum computers. He calls his algorithm CV-QRNN, from

Continuous Variable Quantum Recurrent Neural Network. The structure of the network is shown in figure 6.

Figure 7. Recurrente block (Siemaszko 2023)



The method was tested using in the prediction of values of the Bessel function f(x) ≡ J_0 (x) and in the classification of images using the MNIST dataset using only two digits, 3 and 6 to reduce the problem to a binary classification problem. The results obtained by the quantum method were compared with the results of a classical LSTM and show a convergence to values in the same order of magnitude than the classical model but with a smaller number of steps, which can be considered an advantage.

3.7 Nikoloska 2023

Nikolaska et al. 2023 proposes a model for a time warping-invariant quantum recurrent Neural Network (TWI-QRNN), based on the work of Bausch [1]. The model was tested by prediction of a possibly time-warped cosine function and the dynamics of a system of spins. The results were compared with the QRNN model and with a classical LSTM, achieving lower loss values than these models.

3.5 Oğur 2023

The publication by Oğur (2023) investigates the effects of superposition and quantum entanglement for weather prediction tasks. For this, five different variational quantum circuits are used, using two different layers to generate entanglement during the application of the variational circuit and adding or excluding a superposition layer in the initialization of the circuit.

In the first experiment, the circuit is initialized with uniform superposition and the variational layer is composed of an entanglement layer of the first type constructed in the paper, followed by a rotation layer. In the second experiment, the entanglement layer is removed. The third experiment differs from the first by the removal of the uniform overlay layer at the beginning of the experiment. The fourth experiment replaces the entanglement layer used in the first experiment with an entanglement layer of the second type constructed in the paper, and finally the fifth experiment uses a variational layer composed of a rotation layer followed by an entanglement generation layer of the second type followed by a second rotation layer. Figure 9 shows the basic design of the circuit and the figure 10 shows the entanglement layers used.

Figure 8. Basic circuit design (Oğur 2023)

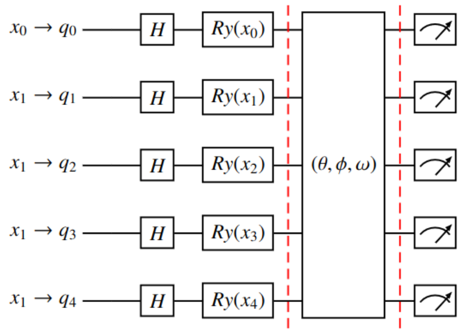
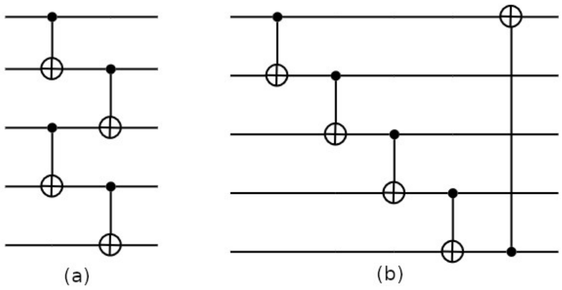


Figure 9. Entanglement layers (Oğur 2023): (a) type 1; (b) type 2.



The obtained results show that the preparation of the initial state of the circuit in uniform superposition reduces the required number of entanglement layers in the variational circuit and that the use of entanglement layers between the variational layers improves the performance of the algorithm.

4. CONCLUSION

We review quantum models of recurrent neural networks and find pure quantum models and hybrid classical-quantum models with different architectures that were tested on different types of data. In general, under certain conditions the results presented by the authors show advantages of quantum models over purely classical algorithms. Among the works reviewed, we consider the work of Oğur et al. the one that best demonstrates the effects of quantum correlations in this type of algorithms, since it tests the effects of superposition and entanglement in five different configurations.

5. REFERENCES

[1] EKMAN, Magnus. Learning Deep Learning: Theory and Practice of Neural Networks, Computer Vision, Natural Language Processing, and Transformers Using TensorFlow. Addison-Wesley Professional; 1st edition (July 19, 2021).

[2] SCHULD, Maria. Machine Learning with Quantum Computers. Springer; 2nd 2021 ed. edição (19 outubro 2022).

[3] BAUSCH, Johannes. Recurrent Quantum Neural Networks. 34th Conference on Neural Information Processing Systems (NeurIPS 2020), Vancouver, Canada.

[4] EMMANOULOPOULOS, Dimitrios; DIMOSKA, Sofia. Quantum Machine Learning in Finance: Time Series Forecasting. ArXiv, Feb 2020.

[5] YEN-CHI, Samuel. Quantum Long Shot-Term Memory. ArXiv, Sep 2020.

[6] TAKAKI, Yuto. Learning temporal data with variatonal quantum recurrent neural network. ArXiv, Dec 2020.

[7] LI, Yanan. Quantum Recurrent Neural Networks for Sequential Learning. ArXiv, Feb 2023.

[8] SIEMASZKO, Michał et al. Rapid training of quantum recurrent neural networks. ArXiv, Mar 2023.

[9] NIKOLOSKA, Ivana et al. Time-Warping Invariant Quantum Recurrent Neural Networks via Quantum-Classical Adaptive Gating., ArXiv, Jun 2023.

[10] OĞUR, Beşir. The effect of superposition and entanglement on hybrid quantum machine learning for weather forecasting. Quantum Information and Computation, Vol. 23, No. 3&4 (2023) 0181-0194.

[11] Cao et al, Quantum Neuron: an elementary building block for machine learning on quantum computers. ArXiv, Nov 2017.