

Quantum Classifier for Natural Language Processing Applications

Shyambabu Pandey¹, Partha Pakray^{1,*}, Riyanka Manna²

¹ National Institute of Technology Silchar, Assam, Silchar, India

² Amrita Vishwa Vidyapeetham, Amaravati, Andhra Pradesh, India

{babushyampandey, parthapakray, riyankamanna16}@gmail.com

Abstract. A deep neural network is a branch of machine learning that is capable of learning and representing complex patterns from a dataset through interconnected multiple layers of neurons. This capability makes it applicable in various fields, such as natural language processing, image processing, and computer vision. Deep learning models show effective performance but face challenges such as complexity and resource demands. On the other hand, quantum machine learning algorithms offer an alternative with potential efficiency compared to their classical counterparts. This paper proposes a Quantum Recurrent Neural Network (QRNN) for natural language processing tasks, which classify text data such as parts of speech, named entity recognition, and sentiment analysis. The proposed method utilizes parameterized quantum circuits that contain the tunable parameters. Our approach uses amplitude encoding to represent classical data into quantum states, partial measurement for label determination, and ancilla qubits to transfer the information from the current state to the next.

Keywords. Quantum computing, quantum machine learning, natural language processing.

1 Introduction

Machine learning [1] is a branch of artificial intelligence [22] that allows computers to learn patterns in a dataset without using explicit programs. The main goal of machine learning algorithms is to be applied in engineering and science fields, with the primary objective of identifying patterns and making decisions according to those patterns. Introducing

deep neural networks [23] enhances the performance and capability of machine learning algorithms, which can find more complex patterns from datasets.

Over the past few years, machine learning algorithms have applied deep learning models to enhance the performance of natural language processing (NLP) applications [24]. Specifically, transformer-based models perform better than traditional neural network models, such as recurrent neural network variants.

However, this superior performance poses different challenges, such as complexity in the model, demands extensive dataset, and raises a considerable amount of computational power, time, and resources [10].

An alternate path is obtaining notice in recent times from quantum computing, which can perform efficient computation for specific problems. Quantum computing [11] is a computing paradigm that follows the law of quantum theory and applies qubits (quantum bits) as processing components.

Qubit is analogous to classical bits capable of existing in the superposition of all the possible states. Additionally, entanglement, a quantum principle, can exploit the correlation between multiple qubits, even if they are physically separated. Such unique characteristics of quantum phenomena make quantum computing solve specific problems faster than their classical solutions. This hope allows us to solve some computational issues efficiently, such as machine learning algorithms.

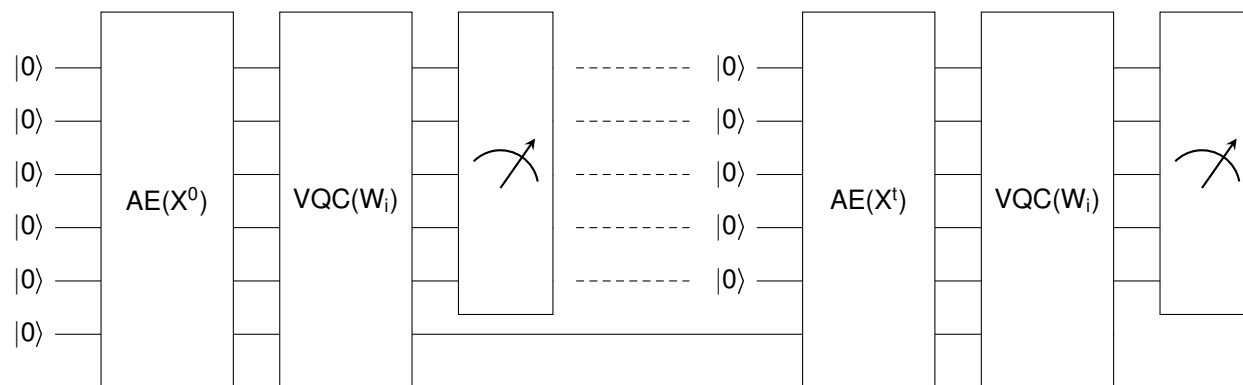


Fig. 1. Architecture of the proposed system

Quantum machine learning (QML) [18], a combination of quantum computing and machine learning to utilize powerful features taken from quantum theory to enhance the computational power of traditional machine learning algorithms. Some quantum machine learning algorithms show remarkable improvement compared to their classical counterparts.

Quantum neural network (QNN) [12] is a subfield of QML algorithms with more learning capabilities than classical neural networks. In recent times, researchers have endeavored to apply a quantum version of classical machine learning algorithms to NLP applications to improve performance.

Most NLP applications classify text data into different labels, such as part-of-speech (POS) tagging, named-entity recognition (NER), and sentimental analysis (SA). To enhance the performance of such NLP tasks, we propose a quantum recurrent neural network (QRNN) as a quantum classifier for the text data.

The proposed system is based on a parameterized quantum circuit consisting of tunable parameters to train the model and employ amplitude embedding to convert each word's word embedding into quantum states.

2 Related Work

Some QML algorithms show exponential speedup compared to their classical counterparts, such

as learning algorithms like quantum principle component analysis [15], and quantum support vector machine [21]. However, these algorithms cannot run on current quantum computers because of the lack of quantum RAM to execute.

To take advantage of current quantum hardware, variational quantum circuits [5] are employed, which consist of a series of parameterized quantum gates. Quantum approximation optimization algorithm [8], hybrid quantum-classical algorithms [3], and QNN are examples of variational quantum algorithms. These algorithms are implemented on current quantum hardware systems, which use parameterized quantum circuits in short-depth.

Researchers are trying to implement hybrid quantum-classical QNN models using quantum computing principles and classical neural network architectures. For example, Liu et al. [14] have introduced a quantum convolutional neural network (QCNN) based on hybrid quantum-classical methodology.

Ceschini et al. [4] proposed a hybrid quantum-classical recurrent neural network to predict time series data of renewable energy. Chen et al. [6] introduced a hybrid model for quantum long short-term memory (QLSTM) that can apply to NISQ devices to handle sequential data. QNN models can be applied in NLP applications to enhance their performance.

Sipio et al. [7] introduced a QLSTM hybrid model and employed it to perform POS tagging. This model demonstrates an attempt to employ

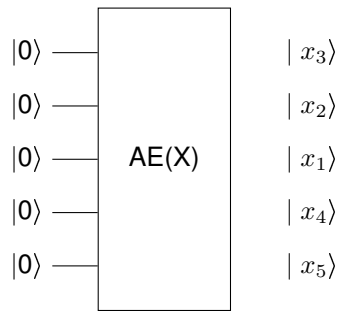


Fig. 2. Circuit for data encoding. Initially, all quantum states are present in default state $|0\rangle$

NLP applications by using QNN. However, this model does not show any significant advantage for POS tagging. Pandey et al. [19] implements a quantum LSTM (QLSTM) for POS tagging of a low-resource Indian language, Mizo.

The authors experimented with different numbers of qubits and performed hyperparameter tuning. However, their experiment's result could be better, showing that current quantum devices do not apply to large datasets. Pandey et al. [18] propose a hybrid quantum-classical QLSTM for POS tagging on code-mixed languages.

This model converts gates of classical LSTM into variational quantum layers. However, the proposed is not able to process large datasets. So, the authors split entire datasets into batches of hundred sentences for the experiments.

The code-mixed dataset consists of nine datasets collected from three social media platforms: Facebook, Twitter, and WhatsApp. The authors [13] propose QRNN to handle sequential data.

The proposed model is employed to predict stock prices and classify the text data that show significant improvements. However, this model applies to small datasets. Quantum natural language processing (QNLP) [16] is another research area to utilize near-term quantum computers for NLP applications. It employs compositional distributional semantics (DisCoCat) that apply the compositional structure of the Pregroup grammars.

It represents the grammatical structure of sentences as a string diagram, encoding a specific

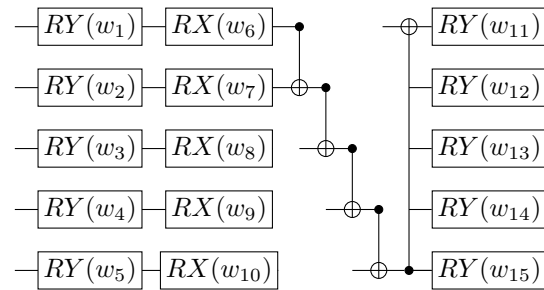


Fig. 3. Circuit for VQC

interaction of words according to the grammar. The DiscoCat converts these string diagrams into quantum circuits to process NLP applications. Various NLP applications have been implemented by the QNLP framework, like question answering [17], grammar-aware classification [17], and sentiment analysis [9]. QNLP demands a massive amount of computational resources, which makes it time-consuming to process NLP applications.

3 System Architecture

We propose a QRNN that presents a quantum counterpart of classical RNN architecture. The QRNN employs variational quantum circuits (VQC) consisting of parameterized quantum gates. These gates imply tunable parameters, which provide flexibility during the model training. In QRNN, each cell of traditional RNN is substituted by VQC. Our main objective is to perform the classification of text into different labels.

So, we perform a measurement of each circuit that provides labels for each input. The structure of the QRNN is divided into three submodules: data encoding, VQC, and measurement. Figure 1 represents the architecture of the proposed system, where AE represents amplitude encoding, and VQC represents a variational quantum circuit.

The given circuit is quantum analogous to classical RNN, where each VQC represent the cells of RNNs. Each time instance, data X_t must be encoded and applied to a VQC to process the data. In the end, partial measurements are employed to determine the labels of corresponding text data.

Again, we initialize the quantum circuit as the initial state and pass the next data instance. We use one ancilla qubit that passes the previous information to the current state, which serves as the hidden state of RNN.

3.1 Encoding Method

We must convert classical information to quantum states to process classical data using the quantum framework. The process of transforming classical data into quantum state is known as data encoding. Various data encoding methods exist, such as basis, amplitude, and angle encoding. In our study, we use amplitude encoding due to its compatibility with our quantum classifier.

Amplitude encoding encodes classical data into the amplitudes of quantum states. It uses the principle of superposition and allows multiple information to be represented in a single quantum state. The choice of amplitude encoding is because it requires fewer qubits to describe the sizeable dimensional dataset. Suppose an N qubit employs amplitude encoding; then it holds 2^n quantum states. Moreover, our systems mainly focus on NLP applications, which always deal with large dimensional datasets.

State-of-the-art (SOTA) QNN models utilize angle encoding for data encoding, where each qubit generates a single quantum state. Handling extensive dimensional data requires more qubits, while current quantum hardware limitations prevent handling large numbers of qubits. As a result, SOTA QNNS can not manage large datasets. Meanwhile, word embedding of each word generates high-dimensional word vectors.

So, our novel approach uses amplitude encoding in QRNN to handle large dimensional word vectors.

3.2 Variational Quantum Circuit

After encoding classical information, the next step applies VQCs to process quantum states. Figure 3 represents a VQC of the proposed system, which consists of Y and X rotational and controlled-not gates. The processing unit of the neural network is a combination of linear and non-linear operations.

So, our proposed system contains rotational gates X and Y, which have parameterized gates consisting of adjustable parameters, where the Y gate can represent non-linear operation [2]. The X gate is a gate that represents linear operations. A controlled-not gate is applied to generate entanglement between different qubits, which increases the circuit's entangling capability.

It helps to identify the patterns between data. The parameters w_1, w_2, \dots, w_{15} are adjustable to train the model. Our goal is to use fewer parameters to learn the model, thereby reducing the overload of many parameters during training. Classical deep learning models demonstrate good performances but require a substantial number of parameters to train the models, which makes them complex models.

On the other hand, quantum computers can find patterns from fewer parameters [20]. This unique property of quantum computers makes them applicable for identifying the patterns in datasets. However, current quantum hardware is limited in handling many parameters. Therefore, our proposed model uses fewer parameters for training, enabling it to be compatible with existing quantum hardware.

3.3 Measurement

Finally, the circuit is measured, generating classical information that will give a label of an observation. The proposed uses partial measurement, which measures all the quantum states except the ancilla qubit. The use of ancilla qubit passes the previous information to the current state.

Tila qubits serve as a hidden state of classical RNN, which maintains the flow of information from the prior state to the current state. The result of the measurement assigns labels of corresponding words. Various measurement methods exist to measure the quantum circuits. However, we use expectation measurement that provides the expected value of each observation.

4 Conclusion & Future Work

We introduce a QRNN as a quantum classifier to perform classification tasks on text data. The proposed system uses quantum mechanics principles to enhance the performance of NLP tasks. The present architecture is a novel approach, applying amplitude encoding to encode classical information and employing partial measurement to determine the label of text data, with an ancilla qubit that passes the previous state information to the current state.

We designed our proposed QRNN to accommodate sizeable dimensional word vectors, maintaining each word's integrity and requiring fewer parameters to train the model, making it compatible with current quantum computers. Our future work will apply the proposed model as a quantum classifier to classify texts such as POS tagging, NER, and text classification.

Acknowledgments

The work presented here falls under Project Entitled "Quantum Machine Learning for Natural Language Processing Applications" (Grant Ref. No. N-21/17/2020-NeGD) supported by MeitY Quantum Computing Applications Lab (QCAL).

References

1. **Baştanlar, Y., Özuysal, M. (2014).** Introduction to machine learning. *miRNomics: MicroRNA biology and computational analysis*, Vol. 1107, pp. 105–128. DOI: 10.1007/978-1-62703-748-8.7.
2. **Bausch, J. (2020).** Recurrent quantum neural networks. *Proceedings of the 34th Conference on Neural Information Processing Systems*, pp. 1368–1379. DOI: 10.48550/ARXIV.2006.14619.
3. **Callison, A., Chancellor, N. (2022).** Hybrid quantum-classical algorithms in the noisy intermediate-scale quantum era and beyond. *Physical Review A*, Vol. 106, No. 1. DOI: 10.1103/PhysRevA.106.010101.
4. **Ceschini, A., Rosato, A., Panella, M. (2022).** Hybrid quantum-classical recurrent neural networks for time series prediction. *Proceedings of the International Joint Conference on Neural Networks*, pp. 1–8. DOI: 10.1109/IJCNN55064.2022.9892441.
5. **Chen, S. Y. C., Huck-Yang, C. H., Qi, J., Chen, P. Y., Ma, X., Goan, H. S. (2020).** Variational quantum circuits for deep reinforcement learning. *IEEE Access*, Vol. 8, pp. 141007–141024. DOI: 10.1109/ACCESS.2020.3010470.
6. **Chen, S. Y. C., Yoo, S., Fang, Y. L. L. (2022).** Quantum long short-term memory. *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 8622–8626. DOI: 10.1109/icassp43922.2022.9747369.
7. **Di-Sipio, R., Huang, J. H., Chen, S. Y. C., Mangini, S., Worring, M. (2022).** The dawn of quantum natural language processing. *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 8612–8616. DOI: 10.1109/icassp43922.2022.9747675.
8. **Farhi, E., Goldstone, J., Gutmann, S. (2014).** A quantum approximate optimization algorithm. *arXiv*. DOI: 10.48550/ARXIV.1411.4028.
9. **Ganguly, S., Morapakula, S. N., Coronado, L. M. P. (2022).** Quantum natural language processing based sentiment analysis using lambeq toolkit. *Proceedings of the 2nd International Conference on Power, Control and Computing Technologies*, pp. 1–6. DOI: 10.1109/ICPC2T53885.2022.9776836.
10. **Guarasci, R., De-Pietro, G., Esposito, M. (2022).** Quantum natural language processing: Challenges and opportunities. *Applied Sciences*, Vol. 12, No. 11, pp. 5651. DOI: 10.3390/app12115651.
11. **Gyongyosi, L., Imre, S. (2019).** A survey on quantum computing technology. *Computer Science Review*, Vol. 31, pp. 51–71. DOI: 10.1016/j.cosrev.2018.11.002.

12. **Jia, Z. A., Yi, B., Zhai, R., Wu, Y. C., Guo, G. C., Guo, G. P. (2019).** Quantum neural network states: A brief review of methods and applications. *Advanced Quantum Technologies*, Vol. 2, No. 7–8. DOI: 10.1002/qute.201800077.
13. **Li, Y., Wang, Z., Han, R., Shi, S., Li, J., Shang, R., Zheng, H., Zhong, G., Gu, Y. (2023).** Quantum recurrent neural networks for sequential learning. *Neural Networks*, Vol. 166, pp. 148–161. DOI: 10.1016/j.neunet.2023.07.003.
14. **Liu, J., Lim, K. H., Wood, K. L., Huang, W., Guo, C., Huang, H. L. (2021).** Hybrid quantum-classical convolutional neural networks. *Science China Physics, Mechanics and Astronomy*, Vol. 64, No. 9, pp. 1–8. DOI: 10.1007/s11433-021-1734-3.
15. **Lloyd, S., Mohseni, M., Rebentrost, P. (2014).** Quantum principal component analysis. *Nature Physics*, Vol. 10, No. 9, pp. 631–633. DOI: 10.1038/nphys3029.
16. **Meichanetzidis, K., Gogioso, S., de-Felice, G., Chiappori, N., Toumi, A., Coecke, B. (2021).** Quantum natural language processing on near-term quantum computers. *Electronic Proceedings in Theoretical Computer Science*, Vol. 340, pp. 213–229. DOI: 10.4204/eptcs.340.11.
17. **Meichanetzidis, K., Toumi, A., de-Felice, G., Coecke, B. (2023).** Grammar-aware sentence classification on quantum computers. *Quantum Machine Intelligence*, Vol. 5, No. 1. DOI: 10.1007/s42484-023-00097-1.
18. **Pandey, S., Basisth, N. J., Sachan, T., Kumari, N., Pakray, P. (2023).** Quantum machine learning for natural language processing application. *Physica A: Statistical Mechanics and its Applications*, Vol. 627, pp. 129123. DOI: 10.1016/j.physa.2023.129123.
19. **Pandey, S., Dadure, P., Nunsanga, M. V. L., Pakray, P. (2022).** Parts of speech tagging towards classical to quantum computing. *Proceedings of the IEEE Silchar Subsection Conference*, pp. 1–6. DOI: 10.1109/silcon55242.2022.10028796.
20. **Phillipson, F. (2020).** Quantum machine learning: Benefits and practical examples. Vol. 2561, pp. 51–56.
21. **Rebentrost, P., Mohseni, M., Lloyd, S. (2014).** Quantum support vector machine for big data classification. *Physical review letters*, Vol. 113, No. 13, pp. 130503. DOI: 10.1103/physrevlett.113.130503.
22. **Russell, S., Russell, S. J., Norvig, P., Davis, E. (2010).** *Artificial intelligence: A modern approach*. Prentice Hall.
23. **Suk, H. I. (2017).** An introduction to neural networks and deep learning. In *Deep Learning for Medical Image Analysis*. Elsevier, pp. 3–24.
24. **Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., Polosukhin, I. (2017).** Attention is all you need. *Advances in Neural Information Processing Systems 30*, Vol. 30, pp. 1–11.

Article received on 02/02/2024; accepted on 14/04/2024.

**Corresponding author is Partha Pakray.*