Title: Uncertainty estimation in quantum extreme reservoir computation towards enabling reliable quantum machine learning

Authors:
Aoi Hayashi[1, 2, 3], Akitada Sakurai[2], Hon Wai Lau[2], William J. Munro[2], and Kae Nemoto[2,3]

[1]SOKENDAI (the Graduate University for Advanced Studies), [2]Okinawa Institute of Science and Technology, [3]National Institute of Informatics

Abstract:
Quantum machine learning has gathered significant attention driven by the expectation that quantum systems may contribute to higher learning performance using fewer physical resources. Quantum Extreme Reservoir Computation (QERC) [1] for instance achieved a testing accuracy rate of 97.1% with only 11 qubits on the MNIST image classification task and outperformed its classical counterpart with the same number of tuning parameters. However, even with such a high testing accuracy rate, we do not know whether the machine learning prediction for unseen data is actually correct. It could be problematic when one embeds a quantum machine learning technique in a decision-making system. How much can we trust the predictions from quantum machine learning?

In classical machine learning, the reliability of neural networks has been extensively explored [2, 3] including Evidential Deep Learning (EDL) [4] based on Subjective Logic [5]. EDL outputs class predictions and the associated uncertainty in classification tasks. It allows us to decide whether the prediction is trustable or not based on its corresponding uncertainty. The proposed training scheme results in a robust neural network to unseen data outside the training data.

In our work, we comprehensively investigated the learning behavior of QERC equipped with this uncertainty estimation method. In classification tasks, we compare the QERC with uncertainty estimation and a classical counterpart for different numbers of classes and features. It is shown that with a fixed number of training parameters, the uncertainty estimation method preferers the quantum feature map in QERC to the one in its classical counterpart in terms of the training stability and the overall performance. We have confirmed the improvement in out-of-distribution detection tasks [6] using the quantum feature map. These results show considerable potential for quantum machine learning to be more reliable with fewer tuning parameters than their classical counterparts.

[1] A. Sakurai, M. P. Estarellas, W. J. Munro, and K. Nemoto, Phys. Rev. Applied **17**, 064044 (2022).
[2] M. Abdar, F. Pourpanah, S. Hussain, D. Rezazadegan, *et al.*, Information Fusion **76**, 243-297 (2021).
[3] J. Gawlikowski, C. R. N. Tassi, M. Ali, J. Lee, *et al*., Artif. Intell. Rev. **56** (Suppl1), 1513-1589 (2023).
[4] M. Sensoy, L. Kaplan, and M. Kandemir, Adv. Neural. Inf. Process. Syst. **31** (2018).
[5] A. Jøsang, *Subjective Logic: A Formalism for Reasoning Under Uncertainty*, Springer, (2016).
[6] J. Yang, K. Zhou, Y. Li, and Z. Liu, arXiv:2110.11334 (2021).