Progetto di ricerca correlato all'assegno

The position is funded by the PRIN 2022 project "understanding the LEarning process of QUantum Neural networks (LeQun)", proposal code 2022WHZ5XH, CUP Master J53D23003890006, CUP J53D23003890006

Abstract of the research project:

Quantum computers promise to be a revolutionary solution to fulfill the increasing need for high-performance computing, and quantum computing has been identified as a key intervention area in the "Programma Nazionale per la Ricerca 2021-27". Determining the problems on which quantum computers can provide major advantages with respect to classical computers is the main theoretical problem of quantum computing and constitutes the challenge that LeQun addresses.

The most promising family of quantum algorithms that can be implemented on the forthcoming generation of quantum computers are variational quantum algorithms, also called quantum neural networks. Despite the promises, there is no problem of practical interest yet where quantum neural networks have a provable advantage over the best classical algorithms. A thorough theoretical study of the trainability, expressibility and generalization properties of quantum neural networks and of their potential advantages with respect to classical computers constitutes the main theoretical challenge of quantum machine learning.

LeQun will tackle this extremely ambitious challenge through the following objectives:

O1:

- To analytically characterize the probability distribution of the functions generated by trained quantum neural networks and determine their trainability and generalization performances.
- To study the training stability against imperfect outcomes of quantum measurements and the query complexity of the entire training process, namely the number of measurements that must be performed to achieve the desired accuracy.

O2: To identify the architectures of quantum neural networks that have the potential to provide major advantages with respect to classical computers and to perform a proof-of-principle validation of the advantages of the identified architectures using real quantum devices and simulators.

LeQun will tackle the challenge with an interdisciplinary approach that connects quantum machine learning with probability theory and quantum many-body physics. The strategy of LeQun is based on the recent breakthrough results in classical machine learning stating that in the mean-field limit of infinite width of the hidden layers, trained deep neural networks are equivalent to Gaussian processes. These results explained the unreasonably good performances of deep neural networks, and LeQun will generalize them to the quantum setting.

LeQun will constitute a successful example of interdisciplinary research and will foster a cross-fertilization among quantum computing, probability theory and many-body quantum physics. The results of LeQun will be highly relevant for all the researchers working on quantum computing and machine learning both in industry and academia, thus contributing to creating value for the whole society, and its results will constitute a fundamental contribution to the main challenge of quantum computing.

Composizione dei membri della commissione dell'eventuale Bando

La valutazione comparativa dei candidati sarà effettuata da una Commissione giudicatrice formata da: - Prof. Giacomo De Palma (Presidente)

Prof. Dario Trevisan (segretario verbalizzante)

Prof. Leonardo Banchi (componente)

Prof Daniele Tantari (eventuale membro supplente)

<u>Requisiti di ammissione</u>

Alle selezioni sono ammessi a partecipare i candidati, anche cittadini di Paesi non appartenenti alla Unione Europea, in possesso di adeguato curriculum scientifico professionale e di:

- Dottorato di ricerca in matematica, fisica, informatica o titolo equivalente corredato da un'adeguata produzione scientifica conseguito in Italia o all'estero;

- Laurea magistrale/specialistica o vecchio ordinamento o titolo equivalente in *nessun ambito specifico* con adeguato curriculum scientifico-professionale.

- è previsto un colloquio;

- in caso di colloquio indicare la modalità: online;

- è prevista una valutazione di competenza della lingua inglese

Piano di Attività

1. Team composition

The team will be composed by the PI Prof. Giacomo De Palma (UniBo), the unit leaders Prof. Leonardo Banchi (UniFi) and Prof. Dario Trevisan (UniPi), postdoc 1 in the UniBo unit jointly supervised by GdP and DT, and postdoc 2 in the UniFi unit supervised by LB.

2. Work plan

The objectives O1 and O2 will be addressed by WP1 and WP2, respectively. Each WP is divided into sub-WPs, which consist of a mathematical and a physical part.

WP1

In WP1, we will characterize the probability distribution of the functions generated by trained quantum neural networks. WP1 consists of the sub-packages WP1.1 and WP1.2.

WP1.1 [months 1-8]

In WP1.1, we will characterize the probability distribution of the functions generated by quantum neural networks with random parameters. WP1.1 consists of a mathematical part (WP1.1a) and a physical part (WP1.1b), which will be carried out in parallel.

WP1.1a [GdP, DT, LB, postdoc 1]

WP1.1a will consist in proving that the output generated by quantum neural networks with random parameters behave as Gaussian processes in the mean-field limit of an input made by infinitely many iid copies of a given quantum state. Depending on whether the input is classical or quantum, we will consider quantum neural networks that have as input many iid copies of either a quantum state that encodes the classical input or the quantum state to be classified. Physical intuition suggests that the dynamics of the network in the aforementioned limit is described by a mean-field theory. We will provide a rigorous proof for this intuition by generalizing to the quantum setting the proof of [Lee 18], which relies on a central limit theorem for arrays of weakly dependent random variables. Our main strategy will be to generalize the techniques in the proof of the quantum Berry-Esseen theorem of [Brandao 15]. We will quantify the goodness of the mean-field approximation through the quantum Wasserstein distance for qubits of [De Palma 21]. We will also consider alternative routes, based on later proofs of the classical result, relying on direct use of characteristic functions or explicit computations of moments [Bracale 20].

WP1.1b [GdP, LB, postdoc 2]

The general aim of WP1.1b is to find and numerically verify physics-inspired methods to design quantum neural networks that admit a mean-field treatment. We will consider many iid copies of the initial state and use second quantization to describe its evolution. By introducing ancillary modes, it is possible to use the second-quantization formalism even when the quantum neural network is not permutationally invariant [Banchi 17]. In this language, the quantum neural network couples all the different copies together and can be described by a Hamiltonian with quartic interactions in the bosonic creation and annihilation operators. These Hamiltonians are routinely studied in the quantum many-body physics literature within mean-field methods [Bender 03]. We will study the evolution of the Hamiltonian parameters via gradient descent, express the gradient using integrals over time evolutions [Banchi 21b] and then perform time-dependent mean field approximations (a la Hartree-Fock-Bogoliubov) to formally approximate the resulting expressions. Furthermore, we will verify the applicability of such approximations through proof-of-principle numerical experiments.

WP1.2 [months 9-16]

In WP1.2, we will extend the results of WP1.1 to trained quantum neural networks. WP1.2 consists of a mathematical part (WP1.2a) and a physical part (WP1.2b), which will be carried out in parallel.

WP1.2a [GdP, DT, postdoc 1]

WP1.2a will consist in extending to trained quantum neural networks the results of WP1.1a. We propose as a strategy to generalize to the quantum setting the proof of [Lee 20] of the equivalence between Gaussian processes and trained deep neural networks. The idea underlying the result of [Lee 20] is that in the limit of infinite width of the hidden layers, the training of the network can be described by the linear model coming from the first-order Taylor expansion in the parameters with respect to their initialization values. This linear model is governed by the neural tangent kernel, which is given by the inner products between the gradients of the generated function with respect to the parameters. The neural tangent kernel is constant over most of the parameter space and does not change during training. We will generalize this proof to quantum neural networks building on the results of [Abedi 22], which provides bounds to the approximation error of the linear model for the training of geometrically local quantum circuits. We will define a quantum counterpart of the neural tangent kernel and prove that it is constant over most of the parameter space and does not change during training. This result will imply that trained quantum neural networks are equivalent to a linear model which can be solved analytically and behave as Gaussian processes. Furthermore, the quantum neural tangent kernel will allow us to determine the mean and the covariance of the function generated by the trained network in terms of the training set and of the cost function.

WP1.2b [GdP, LB, postdoc 2]

In WP1.2b, we will test the equivalence between quantum neural networks and Gaussian processes with numerical experiments and stability analyses. We will study how imperfect estimations of the gradient affect the kernel of the Gaussian process. Since the gradient is reconstructed from measurements on a quantum device, we will study how the number of measurements scales as a function of the number of qubits, of the circuit depth and of the number of interactions in the Hamiltonian describing the neural network. We will generalize the method from [Banchi 21b], which consists in finding bounds on the variance of the gradient estimator and then use concentration inequalities to bound the error. Moreover, we will perform extensive numerical simulations with classical datasets that are publicly available. We will study different encodings of the classical bits into quantum states and verify when the Gaussian-process approximation is valid. We expect that the encoding will not play a significant role in the infinite-width limit, yet for finite width there may be preferred choices. We will try encodings into different orthogonal quantum states and encoding unitary operators whose parameters depend variationally on the classical input. Finally, we will study the generalization properties of trained quantum neural networks and calculate how the von Neumann and the Rényi entropies change after each layer, in order to verify the predicted link between generalization and coarse graining of irrelevant information [Banchi 21a].

WP2 [months 17-24]

In WP2, we will identify the architectures of quantum neural networks that have the potential to provide major advantages with respect to classical computers. WP2 consists of a mathematical part (WP2a) and a physical part (WP2b), which will be carried out in parallel.

WP2a [GdP, DT, postdoc 1]

WP2a consists in analyzing the performances of quantum neural networks and determining which architectures have the potential to outperform classical computers. We will identify the promising architectures through the following criteria:

* The associated quantum neural tangent kernel must be hard to compute on a classical computer as the kernels considered in [Havlíček 19], such that the behavior of the network is hard to simulate. We will check this property by directly inspecting the expression for the quantum neural tangent kernel that we will get in WP1.

* The training time must scale at most polynomially with the size of the input, such that the training can be completed in a realistic time.

* The probably approximately correct Bayesian generalization bounds [Seeger 02] applied to the associated Gaussian process must yield good generalization properties for the quantum neural network.

WP2b [GdP, LB, postdoc 2]

Quantum speed-ups in machine learning are highly dependent on the dataset [Liu 21]. We will look for quantum advantage in a physically motivated problem: Classifying phases of matter. We will start by first studying the quantum phase recognition problem with integrable spin chains that can be exactly solved after a Jordan-Wigner transformation [Banchi 21a] and then move on to study topological phases for which no efficient classical methods exist [Cong 19]. We expect to gain from the integrable case the necessary experience to construct optimal quantum neural network architectures. We will then apply our experience to machine-learning problems with either classical or quantum input and that are widely believed to be hard for conventional methods. Finally, we will study the robustness of quantum classifiers with respect to noisy outcomes from quantum measurements and perform proof-of-principle experiments on real devices using IBM Quantum Experience and Amazon Braket cloud services, already used by LB et al. in [Gentini 21].

[Abedi 22] Abedi, Beigi, Taghavi, Quantum Lazy Training, arXiv:2202.08232

[Banchi 17] Banchi, Burgarth, Kastoryano, Driven quantum dynamics: will it blend?, Phys Rev X 7, 041015 (2017)

[Banchi 21a] Banchi, Pereira, Pirandola, Generalization in quantum machine learning: A quantum information standpoint, PRX Quantum 2, 040321 (2021)

[Banchi 21b] Banchi, Crooks, Measuring analytic gradients of general quantum evolution with the stochastic parameter shift rule, Quantum 5, 386 (2021)

[Bender 03] Bender, Heenen, Reinhard, Self-consistent mean-field models for nuclear structure, Rev Mod Phys 75, 121 (2003)

[Bracale 20] Bracale, Favaro, Fortini, Peluchetti, Large-width functional asymptotics for deep Gaussian neural networks, ICLR 2020

[Brandao 15] Brandao, Cramer, Equivalence of Statistical Mechanical Ensembles for Non-Critical Quantum Systems, arXiv:1502.03263

[Cong 19] Cong, Choi, Lukin, Quantum convolutional neural networks,

Nat Phys 15, 1273 (2019)

[De Palma 21] De Palma, Marvian, Trevisan, Lloyd, The Quantum Wasserstein Distance of Order 1, IEEE Trans Inf Theory 67, 6627 (2021)

[Gentini 21] Gentini, Cuccoli, Banchi, Variational Adiabatic Gauge Transformation on real quantum hardware for effective low-energy Hamiltonians and accurate diagonalization, arXiv:2111.08771

[Havlíček 19] Havlíček et al, Supervised learning with quantum-enhanced feature spaces, Nature 567, 209 (2019)

[Lee 18] Lee, Sohl-Dickstein, Pennington, Novak, Schoenholz, Bahri, Deep Neural Networks as Gaussian Processes, ICLR 2018

[Lee 20] Lee et al, Wide neural networks of any depth evolve as linear models under gradient descent, J Stat Mech 124002 (2020)

[Liu 21] Liu, Arunachalam, Temme, A rigorous and robust quantum speed-up in supervised machine learning, Nat Phys 17, 1013 (2021)

[Seeger 02] Seeger, PAC-Bayesian generalisation error bounds for Gaussian process classification, J Mach Learn Res 3, 233 (2002)