Task- Comment on potential ideas to extend this classical KAN architecture to a quantum KAN and sketch out the architecture in detail.

Taking inspiration from the recent paper *"Quantum Kolmogorov-Arnold Networks (QKAN)"* by Ivashkov et al. (2024) [1], I propose to extend the classical Kolmogorov-Arnold Network (KAN) architecture into the quantum domain by leveraging quantum computational techniques such as Quantum Singular Value Transformation (QSVT) and block encoding. Classical Kolmogorov-Arnold Networks (KANs) have emerged from the Kolmogorov-Arnold representation theorem, which states that any continuous multivariate function can be expressed as compositions and summations of univariate activation functions. While KANs offer enhanced interpretability and accuracy in classical applications, their quantum extension (QKANs) can potentially leverage the power of quantum computation to efficiently model and process multivariate functions encoded directly in quantum states, facilitating operations in high-dimensional Hilbert spaces.

In quantum computing, the efficient representation of classical data is often accomplished using block encoding techniques. Block encoding is a powerful quantum linear algebra technique where classical data, such as vectors or matrices, are encoded within the upper-left block of a larger unitary matrix. Specifically, given a data matrix A, it can be embedded into a quantum state using a unitary operator U, where A occupies a scaled sub-block of U. The encoded quantum state thus holds a normalized representation of the data.

Within the proposed QKAN architecture, activation functions are implemented using Quantum Singular Value Transformation (QSVT). QSVT is a quantum algorithmic technique enabling precise and efficient polynomial transformations of singular values of encoded matrices. In essence, it allows quantum circuits to apply polynomial approximations of nonlinear activation functions directly to the singular values of quantum-encoded data. Specifically, nonlinear activation functions such as sigmoid or tanh, traditionally computationally expensive to implement classically, can be approximated effectively by a finite-order polynomial expressed in the Chebyshev polynomial basis. The Chebyshev polynomial expansions have excellent approximation properties for smooth functions and are represented as follows:

$$f(x) \approx \sum_{j=0}^{d} c_j T_j(x)$$

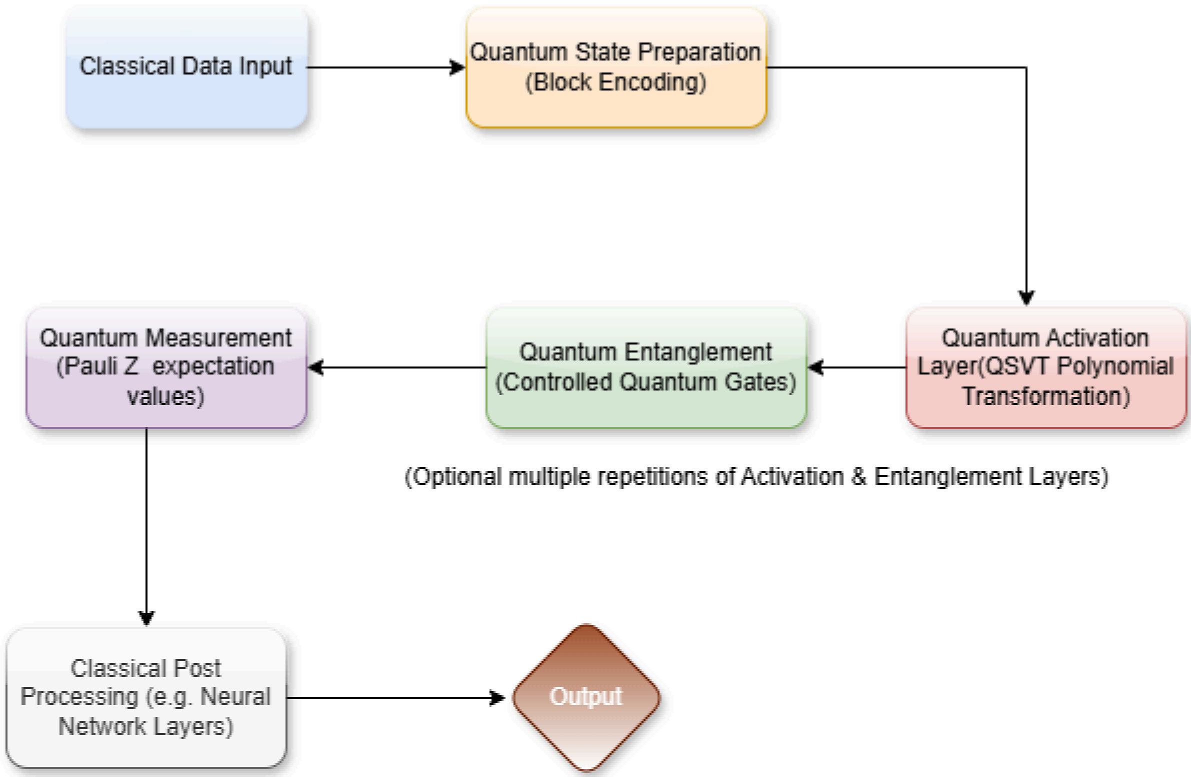
where $T_j(x)$ are Chebyshev polynomials of the first kind and the coefficients c_j become trainable parameters within the quantum circuit. These coefficients are directly encoded into parameterized quantum gates (such as rotations around the X or Y axis, RX, or RY gates), enabling efficient and controlled nonlinear transformations at the quantum circuit level.

To introduce nonlinear activation inside each QKAN layer, we employ quantum circuits using parameterized quantum singular value transformation gates by the said Chebyshev coefficients. A QKAN layer thereby contains a block-encoded input state followed by various quantum rotations and controlled quantum gates. In these, rotation angles directly represent the Chebyshev polynomial coefficients. Controlled operations (e.g., CNOT or controlled rotation gates) enable entanglement between qubits, increasing the representational capacity of the network by capturing high-dimensional, complex interactions between input variables. This quantum entanglement, not classically available, enriches feature representations and potentially leads to better performance when modeling highly nonlinear and multivariate relationships.

The general QKAN architecture consists of several layers, each performing quantum singular value transformations to increasingly extract nonlinear features. After applying these quantum transformations, the ensuing quantum state captures intricate nonlinear interactions effectively. Observables, commonly Pauli operators (e.g., Pauli-Z), are measured to get information out of this quantum state. By evaluating the expectation values of these observables, the quantum outputs are transformed back into classical predictions, suitable for downstream machine learning tasks such as classification or regression. For more complex data reconstruction, quantum state tomography methods may be employed to fully characterize the output quantum state, albeit with higher computational overhead.

One of the major benefits of using quantum singular value transformations and block encoding in QKAN is the effective handling and processing of high-dimensional quantum data. Quantum block encoding significantly lowers memory complexity since data representation in quantum states grows logarithmically with input dimension. Furthermore, using QSVT lowers the computational complexity of applying nonlinear activation functions to a significant degree compared to their classical counterparts, providing a means toward quantum advantage in large-scale machine learning applications.

In practice, constructing a QKAN layer begins with a block encoding the input data into a quantum state. Subsequently, quantum singular value transformation layers perform nonlinear transformations, with quantum gates parameterized by trained polynomial coefficients. Multiple such QKAN layers can be stacked to build deeper architectures, each layer progressively refining the quantum state representation of input data. The trained parameters are optimized using classical gradient-based methods, updating quantum gate parameters iteratively to minimize the loss function.



Sketch of the Quantum KAN architecture

References

1) Ivashkov, P., Huang, P. W., Koor, K., Pira, L., & Rebentrost, P. (2024). Quantum Kolmogorov-Arnold Networks (QKAN). *arXiv preprint arXiv:2410.04435*. https://arxiv.org/abs/2410.04435