Quantum Generative Kernels for Stochastic Classification

Keywords: quantum kernel, quantum hidden Markov model, financial sequences

This study proposes a quantum computing approach to the design of machine learning kernels for classification of stochastic symbolic time series. In stochastic classification, the class of a sequence is defined by a probability distribution.

The model is particularly relevant for several financial machine learning tasks, where due to the nature of the environment, the same event can be probabilistically classified into different classes. For instance, a sequence of trade orders might indicate security price manipulation with certain probability. Similarly, a sequence of security price movements might suggest future liquidity clustering towards buy or sell side, but only with a specific likelihood.

In the area of machine learning, kernels are important components of various similarity-based classification, clustering, and regression algorithms. They are functions, which quantify the similarity between examples in the context of a specific learning task. The conventional method for definitions of kernels is to map the examples to a high-dimensional vector space and use a geometric metaphor of similarity as a dot product.

In the context of a stochastic learning environments, we propose a probabilistic metaphor of similarity. Our kernels map the examples to the density operators of a quantum Hilbert space. The variational, or fidelity-based distances are used to quantify the similarity.

To effectively implement this approach, we leverage a quantum generative model of the example space known as a quantum hidden Markov model (QHMM) [1]. The model reflects the partial observability and quasi-stationarity of the financial environments.

A QHMM is a completely positive trace-preserving (CPTP) map (quantum channel) defined by a set of Kraus operators associated with observed symbols. Examples are described by the application of corresponding Kraus operators resulting in sequences of quantum states referred to as *generative state sequences*. The probability of a sequence is defined by the trace of the final density operator in the sequence. This modeling approach allows a formal learning model in the terms if a stochastic process language. The learning criteria is the divergence between model's and target distributions of sequences.

In an earlier study we have demonstrated that the QHMMs are parsimoniously superior to the corresponding classical HMMs: If a Markovian state process can be modeled in a classical stochastic vector space with N dimensions, then there is an equivalent QHMM in quantum Hilbert space with $\sqrt{N}$ dimensions. The implementation of these models in quantum gate computing framework is computationally feasible due to the smoothness and high autocorrelation of their learning landscapes.

We introduce two types of kernels designed for specific classification tasks. Tasks in which the class of a sequence depends on its future stochastic evolution are referred to as *predictive classification tasks*. Given the assumption of a Markovian process, a sequence's future behavior depends solely on the final state in its generative state sequence. In such scenarios, the kernel evaluates the sequence similarity using a distance measure between their final generative states. Positive semidefinite distance measures, such as the Trace distance or Bures distance, are used to define these kernels. We demonstrate Lipschitz-style continuity between the variational distance of the density operators and the divergence of corresponding distributions of future observables. These kernels are referred to as "predictive kernels".

In another category of tasks, the class of a sequence depends exclusively on its structure, such as the presence of specific patterns. We denote these tasks as 'structural classification tasks'. In such instances, the kernel maps a sequence to the expectation of the density operators in the generative state sequence. This design reflects the assumption that the expectation of the density operators captures the underlying patterns of the sequence: the important features are caused by generative states with stronger impact on the expectation. These kernels are defined as trace or Bures distances between expectations of density operators. Since the expectation of density operators are density operators, the kernels are positive semi-definite. These kernels are referred to as "structural kernels".

We have performed extensive empirical study on the impact of the size of the quantum Hilbert space on the proposed kernels behavior. The study confirms the expectation that increasing of the dimension of the quantum Hilbert space enhances the separability of examples.

In the context of a classification task, we demonstrate that the example distances calculated by the proposed kernels are correlated with the classes of corresponding examples. Specifically, examples at shorter distance have high probability to belong to the same class. These kernel features are important for superior performance on classification tasks.

To compare the performance of the proposed kernels against classical ones, we defined classification tasks using a simplified model of directional movements in a stock market. Three common kernel-based algorithms - Support Vector Machine, k-Nearest Neighbors, and Gaussian Processes - were implemented with classical and quantum kernels. Two well-known classical algorithms - Random Forest and Extreme Gradient Boosting were used as benchmarks. In all structural and predictive task scenarios, the quantum kernels exhibited superior performance at clear confidence levels, compared to their classical counterparts and the benchmarks.

We have calculated the kernel matrix for the market model example described in [1] running projected kernels on *ibm_nazca* device. To boost the effective number of shots we ran multiple circuits simultaneously in parallel on a single chip by combining them on a single circuit effectively using 72 qubits to generate 12 processes in parallel. The experiment yielded excellent agreement with the expected structure of the kernel matrix.

References

[1] V. Markov et al., "Implementation and learning of quantum hidden Markov models," arXiv:2212.03796 (2022).