Quantum Cognition Machine Learning: Financial Forecasting

Ryan Samson[1, *], Jeffrey Berger[1], Luca Candelori[1, 2], Vahagn Kirakosyan[1], Kharen Musaelian[1, 3], and Dario Villani[1, 3]

[1]Qognitive, Inc., Miami Beach, FL
[2]Wayne State University, Department of Mathematics, Detroit, MI
[3]Duality Group, Miami Beach, FL
[*]Primary Investigator (ryan.samson@qognitive.io), remaining authors presented in alphabetical order.

September 20, 2024

1 Introduction

The last decade has seen massive improvements in machine learning techniques that hold promise for breakthrough advances in productivity and technology. At the same time, there is a growing understanding that something fundamental is missing in the current framework of generative AI and statistical machine learning in general [11, 7].

In this paper, we show how a machine learning approach based on the ideas of quantum cognition ([8, 22] and references therein) is capable of capturing features of a system differently from existing machine learning approaches. Additionally, this approach can accommodate a large number of features and concept creation not bound by sharp categories.

2 Shortcomings of Classical ML

Whether implicitly or explicitly, machine learning has always been about learning a joint probability distribution of preprocessed features, whose relevance and significance is ascertained by a human understanding of the domain. The fundamental problem with this classical probabilistic description is that its complexity grows exponentially with the number of features. For example, a probability distribution over N binary features corresponds to a vector of size $2^N - 1$. This also leads to an exponential requirement for the amount of data needed to learn this distribution statistically [22].

In finance, given the non-stationarity of financial time series it is easy to see that even for a small set of features you lack sufficient data to model complex joint distributions. For example, imagine you have available for training 10 years of daily data on 3,000 firms, which would amount to approximately 7.8 million observations. If you wish to model the joint distribution of 10 features bucketed into quintiles, you have 5^{10} (approximately 9.8 million) bins, and statistics is clearly not possible.

You could alleviate the dimensionality problem by using coarser bins, i.e. 2^{10} bins would allow for statistics, but this model is likely far less detailed than desired. Alternatively, in some cases you could expand the amount of training data used by extending the sample much further back in time, but this would incorporate data from when markets were substantially different, which again leads to undesirable (less relevant) results. A surprising answer to these problems lies in using quantum, rather than classical probabilities.

3 Quantum Cognition Machine Learning

The first idea of quantum cognition emerged from the works of Aerts et al [1], Khrennikov [20], and Busemayer et al [8] (see [24] for a recent survey). In these works, it is posited that the state of mind is formally given by a quantum state, i.e., a vector in a Hilbert space, and all questions that can be answered within that state of mind are represented as operators in that Hilbert space.

We demonstrate here how representing data as a vector in a Hilbert space with observables represented by operators (matrices), can lead to a logarithmic reduction in the complexity of representation. This dramatic economy of representation may explain why evolution would select quantum cognition over classical statistical learning. As living creatures, we do not encounter the world as well-structured data relevant to a task at hand. Instead, we are confronted by a barrage of unstructured inputs that need to be made sense of, while focusing on what is important and drawing conclusions by abstracting away what is irrelevant.

Building upon these insights, we demonstrate a new practical form of machine learning, which we call Quantum Cognition Machine Learning (QCML). Our formulation naturally lends itself to implementation on quantum computing hardware, but is also easily implementable on classical hardware at lower Hilbert space dimensions. In our formulation, we define an error Hamiltonian as a sum of a loss function for each observable:

$$H(\mathbf{x}_t, \mathcal{L}, \{\hat{A}_k\}) = \sum_k \mathcal{L}(\hat{A}_k, \mathbf{x}_{t,k}) \tag{1}$$

where $\mathbf{x}_t$ is a data vector with K elements and is one of T total data vectors, $\mathcal{L}$ is a an arbitrary non-negative loss function with a Hermitian output, and $\{\hat{A}_k\}$ is a set of Hermitian observable operators which must be learned. We have flexibility in how we choose $\mathcal{L}$ so long as the result is non-negative and Hermitian, and can choose simple forms inspired by Gaussian loss, i.e.:

$$H(\mathbf{x}_t, \mathcal{L}, \{\hat{A}_k\}) = \sum_k (\hat{A}_k - \mathbf{x}_{t,k} \cdot I)^2 \tag{2}$$

or more complex functions for classification, the learning of context, or other objectives. This error Hamiltonian depends on the data itself, the parameters which comprise the set of observable operators $\{\hat{A}_k\}$, and the choice of loss function $\mathcal{L}$. Model tractability can be improved through careful choice of parameterization of our operators, and we have developed several variants of this.

We learn the operators $\{\hat{A}_k\}$ through the formalism of quasi-coherent states [17, 31, 9]. Recall that in quantum mechanics a state is a vector of unit norm in a Hilbert space, and is represented in bra-ket notation by a ket $|\psi\rangle$. The inner product of two states $|\psi_1\rangle$, $|\psi_2\rangle$ is represented by a bra-ket $\langle\psi_1|\psi_2\rangle$. The expectation value of a Hermitian operator O on a state $|\psi\rangle$ is denoted by $\langle\psi| O |\psi\rangle$, representing the expected outcome of the measurement corresponding to O on the state $|\psi\rangle$.

Training these models involves iterative updates to the ground state $|\psi_t\rangle$[1] of the error Hamiltonian and the observables $\hat{A}_k$ to reduce the ground state energy of H until desired convergence is reached. The specifics of each of these steps depend on the choice of the loss function and how we parameterize $\hat{A}_k$.

Model Training

- Randomly initialize parameters of $\{\hat{A}_k\}$
- Iterate over data and operators until desired convergence:
  1. Generate $H(\mathbf{x}_t, \mathcal{L}, \{\hat{A}_k\})$
  2. Holding $\hat{A}_k$ constant, find the ground state $|\psi_t\rangle$ of $H(\mathbf{x}_t, \mathcal{L}, \{\hat{A}_k\})$
  3. Holding $|\psi_t\rangle$ constant, calculate gradients of $H(\mathbf{x}_t, \mathcal{L}, \{\hat{A}_k\})$ w.r.t $\hat{A}_k$
  4. Update $\hat{A}_k$ via gradient descent

---

[1] The ground state $|\psi_t\rangle$ is the eigenstate associated to the lowest eigenvalue of the error Hamiltonian, and is also called the quasi-coherent state

Having suitably trained a set of operators $\hat{A}_k$, consider an arbitrary set of J inputs x_j, $J \subseteq K$, and the corresponding operators $\hat{A}_j$. After solving for the ground state $|\psi_J\rangle$ of the error Hamiltonian computed over this subset, $H(\mathbf{x}_j, \mathcal{L}, \{\hat{A}_j\})$, the expected value of any operator $B \in \hat{A}_k$ is computed as: $\langle\psi_J| B |\psi_J\rangle$. This allows us to forecast any of our K operators based on the quasi-coherent state computed over any subset of the operators[2].

4 QCML Applied to Financial Forecasting

4.1 Overview

We illustrate how QCML can be used to capture complex joint distributions for financial forecasting, while complementing existing Machine Learning techniques[3]. In order to build confidence in the QCML framework, we begin with a very simple example involving two classic features, Momentum and Value. There are many definitions of Value, here we use Revenues scaled by Market Cap, see section 4.2 below for details. Our goal is to train both a QCML model and a simple Neural Network (NN) to predict returns in excess of the linear returns to Momentum and Value, and to compare the output of the two approaches. We then move to a second example incorporating additional well-known features often used in equity forecasting and risk management.

4.2 Data and Features

We collect daily data on a dynamic set of roughly 1,500 US firms screened for the largest size, liquidity and maturity. We create the following commonly used features[4] using market data sourced from Bloomberg, accounting data sourced from S&P Capital IQ, and securities lending data sourced from S&P Securities Finance:

Accruals: Sign flipped 4 quarter change of (Total Assets - Working Capital - Total Liabilities - Long Term Investments + Long Term Debt), scaled by Total Assets, demeaned by GICS Industry Code and cross-sectionally normalized [27, 29]

Beta: Rolling 252 day time series estimated Beta to S&P 500 Index [28, 13, 12]

EBITDA to TEV: Prior 4 quarters Earnings Before Interest, Taxes, Depreciation, and Amortization, scaled by Total Enterprise Value, demeaned by GICS Industry Code and cross-sectionally normalized [21]

Momentum: Returns from 21 days ago to 252 days ago, cross-sectionally normalized [19, 10]

Operating Efficiency: Prior 4 quarters Revenues, scaled by Total Assets, demeaned by GICS Industry Code and cross-sectionally normalized [30]

Profit Margin: Prior 4 quarters Net Income, scaled by prior 4 quarters Revenues, demeaned by GICS Industry Code and cross-sectionally normalized [30]

Short Utilization: Sign flipped Lender Value on Loan averaged with a 10 day half-life, scaled by Active Lendable Value averaged with a 10 day half-life, cross-sectionally normalized [26, 25]

Size: Log of Market Cap, cross-sectionally normalized [5, 12]

Value: Prior 4 quarters Revenues, scaled by Market Cap, demeaned by GICS Industry Code and cross-sectionally normalized [6]

GICS Dummies: Dummy variables formed based on GICS Industry Group [4, 3]

[2] Note this also provides a natural way for dealing with missing data: as J can be any subset of K, any missing inputs can simply be excluded from the error Hamiltonian calculated for forecasting, rather than being pre-filled in some manner. These missing inputs could also themselves be forecast given $|\psi_J\rangle$

[3] We do not make any claims of superiority in these examples, but wish to show the framework is efficacious and unique

[4] [15] uses their previously proposed q-factor model to test the robustness of various features proposed in the financial literature, and also serves as an excellent survey of proposed features / anomalies

4.3 Model Setup and Training

Models are trained using daily data from January 2008 through August 2013. While both NN and QCML models can be updated online or through rolling / expanding retraining, for simplicity we keep parameters for both models static after initial training.

Our target variable for purposes of model training is 15 day forward log returns, projected[5] away from model input features as well as Beta, Size, and GICS Dummies. After projection, target returns are cross-sectionally normalized. Although forecasts learned from models using such a residualized target variable will likely have small average correlation with all of the features that target returns have been projected away from (which includes the input features), to focus on performance after controlling for linear effects, we project the final aggregate forecasts away from the relevant input and control features.

To create more robust forecasts, we partition stocks into randomized groups of approximately 50, training individual NN and QCML models over each subset, and average forecasts for each stock across 100 different such partitions.

The NN architecture and training approach we use adheres fairly closely to that recommended in [14], other than our more complex approach to ensembling. Our neural network is implemented in PyTorch and contains 3 hidden layers of 32, 16 and 8 nodes respectively, with batch normalization [16] applied prior to ReLU activation [18, 23]. We use a simple mean squared error loss function, and train using stochastic gradient descent and the Adam optimizer.

Our QCML model uses a Hilbert space dimensionality of 32, and operator representation and training techniques proprietary to Qognitive, Inc., but following the general gradient descent approach outlined in section 3, implemented on classical hardware.

4.4 Model Evaluation

To test performance, we use covariance estimated daily from daily returns[6] to produce Markowitz optimal investment portfolios, where portfolio weight $w = V^{-1}f$ given covariance V and forecast f[7]. Given that the forecasts have been projected away from input and control features, these investment portfolios will also have no exposure to input and control features.[8]

Investment portfolios are smoothed over 15 days to proxy for the fact that realistic investment portfolios must trade into new forecasts gradually. We compute daily returns to the smoothed portfolios, and analyze portfolio performance from September 2013 through March 2024.

We do not remove transaction costs as these forecasts are not intended to represent a standalone investment strategy, but rather, as forecasts capturing complex attributes of the joint distribution which could be additive to an existing investment strategy.

4.5 Momentum / Value Model

For our first example, the model input features are Momentum and Value as defined in section 4.2. Thus, target returns and final forecasts are projected away from Momentum, Value, Beta, Size and GICS Dummies.

In Table 1 we show the Sharpe ratios of each of the forecasts, as well as the Sharpe ratios of an equal-risk weighted combination of the two.[9] We find that both the NN and QCML models produce moderately positive signals, which are in excess of the linear returns to the inputs. However, correlation of the return streams generated by the two approaches is only 0.41 over the full test

[5] If you wish to solve for portfolio weights which maximize expected returns, with a penalty for expected portfolio variance, while maintaining zero exposure to a set of controls, then for weights w, forecast f, asset variance V, risk aversion μ, and controls M, you need to solve for w which minimizes $-w^\intercal f + 0.5\mu w^\intercal V w$ such that $w^\intercal M = 0$. The solution is $w = V^{-1}Rf$ where $R = I - M(M^\intercal V^{-1}M)^{-1}M^\intercal V^{-1}$. Thus R is a projection operator that projects away from M, consistent with our desired investment process. Rf is also the residual from an inverse variance weighted regression of f on M

[6] Covariance is estimated using a technique proprietary to Duality Group

[7] Note that forecasts here have already been projected away from inputs and controls as explained in section 4.3 and footnote 5

[8] Given a matrix of input and control features M, we have $f^\intercal M = 0$ and $w^\intercal M = 0$

[9] The risk scalar for each strategy is a simple expanding window sample standard deviation of forecast returns, lagged two days

Table 1: Sharpe Ratios of returns to the NN and QCML forecasts for the Momentum / Value models. Strategy portfolios have zero exposure to Momentum, Value, Beta, Size, and GICS Industry Groups, and thus no linear contribution of those features to returns.

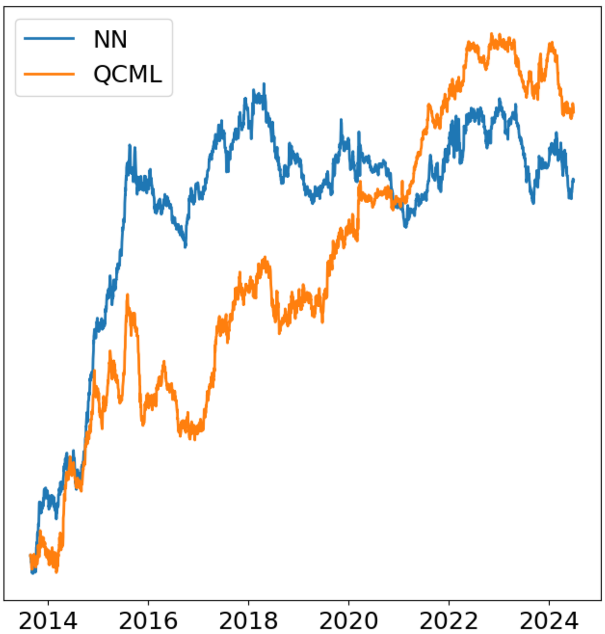
Period	NN	QCML	Combination
Sep 2013 - Jun 2024	0.58	0.69	0.75
Sep 2013 - Apr 2017	1.91	0.80	1.61
May 2017 - Dec 2020	-0.30	0.73	0.27
Jan 2021 - Jun 2024	0.11	0.53	0.33

period, and we see an equal-risk weighted combination of the two approaches outperforms either individual approach.

We plot the cumulative returns to the NN and QCML forecasts in Figure 1. From a visual examination of the performance, it is clear the models are picking up on similar, but distinct, underlying patterns.

As these simple models have only 2 inputs and a single forecast, we can also plot surfaces of the forecasts as a function of the inputs, to visually compare what the NN and QCML models learn. We plot the QCML and NN forecasts relative to Value and Momentum in Figure 2.

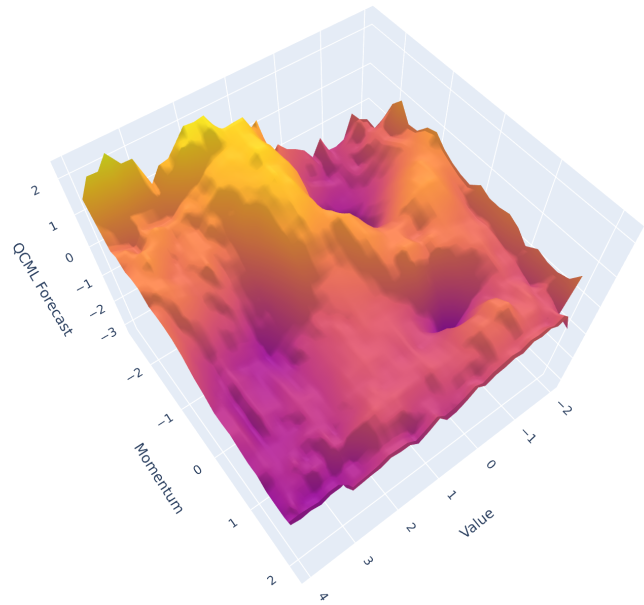
Figure 1: Cumulative returns to the NN and QCML forecasts for the Momentum / Value models. Strategy portfolios have zero exposure to Momentum, Value, Beta, Size, and GICS Industry Groups, and thus no linear contribution of those features to returns. From a visual examination of the performance, it is clear the models are picking up on similar, but distinct, underlying patterns. The correlation of realized returns is 0.41. Returns have been scaled by full sample realized volatility.

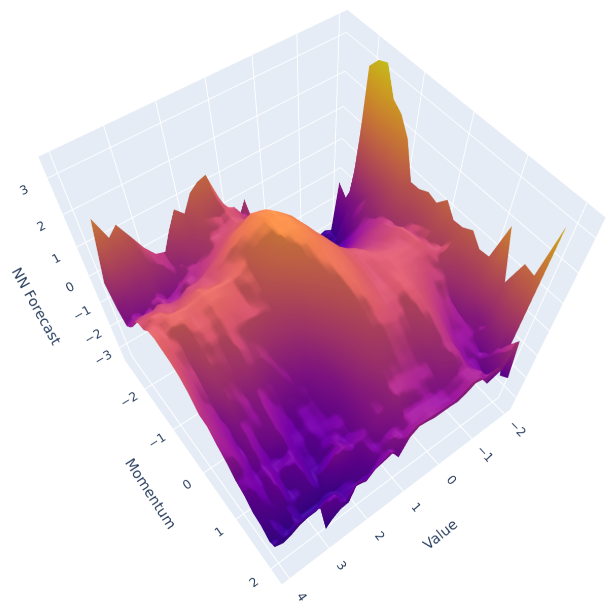


We notice some similarities between the surfaces - both models predict strong returns for moderately high Value / moderately low Momentum stocks and poor returns for high Momentum / high Value stocks as well as low Momentum / low Value stocks. In both cases we could interpret the models as having learned that Value is much stronger among low Momentum (loser) stocks, and that the Momentum effect is weaker among high Value stocks while being stronger among low Value (expensive) stocks.

This is consistent with previously documented findings for patterns seen in the interaction of Value and Momentum stocks. For example, see [2], which documents a similar pattern despite using a different Value metric (Book Value scaled by Market Cap) and having a data sample independent in time relative to our training sample.

Figure 2: QCML (upper) and NN (lower) Momentum / Value Model Forecasts as a function of Momentum and Value





Of course, the details of the NN and QCML forecast surfaces differ, and both approaches likely learn an effect which is noisy relative to the truth. However, the fact that QCML learns something both unique and reasonable when compared to a NN helps us to build confidence in QCML as a forecasting technqiue. Additionally, as the number of features greatly increases, QCML would have advantages as it does not suffer from the curse of dimensionality.

4.6 Extended Model

We now extend to the more interesting case of several input features. For these models, in addition to Momentum and Value, our input features include Accruals, EBITDA to TEV, Operating Efficiency, Profit Margin, Short Utilization, and Size. As before, target returns and final forecasts are projected away from inputs and controls, which in this case is all of the features defined in 4.2.

In Table 2 we show the Sharpe ratios of each of the forecasts, as well as the Sharpe ratios of an equal-risk weighted combination of the two (formed as in the prior section).

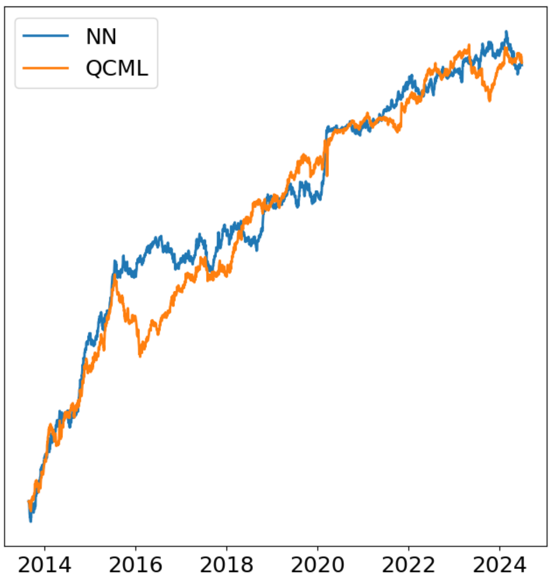
In this case, both the NN and QCML models produce strongly positive signals, which are in excess of the linear returns to the inputs. Additional, correlation of the return streams generated by the two approaches is even lower than in the previous example, 0.32 over the full test period, and we see an equal-risk weighted combination of the two approaches outperforms either individual approach.

We plot the cumulative returns to the NN and QCML forecasts in Figure 3.

Table 2: Sharpe Ratios of returns to the NN and QCML forecasts for the Extended models. Strategy portfolios have zero exposure to all input and control features (all features defined in section 4.2), and thus no linear contribution of those features to returns.

Period	NN	QCML	Combination
Sep 2013 - Jun 2024	1.25	1.25	1.39
Sep 2013 - Apr 2017	1.91	1.72	1.81
May 2017 - Dec 2020	1.05	1.26	1.48
Jan 2021 - Jun 2024	0.64	0.63	0.78

Figure 3: Cumulative returns to the NN and QCML forecasts for the Extended models. Strategy portfolios have zero exposure to all input and control features (all features defined in section 4.2), and thus no linear contribution of those features to returns. The correlation of realized returns is 0.32. Returns have been scaled by full sample realized volatility.



4.7 Combining Nonlinear and Linear Models

We next demonstrate the ability of QCML forecasts to improve upon linear forecasts, beyond the improvement seen from NN forecasts. We do this by taking equal-risk weighted combinations of linear forecasts with the non-linear NN and QCML forecasts.

For each previously discussed model (Momentum / Value and Extended models), we produce linear forecasts from all input features except Size, which is maintained as a control. We produce our linear forecast by taking the linear features and projecting them away from Beta, Size, and GICS Dummies. This is not a statement regarding which linear features should be included in a forecast, how to best form hedged portfolios of linear features, or how to appropriately weight a diverse set of forecasts. It is solely meant to provide a simple demonstration of the ability of QCML forecasts to improve linear models. We show the Sharpe ratios of each of the linear forecasts in table 3.

In Table 4 we show the Sharpe Ratios of returns to equal-risk combined linear forecasts for Momentum and Value features only, those linear forecasts with the NN Momentum / Value forecasts added, those linear forecasts with the QCML Momentum / Value forecasts added, and all four strategies (Momentum, Value, Momentum / Value NN, and Momentum / Value QCML). Both the NN and QCML forecasts improve the Sharpe Ratio of the linear forecasts alone, with a slightly larger improvement seen from adding the QCML forecast.

In Table 5 we show the Sharpe Ratios of returns to equal-risk combined linear forecasts for all features, those linear forecasts with NN Extended Model forecasts added, those linear forecasts with QCML Extended Model forecasts added, and all nine strategies (seven linear forecasts, Extended NN, and Extended QCML). Again, both the NN and QCML forecasts improve the Sharpe Ratio of the linear forecasts alone, with a slightly larger improvement seen from adding the QCML forecast. Here, the best Sharpe Ratio is achieved by using both the NN and QCML forecasts.

We plot the cumulative returns to the linear and enhanced forecasts for Momentum / Value models in Figure 4. We plot the cumulative returns to the linear and enhanced forecasts for Extended models in Figure 5.

Table 3: Sharpe Ratios of returns to linear forecasts (projected away from Beta, Size, and GICS Dummies).

Period	Accruals	EBITDA to TEV	Momentum	Operating Efficiency	Profit Margin	Short Utilization	Value
Sep 2013 - Jun 2024	0.29	0.04	0.65	1.36	0.27	0.93	0.25
Sep 2013 - Apr 2017	0.51	0.86	0.96	1.56	0.85	1.90	1.05
May 2017 - Dec 2020	-0.30	-1.48	0.47	1.12	-0.24	-0.29	-0.92
Jan 2021 - Jun 2024	0.84	1.19	0.53	1.42	0.24	1.30	1.15

Table 4: Sharpe Ratios of returns to equal-risk combined linear forecasts for Momentum and Value, those linear forecasts with NN Momentum / Value forecasts added, those linear forecasts with QCML Momentum / Value forecasts added, and all Momentum / Value strategies. The QCML forecast provides better diversification than the NN forecast.

Period	Linear Strategies	Linear Strategies + NN	Linear Strategies + QCML	All Strategies
Sep 2013 - Jun 2024	0.91	0.99	1.11	1.07
Sep 2013 - Apr 2017	1.84	2.46	1.87	2.25
May 2017 - Dec 2020	-0.72	-0.71	0.02	-0.13
Jan 2021 - Jun 2024	1.54	0.99	1.42	0.97

Figure 4: Cumulative returns to equal-risk combined linear forecasts for Momentum and Value only, those linear forecasts with NN Momentum / Value forecasts added, those linear forecasts with QCML Momentum / Value forecasts added, and all Momentum / Value strategies. Returns have been scaled by full sample realized volatility.

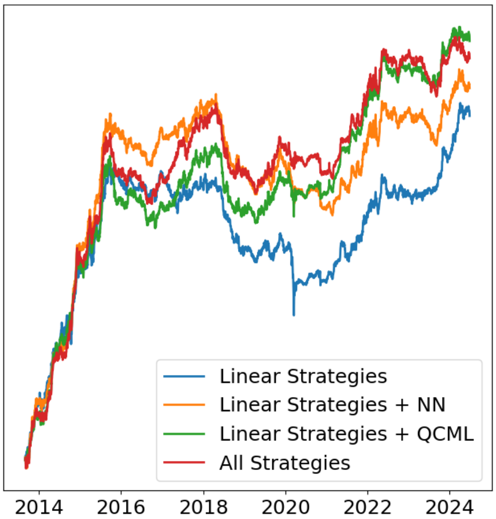
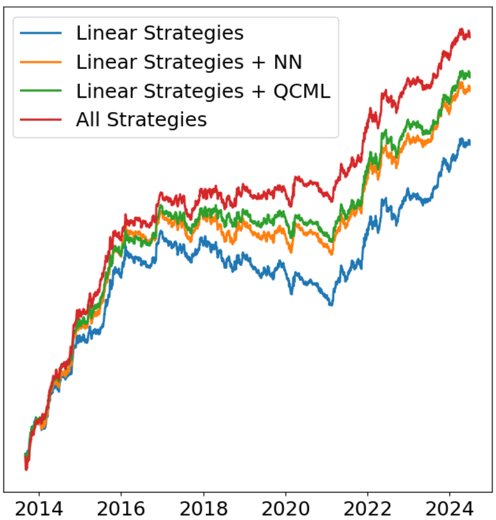


Table 5: Sharpe Ratios of returns to equal-risk combined linear forecasts for all features, those linear forecasts with NN Extended Model forecasts added, those linear forecasts with QCML Extended Model forecasts added, and all Extended strategies. The QCML forecast provides better diversification than the NN forecast, but the best Sharpe Ratio is achieved by using both the NN and QCML forecasts.

Period	**Linear Strategies**	**Linear Strategies + NN**	**Linear Strategies + QCML**	**All Strategies**
Sep 2013 - Jun 2024	1.17	1.38	1.43	1.58
Sep 2013 - Apr 2017	2.31	2.57	2.66	2.82
May 2017 - Dec 2020	-0.66	-0.32	-0.26	0.03
Jan 2021 - Jun 2024	1.77	1.81	1.82	1.83

Figure 5: Cumulative returns to equal-risk combined linear forecasts for all features, those linear forecasts with NN Extended Model forecasts added, those linear forecasts with QCML Extended Model forecasts added, and all Extended strategies. Returns have been scaled by full sample realized volatility.



5 Conclusion

We developed and demonstrated the use of a new machine learning paradigm using principles of quantum cognition, which we call QCML. In a world with an immense proliferation of data sets, the need to address large feature sets paired with small observation sets will become more and more pressing. QCML achieves logarithmic economy of data representation, making it well suited to meet this challenge, as well as the challenge of non-stationarity which reduces the amount of relevant data available.

With classic statistical learning, every time you add a variable to a model the uncertainty grows. This leads to difficulty or instability in forming statistical estimates, and creates a requirement for judicious choices of model variables. In quantum systems, by contrast, the uncertainty of the whole system can be less than the uncertainty of the components ([22] and references therein).

Here, we demonstrated an application of QCML to forecasting stock returns in excess of the linear contribution to returns of the input features. We benchmarked forecasts produced by QCML against forecasts produced by a neural network and showed that, even for relatively simple systems and a classical hardware implementation, QCML can offer an advantage over or complement the forecasts produced by the neural network.

In future work, we plan to continue studying the properties and applications of QCML. [9] shows one such research direction, where the authors extend QCML to apply it to manifold learning, specifically to the estimation of intrinsic dimension of data sets, demonstrating the practicality of the proposed method on synthetic manifold benchmarks as well as real data sets. We also are exploring ways to integrate QCML with existing statistical techniques, and working towards a practical implementation on quantum hardware.

References

[1] Diedrik Aerts and Sven Aerts. Applications of quantum statistics in psychological studies of decision processes. *Foundations of Science*, 1(1):85–97, 1995.

[2] Clifford S. Asness. The interaction of value and momentum strategies. *Financial Analysts Journal*, 53(2):29–36, 1997.

[3] Clifford S. Asness, Andrea Frazzini, and Lasse H. Pedersen. Low-risk investing without industry bets. *Financial Analysts Journal*, 70(4):24–41, 2014.

[4] Clifford S. Asness, Roger B. Porter, and Ross L. Stevens. Predicting stock returns using industry-relative firm characteristics. *Capital Markets eJournal*, 2000.

[5] Rolf W. Banz. The relationship between return and market value of common stocks. *Journal of Financial Economics*, 9(1):3–18, 1981.

[6] William C. Barbee, Sandip Mukherji, and Gary A. Raines. Do sales–price and debt–equity explain stock returns better than book–market and firm size? *Financial Analysts Journal*, 52:56–60, 1996.

[7] Paul Burchard. Linkedin posts, 2023-2024. `https://www.linkedin.com/in/paulburchard`.

[8] Jerome Busemeyer and Peter D. Bruza. *Quantum Models of Cognition and Decision*. Cambridge University Press, 2012.

[9] Luca Candelori et al. Robust estimation of the intrinsic dimension of data sets with quantum cognition machine learning, 2024. `https://arxiv.org/pdf/2409.12805`.

[10] Mark Carhart. On persistence in mutual fund performance. *Journal of Finance*, 52(1):57–82, 1997.

[11] Anna Dawid and Yann LeCun. Introduction to latent variable energy-based models: A path towards autonomous machine intelligence. *Les Houches Summer School Lecture Notes*, 2022.

[12] Eugene Fama and Kenneth French. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56, 1993.

[13] Eugene F. Fama and James D. MacBeth. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81:607 – 636, 1973.

[14] Shihao Gu, Bryan Kelly, and Dacheng Xiu. Empirical Asset Pricing via Machine Learning. *The Review of Financial Studies*, 33(5):2223–2273, 02 2020.

[15] Kewei Hou, Chen Xue, and Lu Zhang. Replicating anomalies. *The Review of Financial Studies*, 33(5):2019–2133, 2020.

[16] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift, 2015.

[17] Goro Ishiki. Matrix geometry and coherent states. *Physical Review D*, 92(4), August 2015.

[18] Kevin Jarrett, Koray Kavukcuoglu, Marc'Aurelio Ranzato, and Yann LeCun. What is the best multi-stage architecture for object recognition? *2009 IEEE 12th International Conference on Computer Vision*, pages 2146–2153, 2009.

[19] Narasimhan Jegadeesh and Sheridan Titman. Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance*, 48(1):65–91, 1993.

[20] Andrei Khrennikov. Quantum-like brain: "interference of minds". *Biosystems*, 84(3):225–241, 2006.

[21] Tim Loughran and Jay W. Wellman. New evidence on the relation between the enterprise multiple and average stock returns. *Journal of Financial and Quantitative Analysis*, 46(6):1629–1650, 2011.

[22] Kharen Musaelian et al. Quantum cognition machine learning: AI needs quantum, 2024. https://www.qognitive.io/QCML%20-%20Qognitive,%20Inc.pdf.

[23] Vinod Nair and Geoffrey E. Hinton. Rectified linear units improve restricted boltzmann machines. In *International Conference on Machine Learning*, 2010.

[24] Emmanuel M. Pothos and Jerome R. Busemeyer. Quantum cognition. *Annual Review of Psychology*, 73(1):749–778, 2022.

[25] Deutsche Bank Global Markets Research. Signal processing: The long and the short of it, January 2011. https://cdn.ihsmarkit.com/www/pdf/0720/DB-Long-and-Short-Of-It.pdf.

[26] Macquarie Equities Research. Quantamentals: Refining short interest, March 2010. https://cdn.ihsmarkit.com/www/pdf/0720/MacQuarie-Quant-Research---Refining-Short-Interest.pdf.

[27] Scott A. Richardson, Richard G. Sloan, Mark T. Soliman, and Irem Tuna. Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics*, 39(3):437–485, 2005.

[28] William Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19(3):425–442, 1964.

[29] Richard G. Sloan. Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review*, 71(3):289–315, July 1996.

[30] Mark T. Soliman. The use of dupont analysis by market participants. *Corporate Finance: Valuation*, 2007.

[31] Harold C. Steinacker. Quantum (matrix) geometry and quasi-coherent states. *Journal of Physics A: Mathematical and Theoretical*, 54(5):055401, January 2021.