FIRE: THE FIRST-YEAR INNOVATION & RESEARCH EXPERIENCE

QUANTUM MACHINE LEARNING

Quantum-Enhanced Forecasting: Leveraging Quantum Gramian Angular Field And CNNs for Stock Return Predictions

A. Bapanapalli, A. Mahmud, C. Nadig, K. Thakker, Dr. S. Jabeen

Introduction - QML

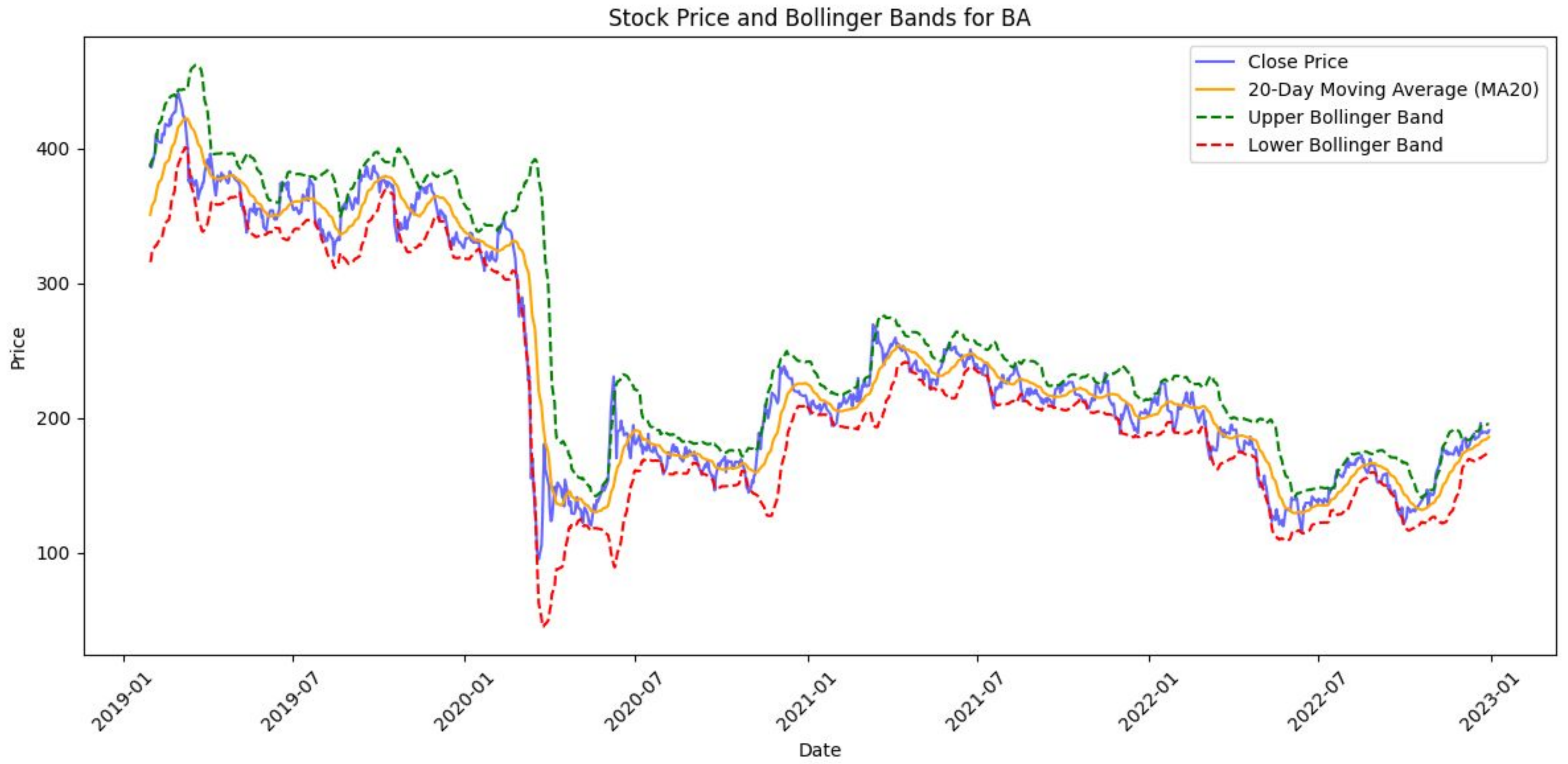
As of recent years, both Machine Learning and Quantum Computing have been rapidly growing, giving rise to the emerging field of Quantum Machine Learning (QML). QML aims to exploit the unique properties of quantum mechanics, such as superposition and entanglement, to enhance the performance of machine learning algorithms. QML algorithms can potentially offer exponential speedup over classical algorithms, enabling the efficient processing of large datasets and the discovery of patterns that may be otherwise hidden. However, despite the promise of QML, its implementation faces significant challenges, such as the limited availability of quantum hardware and the difficulty in training quantum models. Nevertheless, with the rapid advancement of quantum technologies, QML is expected to have a transformative impact on the field of machine learning in the future.

Portfolio Dataset and CNNs

We began by replicating a study using stock data from the Hong Kong Stock Exchange (HKEX) to validate the use of Gramian Angular Fields (GAFs) for capturing trends. This proof of concept guided the creation of a dataset comprising ~75,000 Quantum Gramian Angular Field (QGAF) images from 20 stocks, sourced from Yahoo Finance. The dataset spans diverse market behaviors and temporal patterns, with stock trends labeled as up or down based on price movements.

Time-series stock data, segmented into 50-day overlapping windows with a 1-day step, underwent preprocessing to remove anomalies and ensure integrity. Features such as moving averages and Bollinger Bands were transformed into QGASF and QGADF matrices using quantum circuits and visualized as images. Normalization and feature scaling standardized inputs for use in a CNN, which incorporates both features for refined predictive analysis.

Figure 1: Stock Features on Dataset Graph



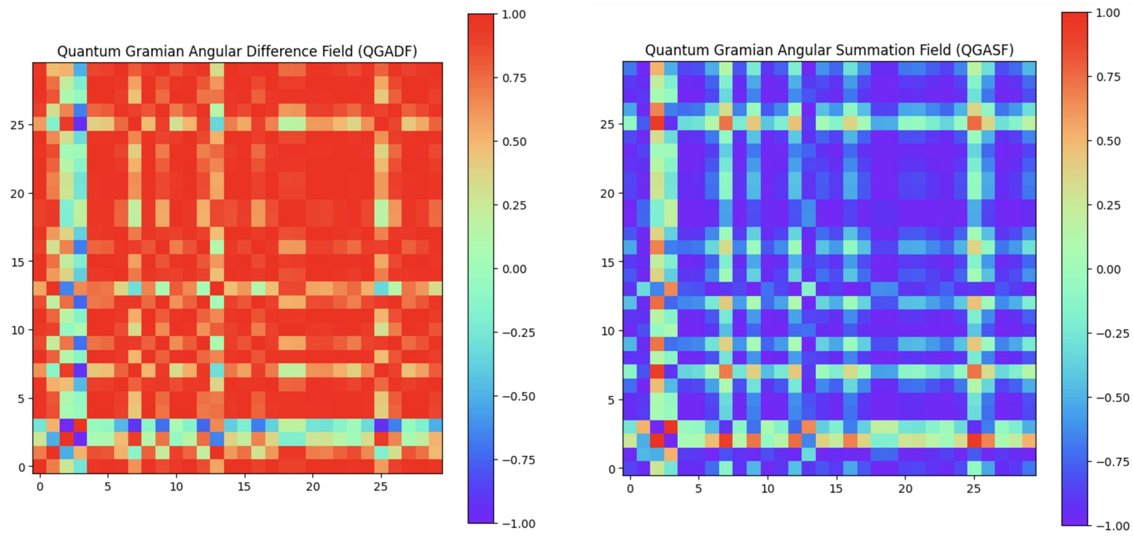


Figure 2: Quantum Gramian Angular Difference Field and Quantum Gramian Angular Summation Field on the Left

CNN Architecture and Pipeline

Our approach employs a conventional CNN architecture optimized for image classification, consisting of convolutional layers for feature extraction, pooling layers for dimensionality reduction, and fully connected layers for classification. Hyperparameters such as filter size, stride, and layer count were carefully tuned to enhance performance on GAF data. A unique feature of our method is quantum data preprocessing, where quantum gates normalize stock time-series data before transformation into images. This step encodes features like price and volatility while capturing subtle temporal relationships often missed by traditional techniques. The resulting Quantum Gramian Angular Fields (QGAFs) extend the concept of GAFs by leveraging quantum normalization to enrich data representation, improving the CNN's ability to identify intricate patterns in stock trends.

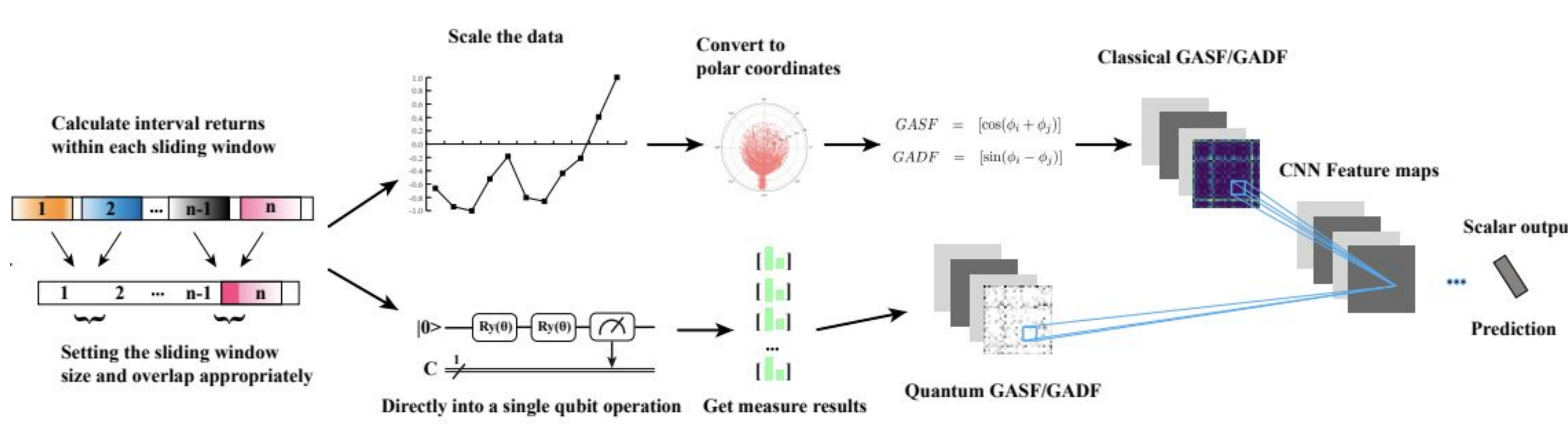


Figure 2: Visual Description of the Classical and Quantum Data Pipeline (Xu et al., 2023)

After retrieving time series data from specific stocks, the data is fed into the quantum circuit shown above, where two rotational gates are applied to try and assume the same effect as calculating cos(a+b) for the first circuit, and sin(a-b) for the second. The computed trigonometric functions provide insights into the angular relationships within the data, which are then translated into pixel intensities to construct the QGAF images. By running these circuits multiple times and averaging the results, we can improve the accuracy of our calculations. These circuits help to identify relationships in time series data and are essential in creating QGAF images.

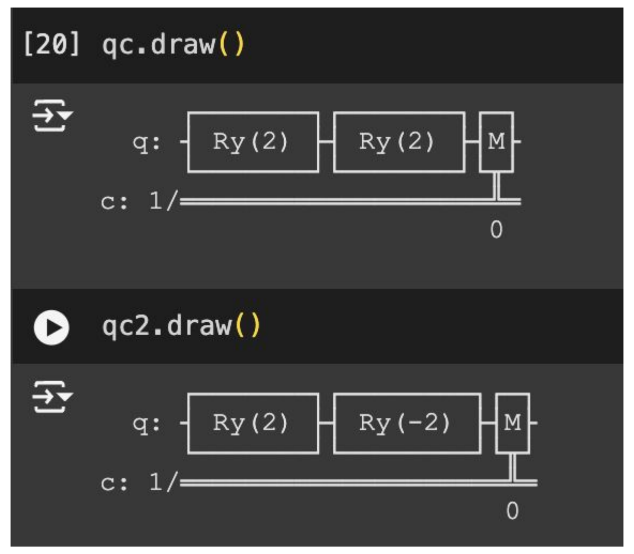


Figure 3: Quantum Circuit with 2 Rotational Gates

QGAF vs GAF

The primary distinction between a Gramian Angular Field (GAF) and a Quantum Gramian Angular Field (QGAF) lies in the preprocessing of the input data. GAFs transform time-series data into angular information, encoding temporal dependencies and relationships as images that are spatially structured for input into a CNN. This is achieved through a polar coordinate transformation of normalized data, preserving chronological order in a matrix representation.

QGAFs enhance traditional GAFs by applying quantum normalization techniques, encoding subtle price and volatility variations. This enriches data representation, enabling QGAFs to capture intricate patterns and provide more robust inputs for machine learning in complex tasks like financial trend analysis.

Procedure and Results

The simulations are conducted using Qiskit's simulator backend to generate QGAF images. Time-series stock price data are broken into segments. These segments' vector representations are fed into quantum circuits that generate QGAF images. The features include the moving average and bollinger bands. Our CNN implements both features simultaneously for the interest of producing more refined results.

Our model achieved an accuracy rating of 70% for training and a validation accuracy of 47%. On top of that, we observed a consistent decrease in training and validation loss over epochs, indicating effective learning. We also ensured the dataset was balanced between up and down labels. The results highlight the potential and limitations of our model. While it achieves solid training accuracy, the lower validation accuracy suggests room for improvement in generalization. These findings validate the use of QGAF transformations for financial forecasting but point to the need for larger datasets, improved architectures, or real Quantum Processing Units to enhance robustness and predictive power.

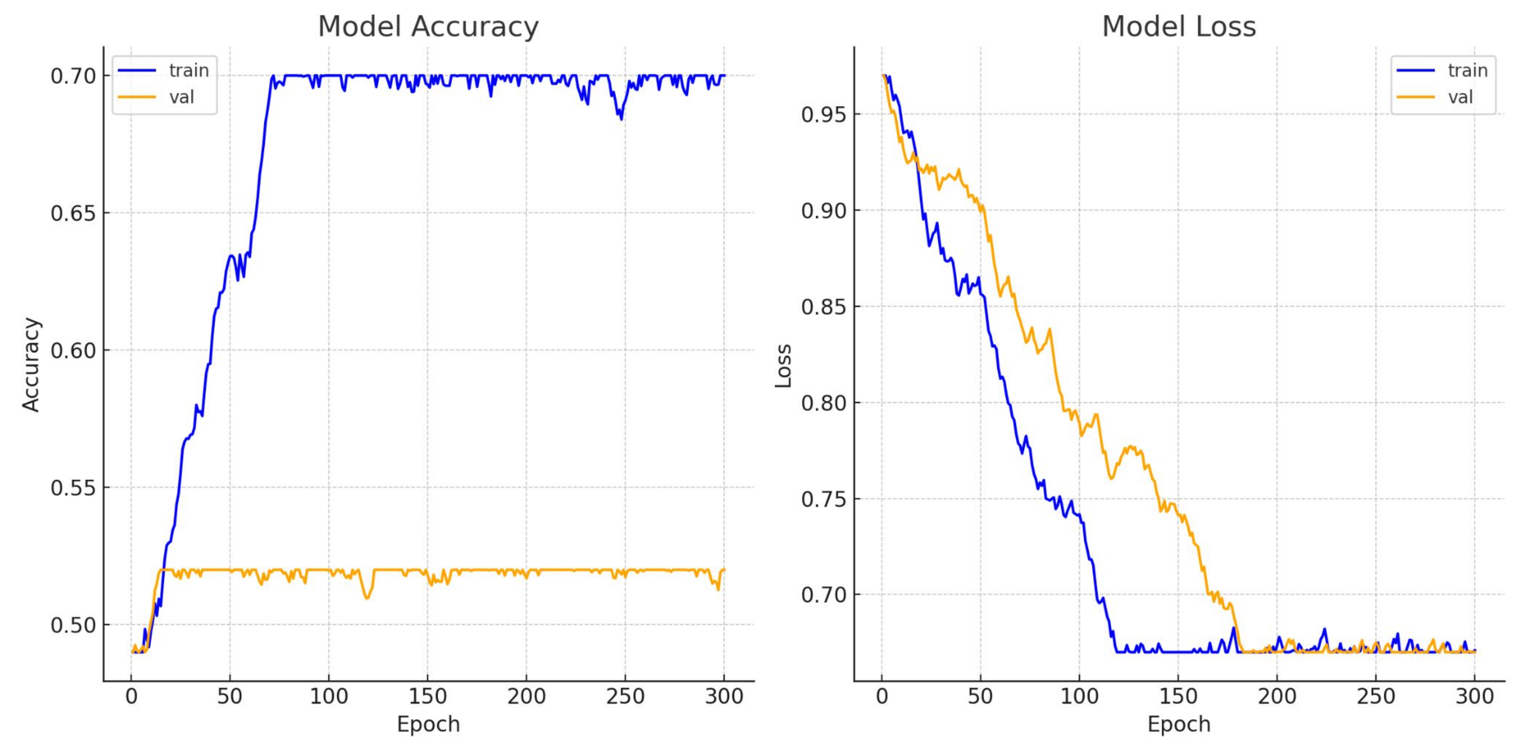


Figure 4: Model Accuracy and Model Loss Graphs During CNN Training

Future Outlook

In summary, we developed a multi-input CNN model using QGAF images to predict stock price movements, highlighting the potential of quantum machine learning in financial forecasting. QGAFs provide unique data representations that could yield deeper insights as quantum hardware advances. Future improvements include incorporating more tickers, extending the date range for enhanced robustness, transitioning to real Quantum Processing Units (e.g., IONQ) for performance evaluation, and exploring advanced neural architectures like QCNNs or pre-trained models (e.g., ResNet, EfficientNet).

References

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv.

https://arxiv.org/abs/1810.03787

Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. arXiv. https://arxiv.org/abs/1506.00327

LurreMcFly. (n.d.). ERP prediction contest. GitHub. https://github.com/LurreMcFly/erp_prediction_contest/tree/master

Schwab. (n.d.). Bollinger Bands®: What they are and how to use them. Charles Schwab.

https://www.schwab.com/learn/story/bollinger-bands-what-they-are-and-how-to-use-them#:~:text=Bollinger%20Bands%C2%AE%20help%20you,potential%20entry%20and%20exit%20points

Spall, J. C. (1998). An overview of the simultaneous perturbation method for efficient optimization. Johns Hopkins University Applied Physics Laboratory. https://www.jhuapl.edu/SPSA/PDF-SPSA/Spall_An_Overview.PDF

Tao, J., Chen, W., Gou, Z., & Zhang, Y. (2023). Quantum-enhanced forecasting for financial market prediction. arXiv. https://arxiv.org/abs/2310.07427