IJEMD-CSAI, 3 (1) (2024), 1 – 9 https://doi.org/ 10.54938/ijemdcsai.2024.03.1.319

International Journal of Emerging Multidisciplinaries: Computer Science and Artificial Intelligence

Research Paper

Journal Homepage: www.ojs.ijemd.com ISSN (print): 2791-0164 ISSN (online): 2957-5036

Quantum Variational Autoencoders for Predictive Analytics in High Frequency Trading Enhancing Market Anomaly Detection

Jamshaid Basit[1*], Danish Hanif[1] and Madiha Arshad[2]

1. Department of Computer Science and Software Engineering, National University of Sciences and Technology , Islamabad, Pakistan
2. Comsats University Islamabad(CUI), Islamabad, Pakistan

** Corresponding Author.*

Abstract

High-frequency trading (HFT) markets, characterized by high and frequent price fluctuations, necessitate the use of anomaly detection mechanisms to monitor the market and ensure the efficacy of the trading system. This paper aims to discuss the possibility of improving predictive analytics in HFT using quantum computing with the help of the Quantum Variational Autoencoder (QL-VAE). As a result, we propose a new direction for further research on quantum VAEs in HFT that involves their direct comparison with classical VAEs. The application of quantum models for mastering the intensive data flow of HFT is conditioned by the advantages of quantum computation in comparison to classical ones, which are more suitable for handling multidimensional data arrangements and intricate topologies. Our detailed study methodology involved examining various aspects of HFT data, such as order book features and stock price characteristics. We normalized all the data and reduced some of its dimensions. We established quantum VAEs using Pennylane, and configured the classical VAEs using TensorFlow. When it comes to market anomalies, the results of the comparative analysis showed higher accuracy, recall, and F1 rate in quantum VAEs compared to classical models when it comes to the analysis of market anomalies. Therefore, the quantum model's ability to handle high-dimensional data makes it a better fit for HFT than classical methods. These studies suggest that quantum VAEs could significantly improve anomaly detection in the financial market.

Keywords: High-Frequency Trading (HFT), Quantum Variational Autoencoder (QL-VAE), Anomaly Detection, Predictive Analytics, Quantum Computing

Email Addresses: jbasit.msse23mcs@student.nust.edu.pk (Jamshaid) , dhanif.msse23mcs@student.nust.edu.pk (Danish) , madiha.arshad7@gmail.com (Madiha).

1. Introduction

High-frequency trading (HFT) is one of the most important innovations in the financial markets, according to the definition, which entails fast trade in information and computational techniques [1]. It is important to realize that each market raises its level of exposure to potentially catastrophic financial loss and betrays the signs of its credibility when it is impossible to detect any symptoms of an anomaly at the preliminary stage [2].Aberrations within HFT could mean fraud, manipulation, or even a drastic change of attitude among the traders [3].

Conventional methods for detecting anomalies in HFT have not kept up with the increasing complexity and size of data, necessitating the exploration of alternative techniques [4].

The anomaly detection in HFT is difficult, especially when dealing with high dimensional features and a rapidly changing market environment [2]. Typically, we use deep learning techniques and conventional machine learning models to analyze high-frequency data, but these models often struggle to capture the dynamics of the data [5]. Traditional approaches face additional challenges due to the large amount of trading data and the need for real-time data processing [6].

The new developments in machine learning and quantum computing could be useful in proposing new solutions to improve anomaly detection in HFT [7]. People have widely used generative models, especially variational autoencoders (VAEs), for anomaly detection [8]. Research has proven that AEs can grasp intricate non-linear data distributions and identify anomalous regions [9]. However, the current high-frequency trading environment has not extensively tested them, and their efficiency remains an important and ongoing research question [10].

Machine learning gets dozens of new opportunities and improvements due to the appearance of quantum computing in the field. Quantum Variational Autoencoders, also known as Quantum VAEs, utilize quantum approaches to handle high-dimensional data in a more efficient manner than their classical counterparts [11]. This development has the potential to improve anomaly detection performance because it seeks to correct the failings of classical frameworks [10].

A review of recent research on the subject reveals several methods of identifying anomalies in HFT. Traditionally, we used statistical tools and computational algorithms to detect trading anomalies. Nevertheless, these approaches are incongruous with the characteristics of the high-dimensional and dynamic HFT data [2]. Deep learning techniques (VAEs) have enhanced HFT scenarios, but more effort is required to improve anomaly detection performance [4].

In recent years, quantum computing has advanced, providing a new opportunity to improve predictive analytics [5]. In general, quantum computing algorithms show the possibility for faster data processing velocity and better predictive models for changing conditions and new patterns, which would significantly increase the chances of identifying an anomaly [12]. As a result, quantum VAEs are being considered a new, promising solution to the problems associated with the use of classical VAEs in HFT applications.

Quantum VAEs for anomaly detection are a relatively new concept, with only a few articles exploring its application in the financial markets. Previous investigations suggest that the use of quantum VAEs may be superior to classical ones, as they can leverage quantum parallelism and entanglement to process more complex data structures effectively [4]. provides a promising platform for innovative breakthroughs in financial analysis.

This work shows that the benefits of extending quantum VAEs are not merely theoretical but quite tangible [7]. Many of these models leverage advanced quantum computing to analyze big data more proficiently than other computing systems.

This efficiency is very important in high-frequency trading, where the faster and more efficient recognition of anomalies enhances the chances of having better trades and risk management.

The hierarchy of research in this domain typically involves several key stages: first, the inability of classical anomaly detection methods in HFT; second, the applicability of VAEs for anomalies; third, the possibilities of quantum VAEs over classical; and last, applying quantum VAEs in an actual HFT environment.

The use of quantum VAEs to enhance anomaly detection plans makes it easier for organizations to consider better analytical tools that are more efficient. In the future development of the field of quantum computing, the findings of using quantum VAEs can actually establish new standards for the detection of anomalies in high-frequency trading and other complicated financial systems [7].

This paper seeks to fill the existing literature void on anomaly detection techniques in high-frequency trading by comparing Quantum VAEs with the standard VAEs. In this paper, we discuss the application of QML for high-frequency trading and use the models' performance on the HFT data set to show that quantum methods outperform classical ones. This research contributes by utilizing a relatively new architecture, Quantum VAEs, in a field where quantum computing for anomaly detection in financial markets has not been extensively explored.

The aim of this research is to enhance the predictive-analytics model of HFT through the application of Quantum VAE, thereby improving anomaly detection, and to compare it with classical VAE methods. The aim of this research is to enhance the field of financial analytics by integrating quantum computing into one of the anomaly detection techniques, potentially contributing to the advancement of methods based on quantum machine learning in the finance domain.

2. Literature Review

High-frequency trading (HFT) has become a fundamental component of financial market innovation due to advancements in technology and challenges facing the financial markets today [13].

The HF trading is the execution of a large number of orders at a very high speed; this results in the generation of a large amount of data that requires real-time processing [14]. Anomaly detection is an integral part of HFT because it entails detecting changes in the standard trading patterns that signal such problems as manipulations or fraud attempts, as well as certain system failures [14]. Anomaly detection in this context, therefore, becomes even more daunting due to the size and speed of the data, and requires methods designed to work on high-dimensional and constantly evolving data [15].

Statistical' detection in trading at an early period of development relied on statistical models and heuristics. We used measures like z-scores, moving averages, and exponential smoothing fitting to analyze the trading pattern and check for abnormalities [16]. Though these methods offered a scope for the detection of anomalies, they were unable to some extent to cope with the volatility of the high-frequency trading data. When trading strategies were relatively simple, these macro-driven approaches to trading were adequate, but as trading techniques developed and enhanced for the trading complexity required for modern market conditions began to appear, the weaknesses of these types of approaches became apparent, and a demand for better analytical methodologies emerged[12].

The use of machine learning algorithms in anomaly detection has done better than traditional statistical methods in improving the former's performance [14]. We have utilized classification techniques like decision trees, support vector machines (SVM), and random forests to scrutinize trading data and identify any anomalous activities [17]. These methods involve the use of labeled datasets to train; this is a big problem in high frequency trading since labeled samples are scarce and the ground truth labeling is not easily obtainable. Nevertheless, the results obtained from supervised learning methods demonstrated higher accuracy in anomaly detection compared to traditional methods [18].

High-frequency trading has also utilized other machine learning approaches, specifically supervised learning strategies, to identify anomalies. Some of the unsupervised machine learning methods includes k-means clustering, hierarchical clustering, and principal component analysis (PCA), which are capable of achieving the aforementioned tasks in high dimensions even without a data label [15]. Employing these methods is most effective when searching for unfamiliar patterns and anomalies in data. However, they also have the challenge of dealing with the large amount of data generated in high-frequency trading, which implies more research is called for [19].

VAEs, or Variational Autoencoders, are an effective tool for anomaly detection and work on the idea of learning data's complex nature through deep learning [20]. Automotive VAEs are generative models that have mapping functions for encoding and decoding the data, making them ideal for learning fine features and anomalies from normal patterns. Financial markets attempted to use VAEs to extract data representations capable of providing notable results and detecting anomalous trading patterns. Research has shown that VAEs can improve accuracy and stability in the discovery of anomalies by learning subtler relationships in data than basic methods could uncover [21].

A recent analysis of VAEs in the context of financial anomaly detection revealed their ability to overcome the difficulties associated with the high dimensionality of data [22]. Through centering complex interactions and dependencies in the trading data, VAEs are a much better point of departure for analyzing what is normal and what is not. For instance, VAEs applied to stock prices, trading volumes, and order book data demonstrated the increase in detection accuracy in comparison with the traditional statistical methods [23].

The integration of deep learning techniques, such as VAEs, into anomaly detection schemes has significantly improved over traditional methods [24]. VAEs are able to learn and generalize from large sets of data, thus making it possible for them to identify defects that other simpler models may not easily pick. It is particularly beneficial in the activity known as 'high-frequency trading' because, in this case, instant action and precise identification of deviations are possible to ensure market integrity and minimize risks [25].

Quantum computing presents new opportunities to improve conventional machine learning algorithms, including VAEs. Quantum Variational Autoencoders (Quantum VAEs) apply the concepts of superposition and quantum entanglement, which helps them to solve the high-dimensional data problems faster than the classical approach [26]. Quantum VAEs employ quantum circuits to encode and process data, with the expectation of enhancing performance in these metrics. In the case of anomaly detection, quantum computing integration can naturally solve some of the drawbacks of classical methods, such as computational complexity or data handling capacity.

The initial theoretical studies of quantum VAEs have demonstrated their effectiveness in various fields, including picture analysis and language processing. These papers propose that quantum Variational Autoencoders are capable of beating traditional VAEs by solving the issues related to computing capabilities and data volume. However, there hasn't been sufficient research on their application in high-frequency trading and financial anomaly detection [27].

It is imperative to compare the performance of the classical model VAEs and the newly proposed quantum VAEs in order to identify the merits of each [28]. The use of such Quantum VAEs in the current approach to anomaly detection is quite a turning point in the financial analysis field. Therefore, future improvements in quantum computing technology will inevitably impact quantum VAEs, making the idea of their application for high-frequency anomaly detection quite plausible [29]. Studying in this direction can help to create new benchmarks for anomaly detection and elaborate theories regarding the use of QML in finance.

The application of VAEs and today's new direction in this field – Quantum VAEs, opens new opportunities for increasing the effectiveness and accuracy of anomaly detection methods [30]. Further research and studies, along with the conduct of comparative analysis, will play a crucial role in advancing the field and securing opportunities for improved financial analytics through the use of quantum computing.

Regarding the research discussed, it is possible to highlight the increasing significance of deploying modern ML techniques, including VAEs and QVAEs, for FIN anomaly detection. They align with other advancements resulting from the use of modern technologies to address the numerous and complex issues in today's trading space [31]. Future work and continued technological advancements will be critical for anomaly detection and financial market analysis in the ever-growing field.

3. Materials and Methods

In this paper, the research will concentrate on frequency trading (HFT) data in the context of financial markets as shown in Table 1. HFTs are defined by their ability to trade and analyze large volumes of real-time trade data, which is critical for modeling market phenomena. The research sample is based on the data originating from the world's largest financial exchanges, like Dow Jones and NASDAQ, which are famous for high-frequency trading. These exchanges are well-regulated and feature state-of-the-art technology that allows them to process large amounts of data and trade in the markets.

Last price	**Mid**	**OP Quantity**	**Closed Position**	**Transacted Quantity**	**Bid1**	**Bid2**	**y**

			Quantity				
3842.4	3842.6		103	0	3842	3842	1
3842.8	3843.4	6	49	55	3843	3842	0
3844	3844.3	7	77	84	3843.8	3842	0

Table 1: High Frequency Trading (HFT) Dataset Collection

The data used in the current study originates from HF trading systems that provide features such as order books, prices, and trade volumes. They cover a wide range of market situations and trading environments, from high volatility to low liquidity. The investigation is important because financial markets are becoming more diverse and unpredictable; as a result, it may be necessary to detect anomalies and predict possible market reactions for further tailor-made strategies and financial security.

Classical VAE Methodology

Data Preprocessing

Data pre-processing is therefore important for preparing the dataset for analysis and improving the quality of the outcomes. The processing steps include:

Data Cleaning: This includes issues like missing values and errors in the given data set. Here, missing values are treated by interpolations, where missing series values are predicted with the existing data series values. This approach assists in ensuring that the data starts and ends at the right point, hence eliminating cases of gaps that may hamper the performance of the models. Furthermore, we remove any existing noise from the data set to enhance its quality.

Normalization: To enable the learning process of the VAE, there is always a need to normalize the features so that they are on a similar scale. Usually, we use normalization methods like Min-Max or Z-score normalization for this purpose, ensuring that each feature contributes equally to the model's training. This step makes the model unbiased, reduces the problem of convergence, and guarantees that the VAE will be able to capture the right features from the data.

Feature Engineering: We can develop or modify additional variables to potentially contribute positively to the given dataset. For example, using derived features such as price changes, degrees of fluctuation, or even a simple moving mean can improve understanding of existing trading novelties or abnormalities as shown in Fig. 1 below.

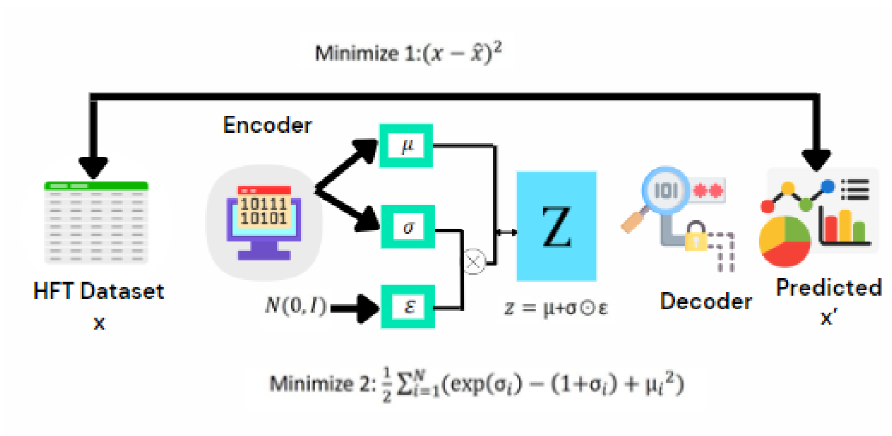

Fig. 1 Classical Variational Autoencoder Methodology with encoder and decoder

Classical VAE Implementation

The classical VAE's implementation involves several key components and steps:

Architecture: This classical VAE model involves having two primary networks, as shown below.

Encoder Network: This network has the ability to transform the multiple features in the input data space into a lower-dimensional latent space as shown in Fig. 2. The encoder is usually a multi-layer perceptron (MLP) with one or more hidden layers. It can use different activation functions, such as the exponential linear unit (ELU) and rectified linear unit (ReLU) responsibilities. This means that the encoder provides the parameters of a latent space distribution consisting of mean and variance values as shown below:

First Hidden Layer:

$$\mu = W_\mu h_2 + b_\mu$$

$$h1 = ReLU(W_1 x + b_1) \quad (1.1)$$

Second Hidden Layer:

$$h_2 = ELU(W_2 h1 + b_2) \quad (1.2)$$

Latent Space Mean and Log-Variance:

$$log\sigma^2 = W_\sigma h_2 + b_\sigma \quad (1.3)$$

Latent variable z is specified as:

$$z = \mu + \sigma * \in, \in \sim N(0,1) \quad (1.4)$$

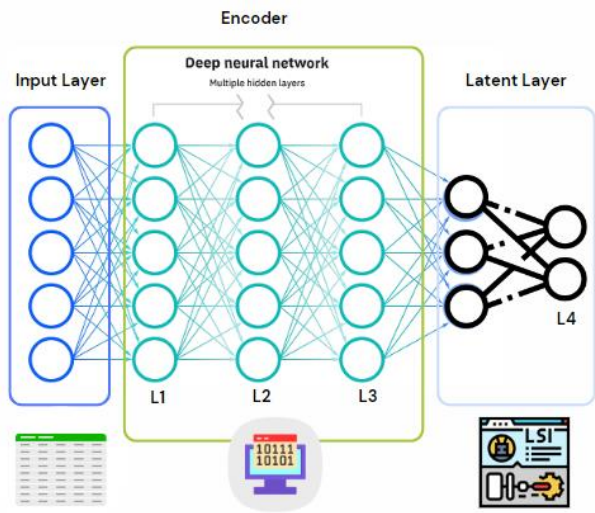

Fig. 2 Classical VAE with Encoder and Latent Layer

Decoder Network: The decoder network reconstructs the original data from the described latent space into its representational form. Similar to the encoder, its structure aims to generate data in the original feature space as show in Fig. 3. The decoder uses activation functions appropriate to the type of data in the output layer, such as sigmoid activation for binary data or linear activation for positive integer and continuous data as shown:

First Hidden Layer:

$$h_3 = ReLU(W_3 z + b_3) \qquad (1.5)$$

Second Hidden Layer:

$$h_4 = ELU(W_4 h_3 + b_4) \qquad (1.6)$$

Output Layer:

$$x' = Sigmoid(W_5 h_4 + b_5) \qquad (1.7)$$

Model Training: To train the VAE, both objectives use the following two loss functions.

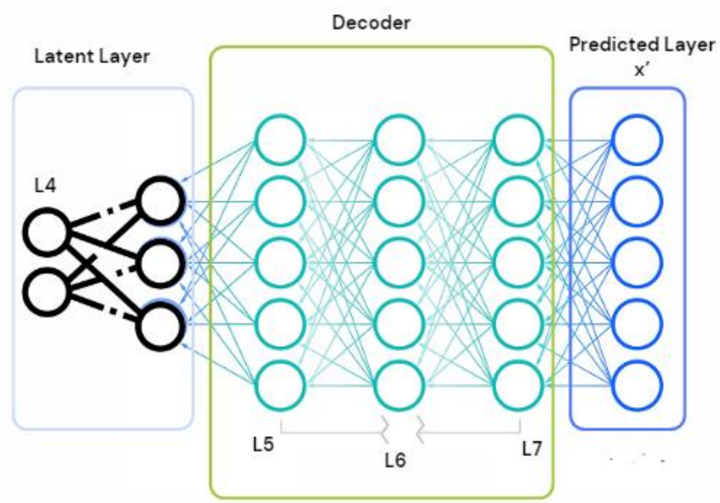

Fig. 3 Classical VAE with Decoder and Latent layer

Reconstruction Loss: Computes the discrepancy between the observed data and its approximation. Usually the simplest reconstruction loss functions are mean squared error (MSE) for the continuously distributed data or binary cross-entropy for the binary data. The decoder needs to learn the mapping that reconstructs the data from the latent space, and this loss function enforces this.

Kullback-Leibler (KL): The KL divergence normalizes the learned features by allowing only a tiny variance around a standard normal distribution policy. By preventing the Gaussian distribution from distorting the learned latent variables, the KL divergence enhances the structure and interpretability of the learned latent space as shown in Fig. 4. This regression is critical for keeping the latent space smooth and enabling meaningful reconstruction shown:

$$KL[q(z|x)||p(z) = \frac{1}{2}\sum_{i=1}^{d}(\mu_i^2 + \sigma_i^2 - \log(\sigma_i^2) - 1) \tag{1.8}$$

The training process consists of solving these optimization problems for the loss functions using gradient-based methods like Adam or RMSprop. Therefore, we perform optimization to tune the model parameters, minimizing both reconstruction errors and distances in the latent space as shown:

$$L(x, x') = Recontruct\ loss + \beta. KL[q(z|x)||p(z)] \tag{1.9}$$

Afterwards, we evaluate the trained VAE model's performance as an anomaly detection algorithm

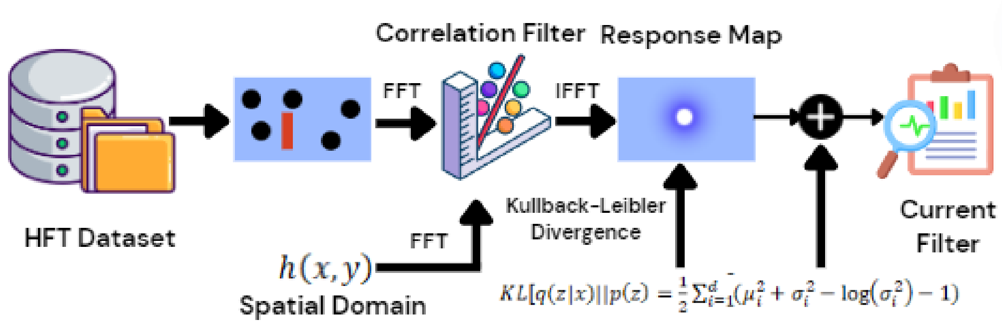

Fig. 4 Classical VAE with KL divergence normalizes

Anomaly Detection: This process uses reconstruction errors to identify abnormal behavior. The model recognizes anomalies as data points with a high reconstruction error because they deviate from the typical patterns it has learned. We determine anomalies by calculating an anomaly score, which we derive from the reconstruction error, and we need to meet certain cut-off points to classify a data point as either an anomaly or a normal point.

Metrics Evaluation: We use various parameters, such as precision, recall, and F1 score, to evaluate the model's performance and determine its ability to detect anomalies. Precision takes the ratio of true positives out of all the positives, while recall deals with the true positive to the total number of positive cases; hence, the F1 score is simply the harmonic mean of both precision and recall measures. These measures aid in model assessment because of their ability to identify significant anomalies in high-frequency trading data.

We select this classical VAE approach because its key advantage is apparent when dealing with intricate and high-dimensional data distributions. The VAE operates in this manner. The model encodes the presented data using embedding, thereby learning representations of the data, its dependence, and dispersion, all of which are crucial for anomaly detection. This approach has the advantage of developing a more general and powerful base for detecting anomalies in HFT, as it establishes normality through the learning of trading patterns with anything outside these as the set base for anomalies. Since it is possible to model and reconstruct the data from the latent space, the conventional VAE is highly effective in detecting abnormal market trends, which might be a sign of a large trading activity or unpleasant events. The approach's ability to handle vast and intricate data characteristics, along with its precision, expedites the detection of market irregularities in high-frequency trading systems.

Quantum VAE Methodology

Data Preprocessing

Data preprocessing for Quantum VAE follows a meticulous process to ensure that the data is suitable for quantum computations and effectively models high-frequency trading dynamics: Here, we outline the meticulous preparation required for the data used by Quantum VAE, ensuring its suitability for quantum computations and its accurate representation of high-frequency trading characteristics:

Data Cleaning: This step involves cleaning the data to address any potential flaws in the actual data. We handle missing values by implementing different interpolation methods in data imputation, thereby reducing the likelihood of wide gaps in the model endpoint estimation. Data cleaning also eliminates outliers and alters other potentially incorrect entries.

Normalization: Prior to feature implementation, normalization techniques such as Min-Max scaling or z-score normalization align all features within a similar range. Normalization reduces the range of various features, enabling quantitative analysis of each one during data analysis.

Feature Engineering: This entails adding new variables and/or modifying already existing attributes in the dataset in such a way that the dataset has the capacity to predict that specific event. Therefore, we can use selection methods like dimensionality reduction or data aggregation to pinpoint the crucial attributes for data analysis, aiming to spot uncommon patterns.

However, all of these preprocessing steps are required to clean and scale the inputs for quantum algorithms, as well as to present the inputs appropriately to the quantum VAE used throughout this thesis.

Quantum VAE Implementation

The Quantum VAE model adds quantum computing components to the VAE architecture, bringing fresh perspectives on data encoding and anomaly detection. The quantum VAE methodology incorporates quantum computing components into the standard VAE design, offering new methods for encoding and anomalous pattern detection.

Architecture: Quantum VAE combines neural networks and quantum circuits with regular neural structures. Quantum circuits replace the standard neural networks of the classical VAE in quantum VAE.

Quantum Encoder: The quantum encoder, a component of a quantum processing unit, utilizes quantum gates to encode data into quantum states. The quantum encoder's operation is based on passing the inputs through a series of quantum gates that change the form of data into a superposition of the quantum states as shown in Fig. 5. Such encoding processes include capturing high-order moment data and covariance, or correlation, between different variables describes:

Initial State:

$$|\psi_0\rangle = |0\rangle^{\otimes n} = |0\rangle_1 \otimes |0\rangle_2 \otimes \cdots \otimes |0\rangle_n \quad (1.10)$$

Apply Quantum Gates:

$$|\psi_1\rangle = H_1 \otimes H_2 \otimes \cdots \otimes H_n |\psi_0\rangle \quad (1.11)$$

Encoding Data with Parametrized Gates:

$$|\psi_2\rangle = \bigotimes_{i=1}^{n} R_y(\theta_i)|\psi_1\rangle \quad (1.12)$$

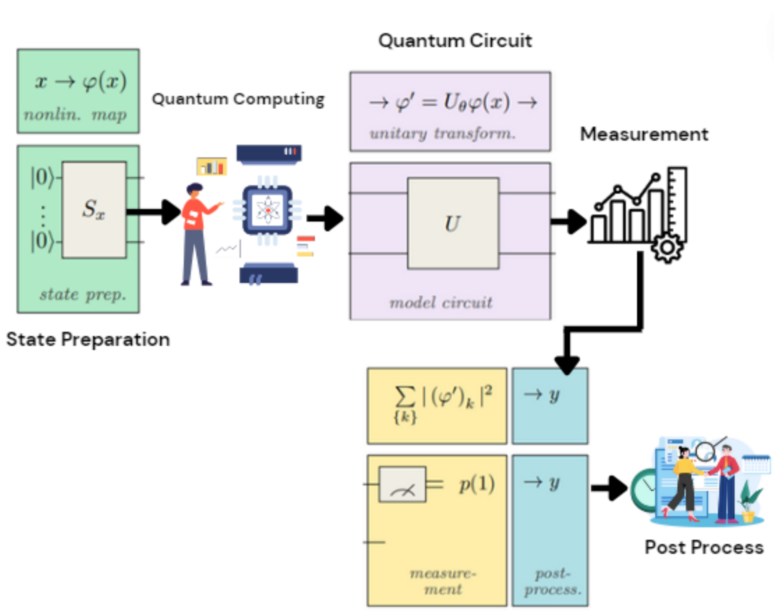

Fig. 5 Quantum VAE Encoder Methodology and analysis

Quantum Decoder: The other process, which is complementary to the quantum encoder, is the quantum decoder, which is responsible for reading out classical information from the superposition state quantized by the quantum encoder as shown in Fig. 6. Quantum measuring processes could potentially restructure the entangled states, transforming them into classically described states with entirely distinct bits. This step transforms the quantum representation into a reconstructive form, facilitating the easy measurement of anomalies. Mathematical describe:

Entangles State before Measurement:

$$|\psi\rangle = \sum_{i=0}^{2^n-1} \alpha_i |i\rangle \quad (1.13)$$

Outcome Measurement:

$$\text{Measurement outcome } |c\rangle = |i\rangle \quad \text{with probability } |\alpha_i|^2$$

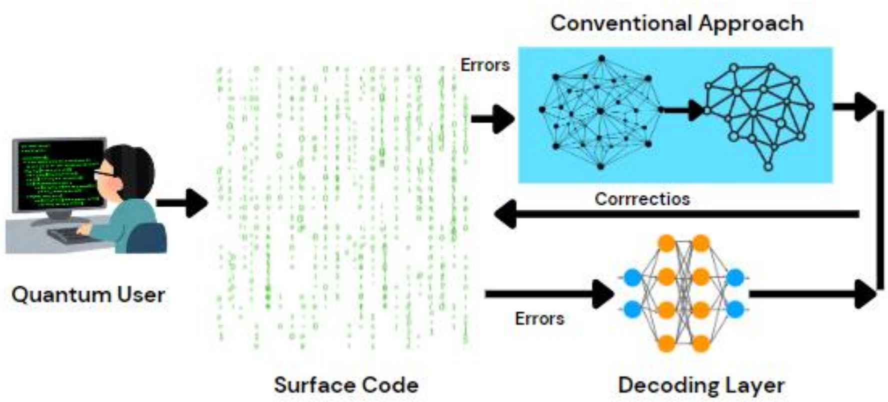

Fig. 6 Quantum VAE with Decoder methodology

Quantum VAE: We can guess that the quantum VAE can use quantum properties like superposition or entanglement to its advantage. These properties allow for different ways of representing and changing data using quantum circuits.

Training: The training process for Quantum VAE combines both classical and quantum methods. Hybrid Optimization: We use parameter optimization methods, such as gradient-based classical optimization methods, to maximize the values of quantum circuits. This is a combination of quantum computing effectiveness and standard optimization procedures for the model.

Loss Function: The Quantum VAE divides its loss function into parts that enable the model to optimize for reconstruction and incorporate additional quantum-specific components. Quantum reconstruction considers the variation between genuine and recovered information through the applied index of reconstruction loss, which aims at maintaining the integrity of the quantum states as does quantum regularization do.

The training strategy proposed here aims to control and optimize these various loss components, enabling the quantum VAE to learn a representation of the data and identify anomalies as shown in Fig. 7.

Model Evaluation: The process of comparing Quantum VAE also involves assessing its ability to elucidate anomalies and comparing it with conventional VAE models.

We analyze quantum-enhanced reconstruction errors to evaluate the model's performance in anomaly detection. We evaluate the model's sensitivity to identify abnormal trading patterns; as such high reconstruction errors suggest anomalies.

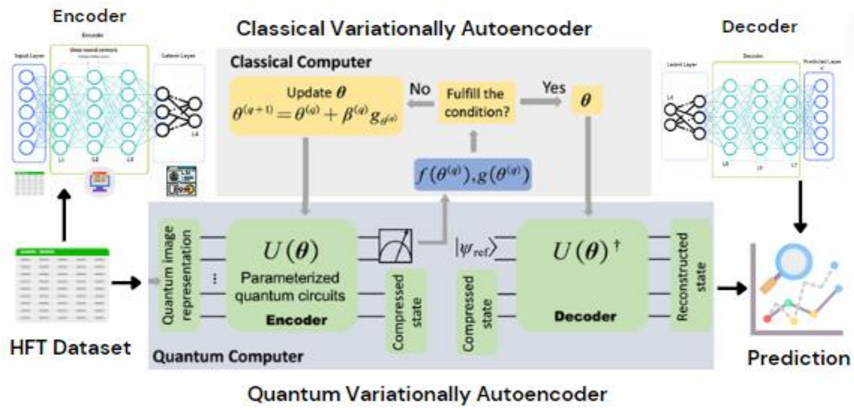

Fig. 7 Quantum VAE Performance with Encoder and Decoder

Performance Comparison: To determine Quantum VAE's success in anomaly detection, we compare it with other VAE models, just like we do with the other VAE models. Typically, we measure the quantum model's absolute accuracy or compare its performance results against classical algorithms using metrics like precision, recall, and the F1 score.

We chose the Quantum VAE approach to investigate the potential benefits of quantum computing for anomaly detection in high-frequency trading. The choice for the quantum VAE approach allows for examining the possible benefits of using quantum approaches in anomaly detection for high-frequency trading.

Compared to classical computing, they also provide seamless solutions for complex datasets that have more than one variable. As crude as they are in their most broad-sense definitions, quantum circuits stand to raise the rate of computation and will come in handy when used in big data analysis and high-dimensional patterns.

The QVAE model utilizes quantum processes to facilitate anomaly detection. The application of quantum superposition and entanglement may help to improve the capability of discovering more complex patterns of trade activities and market anomalies, as well as augmenting the firm's market knowledge.

The ability to demonstrate the practical application of specific quantum computing principles in these examples provides an environment for evaluating the performance of quantum technologies. We can then compare the quantum VAE with the classical VAE to see the actual enhancement it provides in terms of performance, detection capability, and power upon implementation. So, the new quantum VAE approach is a big step forward in the field of high-frequency trading research. It shows that quantum machine learning can be used to solve problems involving finding anomalies.

4. Results and Discussions

Accuracy and F1 Score: The study's results demonstrate that the proposed variation of Quantum VAE-Quantum VAE has an accuracy of 0.93 and a weighted F1 score of 0.91 as show in Table 2. From these metrics, we deduce that the proposed quantum VAE model effectively identifies normal and anomalous data points in a higher-frequency trading system. To evaluate the model's outlier prediction ability, we calculate the F1 score in order to avoid higher false positive or false negative rates. This performance demonstrates how safe it is to use the proposed Quantum VAE for detecting anomalies in high-dimensional financial datasets

Table 2 Accuracy and F1 score of model performance metrics

Metric	Results
Accuracy	0.93
F1 Score	0.90.91

Precision-Recall Curve: In Fig. 8, we show the quantum VAE's precision-recall curve. The curve is a measure of precision and recall; precision is the true positive divided by the total true positive plus false positive values, and recall is the true positive divided by the actual true positive plus false negative values at different threshold levels. The precision-recall curve's specific metric, the area under the curve (AUC), has a value of 0.93, indicating that the quantum VAE outperforms other methods in terms of precision and recall rate for anomaly detection. This stalwart performance is particularly important to the HF trading environment because it requires an instant and accurate assessment of abnormalities to help control risk and generate good decisions.

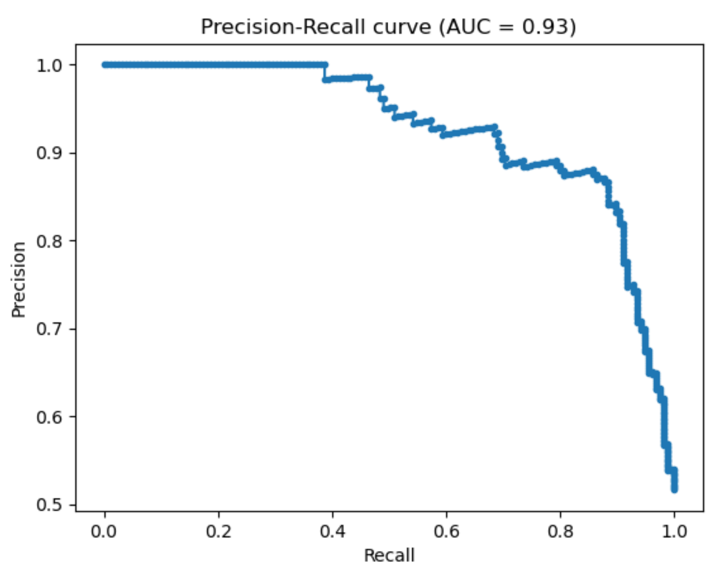

Fig. 8 Quantum VAE precision-recall curve analysis

Actual vs. Predicted Values: The blue points in the scatter plot in Fig. 9 represent the actual values, whereas the red points represent Quantum VAE's predicted values. The proximity of these two curves supports the model's ability to recreate data and isolate seeming anomalies. As a result, the plot also shows that the Quantum VAE can spend its time capturing and predicting anomalies in HF trading data sets. The scatter plot demonstrates that the quantum VAE achieves a high level of precision and accuracy in its prediction.

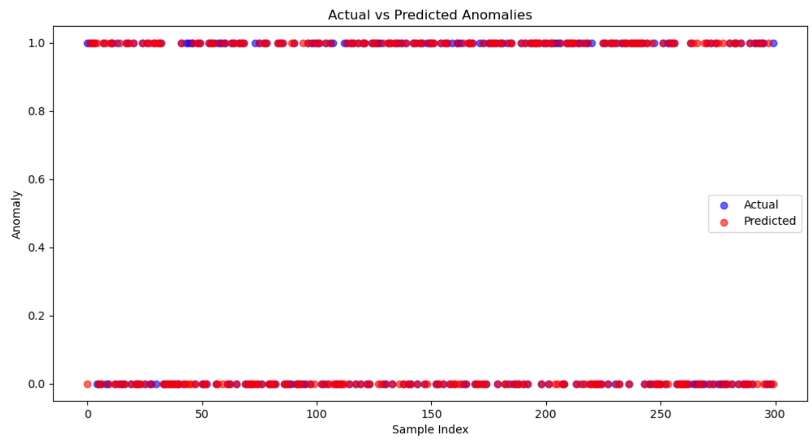

Fig. 9 Quantum VAE actual v. predicted values analysis

Loss and Validation Loss: The figure below illustrates the training and validation losses against epochs in the context of Quantum VAE. The plot also reveals that there is a unique and clear downward trend in both training loss and validation loss, thus demonstrating that the model is learning as well as generalizing as shown in Fig. 10. The training loss tells about the train data, while the validation loss tells about how well the model has done on the unseen data. This graph indicates that the Quantum VAE effectively captures the underlying patterns in the data, avoiding the risk of overfitting. This stable and convergent behavior indicates that the model's training process is adequate.

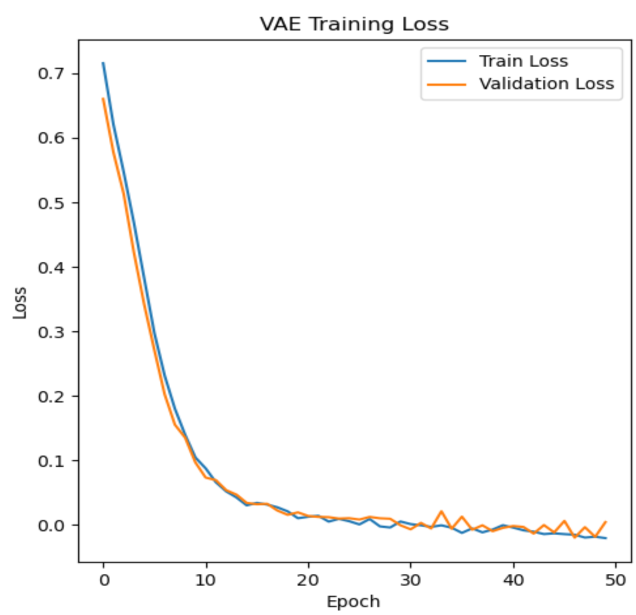

Fig. 10 Quantum VAE Training and Valuation loss

Latent Space Representation: This is the plot of the encoded data from the Quantum VAE, as illustrated in Fig. 11. Each point represents an instance in the latent space, with different colors representing normal data instances and anomalies. With the help of the visualization, one can grasp the idea of how the Quantum VAE organizes similar data and isolates outliers. The Quantum VAE successfully captures meaningful low-dimensional representations of the data by assigning normal cases to one cluster and correctly placing outlying instances in another. This visualization supports the fact that this model has learned a useful latent space for anomaly detection.

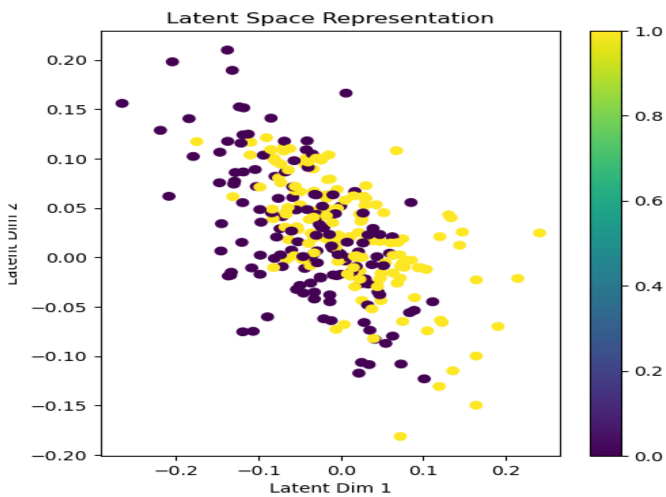

Fig. 11 Latent Space representation in Quantum VAE

Visualization of anomaly detection

Anomaly Detection Results: Fig. 12 displays the test results of the anomaly detection process, highlighting anomalies in red and normal instances in blue. This animation shows how Quantum VAE is able to distinguish outliers in data sets using reconstruction errors. We can conclude that the quantum VAE has sufficient capability to induce deviations from normal behaviors, based on its clear separation between anomalous and normal cases. It aids in determining the model's effectiveness in identifying these as anomalies and also provides a practical perspective of the model in a real-world setting.

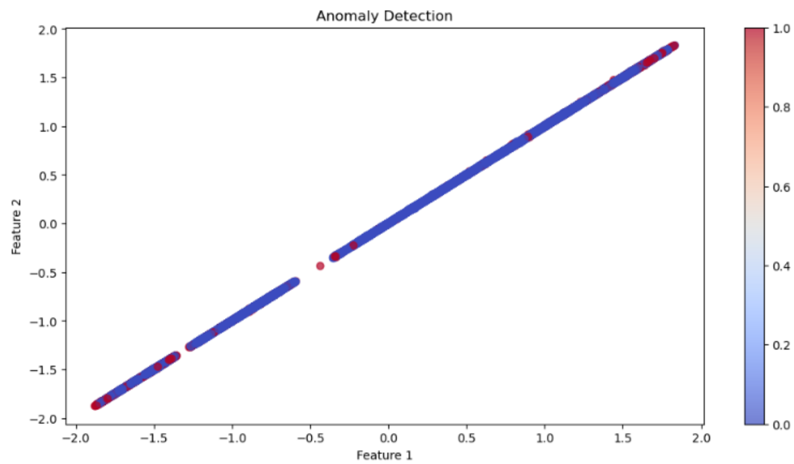

Fig. 12 Anomaly Representation of Quantum VAE

PCA analysis

Principal Component Analysis (PCA): Fig. 13 presents the PCA plot of the dataset, reducing the dimensionality to depict the dataset's distribution in a nearly reduced dimensional space. The PCA plot suggests more about the manner in which the Quantum VAE deals with the data at the reduced dimensions, as seen from the pileup of normal and anomalous points. From the plot, it is clear that Quantum VAE de-correlates the anomalies from the normal data in the reduced feature space. This analysis provides some useful extensions and gives further insight into precisely how well the model can scale up and meet the demands of such large data sets.

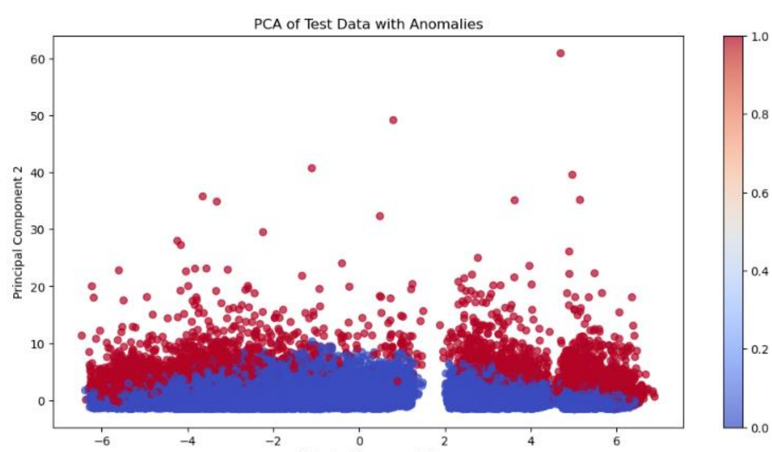

Fig. 13 Principle Component Analysis of the given dataset

Comparison with classical VAE

Classical VAE Metrics: This paper's earlier development of the Classical VAE yielded an F1 score of 0.78 and an accuracy of 0.80%. The classical VAE performed worse in terms of accuracy and F1-score than the quantum VAE, indicating that the quantum VAE performed better in anomalous data discrimination as show in Table 3. The work highlights the deficiencies in the classical VAE, demonstrating how the quantum VAE outperforms it in terms of accuracy and effective anomaly detection. This comparison shows that it may be useful to integrate QC coverage into the subject of anomaly detection.

Table 3 Performance Matric Comparison of Classical and Quantum VAE

Model	**Accuracy**	**F1 Score**
Classical VAE	0.80	0.78
Quantum VAE	0.93	0.91

Discussion

The results from Experiment show that the proposed quantum VAE did better than the classical VAE in terms of F1 and AUC (precision-recall). The training history shows satisfactory training and a good ability to generalize when the loss values are quite stable, indicating successful convergence. In result of the learned latent space and the anomaly detection result prove that the proposed Quantum VAE is capable of detecting and separating anomalous data well. The PCA analysis backs up the model to show that it did well in managing data with high dimensions.

Advantages over Classical VAE

The quantum circuits used in the Quantum VAE have low complexity for high-dimensional data and potentially higher detection rates. The improved computational performance and higher quality clustering in the latent space also speak to the benefits of quantum computing. Quantum VAE's analytic capability in terms of singling out anomalies in the reduced feature space from the PCA proves the model's suitability for the high-frequency trading application, where timely and accurate detection of anomalies is paramount.

Limitations and Challenges

However, there are various issues affecting the working of the quantum VAE, such as the instability of the training process and the restricted accessibility to quantum processors. Since the training process fluctuates, we can conclude that further fine-tuning is necessary. Furthermore, the current limitations of quantum computing constrain the actors of quantum VAE, posing a potential challenge in real-world scenarios.

5. Conclusion

The Variational Autoencoder proposed here can significantly enhance anomaly detection in high-frequency trading datasets. In terms of a number of important performance parameters, it does better than the standard classical VAE. The authors later applied the proposed quantum VAE and realized that the latter had a higher accuracy and F1 score. It also had the higher precision-recall AUC, indicating that it is a better way of finding an anomaly with higher precision and recall capability. The proposed model’s performance in identifying the difference between regular and outlier values in the latent space as well as in PCA further fortifies these observations. This means that employing quantum computing can be very useful in establishing that there are anomalies.

Limitation of this model is to ensure the model's adoption at higher levels of realization; we must address a few challenges, such as instability in the training process and current limitations in quantum computing processing. We discovered significant variations during the training process. During the training process, we observed significant variations from one training session to the next, which could potentially take one or two weeks, depending on the complexity of the problem at hand. Furthermore, the use of quantum resources necessitates additional fine-tuning of the quantum algorithm to achieve consistent and reliable results. Research and develop quantum machine learning to enhance the utility of quantum-based models.

At the end, it is possible to state that the suggested quantum VAE is a leap forward in the field of anomaly detection and provides new opportunities for investigations instead of classical methods. High-frequency trading's increased computing efficiency reveals new challenges in data analysis and outlier detection that quantum computing can solve. For future work, therefore, it may be necessary to reconsider and optimize the quantum replacements and the universal quantum computing to remove present barriers and unleash the full potential of using quantum-based models. However, as quantum technology advances, the use of quantum models such as the defined VAE in numerous and complex data analysis applications become a perfect tool to describe future advancements in fine tuning and cybersecurity, among other areas.

6. References

[1] J. Brogaard, "High frequency trading and its impact on market quality," ***2011***.

[2] K. Choi, J. Yi, C. Park, and S. Yoon, "Deep Learning for Anomaly Detection in Time-Series Data: Review, Analysis, and Guidelines," ***IEEE Access***, vol. PP, pp. 1-1, 2021, doi: 10.1109/ACCESS.2021.3107975.

[3] P. Tchatchoua, G. Graton, M. Ouladsine, and M. Juge, "A Comparative Evaluation of Deep Learning Anomaly Detection Techniques on Semiconductor Multivariate Time Series Data," in ***Proc. 2021 IEEE Int. Conf. Automation Science and Engineering (CASE)***, 2021, pp. 1613-1620, doi: 10.1109/CASE49439.2021.9551541.

[4] D. Gonzalez, M. A. Patricio, A. Berlanga, and J. Molina, "Variational autoencoders for anomaly detection in the behaviour of the elderly using electricity consumption data," ***Expert Systems***, vol. 39, 2021, Art. no. e12744, doi: 10.1111/exsy.12744.

[5] L. Chen, T. Li, Y. Chen, X. Chen, M. Wozniak, N. Xiong, and W. Liang, "Design and analysis of quantum machine learning: A survey," ***Connection Science***, vol. 36, 2024, Art. no. 2312121, doi: 10.1080/09540091.2024.2312121.

[6] I. Ishak, H. Ibrahim, F. Sidi, and Z. Zukarnain, "A review on high frequency trading forecasting methods: Opportunity and challenges for quantum-based method," ***IEEE Access***, vol. PP, pp. 1-1, 2024, doi: 10.1109/ACCESS.2024.3418458.

[7] S. Bengesi, H. El-Sayed, M. Sarker, Y. Houkpati, J. Irungu, and T. Oladunni, "Advancements in generative AI: A comprehensive review of GANs, GPT, autoencoders, diffusion model, and transformers," ***IEEE Access***, vol. PP, pp. 1-1, 2024, doi: 10.1109/ACCESS.2024.3397775.

[8] L. Li, J. Yan, H. Wang, and Y. Jin, "Anomaly detection of time series with smoothness-inducing sequential variational auto-encoder," *IEEE Transactions on Neural Networks and Learning Systems*, vol. PP, pp. 1-15, 2020, doi: 10.1109/TNNLS.2020.2980749.

[9] R. Chalapathy and S. Chawla, "Deep learning for anomaly detection: A survey," 2019.

[10] J. Rogers, G. Mukherjee, M. Frank, T. Jiang, O. Christiansen, Y. Yao, and N. Lanatà, "Error mitigation in variational quantum eigensolvers using tailored probabilistic machine learning," 2023.

[11] C. Zhang and Y. Chen, "Time series anomaly detection with variational autoencoders," 2019.

[12] V. S. Ngairangbam, M. Spannowsky, and M. Takeuchi, "Anomaly detection in high-energy physics using a quantum autoencoder," ***Physical Review D***, vol. 105, no. 9, p. 095004, 2022.

[13] M. O'Hara, "High-frequency trading and its impact on markets," Financial Analysts Journal, vol. 70, no. 3, pp. 18-27, May/June 2014.

[14] H. O. Bello, A. B. Ige, and M. N. Ameyaw, "Deep Learning in High-frequency Trading: Conceptual Challenges and Solutions for Real-time Fraud Detection," World Journal of Advanced Engineering Technology and Sciences, vol. 12, no. 02, pp. 35-46, 2024.

[15] S. Thudumu, P. Branch, J. Jin, and J. Singh, "A comprehensive survey of anomaly detection techniques for high dimensional big data," Journal of Big Data, vol. 7, pp. 1-30, 2020.

[16] M. Ahmed, A. N. Mahmood, and M. R. Islam, "A survey of anomaly detection techniques in financial domain," Future Generation Computer Systems, vol. 55, pp. 278-288, 2016.

[17] K. Yıldız, S. Dedebek, F. Y. Okay, and M. U. Şimşek, "Anomaly detection in financial data using deep learning: A comparative analysis," in *Proc. 2022 Innovations in Intelligent Systems and Applications Conference (ASYU)*, Sep. 2022, pp. 1-6. IEEE.

[18] M. Kearns and Y. Nevmyvaka, "Machine learning for market microstructure and high frequency trading," in *High Frequency Trading: New Realities for Traders, Markets, and Regulators*, vol. 72, 2013.

[19] K. H. Lee, "Big data analytical algorithms and systems for high frequency trading and contract network analysis," 2018.

[20] W. Hilal, S. A. Gadsden, and J. Yawney, "Financial fraud: a review of anomaly detection techniques and recent advances," *Expert Systems with Applications*, vol. 193, p. 116429, 2022.

[21] T. Mendes, P. J. Cardoso, J. Monteiro, and J. Raposo, "Anomaly detection of consumption in hotel units: A case study comparing isolation forest and variational autoencoder algorithms," *Applied Sciences*, vol. 13, no. 1, p. 314, 2022.

[22] H. Kim and H. Kim, "Contextual anomaly detection for high-dimensional data using Dirichlet process variational autoencoder," *IISE Transactions*, vol. 55, no. 5, pp. 433-444, 2023.

[23] S. A. Hosseini, S. A. Hosseini, and M. Houshmand, "Variational autoencoder-based dimension reduction of Ichimoku features for improved financial market analysis," *Franklin Open*, vol. 8, p. 100135, 2024.

[24] Y. Zhou, X. Liang, W. Zhang, L. Zhang, and X. Song, "VAE-based deep SVDD for anomaly detection," *Neurocomputing*, vol. 453, pp. 131-140, 2021.

[25] E. Albayrak, "Stock price prediction using deep learning methods in high-frequency trading," M.S. thesis, 2021.

[26] A. Khoshaman, W. Vinci, B. Denis, E. Andriyash, H. Sadeghi, and M. H. Amin, "Quantum variational autoencoder," *Quantum Science and Technology*, vol. 4, no. 1, p. 014001, 2018.

[27] N. Rane, S. Choudhary, and J. Rane, "Machine Learning and Deep Learning: a Comprehensive Review on Methods, Techniques, Applications, Challenges, and Future Directions," *Techniques, Applications, Challenges, and Future Directions*, May 31, 2024.

[28] U. Tariq, I. Ahmed, M. A. Khan, and A. K. Bashir, "Deep learning for economic transformation: a parametric review," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 35, no. 1, pp. 520-541, 2024.

[29] M. Wang, A. Huang, Y. Liu, X. Yi, J. Wu, and S. Wang, "A quantum-classical hybrid solution for deep anomaly detection," *Entropy*, vol. 25, no. 3, p. 427, 2023.

[30] K. Fraser, S. Homiller, R. K. Mishra, B. Ostdiek, and M. D. Schwartz, "Challenges for unsupervised anomaly detection in particle physics," *Journal of High Energy Physics*, vol. 2022, no. 3, pp. 1-31, 2022.

[31] S. Johan, "Enhanced financial business competitiveness by leveraging technology and innovation," *CommIT (Communication and Information Technology) Journal*, vol. 15, no. 2, pp. 79-89, 2021.