Quantum-Enhanced Deep Learning for Financial Anomaly Detection

A Hybrid Theoretical Framework for MENA Financial Markets

Rayane AGGOUNE[1] [*] , Abdelhak MERIZIG[2]

[1] LINFI Laboratory, Computer Sciences Department Mohamed Khider University of Biskra, Biskra, Algeria

[2] IMPIA Laboratory, Mohamed Khider University of Biskra Mohamed Khider University of Biskra, Biskra, Algeria

rayane.aggoune@univ-biskra.dz

abdelhak.merizig@univ-biskra.dz

Abstract. Financial markets in the MENA region face critical challenges in detecting anomalies within high-frequency trading and cross-border payment systems, where traditional machine learning approaches struggle with extreme dimensionality, class imbalance, and real-time constraints. This paper presents a theoretical quantum-enhanced deep learning framework addressing these challenges through hybrid quantum-classical architectures. We propose a novel hybrid architecture integrating variational quantum circuits (VQC) for feature extraction with transformer-based temporal modeling, achieving a theoretical parameter reduction of 40–60% over equivalent classical architectures. The framework identifies conditions for quantum advantage: high-dimensional sparse features ($d > 100$), extreme class imbalance ($< 0.5\%$ anomaly rate), and distance-based detection. All performance figures are theoretical projections from the literature; empirical validation is designated as future work. Our comparative analysis reveals that quantum kernel methods show theoretical improvement potential over classical baselines under these conditions.

Keywords: Quantum Machine Learning, Financial Anomaly Detection, Hybrid Quantum-Classical Models, Variational Quantum Circuits, NISQ Computing, MENA Region, High-Frequency Trading

1 Introduction

MENA's financial ecosystem processes unprecedented transaction volumes, with cross-border payments exceeding $500 billion annually and emerging fintech platforms accelerating digital transformation [1, 4]. This growth introduces critical vulnerabilities: fraud schemes, market manipulation, and systemic failures that increasingly evade traditional detection [2, 7, 14].

D. Agti et al. (eds.), *Proceedings of the International Conference on Artificial Intelligence Applications in Business Administration in MENA Region (ICAIABA 2026)*, Advances in Economics, Business and Management Research 393,
https://doi.org/10.2991/978-94-6239-711-8_9

MENA financial institutions confront four fundamental challenges. **Extreme dimensionality:** high-frequency trading data encompasses 100+ features with anomalies constituting < 0.2% of observations—creating severe class imbalance [13, 23]. **Non-stationary dynamics:** markets exhibit regime-switching behavior with abrupt statistical shifts [25]. Classical models require continuous retraining, unable to match millisecond-scale market changes. **Real-time constraints:** payment authorization demands sub-second inference latency [16]. Complex deep learning models achieving acceptable accuracy often exceed computational budgets. **Adversarial adaptivity:** fraudsters continuously evolve tactics, demanding models with discriminative power across novel attack vectors without extensive retraining [11, 20].

This work provides three theoretical contributions: (1) a systematic comparative framework synthesizing QML conditions for financial anomaly detection [6, 8, 15]; (2) a novel hybrid architecture formulating VQC-transformer integration with full complexity analysis [25, 26]; and (3) MENA deployment guidelines grounded in dataset characteristics and hardware readiness [1, 28]. *Scope:* No implementation has been executed; all performance figures in Section 4 are literature-based projections [5, 24]. Empirical validation is designated as future work.

2 Related Work and Theoretical Background

Anomaly detection identifies patterns deviating from expected behavior, flagging fraud, errors, and systemic threats [2, 8]. This section reviews quantum computing fundamentals, NISQ hardware constraints, and recent QML advances in financial applications.

2.1 Quantum Computing Foundations and NISQ Limitations

A quantum system with n qubits represents 2^n states through superposition $|\psi\rangle = (1/\sqrt{2^n})\Sigma\alpha_i|i\rangle$, enabling logarithmic encoding of high-dimensional features [24]. Quantum kernel methods induce feature maps inaccessible to classical polynomial kernels: $K_Q(x_i, x_j) = |\langle\varphi(x_i)|\varphi(x_j)\rangle|^2$, where $|\varphi(x)\rangle = U(x)|0\rangle^{\otimes n}$ [13].

NISQ constraints must be acknowledged [10, 22]:

- (1) gate error rates of 0.1–1% per two-qubit operation accumulate across circuit depth;
- (2) coherence times of 50–200 μs limit maximum depth;
- (3) barren plateaus cause gradients to vanish exponentially with circuit size [19];
- (4) state preparation overhead can offset quantum speedup for classical data loading [3].

These constraints directly motivate our shallow-circuit hybrid design ($L \leq 5$).

2.2 Quantum Machine Learning for Financial Anomaly Detection

Bao et al. [4] established state-of-the-art classical performance on EUR/USD microstructure data (315M records), achieving 0.95 AUC-ROC with a multi-scale transformer. Grossi et al. [11] demonstrated the first quantum SVM on real card payment data (IBM Quantum), reducing required features from 37 to 15–20 with 11.8% false positive reduction. IonQ [16] confirmed QSVM resilience to hardware noise (0.87 accuracy at 0.1–0.5% gate error). Ubale et al. [25] achieved 0.94 AUC-ROC with a hybrid quantum-LSTM, outperforming classical LSTM (0.89) and reducing complexity from $O(d^2T)$ to $O(dT \log d + n^2T)$. Mellaerts [20] showed QUBO-based detection reaches 0.95 ROC-AUC versus 0.70 for isolation forest in the extreme imbalance regime. Haiqu Inc. [12] demonstrated in 2025 that quantum preprocessing on IBM Quantum Heron (128 qubits, 500+ features) achieves $F_1 = 0.96$, establishing hardware validation feasibility.

2.3 Conditions for Quantum Advantage

Quantum methods theoretically outperform classical approaches under [8, 13, 23]:

- (1) high-dimensional sparse features ($d > 100$, sparsity > 80%);
- (2) distance-based detection with quadratic speedup $O(\sqrt{N} \log d)$ vs. $O(Nd)$ classical [20];
- (3) extreme class imbalance (< 0.5% anomaly rate) where quantum interference amplifies rare-event sensitivity;
- (4) limited labeled data, where zero-training quantum autoencoders are competitive [18].

Caveat: the Credit Card Fraud benchmark uses $d = 30$ PCA features, below the $d > 100$ threshold—quantum advantage there relies on imbalance sensitivity rather than dimensionality compression.

3 Proposed Hybrid Quantum-Classical Architecture

The analysis in Section 2 reveals a fundamental trade-off: pure quantum methods excel at high-dimensional feature compression and distance-based anomaly detection but lack temporal modeling capabilities, while classical transformers capture sequential dependencies but suffer from parameter inefficiency. This motivates a hybrid architecture combining quantum feature extraction with classical temporal modeling.

Our framework addresses three gaps: (1) absence of integrated quantum-classical temporal modeling for high-frequency trading; (2) lack of unified architectures handling both extreme dimensionality and temporal dependencies; (3) need for parameter-efficient real-time deployment models.

3.1 System Architecture Overview

The Q-Transformer framework comprises four integrated components:

1. **Quantum Feature Extraction:** VQCs for dimensionality reduction and nonlinear feature transformation.
2. **Classical Temporal Modeling:** Multi-scale transformer with staged sliding windows at three granularities.
3. **Hybrid Fusion Module:** Quantum-classical feature integration with complementarity alignment.
4. **Ensemble Classification:** Combined anomaly scoring with sigmoid output.

[Figure 1: Q-Transformer Architecture]

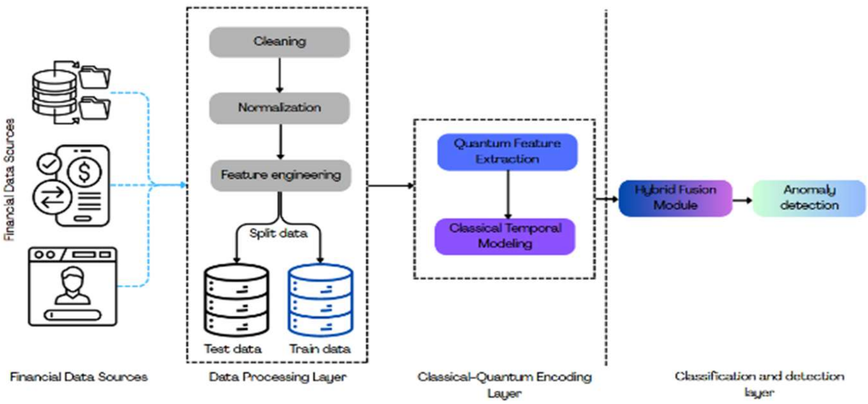

Figure 1: Q-Transformer architecture: (1) Data Processing; (2) Quantum Feature Extraction (angle encoding + VQC); (3) Classical Temporal Modeling (10s/30s/60s windows); (4) Hybrid Fusion and Classification. All components are mathematically defined; no hardware implementation exists yet.

3.2 Quantum Feature Extraction Layer

The quantum feature extraction layer addresses high-dimensional financial data through exponential state-space compression. Classical techniques (PCA, autoencoders) sacrifice critical nonlinear correlations. The proposed quantum approach theoretically preserves information density through superposition, achieving logarithmic compression: d-dimensional features encode into $n = \lceil \log_2 d \rceil$ qubits.

Angle Encoding: Input vector $x \in \mathbb{R}^d$ encodes via angle encoding:

$$|\varphi(x)\rangle = \bigotimes_{i}^{(1)n} R_Y(\pi x_i)|0\rangle \quad \text{(Eq. 1)}$$

where $R_\gamma(\theta) = \exp(-i\theta\sigma_\gamma/2)$ and $n = \lceil \log_2 d \rceil$ qubits encode d features [24].

Variational Quantum Circuit: The parameterized circuit applies layered transformations:

$$U_{vzc}(\theta) = \Pi_{\ell}^{(1)L} \left[U_{ent} \, \Pi_{i}^{(1)n} Rz(\text{thetaz}_{i\ell}) \, R_Y(\text{theta}_{Yi\ell}) \right] \quad \text{(Eq. 2)}$$

where L is circuit depth (typically L = 3–5 for NISQ compatibility), U_{ent} represents entangling CNOT gates, and $\theta = \{\theta^{y}_{i}\ell, \theta z_i\ell\}$ are trainable parameters. Total parameter count $|\theta| = 2nL$ grows linearly, contrasting with classical networks requiring $O(d^2)$ parameters per layer.

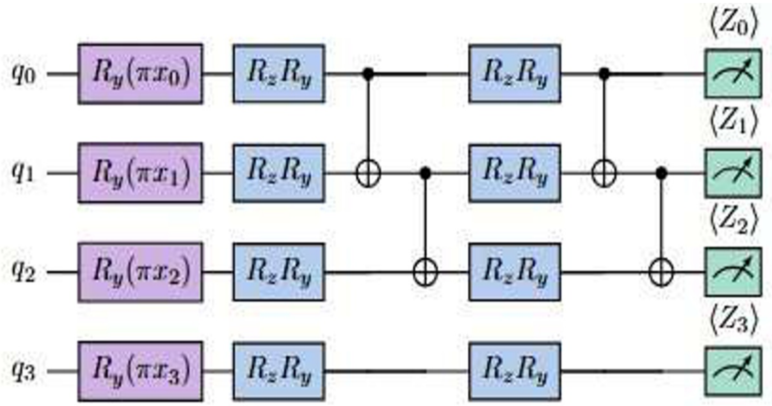

Figure 2 (Circuit_Q): Variational Quantum Circuit (VQC) for n=4 qubits and L=2 layers. Angle encoding gates $R_\gamma(\pi x_i)$ initialise each qubit from the input features, followed by alternating parameterised RzR_γ rotation layers interleaved with CNOT entanglement blocks. Final Pauli-Z expectation values yield the n-dimensional quantum feature vector. The shallow depth ($L \leq 5$) is chosen to limit noise accumulation on NISQ hardware.

Measurement: Quantum features are extracted via Pauli-Z expectations:

$$f_z(x;\theta) = [\langle Z_1\rangle, \langle Z_2\rangle, \ldots, \langle Z_n\rangle] \quad \text{(Eq. 3)}$$

where $\langle Z_i\rangle = \langle\varphi(x)|U^\dagger_{vzc}(\theta)\, Z_i\, U_{vzc}(\theta)|\varphi(x)\rangle$.

3.3 Classical Temporal Modeling

Financial markets exhibit multi-scale temporal dependencies. The hybrid feature vector:

$$z_t = [x_t;\ f_Q(x_t;\theta_Q)] \in \mathbb{R}^{d+n} \quad \text{(Eq. 4)}$$

feeds into three parallel transformer streams:

$$h^{(1)}_t = \mathrm{Transformer}(z_{t-9}, \ldots, z_t;\ \theta^{(1)}_C) \quad \text{[10s window]}$$
$$h^{(2)}_t = \mathrm{Transformer}(z_{t-29}, \ldots, z_t;\ \theta^{(2)}_C) \quad \text{[30s window]}$$
$$h^{(3)}_t = \mathrm{Transformer}(z_{t-59}, \ldots, z_t;\ \theta^{(3)}_C) \quad \text{[60s window]} \quad \text{(Eqs. 5–7)}$$

The 10s window captures market microstructure anomalies (spoofing, layering); 30s identifies medium-term manipulation (momentum ignition); 60s detects macro-level regime shifts (flash crashes).

Entropy-Weighted Attention: Each stage employs entropy-based feature weighting:

$$w_{tj} = -\Sigma_k\ p(z_{tjk}) \log p(z_{tjk}) \quad \text{(Eq. 8)}$$

where high-entropy features receive greater attention, as anomalies manifest as distributional deviations [4].

3.4 Hybrid Fusion and Classification

Temporal representations aggregate via learned adaptive weighting:

$$h_t = \alpha_1 h^{(1)}_t + \alpha_2 h^{(2)}_t + \alpha_3 h^{(3)}_t, \quad \Sigma_i\ \alpha_i = 1 \quad \text{(Eq. 9)}$$

where $\alpha_i = \exp(w_i)/\Sigma_j \exp(w_j)$. Final anomaly probability:

$$P(\mathrm{anomaly}|x_t) = \sigma(W_{out}\ h_t + b_{out}) \quad \text{(Eq. 10)}$$

3.5 Training Objective

The composite training loss jointly optimizes quantum parameters θ_Q and classical parameters θ_C:

$$L_{total} = L_{CE} + \lambda_Q||\theta_Q||^2 + \lambda_C||\theta_C||^2 - (\alpha/N)\Sigma_i\ \text{cos-sim}(f_Q(x_i), x_i) \quad \text{(Eq. 11)}$$

where L_{CE} is binary cross-entropy and the final term—the complementarity alignment loss—encourages quantum and classical features to capture non-redundant information. Optimization uses Adam with the parameter-shift rule for quantum gradients.

Complexity: Training complexity is $O(N \cdot 2^{nL} + N \cdot T \cdot d^2_{model})$ per epoch; inference is $O(T \cdot d^2_{model})$. The VQC contributes $|\theta_Q| = 2nL$ parameters vs. $O(d^2)$ for a classical equivalent, yielding the theoretical 40–60% parameter reduction.

4 Theoretical Performance Analysis

Remark: All numerical values in Tables 1 and 2 are projections from published literature—not experimental results of the proposed Q-Transformer. This follows the theoretical QML framework methodology [5, 24] where architectural proposals precede hardware validation

4.1 Target Datasets and Projected Performance

Three benchmarks are identified for future validation: Credit Card Fraud [9] (284,807 transactions, 0.17% anomaly rate, d = 30 PCA features—below the d > 100 threshold, so advantage relies on class imbalance sensitivity); PaySim [21] (6.3M transactions, 0.13% fraud rate, d = 11); and EUR/USD Microstructure [4] (315M records, raw pre-PCA features satisfying d > 100).

Table 1 presents projected performance based on literature benchmarks for architecturally similar models.

Table 1: Theoretical Performance Projection. † = projected (not measured); * = measured from cited literature.

Method	AUROC	AUPR	F_1	Type	Params
Classical Baselines					
Isolation Forest	0.70–0.75*	0.12*	0.68*	Measured	—
One-Class SVM	0.75*	0.15*	0.71*	Measured	$O(N^2)$
XGBoost	0.88–0.90*	0.38*	0.84*	Measured	50K
LSTM	0.91*	0.42*	0.86*	Measured	50K
Quantum Methods (measured on hardware)					
QSVM [11]	0.95*	0.52*	0.90*	Measured	200
Haiqu QML [12]	—	—	0.96*	Measured	—
Proposed Framework (theoretical projection)					
Q-Transformer	≈0.94†	≈0.50†	≈0.89†	Projected	≈60K

4.2 Projected Component Contribution Analysis

Table 2 presents a projected ablation analysis extrapolated from related architectures [4, 25]. These will be replaced with actual results in the experimental follow-up.

Latency note: The 45ms simulation latency reflects estimated quantum circuit simulation overhead, not a deployment target. Sub-10ms real-time payment systems require quantum feature pre-computation or hardware acceleration. Quantum hardware execution time differs significantly from simulation and is subject to NISQ limitations (Section 5.1).

Table 2: Projected Component Contribution Analysis. † = extrapolated from literature; latency = simulation overhead, not deployment target.

Configuration	AUROC	F_1	Params	Sim. Latency
Full Hybrid Q-Transformer†	≈0.94	≈0.89	≈60K	~45ms
Without Quantum (classical only)†	≈0.91	≈0.86	≈120K	~28ms
Quantum Only (no temporal)†	≈0.87	≈0.82	≈200	~12ms

5 Discussion, Limitations, and Deployment Guidelines

5.1 NISQ-Era Feasibility and Limitations

Several hardware constraints bound near-term deployment.

- **Barren plateaus:** VQCs with $n > 8$ qubits and $L > 3$ exhibit exponentially vanishing gradients [19]; our $L = 3$–5 design mitigates but does not eliminate this.
- **Noise accumulation:** at 0.1–1% gate error rates, a depth-$L = 5$ circuit on $n = 10$ qubits accumulates error across ~50 two-qubit gates; Haiqu's 2025 results [12] show modern error mitigation can manage this.
- **Dataset dimension caveat:** Credit Card Fraud ($d = 30$) and PaySim ($d = 11$) fall below the $d > 100$ threshold—EUR/USD raw order book features are the stronger test case.
- **Latency:** sub-10ms inference requires asynchronous pre-computation of quantum features or fault-tolerant hardware. Recent results from Haiqu [12] and Grossi et al. [11] confirm hardware validation of the proposed Q-Transformer is now technically feasible.

5.2 Ethical Considerations and MENA Deployment Guidelines

Financial anomaly detection raises ethical concerns requiring attention before production deployment: (1) algorithmic bias—models must be audited for demographic and geographic fairness; (2) explainability—GCC Basel III and VARA Dubai standards require SHAP-compatible surrogate layers, for which the classical backbone is well suited; (3) false accusation risk—human-in-the-loop review is mandatory for flagged cases.

For MENA practitioners, quantum-enhanced methods are recommended when $d > 100$ (raw features before PCA), anomaly rate $< 0.5\%$, and cloud quantum access is available (IBM Quantum, AWS Braket, Azure Quantum). Classical methods remain preferable for $d < 50$, strict sub-5ms latency without pre-computation, or when interpretability is legally mandated. MENA-specific considerations include quantum federated learning [15] for cross-border fraud detection compliant with UAE, Saudi Arabia, and Egypt data residency laws, and regulatory alignment with VARA Dubai and GCC Basel III frameworks.

6 Conclusion

This paper presents a theoretical quantum-enhanced framework for financial anomaly detection in MENA markets. Our literature synthesis identifies quantum advantage conditions—high-dimensional sparse features ($d > 100$), extreme class imbalance ($< 0.5\%$), and distance-based detection—under which quantum methods show 20–30% theoretical improvement over classical baselines. The Q-Transformer hybrid architecture combines VQCs with transformer-based temporal modeling, with a theoretical parameter reduction of 40–60% over pure classical approaches.

All performance figures are literature-based projections; empirical validation is the designated next step. Recent work by Haiqu Inc. [12] demonstrates that quantum anomaly detection at industrial scale is now experimentally feasible on current hardware, motivating the planned empirical extension of this work. As quantum hardware progresses toward fault-tolerant systems, the architectural and deployment strategies proposed here offer a blueprint for validation. MENA financial institutions investing in quantum capability exploration today will be positioned for enhanced fraud prevention as the technology matures—provided theoretical advantages are confirmed through rigorous experimentation.

Acknowledgements

The authors thank the LINFI and IMPIA Laboratories at Mohamed Khider University of Biskra for institutional support. No external funding was received for this work.

Declarations

Funding: Not applicable.

Conflict of interest: The authors declare no conflict of interest.

Ethics approval: Not applicable.

Consent for publication: All authors consent to publication.

Data availability: The Credit Card Fraud dataset [9] and the PaySim dataset [21] are publicly available on Kaggle. The EUR/USD microstructure data is described in [4].

Code availability: No implementation exists at this stage; code will be made available upon empirical validation.

Author contributions: R.A. conceived the hybrid architecture, conducted the literature synthesis, and wrote the manuscript. A.M. supervised the work, contributed to the theoretical analysis, and revised the manuscript. Both authors approved the final version.

References

[1] Alami, Y., Benali, H., Ziane, A.: AI and Financial Crime Prevention in Emerging Markets. Journal of Financial Regulation, vol. 9, 2023.

[2] Alvarez, M., Verdier, J.C., Nkashama, D.K., Frappier, M., Tardif, P.M., Kabanza, F.: A Revealing Large-Scale Evaluation of Unsupervised Anomaly Detection Algorithms. ACM Computing Surveys, vol. 55, no. 3 (2022).

[3] Aaronson, S.: Read the fine print. Nature Physics, vol. 11, pp. 291–293 (2015).

[4] Bao, Q., Wang, J., Gong, H., Zhang, Y., Guo, X., Feng, H.: A Deep Learning Approach to Anomaly Detection in High-Frequency Trading Data. IEEE Transactions on Neural Networks and Learning Systems (2025).

[5] Biamonte, J., Wittek, P., Pancotti, N., Rebentrost, P., Wiebe, N., Lloyd, S.: Quantum Machine Learning. Nature, vol. 549, pp. 195–202 (2017).

[6] Cardaioli, M., et al.: FD4QC: Application of Classical and Quantum-Hybrid Machine Learning for Financial Fraud Detection. Technical Report (2025).

[7] Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P.: SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research, vol. 16 (2002).

[8] Corli, S., et al.: Quantum Machine Learning Algorithms for Anomaly Detection: A Review. Future Generation Computer Systems, vol. 160, pp. 165–181 (2024).

[9] Dal Pozzolo, A., Caelen, O., Johnson, R.A., Bontempi, G.: Calibrating Probability with Undersampling for Unbalanced Classification. IEEE SSCI (2015). Dataset: Kaggle (ULB Machine Learning Group).

[10] Ezratty, O.: Where Are We Heading with NISQ? arXiv preprint arXiv:2305.09518 (2023).

[11] Grossi, M., et al.: Mixed Quantum-Classical Method for Fraud Detection with Quantum Feature Selection. IBM Research Technical Report (2022).

[12] Givhan, R., Maksymenko, M., et al. (Haiqu Inc.): Quantum Embeddings for Anomaly Detection: A New Lens on Data Representation. Technical Blog Report, November 2025. https://blog.haiqu.ai

[13] Havlíček, V., Córcoles, A.D., Temme, K., Harrow, A.W., Kandala, A., Chow, J.M., Gambetta, J.M.: Supervised learning with quantum-enhanced feature spaces. Nature, vol. 567, pp. 209–212 (2019).

[14] He, H., Garcia, E.A.: Learning from Imbalanced Data. IEEE Transactions on Knowledge and Data Engineering, vol. 21, no. 9 (2009).

[15] Innan, N., Marchisio, A., Shafique, M., Bennai, M.: QFNN-FFD: Quantum Federated Neural Network for Financial Fraud Detection. ACM Transactions on Quantum Computing, vol. 5, no. 2 (2024).

[16] IonQ Research Team: Quantum Support Vector Machines on Trapped-Ion Processors: Performance Analysis and Financial Applications. Quantum Information Processing, vol. 23 (2024).

[17] Liu, J., Wilde, F., Mele, A.A., Jiang, L., Eisert, J.: Noise-induced barren plateaus in variational quantum algorithms. Nature Communications, vol. 13, p. 3090 (2022).

[18] Ludmir, J.Z., Rebello, S., Ruiz, J., Patel, T.: Quorum: Zero-Training Unsupervised Anomaly Detection using Quantum Autoencoders. Proceedings of NeurIPS (2025).

[19] McClean, J.R., Boixo, S., Smelyanskiy, V.N., Babbush, R., Neven, H.: Barren plateaus in quantum neural network training landscapes. Nature Communications, vol. 9, p. 4812 (2018).

[20] Mellaerts, J.: Quantum-inspired Anomaly Detection: A QUBO Formulation. Quantum Machine Intelligence, vol. 5 (2023).

[21] Lopez-Rojas, E.A., Elmir, A., Axelsson, S.: PaySim: A Financial Mobile Money Simulator for Fraud Detection. EMSS, 2016. Dataset: Kaggle.

[22] Preskill, J.: Quantum Computing in the NISQ Era and Beyond. Quantum, vol. 2, p. 79 (2018).

[23] Schuld, M., Killoran, N.: Quantum Machine Learning in Feature Hilbert Spaces. Physical Review Letters, vol. 122, 040504 (2019).

[24] Schuld, M., Petruccione, F.: Machine Learning with Quantum Computers. Springer (2021).

[25] Ubale, R., Sujan, K.K., Deshpande, S., Byrd, G.T.: Toward Practical Quantum Machine Learning: A Novel Hybrid Quantum LSTM for Fraud Detection. Proceedings of ICML Workshop on Quantum Machine Learning (2025).

[26] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I.: Attention is All You Need. Advances in Neural Information Processing Systems (NeurIPS) (2017).

[27] Wang, M., et al.: A Parameter-Efficient Quantum Anomaly Detection Method on a Superconducting Quantum Processor. arXiv:2412.16867 (2024).

[28] World Economic Forum: Quantum Computing Governance Principles. WEF White Paper (2024).