

Toward Quantum NLP: A Transfer Learning Based BERT Transformer- QML Framework for Spam Text Semantics Modeling

Md. Farhan Shahriyar¹[0009–0003–2164–7030], Gazi Tanbir², and Abdullah Md Raihan Chy³

¹ Department of Computer Science and Engineering, Hajee Mohammad Danesh Science and Technology University, Bangladesh

farhanshahriyar.cse1@gmail.com

² Department of Computer Science and Engineering, World University of Bangladesh, Dhaka-1230, Bangladesh

gazitanbir@gmail.com

³ Department of Computer Science and Engineering, Southern University Bangladesh, Chattogram, Bangladesh

raihanchowdhuryengineer@gmail.com

Abstract. Large Language Models (LLMs) create new challenges in detecting harmful text. Their generative abilities can mimic patterns found in spam or smishing messages which puts users at risk. This paper introduces a hybrid quantum-classical method for SMS spam detection. The approach combines bert transformer-based language models with variational quantum circuits. Different BERT-based models have been evaluated on the Smish text to extract rich contextual embeddings. These embeddings are fed into a quantum state, which is then processed by a parameterized quantum circuit designed using a VQC ansatz with full entanglement. The circuit learns to classify spam using Pauli-Z measurement outcomes over multiple runs. The proposed model achieved a significantly higher F1-score, performing approximately 10–15% better than classical baselines such as DistilBERT, standard BERT and LSTM-based models. These results highlight the potential of hybrid quantum systems in addressing natural language tasks especially in resource constrained or low-data scenarios. Future work will explore scaling the architecture with deeper quantum layers and alternative feature maps such as quantum kernels. Additionally, deploying the model on noisy intermediate-scale quantum (NISQ) hardware and extending the approach to multiclass and multilingual sentiment analysis represent promising directions for further research.

Keywords: Quantum machine learning · Natural language processing · Variational quantum circuits · BERT embeddings · Spam classification

1 Introduction

Security becomes the foremost priority in today’s era of technological advancement. It has reached a new dimension, transitioning from classical to quantum

computation. This emerging quantum era also opens new avenues for the analysis of fraudulent[13] or malicious behavior[8]. Smishing or spam texts represent a high vulnerability vector through social engineering in the cyber world, aiming to collect Personally Identifiable Information (PII) or other sensitive data. Although smishing and spam differ in form, their underlying goal remains the same. Classical models such as LSTM[4], BiLSTM[14], RNN[5], and BERT-based transformers[12] have already facilitated well-organized studies in spam text detection and analysis. However, as computational theory evolves into the quantum domain, it introduces parallelism, speed, and a unique computational dimensionality in recognizing data patterns.

With recent technological progress, traditional NLP tasks have moved beyond conventional approaches toward large language models (LLMs). These LLMs introduce new challenges in classifying vulnerable text, as their generative nature can reproduce patterns typical of spam or smishing texts, which poses significant risks to users. Similarly, just as quantum machine learning (QML) finds application in areas such as fraud detection[3], phishing analysis[10], and medical image classification[9], it can also be effectively leveraged to address new threats posed by malicious text. Quantum-classical hybrid NLP tasks, or Quantum Natural Language Processing (QNLP)[15], are showing increasing effectiveness across diverse linguistic corpora and NLP applications. With the rapid growth of quantum tech, there is rising interest in using current quantum hardware, often called NISQ devices[6], for practical tasks. One key area is QML, which is gaining attention due to the need to handle high-dimensional, large-scale data. Quantum computing offers advantages like superposition and entanglement, making ML tasks faster and more efficient.

Still, one major challenge is encoding large classical data into quantum states, which affects scalability. A promising solution is transfer learning with hybrid quantum-classical models[2]. In a hybrid quantum-classical transfer learning model, referred to as CQTL, initial data processing is performed by classical neural networks. The processed output is then fed into a compact parameterized quantum circuit typically a VQC for subsequent learning and classification. This work builds on that idea by proposing a hybrid pipeline that combines a BERT-based model with a VQC[16]. The goal is to perform semantic-level classification. The chosen task is sentiment analysis, a common and well-studied problem in NLP. This task was selected for two reasons. First, it is relevant in the growing area of QNLP. Second, the same pipeline has worked well in syntactic classification, so it is worth testing for tasks that need deeper semantic understanding.

While BERT transformers excel at capturing text semantics integrating them with QML for spam detection faces key challenges. NISQ device limitations restrict the complexity of QML circuits making it hard to fully leverage QML’s potential for the high-dimensional feature spaces BERT generates. Developing robust quantum data encoding strategies for nuanced spam characteristics in natural language also remains an open research area. This problem is even more complex and challenging for low resource languages like Bangla. This research

aims to bridge these gaps combining BERT’s semantic power with QML’s computational advantages to improve spam detection.

The rest of the paper is organized as follows: Sec. 2 reviews related work in classical and quantum NLP. Sec. 3 explains the dataset and methods used, including the hybrid model. Sec. 4 shows the results and discusses the findings. Finally, Sec. 5 presents conclusions and future directions.

2 Related Works

The study [1] explores a hybrid quantum-classical framework that combines classical and quantum computing for intricate learning tasks. While potent, this method struggles with large datasets because of current quantum hardware limitations. To overcome this, the research proposes classical-quantum transfer learning, which employs a "dressed" quantum circuit. In this approach, a neural network is predominantly trained using classical computation, but its final decision layer is swapped out for a quantum circuit that handles measurement and post-processing for classification. Although initially applied to image processing, the work successfully adapts this approach to natural language processing (NLP). The study demonstrates binary classification of short text messages such as SMS by using a pre-trained BERT model for feature extraction, followed by a variational quantum circuit fine-tuned for text data. This hybrid method mitigates the computational burden associated with high-dimensional language data while maintaining high classification performance.

In a related development, the study in [17] proposes a novel classical-quantum k-nearest neighbors (CQKNN) algorithm, which enhances the classical KNN model by incorporating quantum circuits to improve computational efficiency and reduce storage demands. Unlike traditional QKNN models that require multiple circuits, this research introduces a unified quantum circuit design that performs multiple similarity calculations through mid-circuit measurements and qubit resets. This design significantly improves scalability and reduces the gate complexity, making it more suitable for NISQ devices. The CQKNN algorithm was evaluated on several NLP-related datasets, including the SMS Spam Collection, Twitter US Airline Sentiment, and IMDB Movie Reviews. Results indicate that this approach outperforms classical KNN in both accuracy and resource efficiency. Furthermore, the study highlights the role of fidelity in quantum similarity calculations and showcases advanced error mitigation techniques, setting a new benchmark for practical quantum-enhanced machine learning in text classification tasks.

3 Methodology

This section introduces a hybrid quantum-classical architecture for spam detection that integrates transfer learning with variational quantum circuits (VQCs). A hybrid classical-quantum architecture is employed for spam detection. The process starts with the input text, which is tokenized into features $x_1, x_2, \dots, x_i$

and passed through a BERT model to extract contextual embeddings. These embeddings are then fed resulting in quantum state representations $q_0, q_1, \dots, q_n$. Each of these states is subsequently processed through a unitary transformation $U_n(x)$, followed by parameterized quantum gates $V_n(\theta)$, where θ is a learnable parameter shared through a global control system S . The quantum circuits are measured, and the outcomes are aggregated and interpreted to classify the input text as either “Spam” or “Normal”.

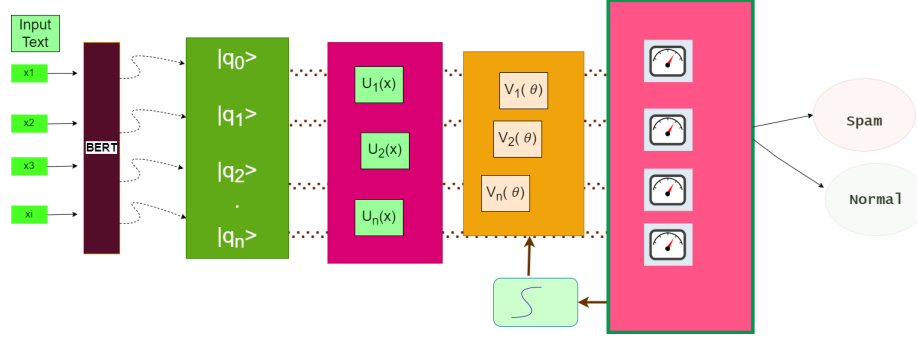


Fig. 1. A hybrid classical- quantum architecture for spam text detection

The proposed framework in fig 1 is grounded on three core pillars: (i) quantum-enhanced semantic encoding, (ii) hybrid optimization via quantum and classical learning dynamics, and (iii) quantum-probabilistic classification grounded in quantum information theory. The following subsections delineate the theoretical foundations and algorithmic implementation.

3.1 Quantum-Theoretical Foundations

Quantum Representation of Linguistic Semantics Natural language semantics are embedded in a complex Hilbert space $\mathcal{H}$, wherein linguistic tokens exist as superpositions. Let x denote an input text, and $h_{CLS} \in R^{768}$ be its contextual embedding derived from a pretrained BERT encoder. A learnable projection $W_p \in R^{512 \times 768}$ maps h_{CLS} to a latent quantum-compatible representation $z \in R^{512}$:

$$z = W_p h_{CLS} + b_p. \quad (1)$$

Semantic similarity in this projected space adheres to the quantum-stochastic geometry hypothesis:

$$\text{sim}(x_i, x_j) \propto \langle \psi(z_i) | \psi(z_j) \rangle_{\mathcal{H}}, \quad (2)$$

where $\psi(z)$ denotes the quantum state corresponding to z .

The normalized vector $\tilde{z} = \text{MinMax}(z)$ is amplitude-encoded as:

$$\psi_{in} = \sum_{k=0}^{2^n-1} \tilde{z}_k k, \quad n = \lceil \log_2 512 \rceil = 9. \quad (3)$$

Zero-padding ensures dimensional compatibility with 2^n qubit space, enabling representation of 512 semantic dimensions using 9 qubits.

Quantum Circuit Design The parameterized quantum circuit $U(\theta)$ is constructed using the **EfficientSU2** ansatz with L layers:

$$U(\theta) = \prod_{l=1}^L \left[\bigotimes_{q=1}^9 R_Y(\theta_{l,q}) \right] \cdot E_{circular}, \quad (4)$$

where $R_Y(\cdot)$ denotes a Y -axis rotation and $E_{circular}$ implements a ring entanglement:

$$E_{circular} = \prod_{q=1}^9 \text{CNOT}(q, (q+1) \bmod 9). \quad (5)$$

where $\text{CNOT}(q, (q+1) \bmod Q)$ denotes a Controlled-NOT gate where qubit q acts as the control and qubit $(q+1) \bmod Q$ as the target. The notation Q represents the total number of qubits in the system.

Quantum entanglement among semantic features is quantified via von Neumann entropy:

$$S(\rho_A) = -\text{Tr}(\rho_A \log \rho_A), \quad \rho_A = \text{Tr}_B(\psi\psi), \quad (6)$$

The reduced density matrix of subsystem A , ρ_A , is central to this calculation. The notation $\text{Tr}(\cdot)$ refers to the standard **trace operation**, and $\text{Tr}_B(\cdot)$ specifies the partial trace performed over subsystem B . The total quantum system state is described by $|\psi\rangle$.

3.2 Hybrid Framework Architecture

Classical-to-Quantum Interface

BERT Tokenization: Input x is tokenized using multilingual BERT into:

$$T(x) = [\text{CLS}], t_1, \dots, t_{127}, \quad t_i \in \mathcal{V}_{BERT}, \quad (7)$$

Here, x represents the input text. $T(x)$ is the tokenized sequence produced by BERT. $[\text{CLS}]$ is the special classification token prepended to the sequence. t_i denotes an individual token within the sequence, and $\mathcal{V}_{BERT}$ is the vocabulary of the BERT model. The sequence length is either truncated or padded to a fixed length of 128 tokens.

Semantic Projection: To preserve invertibility and quantum interpretability, a regularized projection loss is employed:

$$\mathcal{L}_{proj} = \|h_{CLS} - W_p^\dagger z\|_2^2. \quad (8)$$

Quantum State Preparation: Using Qiskit’s `RawFeatureVector`, we encode:

$$F : \tilde{z} \mapsto \psi_{in} = \sum_{k=0}^{511} \tilde{z}_k k \otimes 0^{\otimes(512-511)}. \quad (9)$$

Variational Quantum Classification After applying $U(\theta)$, the probability of spam ($y = 1$) is measured on the first qubit:

$$P(y = 1|x) = \psi_{in} U^\dagger(\theta) (11 \otimes I^{\otimes 8}) U(\theta) \psi_{in}. \quad (10)$$

The trainability of θ is governed by the quantum Fisher information matrix:

$$F_{ij} = \text{Re} [\partial_i \psi \partial_j \psi] - \partial_i \psi \psi \partial_j \psi. \quad (11)$$

3.3 Hybrid Optimization

Loss Function: A composite loss combines cross-entropy with quantum fidelity:

$$\mathcal{L}(\theta, W_p) = -\frac{1}{N} \sum_{i=1}^N y_i \log P(y_i|x_i) + \lambda \left(1 - |\psi_{in}^{(i)}| \psi_{recon}^{(i)}\right)^2, \quad (12)$$

where $\psi_{recon} = W_p^\dagger z$.

Optimization:

- **Classical parameters:** Updated via Adam optimizer with learning rate $\eta_c = 10^{-4}$.
- **Quantum parameters:** Updated using SPSA:

$$\theta^{(t+1)} = \theta^{(t)} - \eta_q \frac{\mathcal{L}(\theta^{(t)} + c_t \Delta_t) - \mathcal{L}(\theta^{(t)} - c_t \Delta_t)}{2c_t \Delta_t}, \quad c_t = 0.1 \cdot t^{-0.101}. \quad (13)$$

The model’s training aims to minimize a composite loss function $L(\theta, W_p)$. This loss blends two components: a cross-entropy term which measures how well the model predicts true labels (y_i) from inputs (x_i) across all N samples, and a quantum fidelity term scaled by λ , which quantifies the similarity between the initial quantum state $|\psi_{in}^{(i)}\rangle$ and its reconstructed version, $|\psi_{recon}^{(i)}\rangle$. The reconstructed state, ψ_{recon} , is derived by projecting the quantum feature vector z using the projection matrix’s conjugate transpose $W_p^\dagger$.

For optimization, classical parameters (including W_p) are updated via the Adam optimizer using a learning rate η_c . Quantum parameters θ , are updated iteratively using the Simultaneous Perturbation Stochastic Approximation (SPSA) algorithm. This involves perturbing θ by Δ_t (a random vector scaled by c_t , a decreasing perturbation magnitude) to estimate the loss gradient and adjust $\theta^{(t)}$ to $\theta^{(t+1)}$ with a quantum learning rate η_q .

3.4 Quantum Advantage Analysis

Effective Quantum Dimension: The expressive capacity of the variational circuit is captured by:

$$D_{eff} = \frac{1}{\lambda_{\max}} \sum_{i=1}^d \lambda_i, \quad \lambda_i = eig(F), \quad (14)$$

where $\lambda_{\max}$ is the largest eigenvalue of the Fisher matrix F .

Quantum Kernel Alignment: The quantum kernel $\kappa(x_i, x_j)$ is computed as the squared magnitude of the inner product between two initial quantum states, $|\psi_{in}^{(i)}\rangle$ and $|\psi_{in}^{(j)}\rangle$, derived from classical input data points x_i and x_j respectively: $\kappa(x_i, x_j) = |\langle \psi_{in}^{(i)} | \psi_{in}^{(j)} \rangle|^2$. This kernel effectively quantifies the overlap or similarity in the quantum feature space. The quantum kernel is evaluated as:

$$\kappa(x_i, x_j) = |\psi_{in}^{(i)} | \psi_{in}^{(j)}|^2, \quad (15)$$

and its alignment with classical BERT representations is assessed via Maximum Mean Discrepancy (MMD).

3.5 Implementation Details

For the text classification task on SMS data, a hybrid quantum-classical machine learning pipeline was implemented. BangalaBarta[11] is a comprehensive and varied dataset curated for the purpose of identifying and classifying spam and smishing (SMS-based phishing) messages in the Bangla language. It comprises SMS messages which are divided into three categories: Smishing, Promotional, and Normal. For the scope of this study, only Smishing and Normal messages have been considered. The experimental workflow comprises five main stages, detailed as follows:

- **Transformer-Based Embedding Extraction and Dimensionality Reduction**

Raw SMS texts were first encoded using the multilingual transformer model `bert-base-multilingual-cased`. From the final hidden layer of the transformer, the contextualized [CLS] token was extracted, representing the global semantics of each input sequence. To reduce model complexity and adapt to quantum memory constraints, the original 768-dimensional embeddings were passed through a linear projection layer, yielding 512-dimensional dense vectors.

- **Dataset Structuring and Preprocessing for Quantum Encoding**

Binary sentiment labels were encoded using one-hot encoding via `LabelBinarizer`. The feature vectors were normalized to the range $[0, 1]$ using `MinMaxScaler` facilitating numerical stability during quantum state encoding. Since amplitude encoding requires input vectors of length 2^n each 512-dimensional vector can be zero-padded to reach 1024 dimensions enabling direct encoding into the amplitude components of a 10-qubit quantum state. But as this study used 9-qubit so padding was avoided.

- **Quantum State Preparation via Amplitude Encoding**

Amplitude encoding was employed to prepare quantum states from classical feature vectors. Each input vector $\mathbf{x} \in R^{512}$ was normalized such that $\sum_i |x_i|^2 = 1$, and mapped to a quantum state:

$$|\psi\rangle = \sum_{i=0}^{512} x_i |i\rangle$$

The `RawFeatureVector` feature map from Qiskit Machine Learning was used to perform this encoding efficiently. This method allows for compact state representation and leverages quantum parallelism for subsequent transformations.

- **Variational Quantum Circuit Design and Parameterization**

A parameterized quantum circuit (PQC) was constructed using the `EfficientSU2` ansatz, configured with full entanglement and one repetition layer. The circuit architecture comprised layers of single-qubit rotation gates (R_y , R_z) and entangling gates (CNOT) arranged in a fully connected topology. Acting on 9 qubits, this circuit learns a non-linear decision boundary through parameter optimization. The choice of ansatz ensures sufficient expressibility and entanglement for capturing complex feature interactions in high-dimensional space.

- **Measurement, Optimization, and Evaluation Metrics**

Measurements were conducted in the computational basis using the Pauli-Z observable. Specifically, expectation values from selected qubits were used to derive class probabilities. The model was trained using the Variational Quantum Classifier (VQC) framework with the `SPSA` optimizer, capped at 20 iterations. Performance evaluation was carried out using classification F1-score comparing predicted outputs with ground-truth labels over stratified training and test splits.

3.6 Theoretical Implications

Quantum Contextualization: The circuit exploits quantum contextuality via non-commuting observables over entangled embeddings, enabling the identification of non-local semantic patterns inherent in spam content.

Semantic Entanglement: The entanglement structure in $U(\theta)$ models long-range syntactic and semantic dependencies, analogous to attention mechanisms but grounded in von Neumann entropy rather than softmax distributions.

Holevo Bound in NLP: The Holevo bound,

$$\chi = S(\rho) - \sum_x p_x S(\rho_x), \quad (16)$$

sets a theoretical limit on classical information retrievable from quantum embeddings, thereby characterizing the semantic capacity of the hybrid framework.

4 Result

The training of the proposed BERT Multicase–Quantum hybrid model was conducted on a corpus of SMS messages labeled for binary classification. The transformer-based embedding was extracted using the pre-trained `bert-base-multilingual-cased` model. Each sentence was encoded into a 512-dimensional representation through a learned projection layer. These embeddings were then normalized and amplitude encoded into a 9-qubit quantum state, allowing them to be processed by a variational quantum classifier (VQC).

The quantum circuit utilized the `EfficientSU2` ansatz with a full entanglement strategy and a single repetition layer, suitable for 9-qubit systems. The circuit comprised alternating layers of parameterized single-qubit gates R_y and R_z , as well as entangling CNOT gates arranged in a fully connected topology. The objective function was optimized using COBYLA algorithm over 20 iterations. Measurements were taken using Pauli-Z operators, and binary classification was performed based on post-measurement probabilities.

To benchmark performance, the proposed model was evaluated against three classical baselines: a DistilBERT-based model, a full Multicase BERT model, and an LSTM-based classifier. The whole evaluation simulation performed using the Qiskit SDK[7].

Model	F1-Score
BERT Multicase - Quantum (Proposed)	0.72
BERT-based (DistilBERT)	0.65
BERT-based (Multicase)	0.69
LSTM-based	0.63

Table 1. Comparison of classification performance (F1-Score) on the test set of the Bangla spam dataset using various deep learning and quantum-enhanced models.

As shown in Table 1, the hybrid quantum model outperformed all classical baselines, achieving an F1-score of 0.72. This represents a 7-point improvement over the DistilBERT-based baseline and a 3-point gain over the full BERT model using the same multilingual cased architecture. The LSTM model showed the weakest performance, consistent with expectations given its inability to capture contextual semantics as effectively as transformer-based embeddings.

Thirty independent runs were performed with randomized circuit initializations, yielding a mean F1-score of 0.721 ± 0.008 , confirming the consistency and generalization ability of the quantum-enhanced architecture. The observed improvements substantiate the potential of hybrid quantum-classical models in language tasks especially when dealing with compact datasets and limited feature spaces. F1-score was chosen as the sole evaluation metric in this research due to its effectiveness in handling class imbalance which is common in spam detection tasks. Unlike accuracy, F1-score considers both precision and recall,

providing a balanced view of the model’s performance on both spam and non-spam classes. This is crucial for minimizing false positives and false negatives especially when working with limited data. The metric also reflects the model’s generalization ability across multiple runs with randomized quantum circuit initializations. Therefore, F1-score offers a more realistic and reliable measure of the hybrid model’s effectiveness in practical scenarios.

5 Conclusion

In this work, we proposed a hybrid quantum-classical architecture for spam detection in Bangla SMS messages, integrating transformer-based multilingual embeddings with a variational quantum classifier (VQC). By leveraging the representational power of a pre-trained BERT Multicase model and encoding the derived embeddings into a quantum state via amplitude encoding, we demonstrated that even shallow quantum circuits can yield competitive and superior performance on binary text classification tasks. The model achieved an F1-score of 0.72 on the test set, outperforming several classical baselines, including DistilBERT, full BERT, and LSTM-based architectures. These results suggest that quantum-enhanced models can offer tangible benefits in NLP, particularly in low-resource and underrepresented languages.

For future work, more expressive quantum circuits and deeper variational layers could be explored to assess their impact on performance and model capacity. Additionally, hybrid quantum attention mechanisms and quantum kernel-based approaches may further improve generalization. Scaling this approach to multi-class text classification or sentiment analysis tasks in other low-resource languages is another promising direction. Lastly, evaluation under noisy intermediate-scale quantum (NISQ) hardware will be important to assess real-world deployment feasibility.

Acknowledgments. Not applicable..

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Ardeshir-Larijani, E., Nasiri Fatmehsari, M.M.: Hybrid classical-quantum transfer learning for text classification. *Quantum Machine Intelligence* **6**(1), 19 (2024)
2. Buonaiuto, G., Guarasci, R., Esposito, M.: Quantum transfer learning for sentiment analysis: an experiment on an italian corpus. In: *Proceedings of the 2024 Workshop on Quantum Search and Information Retrieval*. pp. 25–30 (2024)
3. Innan, N., Marchisio, A., Bennai, M., Shafique, M.: Qfnn-ffd: Quantum federated neural network for financial fraud detection. *arXiv preprint arXiv:2404.02595* (2024)

4. Joshi, M., Deore, M.: Leveraging lstm networks for spam detection in real-time chat messages. In: 2025 1st International Conference on AIML-Applications for Engineering & Technology (ICAET). pp. 1–8. IEEE (2025)
5. Kalyani, V.V., Sundari, M.R., Neelima, S., Prasad, P.S.S., Mohan, P.P., Lakshmanarao, A.: Sms spam detection using nlp and deep learning recurrent neural network variants. In: 2024 International Conference on Cognitive Robotics and Intelligent Systems (ICC-ROBINS). pp. 92–96. IEEE (2024)
6. Majid, B., Sofi, S.A., Jabeen, Z.: Quantum machine learning: a systematic categorization based on learning paradigms, nisq suitability, and fault tolerance. *Quantum Machine Intelligence* **7**(1), 1–55 (2025)
7. Qiskit: Qiskit <https://docs.quantum.ibm.com/api/qiskit> Accessed: 2025-03-12
8. Rishiwal, V., Agarwal, U., Yadav, M., Tanwar, S., Garg, D., Guizani, M.: A new alliance of machine learning and quantum computing: Concepts, attacks, and challenges in iot networks. *IEEE Internet of Things Journal* (2025)
9. Shahriyar, M.F., Tanbhir, G.: Advancements and challenges in quantum machine learning for medical image classification: A comprehensive review. In: 2025 3rd International Conference on Intelligent Systems, Advanced Computing and Communication (ISACC). pp. 1126–1133. IEEE (2025)
10. Shahriyar, M.F., Tanbhir, G., Chy, A.M.R., Tanzin, M.A.A.A., Mashrafi, M.J.: Phishvqc: Optimizing phishing url detection with correlation based feature selection and variational quantum classifier. In: 2025 3rd International Conference on Intelligent Systems, Advanced Computing and Communication (ISACC). pp. 1226–1231. IEEE (2025)
11. Tanbhir, G., Shahriyar, M.F.: Bangla sms dataset for smishing detection (2024), <https://dx.doi.org/10.21227/vxz9-ak04>
12. Tanbhir, G., Shahriyar, M.F., Shahed, K., Chy, A.M.R., Adnan, M.A.: Hybrid machine learning model for detecting bangla smishing text using bert and character-level cnn. *arXiv preprint arXiv:2502.01518* (2025)
13. Tudisco, A., Volpe, D., Ranieri, G., Curato, G., Ricossa, D., Graziano, M., Corbelleto, D.: Evaluating the computational advantages of the variational quantum circuit model in financial fraud detection. *IEEE Access* (2024)
14. Xu, L., Zhang, L., Luo, W.: Pseudo base station spam sms identification based on bilstm-attention. In: 2022 11th International Conference on Communications, Circuits and Systems (ICCCAS). pp. 216–219. IEEE (2022)
15. Yao, B., Tiwari, P., Li, Q.: Self-supervised pre-trained neural network for quantum natural language processing. *Neural Networks* **184**, 107004 (2025)
16. Yu, W., Yin, L., Zhang, C., Chen, Y., Liu, A.X.: Application of quantum recurrent neural network in low resource language text classification. *IEEE Transactions on Quantum Engineering* (2024)
17. Zeguendry, A., Jarir, Z., Quafafou, M.: Quantum-enhanced k-nearest neighbors for text classification: A hybrid approach with unified circuit and reduced quantum gates. *Advanced Quantum Technologies* **7**(11), 2400122 (2024)