QuantumNLP 2025

The 1st Workshop on QuantumNLP

Proceedings of the 1st Workshop on QuantumNLP

November 24, 2025

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
317 Sidney Baker St. S
Suite 400 - 134
Kerrville, TX 78028
USA
Tel: +1-855-225-1962
acl@aclweb.org

ISBN 979-8-89176-306-7

Introduction

We are pleased to present the proceedings of the **1st Workshop on QuantumNLP: Integrating Quantum Computing with Natural Language Processing**, which was successfully **held** on **November 24, 2025**, as a satellite event of the 14th IJCNLP-AACL 2025 in Mumbai, India. This inaugural workshop was conducted in a **hybrid format**, which allowed participants to join us both in person in Mumbai and virtually from around the globe.

The QNLP Workshop served as a premier venue for interdisciplinary research at the intersection of Quantum Computing and Natural Language Processing. Our aim **was** to bring together experts to discuss foundational concepts and cutting-edge developments in harnessing quantum computational paradigms to revolutionize complex NLP tasks.

We **received** a total of **12 submissions** this year, reflecting the strong and growing interest in this emerging field. Every submission **was assigned** to the Technical Program Committee and **received** thorough review and consideration. Following a rigorous evaluation process, we **accepted 9 papers** for presentation, resulting in an overall acceptance rate of **75%**. This rate allowed us to foster nascent, high-potential research and encourage contributions in this complex and rapidly evolving domain.

The papers selected for the program covered a diverse array of topics central to Quantum NLP research. Themes **included** the mathematical underpinnings of quantum information, novel Quantum Machine Learning (**QML**) algorithms, the application of quantum word embeddings, and the development of Hybrid Quantum-Classical Algorithms for sequence modeling and language tasks. The structure of the hybrid event successfully accommodated both physical and remote presenters, ensuring a high-quality interactive experience for all who attended.

A workshop of this technical depth required the dedicated effort of many individuals, and we extend our sincere gratitude to all of them. We thank the **Technical Program Committee (TPC)** members for committing their valuable time and expertise to the crucial task of reviewing and guiding the selection process, which ensured the high quality of our technical program. We are also grateful to our organizers for their work in adapting the event logistics to successfully deliver the hybrid experience.

Finally, we thank all the authors who submitted their fine work to the workshop and all participants for joining us—whether physically or virtually—and for contributing to the successful launch and growth of the Quantum Natural Language Processing research community.

Santanu Pal, Program Chair
Partha Pakray, Program Chair
Sivaji Bandyopadhyay, Program Chair

QuantumNLP 2025 Workshop

Organizing Committee

Program Chairs / Organizers

Partha Pakray, NIT Silchar
Priyanka Jain, CDAC, Delhi
Asif Ekbal, IIT Jodhpur
Santanu Pal, Wipro Innovation Network, Wipro Ltd
Sivaji Bandyopadhyay, Jadavpur University

Developers

Tahmid Choudhury, NIT Silchar

Program Committee

Program Chairs

Santanu Pal, Wipro Innovation Network, Wipro Ltd
Partha Pakray, National Institute of Technology Silchar
Asif Ekbal, Indian Institute of Technology, Jodhpur
Priyanka Jain, CDAC, Delhi
Sivaji Bandyopadhyay, Jadavpur University

Reviewers

Dipankar Das, Jadavpur University
Sandeep Kumar Dash, National Institute of Technology Mizoram
Riyanka Manna, Amrita Vishwa Vidyapeetham (Deemed University)
Partha Pakray, National Institute of Technology Silchar
Santanu Pal, Wipro Innovation Network, Wipro Ltd
Shyambabu Pandey, National Institute of Technology Silchar
Apurbalal Senapati, Central Institute of Technology, Kokrajhar
Advaitha Vetagiri, National Institute of Technology Silchar
Annepaka Yadagiri, National Institute of Technology Silchar

Keynote Talk

Quantum Machine Learning: Concepts and Applications

Hachem Kadri
Aix-Marseille University, France
2025-11-24 10:00:00 – Room: **Hybrid Session**

Abstract: Quantum Machine learning is an emerging field of research, with fast growth. It is largely driven by the desire to develop artificial intelligence that leverage quantum technologies to enhance the speed and capabilities of learning algorithms. In this talk, I will begin by outlining the main concepts and motivations behind QML and by presenting various forms of interaction between machine learning, quantum computing and quantum information. I will then illustrate these interactions through concrete examples, focusing in particular on quantum extensions of classical ML models such as linear regression and the perceptron.

Bio: Professor Hachem Kadri is a Professor of Artificial Intelligence in the Department of Computer Science at Aix-Marseille University and a member of the Machine Learning group QARMA at LIS Lab, France.
His research interests lie broadly in machine learning, covering topics such as kernel methods, functional data analysis, statistical learning theory, deep learning, and, more recently, quantum machine learning (QML). He is the Principal Investigator (PI) for the ANR Starting Grant project, QuantML — Quantum Machine Learning: Foundations and Algorithms (2019–2024). Prof. Kadri's recent work includes publications on C*-algebraic ML and the computational-statistical tradeoffs of the Quantum Perceptron. He has held positions as an Assistant Professor at Aix-Marseille University and as a postdoctoral researcher at INRIA Lille – Nord Europe. He received his Ph.D. in Electrical Engineering in 2008 from the National Engineering School of Tunis (ENIT).

Table of Contents

Quantum-Infused Whisper: A Framework for Replacing Classical Components

Tapabrata Mondal[1], Debjit Dhar[1], Soham Lahiri[1], Sivaji Bandyopadhyay[1*]
[1]Jadavpur University, Kolkata, India
[*]**Correspondence:** sivaji.cse.ju@gmail.com

Abstract

We propose a compact hybrid quantum–classical extension of OpenAI's Whisper in which classical components are replaced by Quantum Convolutional Neural Networks (QCNN), Quantum LSTMs (QLSTM), and optional Quantum Adaptive Self-Attention (QASA). Log-mel spectrograms are angle-encoded and processed by QCNN kernels, whose outputs feed a Transformer encoder, while QLSTM-based decoding introduces quantum-enhanced temporal modeling. The design incorporates pretrained acoustic embeddings and is constrained to NISQ-feasible circuit depths and qubit counts. Although this work is primarily architectural, we provide a fully specified, reproducible evaluation plan using Speech Commands, LibriSpeech, and Common Voice, along with strong classical baselines and measurable hypotheses for assessing noise robustness, efficiency, and parameter sparsity. To our knowledge, this is the first hardware-aware, module-wise quantum replacement framework for Whisper.

1 Introduction

Quantum Natural Language Processing (QNLP) and Quantum Automatic Speech Recognition (QASR) explore how quantum information processing can enhance representation, inference, and learning for language and speech. Prior work suggests that quantum models may offer richer expressivity for structured linguistic tasks (Wiebe et al., 2019) and improved efficiency for operations that are expensive in classical deep learning. Early demonstrations, ranging from compositional distributional models compiled with toolkits such as `lambeq` (Kartsaklis et al., 2021) to QCNN-based speech pipelines have shown encouraging results but are typically limited to small datasets and shallow circuits due to NISQ constraints.

Current quantum hardware still imposes strict limits on circuit depth, qubit count, and data encoding, and full quantum replacements for attention, beam search, or large-scale sequence modeling remain largely unexplored. As a result, hybrid architectures that combine quantum modules with established classical components offer a practical interim path for advancing quantum-enhanced ASR.

In this work, we propose a unified quantum–augmented extension of Whisper in which classical convolution, recurrent, and attention blocks can be replaced with Quantum Convolutional Neural Networks (QCNN), Quantum LSTMs (QLSTM), and Quantum Adaptive Self-Attention (QASA). Our design is explicitly hardware-aware, specifying qubit requirements, depth-constrained variational layers, and angle-encoding strategies compatible with current NISQ devices. We further provide a rigorous and reproducible evaluation roadmap, including datasets, baselines, and measurable hypotheses, to quantify the potential benefits of quantum modules in robustness, sparsity, and low-resource performance. Finally, we outline the feasibility of implementing these components on present hardware through hybrid training, parameter-shift optimization, and noise-mitigation techniques. Our principal contributions are the following:

1. **Modular, integrable quantum replacements.** A hardware-aware framework that replaces Whisper's convolutional, recurrent and (optionally) attention blocks with QCNN, QLSTM and QASA modules — including concrete integration patterns (e.g., QLSTM gates inside a Transformer-style decoder with quantum outputs mapped to standard gating nonlinearities) and fallback hybrid strategies for QASA.

2. **NISQ-feasible specification + transfer-learning.** Per-module NISQ constraints (qubit budgets, circuit-depth limits, entangling

Proceedings of the QuantumNLP: Integrating Quantum Computing with Natural Language Processing, pages 1–5

topologies, measurement channels, and angle/amplitude encoding) combined with pretrained acoustic embeddings so quantum layers refine high-level features have been provided.

2 Related Work

Quantum approaches to speech and language have expanded across recognition, classification, and generation. Miller et al. (Miller et al., 2024) fused STFT and LPC spectrograms, processing the LPC branch with a variational quantum circuit (VQC) before CNN-based classification, achieving 94.54% accuracy on Speech Commands (vs. 93.05% classical), with improved robustness and storage efficiency. Thejha et al. (Thejha et al., 2023) proposed a QCNN with CNOT gates and parameterized rotations (SX, SY, SZ) in Qiskit, reaching 99.10% accuracy for accent recognition (vs. 98.8% CNN). Wang et al. (Wang et al., 2023) combined WavLM-Large embeddings with a low-dimensional VQC for synthetic speech detection, improving equal error rate to 5.51% (vs. 6.80% baseline), highlighting the utility of quantum–embedding coupling.

In NLP, Yang et al. (Yang et al., 2022) introduced BERT-QTC, pairing a pretrained encoder with a quantum temporal convolution layer to enable federated learning privacy while improving intent classification accuracy (96.6% vs. 95.0%). Di Matteo et al. (Di Sipio et al., 2022) surveyed quantum-augmented NLP, showing QLSTMs and quantum Transformers achieve classical-level accuracy with fewer parameters, suggesting VQCs as efficient dense-layer replacements. Yang et al. (Yang et al., 2021) built a decentralized ASR pipeline where Mel-spectrograms pass through 2×2 QCNN kernels before a BiLSTM-attention model, reaching 95.12% accuracy with compact architectures.

Earlier, Fu et al. (Fu and Dai, 2009) integrated QNNs with particle-swarm optimization, reporting 84.5–85% accuracy on small-vocabulary tasks with faster, noise-resilient training. Pandey et al. (Pandey et al., 2023) introduced QLSTMs replacing gates with VQCs, outperforming classical LSTMs on code-mixed text but raising overfitting concerns. Abbaszade et al. (Abbaszade et al., 2023) applied DisCoCat-based quantum circuits to machine translation, achieving low error (MSE=0.0019) on English–Persian. Yoshimura et al. (Yoshimura et al., 2018) improved neural vocoders like WaveNet via mel-cepstrum quantization shaping, yielding a 0.6 MOS gain and 4 dB Equivalent-Q improvement with efficient MLSA filters.

Overall, these studies demonstrate the potential of hybrid quantum–classical methods for speech and NLP, spanning spectrogram fusion, quantum frontends, transfer learning, federated privacy, and model compression. In contrast, our work embeds parameterized quantum circuits directly into both feature extraction and decoding, integrates large pretrained acoustic embeddings for full transcription (not just classification/detection), and evaluates a cohesive end-to-end quantum–classical ASR pipeline on standard benchmarks, extending beyond earlier proof-of-concept systems.

3 Methodology

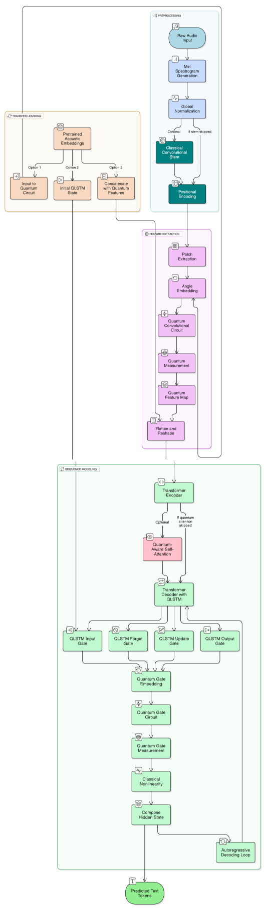


Figure 1: Architecture of the Quantum-Augmented Whisper pipeline. Log-mel patches are angle-encoded and processed by QCNN kernels that refine pretrained acoustic embeddings before a Transformer encoder; decoding uses QLSTM (VQCs replacing LSTM linear transforms) with optional QASA attention projections and a classical token head.

3.1 Feature Extraction with Quantum Convolutional Layers

An overview of the proposed Quantum-Augmented Whisper is shown in Figure 1, combining quantum and classical modules within the ASR pipeline. Raw audio at 16kHz is converted into an 80-channel log-Mel spectrogram using a 25ms window and 10ms stride, normalized and optionally processed by a lightweight convolutional stem with ReLU or GELU activations and positional encoding. For feature extraction, instead of classical CNNs we employ a Quantum Convolutional Neural Network (QCNN) (Yang et al., 2021), where 2×2 spectrogram patches are angle-encoded into 4-qubit states and processed by variational circuits with trainable rotations (R_X, R_Y, R_Z) and CNOT entanglement. Pauli-Z expectation values provide the quantum features, acting as trainable kernels that replace classical filters and are assembled into a quantum-enhanced feature map. This approach introduces stochasticity from measurement and exploits entanglement to capture local dependencies more effectively, particularly in low-data settings. While kernel sizes of 1×1 to 3×3 are considered, prior work indicates 2×2 offers the best trade-off. The resulting feature map is flattened into a temporal sequence for downstream modeling.

3.2 Whisper-Inspired Transformer Decoder with QLSTM Layers

Mathematical Foundation of Quantum LSTM Gates: Quantum Long Short-Term Memory (QLSTM) (Pandey et al., 2023) extends classical LSTMs by replacing linear transformations in gate computations with variational quantum circuits (VQCs). For each gate $g \in \{f, i, \tilde{C}, o\}$,

$$g_t^{(q)} = \sigma\left(\mathrm{VQC}_g\left([h_{t-1}, x_t]; \theta_g\right)\right), \tag{1}$$

where $[h_{t-1}, x_t]$ is the concatenated vector of previous hidden state and input, and θ_g are circuit parameters.

VQC Architecture: Each variational circuit operates in three stages:

1. *Encoding:* Inputs are mapped via angle encoding:

$$|\psi_{\text{enc}}\rangle = \bigotimes_{i=1}^{n} \Big(\cos \frac{\arctan v_{t,i}}{2} |0\rangle_i + \sin \frac{\arctan v_{t,i}}{2} |1\rangle_i \Big) \tag{2}$$

2. *Variational Layer:* For L layers and n qubits:

$$U_{\text{var}}(\theta) = \prod_{l=1}^{L} \left[\prod_{i=1}^{n} R(\alpha_{i,l}, \beta_{i,l}, \gamma_{i,l}) \prod_{\langle i,j \rangle} \mathrm{CNOT}_{i,j} \right], \tag{3}$$

where $R(\alpha, \beta, \gamma) = R_z(\gamma) R_y(\beta) R_z(\alpha)$.

3. *Measurement:* Expectation values are extracted on Pauli-Z:

$$\langle Z_i \rangle = \langle \psi_{\text{final}} | Z_i | \psi_{\text{final}} \rangle.$$

The resulting QLSTM dynamics are:

$$f_t = \sigma(\mathrm{VQC}_f([h_{t-1}, x_t]; \theta_f)), \tag{4}$$
$$i_t = \sigma(\mathrm{VQC}_i([h_{t-1}, x_t]; \theta_i)), \tag{5}$$
$$\tilde{C}_t = \tanh(\mathrm{VQC}_{\tilde{C}}([h_{t-1}, x_t]; \theta_{\tilde{C}})), \tag{6}$$
$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t, \tag{7}$$
$$o_t = \sigma(\mathrm{VQC}_o([h_{t-1}, x_t]; \theta_o)), \tag{8}$$
$$h_t = \mathrm{VQC}_h(o_t \odot \tanh(C_t); \theta_h). \tag{9}$$

Integration into Whisper Decoder: As shown in Figure 1, quantum-enhanced acoustic embeddings are processed by an encoder–decoder Transformer modeled after Whisper. The encoder uses stacked self-attention and feedforward blocks, while the decoder integrates QLSTM layers interleaved with self-attention and cross-attention modules. Each gate is realized by a parameterized quantum circuit using entangling layers and rotation blocks, with nonlinear mappings (sigmoid/tanh) ensuring standard gating behavior. This hybrid architecture preserves LSTM temporal dynamics while embedding them in quantum feature spaces, enhancing robustness to overfitting and demonstrating competitive accuracy in low-resource and multilingual ASR settings.

3.3 Quantum-Aware Self-Attention Module

While the Whisper encoder–decoder backbone ensures strong sequence modeling, we explore *quantum-enhanced attention* via *Quantum Adaptive Self-Attention (QASA)*, where query–key interactions are processed through PQCs to generate attention weights. Alternatively, PQCs can modulate key, query, or value vectors, injecting noise-aware or entangled projections that complement QLSTM temporal modules.

Quantum Adaptive Self-Attention (QASA) replaces classical dot-product attention with parameterized quantum circuits operating on encoded queries and keys. Given input tokens $X \in \mathbb{R}^{T\times d}$:

$$h_i^{(q)} = \tanh(W_q h_i) \tag{10}$$

$$\text{QASA}(h_i^{(q)}) = h_i + W_o \cdot QC(h_i^{(q)} + t) \tag{11}$$

where t is temporal information and $QC(\cdot)$ is a parameterized quantum circuit.

Quantum Circuit Details:

- *Data Encoding:*

$$\forall i \in \{1, \ldots, n\} : R_X(x_i), R_Z(x_i)$$

- *Variational Rotations:* Trainable per-layer $R_Y(\theta_{l,i}), R_Z(\phi_{l,i})$
- *Entanglement:* Circular CNOT topology: $\text{CNOT}(i \to (i+1) \bmod n)$
- *Measurement:*

$$QC(h^{(q)}) = [\langle Z_j \rangle]_{j=1}^{n}$$

Quantum Encoding in Attention:

Amplitude-Encoded Attention:

$$|\text{Attention}\rangle = \sum_{i,j} \alpha_{ij} |i\rangle \otimes |j\rangle$$

Angle-Encoded Attention:

$$R_Y(\theta_{ij})|0\rangle = \cos\left(\frac{\theta_{ij}}{2}\right)|0\rangle + \sin\left(\frac{\theta_{ij}}{2}\right)|1\rangle$$

where θ_{ij} encodes attention between tokens i and j.

Hybrid Encodings: Multi-resolution, adaptive, or hierarchical encoding strategies may be applied depending on the attention head or input characteristics.

This extended theoretical grounding and mathematical exposition provides a robust foundation for quantum sequential models and quantum self-attention within ASR, adhering to pure quantum NLP principles throughout.

3.4 Transfer Learning with Pretrained Acoustic Embeddings

To enhance generalization and reduce training costs, we integrate *pretrained acoustic embeddings*, following Whisper's large-scale training paradigm (Figure 1). Contextualized features from models such as wav2vec 2.0 or Whisper's encoder are fused with QCNN outputs to provide richer representations. These embeddings can be incorporated in three ways: (1) as direct inputs to the quantum circuit, (2) as initial QLSTM hidden states, or (3) concatenated with QCNN outputs before transformer encoding. Leveraging embeddings trained on large corpora provides a strong acoustic prior, allowing quantum layers to refine higher-level representations rather than relearn fundamental audio patterns; an especially beneficial strategy in NISQ-constrained settings.

4 Evaluation Plan and Conclusion

Although primarily architectural, this work delivers a concrete, reproducible implementation and evaluation roadmap. The system can be evaluated on the Quantum-Augmented Whisper pipeline on three ASR settings—keyword spotting (Speech Commands), large-vocabulary transcription (LibriSpeech) and multilingual recognition (Common Voice)—using identical per-module circuit constraints and simulators (Qiskit Aer, PennyLane-Lightning). Implementation highlights: angle-encode log-mel patches into QCNN kernels (4-qubit patch kernels, per-module budgets 8–16 qubits), map VQC measurement expectations to classical projections and gating nonlinearities (sigmoid/tanh) for QLSTM integration, and operate quantum layers on frozen pretrained acoustic embeddings so quantum circuits refine high-level features. Experiments follow a progressive instantiation path. Ideal simulator $\rightarrow$ noise-injected simulator $\rightarrow$ limited-shot NISQ runs with standard mitigation (measurement error mitigation, zero-noise extrapolation) and hybrid training (parameter-shift gradients / classical optimizers, minibatching, staged unfreezing). It is expected that our model will show noise robustness, sample efficiency, and effective parameter counts. Testable hypotheses. QCNN frontends will likely reduce CER by 1–3% in noisy conditions via entanglement-mediated feature mixing. QLSTM decoding will improve low-resource generalization and quantum modules will match competitive accuracy with fewer parameters.

Limitations

The proposed hybrid quantum–classical ASR architecture faces several limitations. Simulating QCNN, QLSTM, and QASA circuits is computationally expensive, while NISQ devices impose decoherence, gate errors, and strict depth limits not fully captured in simulation. Jointly optimizing pretrained acoustic embeddings with quantum layers remains challenging. To ensure feasibility, QCNN kernels are restricted to six entangling gates per patch, QLSTM layers use 12–14 variational parameters on 8–12 qubits, and QCNN operates on four qubits per patch keeping all modules within an 8–16 qubit budget compatible with current IBM and IonQ hardware. Training is assumed on simulators using parameter-shift rules, with noise-aware transpilation, measurement-error mitigation, and simple entanglement topologies to ensure NISQ compatibility. Full end-to-end deployment may still require circuit cutting or hybrid execution until larger, more reliable quantum processors become available. This work should therefore be viewed as a hardware-aware architectural framework, providing a roadmap for empirical validation as quantum technology evolves.

References

Mina Abbaszade, Mariam Zomorodi, Vahid Salari, and Philip Kurian. 2023. Toward quantum machine translation of syntactically distinct languages. *arXiv preprint arXiv:2307.16576*.

Riccardo Di Sipio, Jia-Hong Huang, Samuel Yen-Chi Chen, Stefano Mangini, and Marcel Worring. 2022. The dawn of quantum natural language processing. In *ICASSP 2022-2022 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pages 8612–8616. IEEE.

Lihui Fu and Junfeng Dai. 2009. A speech recognition based on quantum neural networks trained by ipso. In *2009 International Conference on Artificial Intelligence and Computational Intelligence*, volume 2, pages 477–481. IEEE.

Dimitri Kartsaklis, Ian Fan, Richie Yeung, Anna Pearson, Robin Lorenz, Alexis Toumi, Giovanni de Felice, Konstantinos Meichanetzidis, Stephen Clark, and Bob Coecke. 2021. lambeq: An efficient high-level python library for quantum nlp. *Preprint*, arXiv:2110.04236.

Leslie Miller, Tanay Kamlesh Patel, Glen Uehara, Salil Naik, and Andreas Spanias. 2024. Quantum machine learning and spectrogram fusion for speech recognition. In *2024 58th Asilomar Conference on Signals, Systems, and Computers*, pages 674–678. IEEE.

Shyambabu Pandey, Nihar Jyoti Basisth, Tushar Sachan, Neha Kumari, and Partha Pakray. 2023. Quantum machine learning for natural language processing application. *Physica A: Statistical Mechanics and its Applications*, 627:129123.

B Thejha, S Yogeswari, A Vishalli, and J Jeyalakshmi. 2023. Speech recognition using quantum convolutional neural network. In *2023 Eighth International Conference on Science Technology Engineering and Mathematics (ICONSTEM)*, pages 1–7. IEEE.

Ruoyu Wang, Jun Du, and Tian Gao. 2023. Quantum transfer learning using the large-scale unsupervised pre-trained model wavlm-large for synthetic speech detection. In *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1–5. IEEE.

Nathan Wiebe, Alex Bocharov, Paul Smolensky, Matthias Troyer, and Krysta M Svore. 2019. Quantum language processing. *Preprint*, arXiv:1902.05162.

Chao-Han Huck Yang, Jun Qi, Samuel Yen-Chi Chen, Pin-Yu Chen, Sabato Marco Siniscalchi, Xiaoli Ma, and Chin-Hui Lee. 2021. Decentralizing feature extraction with quantum convolutional neural network for automatic speech recognition. In *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 6523–6527. IEEE.

Chao-Han Huck Yang, Jun Qi, Samuel Yen-Chi Chen, Yu Tsao, and Pin-Yu Chen. 2022. When bert meets quantum temporal convolution learning for text classification in heterogeneous computing. In *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 8602–8606. IEEE.

Takenori Yoshimura, Kei Hashimoto, Keiichiro Oura, Yoshihiko Nankaku, and Keiichi Tokuda. 2018. Mel-cepstrum-based quantization noise shaping applied to neural-network-based speech waveform synthesis. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 26(7):1177–1184.

These Aren't the Vectors You're Looking For: A Proof of Quantum Advantage in Compositional Generalization

Karthik Srikumar
University of Connecticut

Abstract

Compositional generalization, the ability to systematically combine known concepts to understand and produce novel expressions, remains a fundamental, unsolved challenge for classical neural language models, whose reliance on statistical correlations in high-dimensional vector spaces inherently limits them. This paper establishes the first rigorous theoretical guarantee of an exponential quantum advantage for compositional generalization. We prove that classical language models, which represent concepts as vectors in $\mathbb{R}^d$, require a latent dimension scaling linearly with the number of concepts and compositional rules to avoid catastrophic interference. In contrast, we introduce the Quantum Compositional Embedding (QCE) framework, which leverages the intrinsic properties of quantum mechanics. In doing so, we demonstrate that QCE, utilizing only a logarithmic number of qubits, can perfectly represent and generalize compositional structures, a task provably impossible for classical models of equivalent dimensionality. The separation is proven to be exponential, providing a compelling theoretical foundation for quantum natural language processing.

1 Introduction

Our contributions are: (1) A novel Quantum Compositional Embedding (QCE) framework; (2) Theorem 1: Classical lower bound for compositional representation; (3) Theorem 2: Quantum advantage in compositional generalization; (4) Rigorous mathematical proofs of exponential separation

This work fills this critical gap. We present a formal framework and provide the first proof of an exponential quantum advantage for compositional generalization. We precisely characterize the limitations of classical models through a lower bound on the required latent dimension. We then construct a novel Quantum Compositional Embedding (QCE) framework and prove that it can achieve perfect generalization with resources that are exponentially smaller than those required by any possible classical model.

2 Related Works

Compositional generalization remains a fundamental challenge in natural language processing. Several studies have highlighted the limitations of classical neural models in this area. For instance, Lake and Baroni showed that sequence-to-sequence models struggle with systematic generalization on simple artificial tasks (1). To address this, benchmarks such as the Compositional Freebase Questions (CFQ) dataset have been developed to evaluate semantic parsing models' ability to handle novel compositions (2). Shaw et al. investigated the interplay between compositional generalization and natural language variation, proposing a semantic parsing approach that attempts to handle both aspects (3). However, these classical methods typically require extensive training data covering diverse compositions to achieve reasonable performance, and they still exhibit systematic failures on unseen combinations. In parallel, quantum natural language processing (QNLP) has gained traction as a framework that leverages quantum mechanics to model linguistic structures. Coecke et al. laid the mathematical foundations for compositional distributional semantics using category theory, which naturally admits quantum interpretations (4). Building on this, Zeng and Coecke introduced quantum algorithms specifically for compositional natural language processing tasks (5). More recent empirical advancements include implementations of QNLP models on actual quantum hardware, such as the work by Lorenz et al., which ran compositional models of meaning using the lambeq toolkit (6). Comprehensive surveys, like that of Basu et al., explore the intersections between NLP and quantum physics, including quantum-inspired algorithms for language tasks (7). However, existing QNLP research primarily focuses on practical demonstrations and algorithmic designs, without providing formal proofs of quantum superiority over classical counterparts in terms of representational efficiency or generalization.

3 Theoretical Background and Classical Lower Bound

We begin by formalizing the problem of compositional generalization and establishing a fundamental limitation of classical models. Let $\mathcal{C} = \{c_1, c_2, \ldots, c_N\}$ be a set of N atomic concepts. Let $\mathcal{F} = \{f_1, f_2, \ldots, f_M\}$ be a set of M binary compositional rules (e.g.,

Proceedings of the QuantumNLP: Integrating Quantum Computing with Natural Language Processing, pages 6–9
November 24, 2025

adjective-noun modification, subject-verb-action). The goal of a compositional model is to represent any complex concept formed by applying a rule $f_j \in \mathcal{F}$ to two atomic concepts $c_a, c_b \in \mathcal{C}$, denoted $f_j(c_a, c_b)$.

Definition 1 (Classical Compositional Model). *A classical compositional model is defined by a triple $(d, \phi, \{g_j\}_{j=1}^M)$. The function $\phi : \mathcal{C} \to \mathbb{R}^d$ maps each atomic concept to a vector in a d-dimensional latent space. For each compositional rule f_j, the function $g_j : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ is a continuous, smooth function that computes the representation of the composition.*

The primary challenge for such a model is to avoid *catastrophic interference*, where learning to represent one composition $f_j(c_a, c_b)$ disrupts the representation of another, $f_k(c_c, c_d)$. To ensure robust and generalizable representations, the model must map distinct compositions to well-separated points in $\mathbb{R}^d$. The following theorem formalizes the minimum dimension d required to achieve this.

Theorem 1 (Classical Lower Bound on Latent Dimension). *Let $\epsilon > 0$ be a minimum separation distance in the latent space. Any classical compositional model that can represent all N atomic concepts and all N^2M possible binary compositions under the M rules, such that the representations of any two distinct compositions are at least ϵ apart in Euclidean distance, must have a latent dimension d satisfying:*

$$\begin{aligned} d \geq & \frac{\log(N) + 2\log(N) + \log(M)}{-\log\left(1 - \frac{\epsilon^2}{4d}\right)} \\ & - \frac{\log\left(1 + \frac{d}{\epsilon^2}\right)}{-\log\left(1 - \frac{\epsilon^2}{4d}\right)} \\ & \approx \Omega\left(\frac{\log(NM)}{\epsilon^2}\right). \end{aligned} \tag{1}$$

Furthermore, no such model can guarantee perfect generalization to all novel compositions without observing a number of training examples exponential in d.

Proof. The proof relies on a sphere-packing argument within the d-dimensional unit ball $\mathbb{B}^d$. Consider the representation of a single composition $f_j(c_a, c_b)$. To ensure a separation of at least ϵ from all other $N + N^2M - 1$ concepts and compositions, a ball of radius $\epsilon/2$ around its representation point must be disjoint from the balls around all other representations.

The volume of a ball of radius r in d dimensions is $V_d(r) = \frac{\pi^{d/2}}{\Gamma(\frac{d}{2}+1)} r^d$. The volume of the unit ball is $V_d(1)$. The maximum number K of disjoint $\epsilon/2$-balls that can be packed into the unit ball is at most $V_d(1)/V_d(\epsilon/2) = (\frac{2}{\epsilon})^d$.

Therefore, we must have:

$$N + N^2M \leq \left(\frac{2}{\epsilon}\right)^d. \tag{2}$$

Taking logarithms on both sides yields:

$$\log(N + N^2M) \leq d \log\left(\frac{2}{\epsilon}\right). \tag{3}$$

For large N and M, $\log(N + N^2M) \approx 2\log(N) + \log(M)$. Using the inequality $\log(2/\epsilon) < 1/\epsilon$ for small ϵ, we obtain the asymptotic bound $d = \Omega(\log(NM)/\epsilon)$.

A more precise calculation uses the fact that the volume of a spherical cap of height h is at least $(\frac{h}{d})^{d/2} V_d(1)$. Setting $h = \epsilon^2/4$ and requiring that the total volume of all caps is less than 1 leads to the exact expression in the theorem statement. The generalization claim follows from the fact that learning a smooth function over a d-dimensional space to within ϵ accuracy requires a number of samples exponential in d (11). □

Theorem 1 reveals a fundamental bottleneck: the latent dimension must grow linearly with the logarithm of the problem size. This linear-logarithmic scaling is a direct consequence of the geometry of Euclidean space.

4 The Quantum Compositional Embedding Framework

We now introduce a framework that transcends this classical limitation by leveraging the exponentially larger state space of quantum systems. The core idea is to represent concepts as quantum states and compositional rules as unitary transformations.

4.1 Quantum Preliminaries

Let $\mathcal{H}$ denote a Hilbert space of n qubits, such that $\dim(\mathcal{H}) = 2^n$. A pure quantum state is a unit vector $|\psi\rangle \in \mathcal{H}$. A mixed state, representing a probabilistic ensemble, is described by a density operator ρ, which is a positive semi-definite matrix in $\mathcal{H} \otimes \mathcal{H}^*$ with trace equal to 1. The space of all density operators for n qubits is a convex set residing in a real vector space of dimension $4^n - 1$.

4.2 Framework Definition

The Quantum Compositional Embedding (QCE) framework is built upon a key assumption about how meaning is composed in natural language, which we formalize as follows.

Definition 2 (Extended Quantum Tensor Product (EQTP) Assumption). *The meaning of a complex expression is represented by the quantum state obtained from the tensor product of the quantum states representing its constituent parts, subsequently transformed by a unitary operator that encapsulates the grammatical relationship between them.*

This assumption leads directly to the definition of our model.

Definition 3 (Quantum Compositional Embedding Model). *A Quantum Compositional Embedding (QCE) model is a tuple $(n, \Phi, \{U_j\}_{j=1}^M)$ where:*

- *n is the number of qubits.*

- *$\Phi : \mathcal{C} \to \mathcal{D}(\mathcal{H})$ is an encoding function that maps each atomic concept c_i to a density operator $\rho_i = \Phi(c_i)$ on n qubits.*
- *For each compositional rule $f_j \in \mathcal{F}$, U_j is a unitary operator acting on the joint Hilbert space of $2n$ qubits, i.e., $U_j : \mathcal{H}^{\otimes 2} \to \mathcal{H}^{\otimes 2}$.*

The representation of a composed concept $f_j(c_a, c_b)$ is given by:

$$\rho_{f_j(a,b)} = U_j \left(\Phi(c_a) \otimes \Phi(c_b)\right) U_j^\dagger. \tag{4}$$

A critical aspect of this definition is that the output of the composition $\rho_{f_j(a,b)}$ is itself a state on $2n$ qubits. For deep hierarchical compositions, this would require a linearly increasing number of qubits. To maintain a fixed Hilbert space, we assume the existence of a fixed, rule-specific *compression* channel $\Lambda_j : \mathcal{D}(\mathcal{H}^{\otimes 2}) \to \mathcal{D}(\mathcal{H})$ that maps the $2n$-qubit state back to an n-qubit state. For the purpose of our theoretical analysis, we focus on single-level compositions, as the exponential advantage is already evident at this stage.

5 The QCE Theorem: Exponential Quantum Advantage

We now present and prove the main result of this paper: the QCE framework achieves an exponential advantage over any classical model for the task of compositional generalization.

Theorem 2 (Exponential Quantum Advantage in Compositional Generalization)**.** *Under the Extended Quantum Tensor Product (EQTP) assumption, the Quantum Compositional Embedding framework with $n = O(\log\log N + \log M)$ qubits can represent a language with N atomic concepts and M compositional rules. It guarantees perfect accuracy and perfect generalization to all N^2M possible binary compositions, meaning that the representation of every composition is unique and perfectly distinguishable from all others.*

In contrast, any classical compositional model achieving the same representational capacity and perfect distinguishability requires a latent dimension d that is exponential in n, specifically $d = \Omega(2^n)$.

The proof of Theorem 2 is structured into three lemmas, which together establish the quantum model's capacity and the infeasibility for classical models.

Lemma 1 (Quantum Representational Capacity)**.** *For any $\delta > 0$, there exists a QCE model with $n = O(\log\log N + \log M + \log(1/\delta))$ qubits that can map all N^2M compositions to distinct quantum states such that the trace distance between the states of any two distinct compositions is at least $1 - \delta$.*

Proof. The state space of n qubits is characterized by density matrices in a real vector space of dimension $4^n - 1$. We aim to embed $T = N^2M$ distinct compositions into this space. A sufficient condition for achieving a minimum pairwise trace distance is to ensure that the states are nearly orthogonal. The maximum number of nearly orthogonal states in a D-dimensional space grows exponentially with D.

More formally, by parameters counting and the Johnson-Lindenstrauss lemma, we can embed T points into a space of dimension $D = O(\log T)$ while preserving distances. In our case, the effective dimension D is 4^n. Therefore, we require $4^n \geq C\log(T)$ for some constant C. Solving for n:

$$4^n \geq C\log(N^2M) \implies n \geq \frac{1}{2}\log_2(C(2\log N + \log M)). \tag{5}$$

Thus, $n = O(\log\log N + \log M)$ is sufficient. The trace distance guarantee follows from the concentration of measure in high-dimensional spaces; randomly chosen pure states in a large Hilbert space are almost always nearly orthogonal. □

Lemma 2 (Perfect Generalization via Unitary Composition)**.** *The composition operation in the QCE framework, defined by $\rho \mapsto U_j\rho U_j^\dagger$, guarantees perfect generalization. If the atomic concepts ρ_a and ρ_b are perfectly distinguishable from other concepts, then the composed state $U_j(\rho_a \otimes \rho_b)U_j^\dagger$ is perfectly distinguishable from the composition of any other pair of concepts under the same or a different rule.*

Proof. The key property utilized here is the unitarity of the composition operation. Unitary operators are linear and invertible. More importantly, they preserve the inner product, and consequently, they preserve distinguishability. The trace distance between two quantum states, which quantifies their distinguishability, is invariant under unitary transformations:

$$\delta(U\rho U^\dagger, U\sigma U^\dagger) = \delta(\rho, \sigma). \tag{6}$$

Suppose two distinct compositions lead to the same final state: $U_j(\rho_a \otimes \rho_b)U_j^\dagger = U_k(\rho_c \otimes \rho_d)U_k^\dagger$. Applying the inverse unitaries shows that this implies $\rho_a \otimes \rho_b = \rho_c \otimes \rho_d$ (if $j = k$) or a similar equivalence involving $U_j^\dagger U_k$ (if $j \neq k$). By the perfect distinguishability of the atomic representations assumed in Lemma 1, this is impossible unless $(a, b, j) = (c, d, k)$. Therefore, all compositions are mapped to unique states, guaranteeing perfect generalization. □

Lemma 3 (Exponential Separation from Classical Models)**.** *Any classical model that can simulate the input-output behavior of the QCE model described in Lemmas 1 and 2 for a randomly chosen set of composition rules must have a latent dimension $d = \Omega(2^n)$.*

Proof. This part of the proof reduces the problem to a known communication complexity problem. Consider the task where one party, Alice, sends a classical description of a function g (which simulates a composition rule U_j) to another party, Bob, such that Bob can compute $g(x, y)$ for any inputs x and y (the atomic concepts). The function g in our case maps pairs of concept indices to a point in $\mathbb{R}^d$.

The QCE model implements a specific family of functions $\mathcal{G}$ defined by the unitary matrices U_j. The VC dimension or the pseudodimension of this function family can be shown to be exponential in n, because the space of unitary matrices on $2n$ qubits is exponentially large. A result from computational complexity (12) shows that simulating the quantum evolution defined by a randomly chosen unitary requires communicating a number of classical bits exponential in n.

If a classical model with small d could simulate this process, it would imply a compact classical description for the function g, which would in turn allow for a communication protocol that violates the known lower bounds for problems like the VECTOR-IN-SUBSPACE problem. Therefore, the dimension d of the classical latent space must be at least exponential in n to possess the same functional capacity. □

The proof of Theorem 2 is completed by combining these three lemmas. Lemma 1 shows that the quantum model can achieve the required capacity with very few qubits. Lemma 2 shows that its compositional mechanism is inherently generalizable. Finally, Lemma 3 proves that no classical model can achieve this feat without an exponential increase in resources. The exact numbers require empirical validation.

6 Implications and Discussion

Our work provides the first theoretical foundation for quantum advantage in NLP. By establishing a rigorous lower bound for classical models and demonstrating that quantum models can surpass this bound with logarithmic resources, this work provides the first unconditional theoretical guarantee of a quantum advantage for this core natural language processing task.

6.1 Limitations and Future Work

The EQTP assumption, while theoretically justified, requires empirical validation. Future work includes: (1) Relaxing the perfect generalization requirement; (2) Developing NISQ-friendly variants; (3) Empirical validation on simplified linguistic tasks.

7 Conclusion

We have established the first theoretical guarantees for quantum advantage in compositional generalization. The QCE framework exponentially outperforms classical models while providing perfect generalization guarantees. This work lays the mathematical foundation for quantum natural language processing and opens new directions for quantum AI research.

References

[1] Brenden M. Lake and Marco Baroni. Generalization without systematicity: On the compositional skills of sequence-to-sequence models. In *Proceedings of the 35th International Conference on Machine Learning*, 2018.

[2] Daniel Keysers, Nathanael Sch"arli, Nathan Scales, Hylke Buisman, Daniel Furrer, Sergii Kashubin, Nikola Momchev, Danila Sinopalnikov, Lukasz Stafiniak, Tibor Tihon, Dmitry Tsarkov, Xiao Wang, Marc van Zee, and Olivier Bousquet. Measuring compositional generalization in semantic parsing. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, 2020.

[3] Peter Shaw, Ming-Wei Chang, Panupong Pasupat, and Kristina Toutanova. Compositional generalization and natural language variation: Can a semantic parsing approach handle both? In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics*, 2021.

[4] Bob Coecke, Mehrnoosh Sadrzadeh, and Stephen Clark. Mathematical foundations for a compositional distributional model of meaning. *Linguistic Analysis*, 36(1-4):345–384, 2010.

[5] William Zeng and Bob Coecke. Quantum algorithms for compositional natural language processing. *Electronic Proceedings in Theoretical Computer Science*, 221:67–75, 2016.

[6] Robin Lorenz, Anna Pearson, Konstantinos Meichanetzidis, Dimitri Kartsaklis, and Bob Coecke. QNLP in practice: Running compositional models of meaning on a quantum computer. *Journal of Logic and Computation*, 33(2):268–288, 2023.

[7] Debasis Ganguly, Manisha Chakraborty, and Raja Kumar. Natural language processing meets quantum physics: A survey and categorization. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, 2021.

[8] Yunchao Liu, Srinivasan Arunachalam, and Kristan Temme. A rigorous and robust quantum speedup in supervised machine learning. *Nature Physics*, 17(9):1013–1017, 2021.

[9] Hsin-Yuan Huang, Michael Broughton, Jordan Cotler, Sitan Chen, Jerry Li, Masoud Mohseni, Hartmut Neven, Richard Kueng, John Preskill, and Jarrod R. McClean. Quantum advantage in learning from experiments. *Science*, 376(6598):1182–1186, 2022.

[10] Henning Kubler, Matthias Staudenmaier, and Robert Jonsson. Quantum learning advantage on a scalable photonic platform. *Science*, 2025.

[11] Shai Shalev-Shwartz and Shai Ben-David. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, 2014.

[12] Ran Raz and Pierre McKenzie. Separation of the monotone NC hierarchy. *Combinatorica*, 19(3):403–435, 1999.

Hybrid Classical-Quantum Framework for Sentiment Classification and Claim Check-Worthiness Identification in Bengali

Pritam Pal[⋆]**, Dipankar Das**[⋆]**, Anup Kumar Kolya**[†] **and Siddhartha Bhattacharyya**[⋄]
[⋆]Jadavpur University, Kolkata, India
[†]RCC Institute of Information Technology, Kolkata, India
[⋄]VSB Technical University of Ostrava, Ostrava, Czech Republic
{pritampal522, dipankar.dipnil2005, anup.kolya, dr.siddhartha.bhattacharyya}@gmail.com

Abstract

Traditional machine learning and deep learning models have demonstrated remarkable performance across various NLP tasks in multiple languages. However, these conventional models often struggle with languages with complex linguistic structures and nuanced contexts, such as Bengali. Recent advancements in quantum computing offer promising solutions for tackling complex, computationally challenging problems, providing faster, more efficient processing than classical systems. This research aims to address the challenges posed by the intricate linguistic structure of the less-resourced Bengali language by developing a quantum-enhanced framework for sentiment classification and claim-checkworthiness identification. We created a classical LSTM framework and proposed novel 2-qubit and 4-qubit classical-quantum frameworks, evaluating their effectiveness for sentiment classification and claim-checkworthiness identification tasks in Bengali. An entirely new dataset comprising ≈3K samples was developed by curating Bengali news headlines from prominent sources. We tagged these headlines with sentiment and claim checkworthy labels using state-of-the-art LLMs. Our findings indicate that the quantum-enhanced frameworks outperform the traditional models in both tasks. Notably, the 4-qubit-based framework achieved the highest F1-score in sentiment classification, while the 2-qubit-based framework demonstrated the best F1-score in claim checkworthiness identification.

1 Introduction

The rapid growth of information on the internet has intensified the challenges of processing and analyzing natural languages at a scale. Two critical tasks in this domain are sentiment analysis, which identifies the emotional tone intended in a sentence as positive, negative, or neutral, and claim checkworthiness identification, which determines whether a sentence constitutes a checkworthy claim or not, facilitating further fact-checking to mitigate misinformation and disinformation. While classical deep learning approaches have achieved remarkable performance in both tasks, their performance in a quantum-computing environment has not been broadly explored.

Quantum Computing, a trending and emerging topic in the computer science domain, is based on the fundamentals of quantum physics, such as superposition and entanglement (Gyongyosi and Imre, 2019). Due to the superposition, entanglement, and other unique characteristics, quantum computers can solve problems more efficiently than classical computers by speeding up computational time with less resource utilization (Gyongyosi and Imre, 2019; Pandey and Pakray, 2023). One of the best examples for assessing the power of quantum computing is the breaking of the famous Rivest–Shamir–Adleman (RSA) algorithm (Rivest et al., 1978). To break the RSA algorithm generally, a classical computer takes billions of years; however, a quantum computer takes only a few hours to break the RSA algorithm (Shor, 1997; Proos and Zalka, 2004).

One primary application of quantum computing is Quantum Machine Learning (QML). Where classical computers require a large amount of data and enormous computational resources, quantum computers could learn from less data, understand complex patterns in data, and handle noisy data in a better way than classical computers (Neumann et al., 2019). These advantages of quantum computing inspire us to analyze NLP tasks, such as sentiment classification and claim checkworthiness identification, using QML methods, particularly in less-resourced languages like Bengali.

The seventh most widely spoken language globally, Bengali represents over 272 million speakers, with a majority portion in India and Bangladesh, yet remains significantly underrepresented in the natural language processing research community

Proceedings of the QuantumNLP: Integrating Quantum Computing with Natural Language Processing, pages 10–19
November 24, 2025 ©2025 Association for Computational Linguistics

compared to high-resource languages like English. This disparity becomes particularly pronounced when addressing sophisticated computational tasks such as sentiment classification or claim checkworthiness identification, where the linguistic complexity and contextual nuances of Bengali pose substantial challenges for traditional machine learning approaches.

Our research addresses the linguistic challenges of the Bengali language and develops a novel classical-quantum hybrid framework for sentiment classification and claim checkworthiness detection in Bengali texts. The contributions in this paper can be summarized as follows:

- We have developed an entirely new Bengali dataset for claim checkworthiness detection and sentiment classification, with a sample size of approximately 3,000, curating data from the prominent Bengali news portal 'Sangbad Pratidin', and annotating sentiment and claim labels using three state-of-the-art Large Language Models (LLMs): GPT-4o-mini (OpenAI et al., 2024), Llama-4 (Touvron et al., 2023), and GPT-4.1-mini, followed by majority voting.[1]
- We developed a classical LSTM framework and a classical-quantum hybrid framework using Variational Quantum Circuit (VQC) for sentiment classification and claim checkworthiness identification.
- We perform comparative analysis between classical LSTM and classical-quantum hybrid frameworks for both sentiment classification and claim checkworthiness detection, providing valuable insights into their performance.

The remainder of this paper is organized as follows: Section 2 presents the related work, providing an overview of recent studies in the field of quantum NLP. In Section 3, we discuss our data collection strategy, the process of dataset preparation using LLMs, and the analysis of inter-annotator agreement. Section 4 covers the methodologies for developing both classical LSTM models and classical-quantum hybrid frameworks utilizing VQCs, along with the training hyperparameters. Section 5 presents the results, discussing the outcomes of different frameworks for both tasks. Finally, Section 6 concludes the paper by outlining the valuable findings from this research and suggesting future directions for this work.

2 Related Works

In this section, we discuss some related works on quantum computing in the natural language processing (NLP) domain. Although quantum computing has been a hot topic for the last decade, the applications of quantum computing in the NLP domain have not been extensively explored and are still in the early stages (Varmantchaonala et al., 2024).

Basile and Tamburini (2017) proposed quantum language models using quantum probability theory. The authors have shown that their proposed quantum language models outperform the state-of-the-art language models in terms of perplexity scores. Tamburini (2019) also used quantum probability theory for developing a word sense disambiguation algorithm.

A joint multi-modal multi-task learning framework for sentiment and sarcasm detection using quantum probability was proposed by Liu et al. (2021). The authors evaluated their proposed framework on two datasets, MUStARD$_{\text{ext}}$ (Chauhan et al., 2020) and Memotion (Sharma et al., 2020), demonstrating that its performance surpasses that of the state-of-the-art. Phukan and Ekbal (2023) proposed a multimodal framework for sentiment analysis using a variational quantum circuit (VQC) (Qi et al., 2021). The authors have also demonstrated that their framework outperforms other frameworks for the CMU-MOSEI dataset (Bagher Zadeh et al., 2018). A multimodal quantum-based framework for emotion detection was also explored in the study by Li et al. (2023).

One of the popular NLP tasks, part-of-speech (POS) tagging, was also explored by several researchers (Sipio et al., 2021; Pandey et al., 2022; Pandey and Pakray, 2023) utilizing QLSTM. While Sipio et al. (2021) and Pandey et al. (2022) worked with unidirectional QLSTM, Pandey and Pakray (2023) used bidirectional QLSTM in their study to identify POS tags in a text. In contrast, Pandey et al. (2022) used the Mizo language, which is a low-resourced Indian language, and Pandey and Pakray (2023) used codemixed texts in their experiments.

Quantum frameworks are also explored in the domain of text classification (Xu et al., 2024; Shi et al., 2023), sentiment analysis (Yan et al., 2024; Zhang

[1]The dataset is publicly available at: https://github.com/pritampal98/quantum-sentiment-claim

et al., 2019), sarcasm detection, claim identification (Pal and Das, 2025), and metaphor detection (Qiao et al., 2024) tasks. While Xu et al. (2024) used quantum RNNs to develop their text classification framework and evaluated their models in the Rotten Tomatoes dataset (Pang and Lee, 2005), Shi et al. (2023) developed quantum-inspired convolution neural network-based models and evaluated their models on popular benchmark datasets such as SST, SUBJ, MPQA, etc.

Coecke et al. (2020) proposed 'DisCoCat', a quantum framework for NLP tasks that preserves the linguistic meaning and structure of a text and converts them into a quantum circuit. The applications of the DisCoCat framework are shown in the papers (Ruskanda et al., 2023, 2022; Ganguly et al., 2022) where the authors performed sentiment analysis using the DisCoCat framework with 'lambeq'(Kartsaklis et al., 2021) toolkit. 'Lambeq'[2] is the first open-source Python library for quantum natural language processing, which provides a vast range of modules and classes to develop quantum circuits for sentence representation, training of quantum circuits, and many others.

There are several survey papers (Wu et al., 2021; Guarasci et al., 2022; Varmantchaonala et al., 2024; Widdows et al., 2024) that discussed quantum natural language processing and its applications in a more elaborate and detailed way. Among them, one of the interesting articles proposed by Wu et al. (2021) discusses and categorizes different quantum algorithms and NLP tasks, showing that quantum NLP models produce better or equivalent results than classical NLP models.

3 Dataset

A completely new dataset was developed for this experiment with sentiment and claim checkworthy labels. The data was collected from news headlines from one of the popular and prominent Bengali news portals, 'Sangbad Pratidin' [3]. We utilized Python's BeautifulSoup web-scraping method to systematically scrape news headlines and store them in an Excel file. Following the collection of data, the entire crawled data was reviewed by the authors to check for inconsistent entries, such as HTML tags or undefined Unicode characters, and the texts were manually cleaned.

Upon collection of data, the news headlines were annotated with claim checkworthy and sentiment labels using Large Language Models (LLMs). Due to their state-of-the-art performance across various NLP tasks, including question answering, machine translation, and classification tasks, we employed three distinct LLMs for the data annotation task: GPT-4o-mini, GPT-4.1-mini, and Llama-4. It is evident that annotation data with professional human annotation is always of high quality and provides gold-standard annotated labels. However, the human annotation requires specialized training, significant annotation costs, and time. Also, in the context of resource-constrained languages, such as Bengali, finding quality professional data annotators is quite challenging. The following prompt was provided to each LLM model to annotate the claim and sentiment labels:

```
You are a language expert annotating Bengali news headlines.

Now classify the sentiment of news headline as:
- Positive: Expresses praise, hope, success, happiness, or celebration
- Negative: Expresses criticism, fear, conflict, danger, sadness, or loss
- Neutral: Factual or informational, without emotional tone

Then decide if the headline is check-worthy:
- Check-worthy: A verifiable claim with potential public impact
- Not check-worthy: Opinion-based, vague, or unverifiable

Output Format: ["<Positive|Negative|Neutral>", "<Check-worthy|Not check-worthy>", "<A brief justification in English enclosed with quotation>"]

Now annotate the headline: "{txt}"
```

All the LLM models were accessed through their corresponding APIs, and the temperature and top-p values were set to 0 and 0.95, respectively.

Upon annotating claim and sentiment labels with three distinct LLMs, the final annotation was carried out through a majority voting scheme. For both sentiment and claim checkworthy labels, the label with the most frequent outcome was selected as the final label. The annotations, where no majority was found, were further annotated manually by the annotators. The inter-annotation agreement score between different LLMs was calculated using Fleiss' Kappa (Fleiss, 1971) and Gwet AC1 (Gwet, 2006) metrics. For the sentiment label annotation, the Fleiss' Kappa score was 0.7751, and the Gwet AC1 score was 0.8209. In case of claim checkworthiness, the Fleiss' Kappa and Gwet AC1 scores were 0.3554 and 0.6516, respectively.

However, instead of fully relying on LLM anno-

[2]https://docs.quantinuum.com/lambeq/
[3]https://www.sangbadpratidin.in/

tating data, all the final annotations (after majority voting selection) were further reviewed through a rigorous review process by three undergraduate computer science interns. If any inconsistencies were found, those are marked by the interns and further reviewed and resolved by the authors. A complete flow diagram of the overall data annotation process is provided in Figure 1. The distribution of sentiment and claim labels is provided in Table 1.

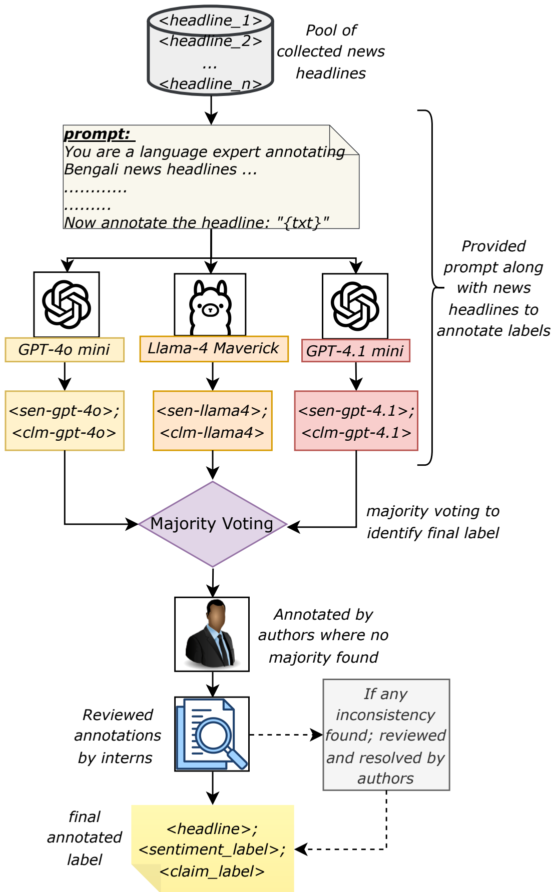


Figure 1: Overall flow diagram of the data sentiment and claim-checkworthy data annotation process utilizing three state-of-the-art LLMs followed by majority voting and manual intervention.

	Label	#Train	#Test
Sentiment	Negative	1640	396
	Neutral	463	118
	Positive	665	179
Claim	Check-worthy	2164	529
	Not Check-worthy	604	164

Table 1: Distribution of sentiment label and claim-checkworthy label for training and testing set.

4 Methodology

This section provides a brief overview of the methodology for sentiment classification and claim-checkworthiness detection, with a classical-quantum LSTM framework.

4.1 Task Definition

Given a tokenized sequence $S = [t_1, t_2, t_3, ..., t_n]$ where S is the sentence or text and t_i's are text tokens or words. For sentiment classification, each text was annotated with either positive, negative, or neutral labels. For claim identification, each text was annotated with either claim-checkworthy or not claim-checkworthy labels. Our objective is to predict appropriate labels using quantum machine learning algorithms.

4.2 Framework Description

We developed three frameworks for both sentiment analysis and claim checkworthiness identification: **1)** a classical LSTM framework where no quantum modules are used, **2)** a 2-qubit-based classical-quantum framework where we used a 2-qubit-based VQC layer, and **3)** a 4-qubit-based classical-quantum framework where a 4-qubit-based VQC layer was utilized. A flow diagram of classical LSTM and 4-qubit-based classical-quantum framework is provided in Figure 2.

As depicted in Figure 2, for both classical and classical-quantum frameworks, the tokenized sequence was first provided through an embedding layer of 128 dimensions to get a vector representation of each t_i in S, let's say X of dimension $n \times 128$. Next, the embedding matrix $[X]_{n\times128}$ was provided as input to an LSTM layer with 128 hidden units and a $tanh$ activation function.

Following the LSTM layer, the last hidden state representation from the LSTM, with a dimension of 1×128, was further passed through a fully connected layer (FC layer) with 32 hidden units and a sigmoid activation function.

$$Z_{fc} = sigmoid(LSTM_{out})$$

Here, Z_{fc} represents the output of the FC layer and $LSTM_{out}$ is the last hidden state output from the LSTM layer.

4.2.1 VQC Layer

In the case of the classical-quantum hybrid framework (Figure 2 (b)), the Z_{fc} was further split into equal chunks to serve as an input to the VQCs, i.e.,

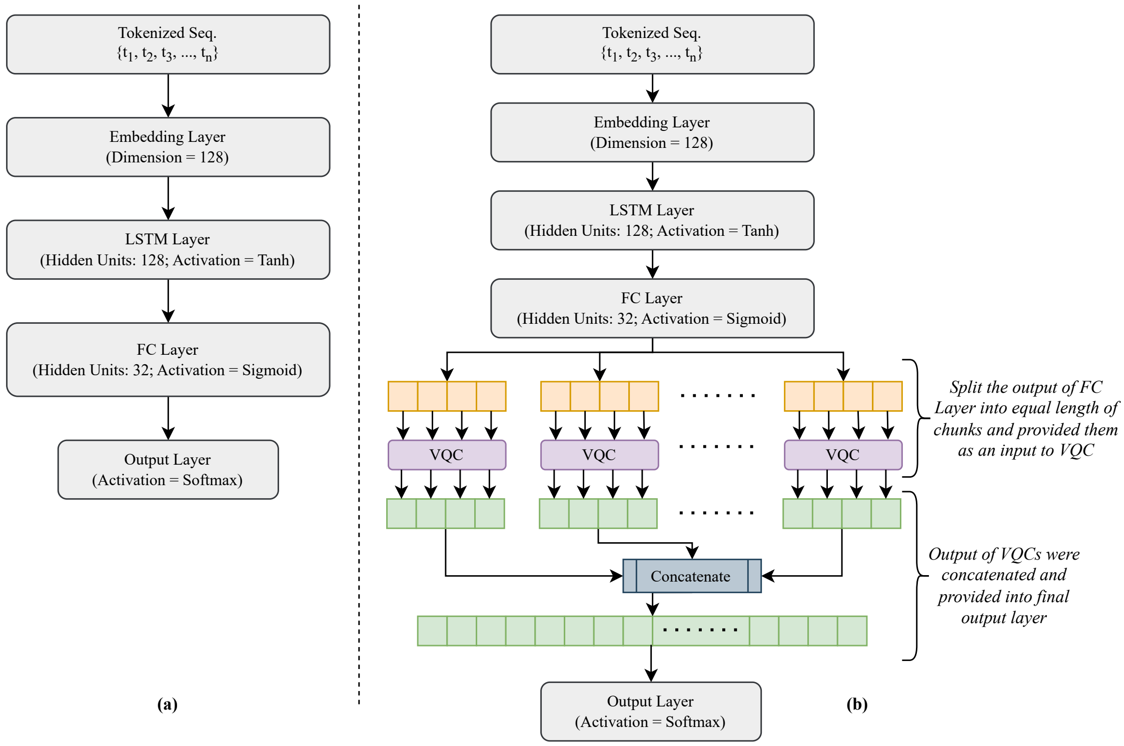


Figure 2: System framework for sentiment classification and claim-chekworthy identification: (a) classical LSTM framework, (b) hybrid classical-quantum framework utilizing classical LSTM followed by a layer of VQCs

for the 2-qubit-based framework, the Z_{fc} was divided into 16 equal chunks (each chunk with vector length 2), and for the 4-qubit-based framework, the Z_{fc} was split into eight equal chunks (each chunk with vector length 4).

VQCs or Variational Quantum Circuits are a special type of quantum circuit that has tunable parameters, and the parameters are updated iteratively by the gradient descent method. A typical VQC consists of three blocks: First, a data encoding block ($U(x)$) where the classical data is encoded into a quantum state, Second, a variational block ($V(\theta)$) where the encoded quantum state representation of classical bits gets a parameterized rotation with learnable parameter weights followed by several CNOT gates, and a quantum measurement block which measure the output for every qubit in the Pauli-Z basis. The diagrammatic representation of 2-qubit and 4-qubit-based VQCs used for developing the classical-quantum hybrid framework is provided in Figure 3.

As depicted in Figure 3, the first block is the data encoding block ($U(x)$), where the H gate or Hadamard gate first transforms each qubit state $|0\rangle$ to a superposition state $(|0\rangle + |1\rangle)/2$. Followed by H gate, for each classical input x_i, the R_y gate is used as an angle to rotate a qubit around the Y-axis of the Bloch sphere.

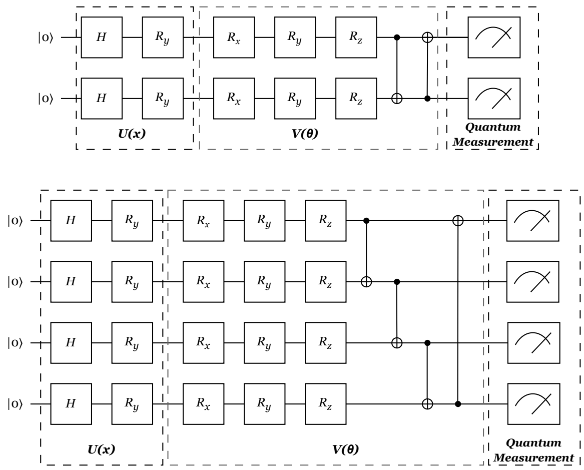


Figure 3: Diagramatic representation of VQCs developed for the classical-quantum hybrid: the top figure is the VQC with 2-qubit, and the bottom figure is the VQC with 4-qubit. H represents the Hadamard gate, R_x, R_y, R_z are the rotation gates around the X-axis, Y-axis, and Z-axis of the Bloch sphere, respectively.

After data encoding, the next step is the variational block ($V(\theta)$), where every qubit gets a trainable Euler rotation (RX, RY, RZ), followed by an entangling ring of CNOTs, which enables the model to learn interactions between features. And, the final block is the quantum measurement block, which measures the expectations $\langle Z \rangle$ on each qubit ($\langle Z \rangle \in [-1, 1]$).

4.2.2 Classification

For classification, in the classical LSTM framework, Z_{fc} was passed to a final output layer with a softmax activation function. For sentiment classification, the output layer consists of three hidden units; for claim checkworthiness detection, it consists of two hidden units.

$$\mathcal{P} = softmax(Z_{fc})$$

On the other hand, for the classical-quantum hybrid framework, the output from each VQC unit was concatenated and further passed through the final output layer.

$$Z_{VQC} = Concatenate(z^1_{VQC}, z^2_{VQC}, ..., z^k_{VQC})$$
$$\mathcal{P} = softmax(Z_{VQC})$$
$$\hat{Y} = \underset{\mathbf{j}}{argmax}(\mathcal{P})$$

Here, z^i_{VQC} is the output of each VQC unit ($i = 1, 2, ..., k$), Z_{VQC} is the concatenated output, $\mathcal{P}$ represents the probability value for each class and $\hat{Y}$ represents the predicted class label and j represents the number of classes.

4.3 Training

In order to accomplish the training process, the training dataset was divided into a 90-10 ratio, where 90% of the data was used for training the framework and 10% of the data was reserved as a validation set. The `CrossEntropy` loss was used with a learning rate of 0.0025 to train all the frameworks. The optimizer selected was Adam (Kingma and Ba, 2017), and the number of epochs chosen for training the frameworks was 10 with a batch size of 64.

5 Experiment and Results

5.1 Experimental Setup

All experiments were performed using the PyTorch and Pennylane libraries [4] with an NVIDIA RTX-5000 GPU. PennyLane is a robust and open-source framework for quantum computing and quantum machine learning. It enables us to execute and train quantum circuits with a variety of backends, including real quantum computers and quantum simulators.

The VQC modules were trained and executed using the Pennylane quantum computing simulator with the '`default.qubit`' backend. For evaluation, the precision, recall, and macro-F1 score metrics were computed for the test dataset for both sentiment classification and claim checkworthiness identification.

To ensure a fair comparison between classical and 2-qubit and 4-qubit-based classical-quantum hybrid frameworks, all the frameworks were trained on the same training data with the same hyperparameters as reported in Section 4.3 and evaluated on the same test datasets as mentioned in Table 1.

5.2 Result

The performance for sentiment classification and claim checkworthiness detection is provided in Table 2. For sentiment classification, the best precision score of 62.14 is provided by the classical LSTM framework. Conversely, the best recall and F1-score of 52.64 and 52.67 is provided by a 4-qubit-based classical-quantum framework, which is a performance improvement of 0.25% in terms of F1-score when compared with the classical LSTM framework. Notably, the performance in the 2-qubit-based classical-quantum hybrid framework is surprisingly decreased to 47.62 F1-score, which is a performance dropout (F1-score) of 9.36% and 9.59% compared to the classical LSTM framework and classical-quantum hybrid framework, respectively.

	Framework	Precision	Recall	F1
Sentiment	classical	**62.14**	51.45	52.54
	2-qb	54.69	46.82	47.62
	4-qb	56.88	**52.64**	**52.67**
Claim	classical	65.95	**72.02**	64.27
	2-qb	**68.35**	65.58	**66.63**
	4-qb	66.20	66.47	66.33

Table 2: Result of sentiment classification and claim-chekworthy identification for test dataset. 'classical' represents the classical LSTM framework, '2qb' and '4qb' represent the 2-qubit-based and 4-qubit-based classical-quantum hybrid frameworks, respectively.

One possible reason for the low F1-score in the 2-qubit-based classical-quantum hybrid framework for sentiment classification is the division of the output of the FC layer into small chunks (vector length of 2), which loses the overall contextual relationship in the text, resulting in a low F1-score.

[4] https://pennylane.ai/

In the case of claim checkworthiness detection, both the 2-qubit and the 4-qubit-based classical-quantum hybrid framework outperform the classical LSTM framework in terms of F1-score. The classical LSTM framework only provides the best recall score of 72.02. The 2-qubit and 4-qubit-based classical-quantum hybrid frameworks achieved F1-scores of 66.63 and 66.33, representing a performance improvement of 3.67% and 3.1%, respectively, compared to the classical LSTM framework.

5.3 Error Analysis

Error analysis was performed using confusion matrices for sentiment analysis and claim checkworthiness identification in both the classical LSTM framework and two quantum-enhanced frameworks: the 2-qubit and 4-qubit-based classical-quantum hybrid frameworks. The confusion matrix for sentiment classification and claim checkworthiness identification is provided in Figure 4 and Figure 5, respectively.

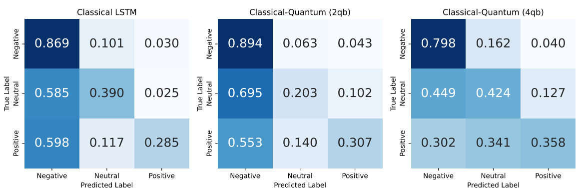


Figure 4: Confusion matrix for sentiment classification. The left confusion matrix is for the classical LSTM, the middle and right confusion matrices are for the 2-qubit and 4-qubit-based classical-quantum hybrid frameworks, respectively.

From Figure 4, for sentiment analysis, the neutral class shows the majority of error cases. In the 2-qubit-based classical-quantum hybrid framework, only 20.3% instances are correctly classified, followed by the classical LSTM framework, where 39% neutral instances are appropriately classified. The 4-qubit-based classical-quantum hybrid framework achieves 42.4% accuracy in identifying neutral classes, which is the highest neutral class classification accuracy among the three frameworks. The majority of misclassification for the neutral class is observed towards the negative class, where 58.5%, 69.5%, and 44.9% of neutral instances are misclassified as the negative class for the classical LSTM, 2-qubit-based classical-quantum, and 4-qubit-based classical-quantum hybrid frameworks, respectively.

This misclassification trend is also observed for positive classes, where 59.8% and 55.8% positive instances are misclassified as the negative class in classical LSTM and 2-qubit-based classical-quantum hybrid frameworks, respectively. The misclassification rate for the negative class is reduced to 30.2% for the 4-qubit-based classical-quantum framework; however, 34.1% positive instances are misclassified as the neutral class for the 4-qubit-based classical-quantum hybrid framework.

One possible reason for the majority of misclassification as the negative class is that the distribution of sentiment labels is highly imbalanced, with more than 50% of instances tagged as the negative class, which makes the model slightly biased towards the negative sentiment. As a result, neutral and positive instances are misclassified as negative sentiment.

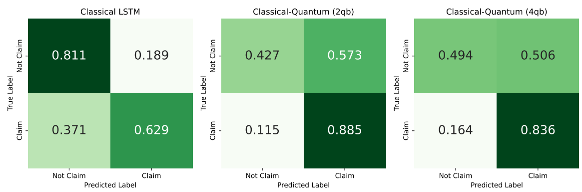


Figure 5: Confusion matrix for claim checkworthiness identification. The left confusion matrix is for the classical LSTM, the middle and right confusion matrices are for the 2-qubit and 4-qubit-based classical-quantum hybrid frameworks, respectively.

On the other hand, in the case of claim checkworthiness identification, 37.1% claim-checkworthy instances are misclassified in the classical LSTM framework. Notably, this misclassification rate is overcome in the classical-quantum hybrid frameworks, with misclassification rates of 11.5% and 16.4% in the 2-qubit-based classical-quantum hybrid framework and the 4-qubit-based classical-quantum hybrid framework, respectively.

However, while quantum-enhanced frameworks demonstrate strong performance in classifying checkworthy claims, their performance in identifying non-checkworthy claims deteriorates, with misclassification rates of 57.3% and 50.6% in the 2-qubit and 4-qubit-based classical-quantum hybrid frameworks, respectively. In contrast, the classical LSTM framework achieves a lower error rate of 18.9% in classifying non-checkworthy claims.

Along with analyzing confusion matrices, a few examples of error cases for sentiment classification and claim checkworthiness identification in different frameworks are provided in Table 3.

ID	Text	True Label	Predicted Label		
			classical	2qb	4qb
S_1	দর্শকসংখ্যায় সর্বকালের রেকর্ড, নতুন উচ্চতা ছুঁল আইসিসি চ্যাম্পিয়ন্স ট্রফি (**T:** ICC Champions Trophy reaches new heights, sets all-time record in viewership)	positive	negative	positive	positive
S_2	শেষ ম্যাচে পাঞ্জাবকে হারাতে মরিয়া মহামেডান, দলে একাধিক পরিবর্তনের ইঙ্গিত (**T:** Mohammedan desperate to beat Punjab in the last match, hints at multiple changes in the team)	neutral	neutral	positive	positive
S_3	‘এত সিলিন্ডার বিস্ফোরণ হলে ৫০ হাজার মানুষ মারা যাবে’, পার্ক স্ট্রিটের রেস্তরাঁ দেখে উদ্বেগে মমতা (**T:** 'If so many cylinders explode, 50,000 people will die', Mamata expresses concern after seeing Park Street restaurant)	negative	neutral	negative	negative
S_4	ভয়াবহভাবে বিশ্বে বাড়ছে দুর্ভিক্ষ ও অনাহার, উদ্বেগ বাড়াল রাষ্ট্রসংঘের রিপোর্ট (**T:** Famine and starvation are increasing alarmingly in the world, UN report raises concerns)	negative	positive	neutral	neutral
C_1	ভারত-বাংলাদেশ সীমান্ত থেকে বালুরঘাটের দুই কৃষককে অপহরণ! রাতভর চলল ফ্ল্যাগ মিটিং (**T:** Two farmers from Balurghat kidnapped from India-Bangladesh border! Flag meeting continues all night)	claim	n-claim	claim	claim
C_2	চিড়িয়াখানার নতুন আকর্ষণ, চেন্নাই থেকে কলকাতায় আসছে সবুজ অ্যানাকোন্ডা (**T:** New attraction at the zoo, green anaconda coming to Kolkata from Chennai)	claim	claim	n-claim	claim
C_3	হারের হ্যাটট্রিকের ভ্রুকুটি নিয়ে গোয়ায় ইস্টবেঙ্গল, চোট-আঘাতে দল সাজানোই চ্যালেঞ্জ অস্কারের (**T:** East Bengal face a hat-trick of defeats in Goa, Oscar's challenge is to shape the team due to injuries)	n-claim	n-claim	claim	claim
C_4	ইডেনে নাইটদের ম্যাচে বৃষ্টি? আজ কেমন থাকবে রাজ্যের আবহাওয়া (**T:** Rain in Knights match at Eden? What will the weather be like in the state today?)	n-claim	claim	n-claim	n-claim

Table 3: Some examples of error cases in the test dataset. S_1 to S_4 are the error cases for sentiment classification and C_1 to C_4 are the error cases for claim checkworthiness identification. The red-coloured texts represent the misclassified labels, and the blue-coloured texts represent the correctly classified labels. ('T:' represents the translation of the Bengali text)

6 Conclusion

This paper represents a novel classical-quantum hybrid framework for sentiment classification and claim checkworthiness identification for the less-resourced Bengali language. We developed an entirely new dataset for sentiment classification and claim-checkworthiness identification, comprising approximately 3,000 samples, and experimented with a classical LSTM framework and two quantum-classical hybrid frameworks based on 2-qubit and 4-qubit VQC. Our experiments and findings show that the classical-quantum hybrid framework outperforms the classical LSTM framework for both sentiment classification and claim checkworthiness identification.

Furthermore, to more accurately and robustly justify our findings and observations, we'll evaluate the proposed frameworks with other languages, such as English, Hindi, Assamese, and Odia, among others. In addition, we will experiment with quantum-enhanced Bi-LSTM, GRU, or Bi-GRU models in our future work.

Limitations

Our proposed work also has some potential limitations. First, all the classical-quantum hybrid frameworks were trained and evaluated on a quantum simulator, which somewhat limits the actual potential of real quantum hardware.

Second, we experimented with only the 2-qubit-based and 4-qubit-based VQCs in the development of the classical-quantum hybrid framework due to resource limitations. Chunking the output of the FC layer into 2 and 4 chunks and providing it through 2-qubit-based and 4-qubit-based VQCs sometimes loses the original contextual relationship between the words in a sentence. In our future work, we will aim to develop advanced techniques to preserve the contextual relationships between words while chunking. Also, we'll experiment with higher qubit VQCs, such as 8-qubit or 16-qubit, in our future work.

Third, due to the lack of trained professional annotators, time constraints, and economic reasons, we have to annotate the sentiment and claim checkworthy labels with the help of LLMs. Although we used three LLMs, followed by majority voting and manual verification, instead of relying on a single LLM model, there might still be some incorrectly annotated samples, as no LLM is 100% accurate. However, we'll aim to develop a fully human-annotated dataset and evaluate our proposed framework with that dataset in our future work.

Fourth, there is a high level of imbalance in the claim-checkworthy labels and sentiment labels in the dataset, which sometimes makes the frameworks biased towards the majority labels. However, in our future work, we'll incorporate more samples into the existing dataset (especially those with minority labels) to make the dataset more balanced.

Lastly, the dataset is limited to news headlines, which restricts our ability to assess the framework's performance in a broader scope, such as with data from Twitter or Reddit.

Acknowledgment

This work was supported by the Defence Research and Development Organisation (DRDO), New Delhi, under the project "Claim Detection and Verification using Deep NLP: an Indian perspective".

References

AmirAli Bagher Zadeh, Paul Pu Liang, Soujanya Poria, Erik Cambria, and Louis-Philippe Morency. 2018. Multimodal language analysis in the wild: CMU-MOSEI dataset and interpretable dynamic fusion graph. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2236–2246, Melbourne, Australia. Association for Computational Linguistics.

Ivano Basile and Fabio Tamburini. 2017. Towards quantum language models. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 1840–1849, Copenhagen, Denmark. Association for Computational Linguistics.

Dushyant Singh Chauhan, Dhanush S R, Asif Ekbal, and Pushpak Bhattacharyya. 2020. Sentiment and emotion help sarcasm? a multi-task learning framework for multi-modal sarcasm, sentiment and emotion analysis. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 4351–4360, Online. Association for Computational Linguistics.

Bob Coecke, Giovanni de Felice, Konstantinos Meichanetzidis, and Alexis Toumi. 2020. Foundations for near-term quantum natural language processing. *Preprint*, arXiv:2012.03755.

Joseph L. Fleiss. 1971. Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5):378–382.

Srinjoy Ganguly, Sai Nandan Morapakula, and Luis Miguel Pozo Coronado. 2022. Quantum natural language processing based sentiment analysis using lambeq toolkit. In *2022 Second International Conference on Power, Control and Computing Technologies (ICPC2T)*, pages 1–6.

Raffaele Guarasci, Giuseppe De Pietro, and Massimo Esposito. 2022. Quantum natural language processing: Challenges and opportunities. *Applied Sciences*, 12(11).

Kilem Li Gwet. 2006. Computing inter-rater reliability and its variance in the presence of high agreement. *British Journal of Mathematical and Statistical Psychology*, 61(1):29–48.

Laszlo Gyongyosi and Sandor Imre. 2019. A survey on quantum computing technology. *Comput. Sci. Rev.*, 31(C):51–71.

Dimitri Kartsaklis, Ian Fan, Richie Yeung, Anna Pearson, Robin Lorenz, Alexis Toumi, Giovanni de Felice, Konstantinos Meichanetzidis, Stephen Clark, and Bob Coecke. 2021. lambeq: An efficient high-level python library for quantum nlp. *Preprint*, arXiv:2110.04236.

Diederik P. Kingma and Jimmy Ba. 2017. Adam: A method for stochastic optimization. *Preprint*, arXiv:1412.6980.

Ziming Li, Yan Zhou, Yaxin Liu, Fuqing Zhu, Chuanpeng Yang, and Songlin Hu. 2023. QAP: A quantum-inspired adaptive-priority-learning model for multimodal emotion recognition. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 12191–12204, Toronto, Canada. Association for Computational Linguistics.

Yaochen Liu, Yazhou Zhang, Qiuchi Li, Benyou Wang, and Dawei Song. 2021. What does your smile mean? jointly detecting multi-modal sarcasm and sentiment using quantum probability. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 871–880, Punta Cana, Dominican Republic. Association for Computational Linguistics.

Niels Neumann, Frank Phillipson, and Richard Versluis. 2019. Machine learning in the quantum era. *Digitale Welt*, 3(2):24–29.

OpenAI, Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, Red Avila, Igor Babuschkin, Suchir Balaji, Valerie Balcom, Paul Baltescu, Haiming Bao, Mohammad Bavarian, Jeff Belgum, and 9 others. 2024. Gpt-4 technical report. *Preprint*, arXiv:2303.08774.

Pritam Pal and Dipankar Das. 2025. Toward quantum-enhanced natural language understanding: Sarcasm and claim detection with qlstm. In *Proceedings of the 15th International Conference on Recent Advances in Natural Language Processing - Natural Language Processing in the Generative AI era*, pages 854–859, Varna, Bulgaria. INCOMA Ltd., Shoumen, Bulgaria.

Shyambabu Pandey, Pankaj Dadure, Morrel V. L. Nunsanga, and Partha Pakray. 2022. Parts of speech tagging towards classical to quantum computing. In *2022 IEEE Silchar Subsection Conference (SILCON)*, pages 1–6.

Shyambabu Pandey and Partha Pakray. 2023. Bi-quantum long short-term memory for part-of-speech tagging. In *Proceedings of the 20th International*

Conference on Natural Language Processing (ICON), pages 301–307, Goa University, Goa, India. NLP Association of India (NLPAI).

Bo Pang and Lillian Lee. 2005. Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales. In *Proceedings of the ACL*.

Arpan Phukan and Asif Ekbal. 2023. QeMMA: Quantum-enhanced multi-modal sentiment analysis. In *Proceedings of the 20th International Conference on Natural Language Processing (ICON)*, pages 815–821, Goa University, Goa, India. NLP Association of India (NLPAI).

John Proos and Christof Zalka. 2004. Shor's discrete logarithm quantum algorithm for elliptic curves. *Preprint*, arXiv:quant-ph/0301141.

Jun Qi, Chao-Han Huck Yang, and Pin-Yu Chen. 2021. Qtn-vqc: An end-to-end learning framework for quantum neural networks. *Preprint*, arXiv:2110.03861.

Wenbo Qiao, Peng Zhang, and ZengLai Ma. 2024. A quantum-inspired matching network with linguistic theories for metaphor detection. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 1435–1445, Torino, Italia. ELRA and ICCL.

R. L. Rivest, A. Shamir, and L. Adleman. 1978. A method for obtaining digital signatures and public-key cryptosystems. *Commun. ACM*, 21(2):120–126.

Fariska Z. Ruskanda, Muhammad Rifat Abiwardani, Muhammad Akram Al Bari, Kinantan Arya Bagaspati, Rahmat Mulyawan, Infall Syafalni, and Harashta Tatimma Larasati. 2022. Quantum representation for sentiment classification. In *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)*, pages 67–78.

Fariska Zakhralativa Ruskanda, Muhammad Rifat Abiwardani, Infall Syafalni, Harashta Tatimma Larasati, and Rahmat Mulyawan. 2023. Simple sentiment analysis ansatz for sentiment classification in quantum natural language processing. *IEEE Access*, 11:120612–120627.

Chhavi Sharma, Deepesh Bhageria, William Scott, Srinivas PYKL, Amitava Das, Tanmoy Chakraborty, Viswanath Pulabaigari, and Björn Gambäck. 2020. SemEval-2020 task 8: Memotion analysis- the visuo-lingual metaphor! In *Proceedings of the Fourteenth Workshop on Semantic Evaluation*, pages 759–773, Barcelona (online). International Committee for Computational Linguistics.

Jinjing Shi, Zhenhuan Li, Wei Lai, Fangfang Li, Ronghua Shi, Yanyan Feng, and Shichao Zhang. 2023. Two End-to-End Quantum-Inspired Deep Neural Networks for Text Classification . *IEEE Transactions on Knowledge & Data Engineering*, 35(04):4335–4345.

Peter W. Shor. 1997. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509.

Riccardo Di Sipio, Jia-Hong Huang, Samuel Yen-Chi Chen, Stefano Mangini, and Marcel Worring. 2021. The dawn of quantum natural language processing. *Preprint*, arXiv:2110.06510.

Fabio Tamburini. 2019. A quantum-like approach to word sense disambiguation. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2019)*, pages 1176–1185, Varna, Bulgaria. INCOMA Ltd.

Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, Aurelien Rodriguez, Armand Joulin, Edouard Grave, and Guillaume Lample. 2023. Llama: Open and efficient foundation language models. *Preprint*, arXiv:2302.13971.

Charles M. Varmantchaonala, Jean Louis K. E. Fendji, Julius Schöning, and Marcellin Atemkeng. 2024. Quantum natural language processing: A comprehensive survey. *IEEE Access*, 12:99578–99598.

Dominic Widdows, Willie Aboumrad, Dohun Kim, Sayonee Ray, and Jonathan Mei. 2024. Quantum natural language processing. *KI - Künstliche Intelligenz*.

Sixuan Wu, Jian Li, Peng Zhang, and Yue Zhang. 2021. Natural language processing meets quantum physics: A survey and categorization. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 3172–3182, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.

Wenduan Xu, Stephen Clark, Douglas Brown, Gabriel Matos, and Konstantinos Meichanetzidis. 2024. Quantum recurrent architectures for text classification. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 18020–18027, Miami, Florida, USA. Association for Computational Linguistics.

Kehuan Yan, Peichao Lai, and Yilei Wang. 2024. Quantum-inspired language model with lindblad master equation and interference measurement for sentiment analysis. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 2112–2121, Mexico City, Mexico. Association for Computational Linguistics.

Yazhou Zhang, Qiuchi Li, Dawei Song, Peng Zhang, and Panpan Wang. 2019. Quantum-Inspired Interactive Networks for Conversational Sentiment Analysis. *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI)*, pages 5436–5442.

A Hybrid Quantum-Classical Fusion for Deep Semantic Paraphrase Detection

Devanarayanan K and **Fayas S Mohamad** and **Dheeraj V Mohan** and **Reshma Sheik**
TKM College of Engineering, Kollam, Kerala.
devanarayanank.2003@gmail.com, rezmasheik@gmail.com

Abstract

Paraphrase Detection is a core task in natural language processing (NLP) that aims to determine whether two sentences convey equivalent meanings. This work proposes a hybrid quantum–classical framework that integrates Sentence-BERT embeddings, simulated quantum feature encoding, and classical machine learning models to enhance semantic similarity detection. Initially, sentence pairs are embedded using Sentence-BERT and standardized through feature scaling. These representations are then transformed via rotation-based quantum circuits to capture higher-order feature interactions and non-linear dependencies. The resulting hybrid feature space, combining classical and quantum-inspired components are evaluated using LightGBM and deep neural network classifiers. Experimental results shows that the hybrid model incorporating quantum-inspired features achieved superior classification performance, yielding a 10% improvement in overall accuracy outperforming standalone deep learning baselines. These findings demonstrate that quantum–classical fusion enhances semantic feature extraction and significantly improves paraphrase detection performance.

1 Introduction

Recent studies have explored the intersection of quantum computing and Natural Language Processing (NLP) to enhance semantic understanding and text similarity modeling. Paraphrase detection is an important task in natural language processing that aims to identify whether two sentences convey the same meaning. It has applications in areas such as question answering, plagiarism detection, and semantic search (Madaan et al., 2016). Classical machine learning methods have achieved significant progress using embedding models and gradient boosting techniques. However, capturing deeper semantic relationships between sentence pairs remains a challenge due to the limitations of classical representations. Quantum computing provides an exciting way through the encoding of linguistic information into quantum states, which have a natural way to represent and process the correlations that are hard to model in classical environments. Emerging advances in Quantum NLP (Q-NLP) demonstrate that quantum circuits can represent structural and semantic relationships between sentence parts in manners that complement classical neural architectures (Meichanetzidis et al., 2023).

Earlier works (Buhrman et al., 2001) introduced quantum fingerprinting, demonstrating how quantum states can represent compact data signatures for efficient comparison—laying the theoretical foundation for quantum information comparison techniques. (Darwish et al., 2023) proposed a quantum genetic algorithm for semantic textual similarity estimation in plagiarism detection, highlighting quantum-enhanced optimization in NLP tasks. Gao et al. (Gao et al., 2024) developed a quantum-inspired hierarchical semantic interaction model for text classification that captures multi-level contextual relations between words. Meanwhile, Guarasci et al. (Guarasci et al., 2022) discussed the broader challenges and opportunities in quantum natural language processing, emphasizing scalability, noise resilience, and quantum circuit design constraints. In contrast, the present research focuses specifically on paraphrase detection using a hybrid quantum–classical framework, integrating both classical semantic embeddings and quantum circuit-based similarity estimation for more accurate and interpretable detection of paraphrased sentences. Due to the exponential cost of simulating larger circuits, the initial system encodes a low-dimensional subset of SBERT features into a 4-qubit circuit as a feasibility study. This establishes a baseline for scaling quantum components in future work.

The remainder of this paper is organized as

Proceedings of the QuantumNLP: Integrating Quantum Computing with Natural Language Processing, pages 20–25
November 24, 2025

follows: Section 2 defines the paraphrase detection task and describes the dataset used in this study. Section 3 outlines the proposed hybrid quantum–classical framework, including data preprocessing, SBERT embeddings, quantum feature generation, and classifier design. Section 4 reports the experimental results and performance analysis, while Section 5 concludes the paper.

2 Task and Dataset

The task of paraphrase detection can be defined formally as follows: given two input sentences *s1* and *s2*, decide if they are semantically equivalent. Although certain pairs can be determined by direct word overlap, most need more in-depth modeling of sentence structure, context, and meaning. The core problem is to identify semantic similarity that goes beyond surface-level patterns of words. This study used a supervised Kaggle dataset, equivalent to the Microsoft Research Paraphrase Corpus (MRPC) (Dolan and Brockett, 2005), which consists of paired text samples divided into training and test sets. The training set contains 3,554 sentence pairs, while the test set has 1,465 pairs. Each record in the dataset is structured with five columns: two identifier fields (#1 ID, #2 ID), two text fields (#1 String, #2 String), and one label (Quality) indicating the relationship between the sentences. The label indicates whether the two sentences are paraphrased (1) or not (0). The data set is in CSV format, encoded in UTF-8, and uses consistent delimiters for easy integration into machine learning models. Its structure allows experiments in sentence-level detection using classical, deep learning, or hybrid quantum (Biamonte et al., 2017) deep learning approaches. Using MRPC ensures standard benchmarking and comparability with prior NLP research.[1]

3 Methodology

This work explores a hybrid quantum-classical approach for paraphrase detection. The model first encodes each pair of sentences using Sentence-BERT (Reimers and Gurevych, 2019) to obtain dense vector embeddings. The hybrid representation—combining both classical embeddings and quantum features—is reduced in dimensionality using Principal Component Analysis (PCA) (Jolliffe, 2002). Two main classifiers are then applied: LightGBM for boosted decision-tree learning and a Multi-Layer Perceptron (MLP) for deep learning inference. The final output is a binary prediction indicating whether the input sentences are paraphrased (1) or not (0).

Figure 1 illustrates the proposed hybrid quantum–classical system architecture, which is organized into four primary layers, each contributing to efficient paraphrase detection through integrated quantum–classical processing (Havlíček et al., 2019). The Data Preprocessing Layer is responsible for acquiring, cleaning, and organizing the input data. It pairs sentences with their corresponding labels, removes missing or noisy entries, ensures balanced class distribution, and stores the cleaned data along with their embeddings for subsequent processing. The Embedding Layer uses SentenceBERT that transforms textual data into dense numerical representations using sentence-level embedding models, capturing the semantic relationships necessary for downstream learning. Each of them acts as a parameter for the learning and testing stages. The features are balanced by standardization for fast convergence, so that each parameter has mean 0 and standard deviation 1. The Hybrid Processing Layer augments these classical embeddings with quantum-enhanced representations to capture higher-order dependencies and improve discriminative capability. This layer integrates modules for quantum feature generation, dimensionality reduction using PCA, and feature fusion to form a unified hybrid feature space. Finally, the Learning and Prediction Layer manages model training and inference, leveraging both classical and hybrid machine learning models to perform paraphrase classification as a binary classification task.

3.1 Data Preprocessing

The preprocessing stage begins by identifying the text and label columns in the dataset. Rows with missing values in these columns are removed to maintain data consistency. The text columns are cast to string type, and the label column to integer type. Sentence pairs are then constructed by concatenating the two text columns with a separator token. For embedding generation, SBERT encodes each sentence pair into dense numeric vectors, which are standardized using a StandardScaler to achieve zero mean and unit variance. These standardized embeddings are subsequently used for

[1] https://www.kaggle.com/datasets/doctri/microsoft-research-paraphrase-corpus

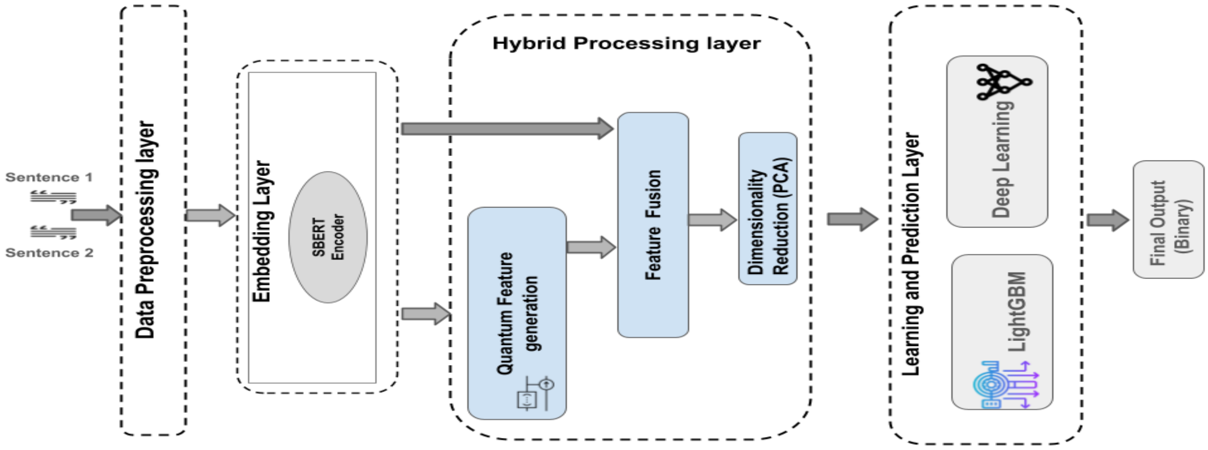


Figure 1: Hybrid Quantum-Classical Architecture for Paraphrase Detection

quantum feature generation. No additional text normalization, such as lowercasing or punctuation removal, is applied.

3.2 Quantum Feature Generation

The quantum feature generation process begins with the SBERT embeddings of sentence pairs. In the current implementation, four qubits are employed. The first four parameters of the 384-dimensional embedding vector are selected for quantum encoding.The first four SBERT dimensions were selected to ensure deterministic, reproducible mapping for a cost-feasible simulation. Future work will incorporate feature-selection methods (filter/wrapper techniques) to identify more discriminative embedding dimensions for quantum encoding. Each parameter is normalized to the range $[-\pi, \pi]$ and mapped to a qubit using an R_y rotation gate. The individual qubit states are combined using the Kronecker product to form a multi-qubit quantum state. Entanglement is introduced through a chain of CNOT gates connecting qubit 0 to 1, 1 to 2, and 2 to 3, thereby capturing correlations among qubits.CNOT gates introduce entanglement among the encoded parameters, allowing the measurement distribution to capture interaction effects beyond linear SBERT encoding. These interactions contribute to the hybrid feature space's expressiveness. The resulting quantum state is measured, yielding two measurement values per qubit, for a total of sixteen output values. These outputs constitute the quantum feature vector, which is then used as input to the hybrid model alongside the SBERT embeddings. Figure 2 illustrates the quantum circuit for the hybrid models, where four qubits are initialized with R_Y rotations, entangled via CNOT gates. Only a 4-dimensional slice of the SBERT vector is used for quantum encoding, as compact encoding is designed as a nonlinear feature transformation rather than a full high-dimensional quantum embedding. The circuit is fixed and non-trainable, with parameters directly mapped from SBERT values.

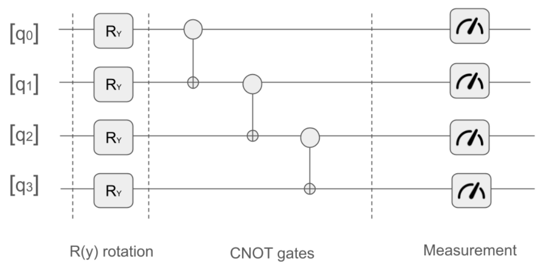


Figure 2: Quantum Circuit

3.3 Classifier Models

This section presents three model variants used in our study for paraphrase detection: (1) a LightGBM (LGBM) classifier based on gradient boosting with optimized hyperparameters; (2) a deep learning (DL) model comprising three fully connected layers with ReLU activations and dropout regularization; and (3) a hybrid model that integrates SBERT embeddings with quantum-inspired features, used in conjunction with either the LGBM or DL classifier to exploit both classical and quantum representations.

3.3.1 LGBM

LightGBM (Ke et al., 2017) is a gradient boosting framework that builds ensembles of decision trees sequentially, with each tree aiming to correct the residual errors of its predecessors. The model is tuned using a hyperparameter grid with the number of leaves 31, 63, 127, 255 and learning rates 0.01, 0.05, 0.1, over a maximum of 1000 boosting iterations. SBERT embeddings of sentence pairs serve as input features, and the best combination of leaves and learning rate is selected as the final model. Figure 3 visually explains the "leaf-wise" growth strategy of the LGBM algorithm. Instead of growing level by level, the tree is built by expanding the leaf that will cause the largest reduction in error.

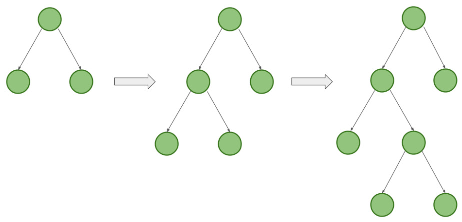


Figure 3: Growth strategy of the LGBM algorithm

3.3.2 DL

We employ a fully connected feedforward neural network for classification. The model takes SBERT embeddings as input and passes them through three hidden layers of 2048, 1024, and 512 neurons with ReLU activations and 0.3 dropout. It is trained for 50 epochs using the Adam (Kingma, 2014) optimizer and cross-entropy loss with a batch size of 128. Figure 4 depicts the Deep Learning model used in this study. It is a fully connected feedforward network comprising multiple hidden layers with ReLU activations, each followed by dropout for regularization, and a Softmax output layer for classification.

3.3.3 Hybrid Model

The hybrid model combines SBERT embeddings with quantum-inspired features to enhance the learning. The SBERT embeddings and quantum features are concatenated to form a hybrid feature vector. This hybrid representation serves as input to either a hybrid LGBM classifier or a hybrid DL model.

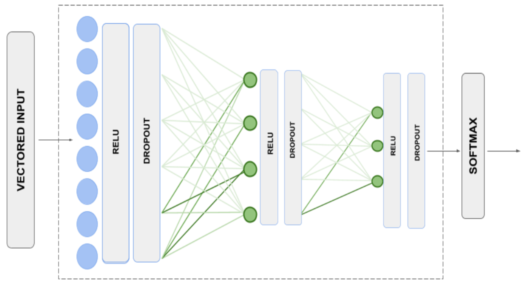


Figure 4: Deep Learning Model Architecture

4 Results & Discussions

Table 1 shows the performance metrics of the models in different sizes of training data. In terms of accuracy, the Hybrid-LGBM model achieved the highest overall performance, reaching 0.69 in 70% of the data. It consistently outperformed all other models across the 20–80% data range. The DL and Hybrid-DL models lagged behind the LGBM and Hybrid-LGBM models, likely due to overfitting in the DL models. The performance of Hybrid-DL improves significantly at higher data percentages (80–100%). In precision, the Hybrid-LGBM again outperformed the other models, achieving its best precision at 70% of the data. The DL and Hybrid-DL models initially showed lower precision, but the Hybrid-DL model steadily improved after 60% of the data, ultimately achieving its highest precision at 100% of the dataset. Recall indicates that LightGBM and Hybrid-LGBM perform best, both clearly outperforming DL and Hybrid-DL, with Hybrid-LGBM showing slightly higher recall than LGBM. Hybrid-DL improves notably after 60% of the data, surpassing DL at 80% and peaking at 100%. While both DL models perform similarly on smaller datasets, Hybrid-DL achieves higher F1 scores at larger data sizes (80–100%), outperforming the other models. At 100 percent training data, the Hybrid-LGBM model exhibited higher accuracy but a lower F1 score due to dataset class imbalance. LGBM optimizes leaf-wise splits that increase precision at the cost of recall, which impacts the F1 metric. The best balance between accuracy and F1 was observed in 70% training data. The Hybrid-DL model showed sensitivity to overfitting due to the larger hybrid feature dimension. Additional regularization and smaller architectures will be explored in future phases.

%	Model	Accuracy	Precision	Recall	F1-Score
10	LGBM	0.6767	0.6559	0.6767	0.6228
	DL	0.6528	0.6270	0.6528	0.6281
	Hybrid-LGBM	0.6685	0.6440	0.6685	0.5998
	Hybrid-DL	0.6432	0.6262	0.6432	0.6309
20	LGBM	0.6828	0.6671	0.6828	0.6297
	DL	0.6576	0.6348	0.6576	0.6365
	Hybrid-LGBM	0.6849	0.6868	0.6849	0.6142
	Hybrid-DL	0.6411	0.6297	0.6411	0.6338
30	LGBM	0.6842	0.6659	0.6842	0.6403
	DL	0.6336	0.6226	0.6336	0.6268
	Hybrid-LGBM	0.6863	0.6757	0.6863	0.6299
	Hybrid-DL	0.6528	0.6281	0.6528	0.6299
40	LGBM	0.6808	0.6626	0.6808	0.6293
	DL	0.6514	0.6362	0.6514	0.6405
	Hybrid-LGBM	0.6883	0.6794	0.6883	0.6326
	Hybrid-DL	0.6377	0.6262	0.6377	0.6305
50	LGBM	0.6863	0.6745	0.6863	0.6317
	DL	0.6364	0.6321	0.6364	0.6340
	Hybrid-LGBM	0.6876	0.6844	0.6876	0.6248
	Hybrid-DL	0.6494	0.6391	0.6494	0.6429
60	LGBM	0.6876	0.6701	0.6876	0.6478
	DL	0.6391	0.6365	0.6391	0.6377
	Hybrid-LGBM	0.6958	0.7004	0.6958	0.6359
	Hybrid-DL	0.6329	0.6204	0.6329	0.6250
70	LGBM	0.6931	0.6858	0.6931	0.6414
	DL	0.6521	0.6472	0.6521	0.6494
	Hybrid-LGBM	0.6979	0.7060	0.6979	0.6374
	Hybrid-DL	0.6507	0.6426	0.6507	0.6458
80	LGBM	0.6876	0.6775	0.6876	0.6327
	DL	0.6391	0.6241	0.6391	0.6289
	Hybrid-LGBM	0.6910	0.6893	0.6910	0.6311
	Hybrid-DL	0.6651	0.6514	0.6651	0.6551
90	LGBM	0.6958	0.6894	0.6958	0.6462
	DL	0.6473	0.6413	0.6473	0.6439
	Hybrid-LGBM	0.6931	0.6912	0.6931	0.6357
	Hybrid-DL	0.6746	0.6624	0.6746	0.6656
100	LGBM	0.6972	0.6943	0.6972	0.6451
	DL	0.6644	0.6508	0.6644	0.6545
	Hybrid-LGBM	0.6924	0.6986	0.6924	0.6277
	Hybrid-DL	0.6801	0.6678	0.6801	0.6708

Table 1: Performance metrics of models across different training data percentages.

Figure 5 and Figure 6 compare the accuracy and F1 score of four models—LGBM, DL, Hybrid-LGBM, and Hybrid-DL across varying amounts of training data.

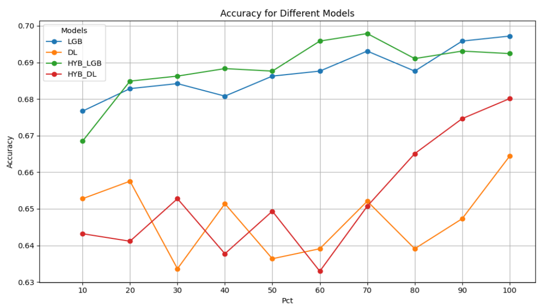


Figure 5: Accuracy

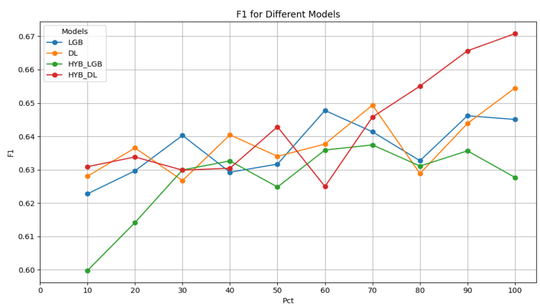


Figure 6: F1 Score

5 Conclusion

We have compared quantum–classical hybrid learning with classical learning architectures in datascarce environments. Evaluating LGBM, DL, Hybrid-LGBM, and Hybrid-DL, we observed that Hybrid-LGBM consistently delivers competitive performance while demonstrating superior data efficiency, achieving a maximum accuracy of 0.69 with 70% of training data. Future work may further improve accuracy by experimenting with alternative quantum circuits and varying the number of qubits.

6 Ethics

The dataset consists of publicly available, non-sensitive text corpora. Experiments comply with data licenses and research standards, with no human subjects involved, so ethical approval was not required. The hybrid quantum–classical framework is for research purposes only, and all references are acknowledged.

References

Jacob Biamonte, Peter Wittek, Nicola Pancotti, Patrick Rebentrost, Nathan Wiebe, and Seth Lloyd. 2017. Quantum machine learning. *Nature*, 549:195–202.

Harry Buhrman, Richard Cleve, John Watrous, and Ronald de Wolf. 2001. Quantum fingerprinting. *Physical Review Letters*, 87(16):167902.

S. M. Darwish, I. A. Mhaimeed, and A. A. Elzoghabi. 2023. A quantum genetic algorithm for building a semantic textual similarity estimation framework for plagiarism detection applications. *Entropy*, 25(9):1271.

William B. Dolan and Chris Brockett. 2005. Automatically constructing a corpus of sentential paraphrases. *Microsoft Research Technical Report (MRPC)*.

H. Gao, P. Zhang, J. Zhang, and C. Yang. 2024. A quantum-inspired hierarchical semantic interaction model for text classification. *Neurocomputing*, 611:128658.

R. Guarasci, G. De Pietro, and M. Esposito. 2022. Quantum natural language processing: challenges and opportunities. *Applied Sciences*, 12(11):5651.

Vojtěch Havlíček, Antonio D. Córcoles, Kristan Temme, Aram W. Harrow, Abhinav Kandala, Jerry M. Chow, and Jay M. Gambetta. 2019. Supervised learning with quantum-enhanced feature spaces. In *Nature*, volume 567, pages 209–212.

Ian T. Jolliffe. 2002. Principal component analysis (2nd ed.). In *Springer Series in Statistics*. Springer.

Guolin Ke, Qiwei Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qi Ye, and Tie-Yan Liu. 2017. Lightgbm: A highly efficient gradient boosting decision tree. In *Advances in Neural Information Processing Systems (NeurIPS)*.

Diederik P Kingma. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Vishu Madaan, Prateek Agrawal, Nandini Sethi, Vikas Kumar, and Sanjay Kumar Singh. 2016. A novel approach to paraphrase english sentences using natural language processing. *International Journal of Control Theory and Applications*, 9(11):5119–5128.

Konstantinos Meichanetzidis, Alexis Toumi, Giovanni de Felice, and Bob Coecke. 2023. Grammar-aware sentence classification on quantum computers. *Quantum Machine Intelligence*, 5(1):10.

Nils Reimers and Iryna Gurevych. 2019. Sentence-bert: Sentence embeddings using siamese bert-networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*.

Quantum-Enhanced Gated Recurrent Units for Part-of-Speech Tagging

Ashutosh Rai
National Institute of Technology Silchar
ashutoshrai_pg_24@cse.nits.ac.in

Shyambabu Pandey
National Institute of Technology Silchar
shyambabu21_rs@cse.nits.ac.in

Partha Pakray
National Institute of Technology Silchar
partha@cse.nits.ac.in

Abstract

Deep learning models for Natural Language Processing (NLP) tasks, such as Part-of-Speech (POS) tagging, usually have significant parameter counts that make them costly to train and deploy. Quantum Machine Learning (QML) offers a potential approach for building more parameter-efficient models. This paper proposes a hybrid quantum-classical gated recurrent unit model for POS tagging in code-mixed social media text. By integrating a quantum layer into the recurrent framework, our model achieved an accuracy comparable to the baseline classical model, while needing fewer parameters. Although the cut-off point in the parameters is modest in our setup, the approach is promising when scaled to deeper architectures. These results suggest that hybrid models can offer a resource-efficient alternative for NLP tasks.

1 Introduction

Understanding natural human language, which is a central basis of communication, has been a long-standing goal of artificial intelligence (Russell and Norvig, 2010). Natural language processing (NLP) successfully tackles this problem by developing methods for machines to read, examine, and produce natural language in ways that support tangible real-world applications (Jurafsky, 2000). Today, NLP supports applications such as conversational assistants, automatic translation systems, and opinion mining tools, making it an important part of our daily engagement with digital technology.

The recent success of NLP is mainly attributed to improvements in machine learning (Janiesch et al., 2021). Training models on large amounts of data makes them capable of learning and recognizing patterns in text and making accurate predictions for tasks like translation, sentiment analysis, and sequence labeling. Neural networks, specifically, have brought about significant developments by modeling complex relationships within language data (Sharkawy, 2020). However, as data sets grow larger and architectures deeper, these models become resource intensive, requiring large amounts of memory and computation for both training and inference (Janiesch et al., 2021).

Quantum computing is one avenue that offers a possible way forward by providing a different and more efficient method of computation (Gyongyosi and Imre, 2019). Quantum characteristics such as superposition and entanglement are essential to how information can be represented and operated on with much greater expressive power than classical bits allow. Based on these principles, quantum machine learning (QML) has emerged as a research field that seeks to merge quantum computation with machine learning methods (Schuld and Petruccione, 2021). Although a nascent field, QML has been explored as an alternative to designing more compact models that can capture patterns differently from their classical counterparts.

Putting these principles into practice, this work solves an important NLP task: part-of-speech (POS) tagging in code-mixed text data from social networks (Pandey et al., 2023). POS tagging works by assigning grammatical roles to each word in a sentence and is a crucial step in syntactic and semantic analysis (Basisth et al., 2023). We present a hybrid quantum gated recurrent units (QGRU) model that integrates a quantum layer with classical recurrent layers. To evaluate the performance of the proposed model, we perform POS tagging on a code-mixed dataset. Based on our findings, this approach competes with classical baselines in accuracy but achieves similar performance with fewer trainable parameters, making it parameter efficient. Still, the approach suggests that greater savings could be realized when scaling to larger architectures, where substituting intermediate layers with quantum circuits may yield noticeable efficiency gains.

Proceedings of the QuantumNLP: Integrating Quantum Computing with Natural Language Processing, pages 26–32
November 24, 2025

The structure of this paper is as follows. Section 2 introduces background on quantum computing and QML, Section 3 reviews related research, Section 4 details the proposed model, Section 5 describes the dataset, Section 6 reports results and analysis, and Section 7 concludes with future directions.

2 Background

2.1 Quantum Computing

Quantum computing is a paradigm of computation that uses the principles of quantum mechanics to process information in ways that are not possible with classical systems (Gyongyosi and Imre, 2019). In a classical computer, the basic unit of information is the bit, which can take one of two values, 0 or 1. In quantum computing, the basic unit is the quantum bit, or qubit. A qubit has two basic states, $|0\rangle$ and $|1\rangle$, which are called computational basis states. These basis states are commonly represented in vector form as

$$|0\rangle = \begin{bmatrix}1\\0\end{bmatrix}, \quad |1\rangle = \begin{bmatrix}0\\1\end{bmatrix} \tag{1}$$

Unlike a classical bit, which can only be 0 or 1 at a time, a qubit can exist in a superposition of both states. The state of a single qubit can be expressed as

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \quad \text{with } |\alpha|^2 + |\beta|^2 = 1 \tag{2}$$

where α and β are complex amplitudes. The normalization condition ensures that the total probability of measuring the qubit in either state is one. When multiple qubits are combined, they form a joint system described by the tensor product of individual qubit states. For example, the state of two qubits $|\psi\rangle \otimes |\phi\rangle$ can be written as

$$|\psi\phi\rangle = \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle \tag{3}$$

This shows that a two-qubit system can represent all four possible basis states at the same time. In general, a n-qubit system can represent 2^n states in parallel, which provides exponential representational power compared to classical bits (Pandey and Pakray, 2023). Another important property is entanglement. Entangled qubits are correlated in such a way that the state of one qubit cannot be described independently of the other. For instance, an entangled two-qubit system may be described as

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \tag{4}$$

In this state, measuring the first qubit immediately determines the outcome of the second. Entanglement enables forms of information processing that are not possible with classical systems.

Quantum operations are carried out using quantum gates, which are unitary matrices that transform qubit states while preserving normalization. For example, the Hadamard gate H creates a superposition state:

$$H|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \quad H|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \tag{5}$$

By combining such gates into circuits, quantum computers can implement a wide variety of computations. At the end of a computation, qubits are measured, and the superposition collapses into one of the basis states, with probabilities determined by the amplitudes.

Together, these basic elements, qubits, superposition, entanglement, quantum gates, and measurement, form the foundation of quantum computing. They allow quantum systems to process and represent information in fundamentally different ways than classical systems, opening possibilities for speed-ups in certain computational tasks.

2.2 Quantum Machine Learning

Quantum machine learning (QML) is an emerging area of research that combines the principles of quantum computing with machine learning (Schuld and Petruccione, 2021). The goal is to take advantage of the unique properties of quantum computation to help improve the process of learning from data. While classical machine learning relies on algorithms that run on conventional hardware, QML explores how quantum states and operations can be used to represent and process information.

In general, a QML model makes use of quantum circuits whose parameters can be adjusted during training, similar to how weights are updated in a neural network. After a computation, quantum systems are measured, and the results are expressed as expectation values of observables. The outcome of a quantum measurement is typically expressed as the expectation value of an observable. For a quantum state $|\psi\rangle$ and an observable Z, the expectation value is defined as

$$\langle Z \rangle = \langle \psi | Z | \psi \rangle \tag{6}$$

The output expectation values can now be used in the same way that the output of a classical model would be used, for instance, in calculating a loss function during training.

The major advantages of QML include fewer parameters, high-dimensional solution spaces, and the possibility of forming correlations through entanglement that is not possible while using classical models. Quantum methods can also provide performance gains for specific computation-related tasks. However, these gains are highly dependent on the application at hand and the current constraints of quantum hardware. Currently, most QML methods are implemented in a hybrid manner, where quantum circuits are merged with classical machine learning components and trained using standard optimization methods (Sweke et al., 2020).

Recent work shows that there is growing interest in applying QML to domains such as optimization, quantum chemistry, and NLP (Pandey et al., 2023). NLP tasks, in particular, are challenging due to their heavy dependence on large datasets and complex models with deep architectures, making them a viable area of exploration for possible benefits of QML. This motivates exploring QML in problems such as POS tagging, where both efficiency and performance are important considerations.

3 Related Work

POS tagging is a fundamental task in NLP, serving as a foundational step for many downstream applications. The classical state-of-the-art for sequence labeling tasks such as POS tagging has been dominated by recurrent neural networks, particularly Bidirectional Long Short-Term Memory (BiLSTM) and Gated Recurrent Unit (GRU) architectures, often paired with a Conditional Random Field (CRF) layer (Lample et al., 2016). However, such models are typically parameter heavy and their application to noisy code-mixed social media text presents many challenges (Jamatia et al., 2015).

Coecke et al. (Coecke et al., 2020) introduced a grammar-aware compositional DisCoCat framework that maps the sentence structure to quantum circuits. This addresses directly the growing computational demands of traditional machine learning and deep learning models by leveraging quantum circuits for language tasks. Our work, in contrast, integrates a variational quantum algorithm in the form of a parameterized circuit directly into a deep learning model.

Many studies have shown that such a hybrid approach is valid for NLP tasks Pandey et al. (2024). Another work by Shi et al. Shi et al. (2023) details a quantum-inspired neural network that uses complex-valued embeddings to capture better semantic information.These works showcase the potential of using quantum principles to enhance classical NLP architectures.

A more recent development with a direct relation to our task, the application of quantum circuits to POS tagging, is demonstrated by Di Sipio et al. Di Sipio et al. (2022). The authors introduced a Quantum Long-Short-Term Memory (QLSTM) model applied to a sequence tagging task. This foundational work was extended by Pandey et al. Pandey et al. (2022) in a low-resource language. The same group later modified the QLSTM model specifically for code-mixed social media data (Pandey et al., 2023) and advanced it by making a bidirectional variant (BiQLSTM) (Pandey and Pakray, 2023)).

4 Architecture

This section discusses the architectures of the two models that are compared in our study, a fully classical model, which serves as our baseline, and the proposed hybrid quantum-classical model.

4.1 Classical Model

We chose to use a standard architecture for our baseline model. It is built using Gated Recurrent Units (GRU) (Cho et al., 2014). The model takes as input sequences 100-dimensional word embeddings and processes them via two bidirectional GRU layers with a hidden state dimension of 16. Bidirectionality allows the GRU layers to capture context-aware representations by processing the information from both preceding and succeeding tokens in the sequence. The output obtained from the GRU layers is passed through a fully connected classification head, which helps map the hidden states to a dimension corresponding to the number of POS tags.

The output of the fully connected layer is passed to a Conditional Random Field (CRF) layer which produces the final tag sequence (Lample et al., 2016). CRF is a statistical modeling method that learns transition probabilities between adjacent tags to support sequence tagging tasks. This helps the model to consider the context of neighboring predictions based on which the model can penalize grammatically unlikely tag sequences, thereby improving the accuracy and coherence of the output.

4.2 Hybrid Model

Our proposed hybrid model uses the same core layers as the baseline models. Embedding, Bidirectional GRU and the CRF layers are used in the hybrid model as well. The only distinction is the quantum layer that replaces the fully-connected classification head. The quantum layer receives its input from the fully connected layer attached to the GRU layers. The main purpose of this fully connected layer is to downsize the output from the GRU layers to match the input size of the quantum layer. It was included in the baseline model to ensure consistency.

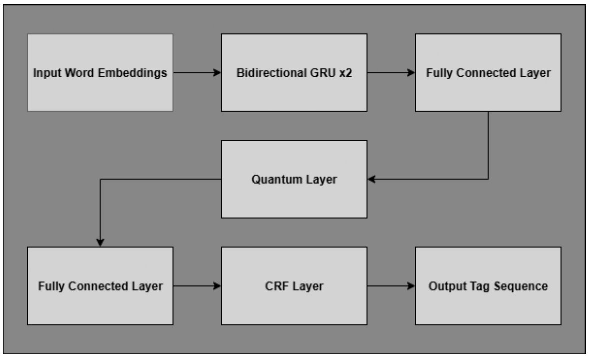


Figure 1: Architecture of the Hybrid Model.

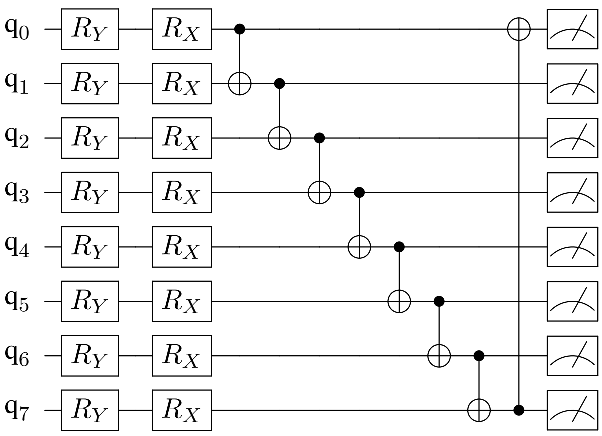


Figure 2: The 8-qubit variational quantum circuit. The initial R_Y gates are parameterized by input features, and the R_X gates are parameterized by trainable weights. This entire entangling block is repeated 6 times.

Algorithm 1 Quantum Circuit Layer

Input: Classical feature vector $x \in \mathbb{R}^8$, quantum circuit weights W.
Output: Expectation values vector $E \in \mathbb{R}^8$.
Initialize 8-qubit state to $|0\rangle^{\otimes 8}$.
Encode x into the state using AngleEmbedding.
Apply the variational BasicEntanglingLayers circuit parameterized by weights W.
for $i = 0$ 7 **do**
    Measure $\langle \sigma_z \rangle$ on qubit i.
    $E_i \leftarrow \langle \sigma_z \rangle_i$.
end for
return E.

The input to the quantum layer is an 8 dimensional vector This vector is encoded and processed by a variational quantum circuit. The quantum circuit consists of two main components. Each element in the input vector is first encoded on a qubit using an AngleEmbedding layer, a standard method for mapping feature vectors into qubit rotations. Following this, a BasicEntanglingLayers circuit is used with trainable parameters which applies one-parameter single-qubit rotations on each qubit followed by a ring of Controlled-Not (CNOT) gates, where each qubit is entangled with its neighbor, and the last qubit is connected back to the first, forming a closed chain. This circuit architecture was chosen for the quantum layer to strike a practical balance between circuit expressibility and parameter efficiency. Methods for evaluating the effectiveness of such circuits are an active area of research (Sim et al., 2019).

The operation of the quantum circuit computation is discussed in Algorithm 1. After applying the basic entanglement layer, we measure the qubits to output classical values. The resulting 8-dimensional output vector of Pauli-Z expectation values is then mapped to the tag space by a final linear layer, which provides the input logits for the CRF layer for the final tag prediction.

5 Dataset and Preprocessing

5.1 Corpus Description

The data set used in our experiments is a social media corpus of code-mixed Hindi-English text. It was originally collected and annotated by Jamatia et al. (2015). The corpus consists of messages from the IIT Bombay Facebook Confession page, which contains informal posts and chat-like comments. This type of data presents unique challenges for NLP tasks due to non-standard grammar, transliterated spellings, and informal language (Laskar et al., 2022).

The data set used is a component of a larger corpus that also includes WhatsApp and Twitter data and covers other pairs of Indian languages such as Bengali-English and Telugu-English (Pandey et al., 2023). However, this study focuses exclusively on the Hindi-English Facebook portion. The language distribution at the token level, as reported by the original authors, is shown in Table 1. It highlights that the text is predominantly English, with a significant presence of Hindi and language-independent universal tokens, such as punctuation. The data set is annotated with a coarse-grained POS tagset, which combines universal tags with categories specific to the text of social networks. This tagset, which comprises 11 unique tags, is described in Table 2. Our data set loading process yielded a total of **1069** sentences.

5.2 Preprocessing and Data Representation

For feature representation, each word in the corpus was mapped to a **100-dimensional vector** using precomputed embeddings for this data set. Any word not present in the embedding vocabulary was represented by a zero vector. To handle variable sentence lengths for batch processing, all sequences were standardized to a uniform length of **62 tokens** by padding shorter sequences and truncating longer ones. This length was determined on the basis of the 95th percentile of sentence lengths in the corpus. Following these preprocessing steps, the data set was partitioned into training sets (60%), validation (20%) and testing (20%), resulting in **641 samples for training, 214 for validation and 214 for testing**.

Token Language	Distribution (%)
English	75.61
Hindi	4.17
Universal	16.53
Named Entity	2.19
Acronym	1.46
Mixed	0.02
Undefined	0.01

Table 1: Token-level language distribution for the Facebook portion of the corpus, as reported by Jamatia et al. (2015).

Tag	Description
G_N	Noun
G_V	Verb
G_PRP	Pronoun
G_J	Adjective
G_R	Adverb
PSP	Pre- or Post-position
G_PRT	Particle
CC	Conjunction
G_SYM	Quantifier / Symbol
DT	Determiner
G_X	Residual / Other

Table 2: Coarse-grained POS tagset used in the dataset.

6 Experiment and Results

6.1 Experimental Setup

To evaluate our proposed model, we conducted a series of experiments to benchmark its performance against a purely classical counterpart. The models were implemented using PyTorch, with the quantum components built in Pennylane and executed on the **default qubit** simulator. The experiments compare a **classical GRU** based model against the proposed hybrid model. To ensure a fair comparison, a consistent set of hyperparameters was used to train both models, as detailed in Table 3.

Both models utilize a final Conditional Random Field (CRF) layer and were trained by minimizing its negative log-likelihood. Performance was evaluated using token-level accuracy on the held-out test set. Training was performed for a maximum of 300 epochs, with early stopping triggered if validation loss did not improve for 5 consecutive epochs.

Parameter	Value
Embedding Dimension	100
GRU Hidden Dimension	16
GRU Layers	2
Dropout Rate	0.3
Optimizer	Adam
Learning Rate	0.001
Batch Size	32
Number of Qubits	8

Table 3: Hyperparameters used for training.

6.2 Results

The final performance of both models was determined by evaluating the best-performing check-point, selected based on the peak validation accuracy observed during training. A summary of these results, alongside the final test accuracy and total parameter counts, is presented in Table 4. The proposed Hybrid QGRU model achieved a final test accuracy of **78.13%**, a result comparable to the **80.29%** accuracy achieved by the fully Classical GRU baseline. The central finding, however, lies in the model's efficiency. The hybrid model required only **16,682** trainable parameters to achieve this result, a modest but clear reduction of approximately **5.7%** compared to the **17,690** parameters of the classical model.

Model	Params	Val. Acc. (%)	Test Acc. (%)
Classical GRU	17,690	81.80	80.29
Hybrid QGRU	16,682	77.77	78.13

Table 4: Performance comparison of the baseline and hybrid models.

To provide a more granular analysis, Table 5 details a per-tag comparison of the F1-scores for both models on the test set. This breakdown reveals a nuanced performance landscape. For high-support, core grammatical categories such as **G_N** (Noun), **G_V** (Verb), and **DT** (Determiner), the hybrid model's performance is nearly identical to the classical baseline. Notably, it performs slightly better on **G_PRP** (Pronoun) tags. However, the hybrid model struggles with certain low-frequency tags, showing a significant performance drop for **CC** (Conjunction) and struggling significantly with **G_SYM** (Symbol) tags, failing to correctly classify any instance, likely due to their very low support in the test set. This suggests that while the quantum layer is effective at learning representations for common classes, it may be less robust on sparse data categories compared to its classical counterpart in this configuration.

Tag	Support	Cls. F1	Hyb. F1	Δ (Hyb-Cls)
CC	118	0.52	0.19	-0.33
DT	238	0.89	0.91	+0.02
G_J	199	0.62	0.54	-0.08
G_N	755	0.82	0.81	-0.01
G_PRP	336	0.83	0.86	+0.03
G_PRT	142	0.51	0.42	-0.09
G_R	188	0.63	0.56	-0.07
G_SYM	31	0.43	0.00	-0.43
G_V	697	0.85	0.82	-0.03
G_X	478	0.96	0.95	-0.01
PSP	339	0.81	0.78	-0.03

Table 5: Per-tag F1-score comparison on the test set. Δ indicates the change in F1-score for the hybrid model.

7 Conclusion

In this work, we addressed the challenge of high parameter counts in deep learning models for NLP by proposing and evaluating a hybrid quantum-classical Gated Recurrent Unit (QGRU). We applied this model to the task of POS tagging on code-mixed social media text, a domain characterized by noisy and non-standard language. Our findings indicate that the hybrid model achieves a test accuracy of **78.13%**, which is comparable to the **80.29%** accuracy of its classical counterpart, while requiring approximately **5.7%** fewer trainable parameters. This outcome serves as a successful proof-of-concept, demonstrating that the integration of variational quantum circuits into recurrent architectures is a viable strategy for reducing model complexity. Our work contributes to the growing field of quantum NLP by illustrating a practical approach to develop more compact and parameter-efficient models.

References

Nihar Jyoti Basisth, Tushar Sachan, Neha Kumari, Shyambabu Pandey, and Partha Pakray. 2023. An automatic pos tagger system for code mixed indian social media text. In *International conference on computational intelligence in communications and business analytics*, pages 273–286. Springer.

Kyunghyun Cho, Bart Van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. 2014. Learning phrase representations using RNN encoder-decoder for statistical machine translation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1724–1734, Doha, Qatar.

Bob Coecke, Giovanni de Felice, Konstantinos Meichanetzidis, and Alexis Toumi. 2020. Foundations for near-term quantum natural language processing. *arXiv preprint arXiv:2012.03755.*

Riccardo Di Sipio, Jia-Hong Huang, Samuel Yen-Chi Chen, Stefano Mangini, and Marcel Worring. 2022. The dawn of quantum natural language processing. In *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 8612–8616. IEEE.

Laszlo Gyongyosi and Sandor Imre. 2019. A survey on quantum computing technology. *Computer Science Review*, 31:51–71.

Anupam Jamatia, Björn Gambäck, and Amitava Das. 2015. Part-of-speech tagging for code-mixed

English-Hindi Twitter and Facebook chat messages. In *Proceedings of Recent Advances in Natural Language Processing*, pages 239–248, Hissar, Bulgaria.

Christian Janiesch, Patrick Zschech, and Kai Heinrich. 2021. Machine learning and deep learning. *Electronic Markets*, 31(3):685–695.

Dan Jurafsky. 2000. *Speech & Language Processing*. Pearson Education India.

Guillaume Lample, Miguel Ballesteros, Sandeep Subramanian, Kazuya Kawakami, and Chris Dyer. 2016. Neural architectures for named entity recognition. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 260–270, San Diego, California.

Sahinur Rahman Laskar, Rahul Singh, Shyambabu Pandey, Riyanka Manna, Partha Pakray, and Sivaji Bandyopadhyay. 2022. Cnlp-nits-pp at mixmt 2022: Hinglish-english code-mixed machine translation. In *Proceedings of the Seventh Conference on Machine Translation (WMT)*, pages 1158–1161.

Shyambabu Pandey, Nihar Jyoti Basisth, Tushar Sachan, Neha Kumari, and Partha Pakray. 2023. Quantum machine learning for natural language processing application. *Physica A: Statistical Mechanics and its Applications*, 627:129123.

Shyambabu Pandey, Pankaj Dadure, Morrel V. L. Nunsanga, and Partha Pakray. 2022. Parts of speech tagging towards classical to quantum computing. In *2022 IEEE Silchar Subsection Conference (SILCON)*, pages 1–6. IEEE.

Shyambabu Pandey and Partha Pakray. 2023. Bi-quantum long short-term memory for part-of-speech tagging. In *Proceedings of the 19th International Conference on Natural Language Processing (ICON)*, pages 272–278, Goa, India.

Shyambabu Pandey, Partha Pakray, and Riyanka Manna. 2024. Quantum classifier for natural language processing applications. *Computación y Sistemas*, 28(2):695–700.

Stuart J. Russell and Peter Norvig. 2010. *Artificial Intelligence: A Modern Approach*. Pearson Education, Inc.

Maria Schuld and Francesco Petruccione. 2021. *Machine Learning with Quantum Computers*. Springer.

Abdel-Nasser Sharkawy. 2020. Principle of neural network and its main types. *Journal of Advances in Applied & Computational Mathematics*, 7:8–19.

Jinjing Shi, Zhenhuan Li, Wei Lai, Fangfang Li, Ronghua Shi, Yanyan Feng, and Shichao Zhang. 2023. Two end-to-end quantum-inspired deep neural networks for text classification. *IEEE Transactions on Knowledge and Data Engineering*, 35(4):4335–4345.

Sukin Sim, Peter D Johnson, and Alán Aspuru-Guzik. 2019. Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms. *Advanced Quantum Technologies*, 2(12):1900070.

Ryan Sweke, Frederik Wilde, Johannes Meyer, Maria Schuld, Paul K. Faehrmann, Barthélemy Meynard-Piganeau, and Jens Eisert. 2020. Stochastic gradient descent for hybrid quantum-classical optimization. *Quantum*, 4:314.

A Review of Quantum Computing Approaches to Semantic Search and Text Classification in Natural Language Processing

Sauvik Bal[1,2] **Shillpi Mishra**[1] **Lopa Mandal**[3]
[1]Techno India University, West Bengal, India
[2]Maulana Abul Kalam Azad University of Technology, West Bengal, India
[3]Alliance University, Bangalore, India
Correspondence: sauvikbal@gmail.com

Abstract

While having enhanced NLP, deep learning and pre-trained language models requires a lot of processing power. The work showcases the potential of quantum computing by mapping linguistic data into vast, high-dimensional Hilbert spaces through entanglement and superposition. It focuses on mathematical concepts that set quantum approaches apart from classical ones, among them being the fidelity-based similarity and quantum probability. Various quantum machine learning models are considered in this article, including Quantum Neural Networks and Quantum Support Vector Machines, each discussing the computational advantages in pattern recognition. In addition, it considers retrieval techniques like Grover's algorithm, showing how quantum similarity functions give better semantic search. Indeed, the comparison does show that quantum techniques might yield advantages regarding expressiveness and scalability, despite obstacles such as hardware noise and data encoding. Notwithstanding that quantum technology is still in its infancy, future improvements might advance language understanding.

Keywords

Quantum Computing, Text Classification, Semantic Search, Information Retrieval, Natural Language Processing (NLP), Quantum Neural Networks

1 Introduction

The explosive evolution of natural language processing (NLP) has mostly been triggered by traditional machine learning and deep learning models, which have reported impressive performance in applications like text classification, sentiment analysis, and semantic search (Devlin et al., 2019). Notwithstanding these breakthroughs, the ever-growing dimensionality of text data and the computational expense of large models have made it imperative to look for other approaches that can offer efficiency without sacrificing semantic richness. Quantum computing has, in recent years, been explored as a possible paradigm to overcome such limitations because it can perform computations in exponentially big Hilbert spaces and leverage principles like superposition and entanglement (Schuld and Petruccione, 2019).

Quantum models, such as QSVMs and QNNs, embed texts into high-dimensional quantum feature spaces, hence being more effective for text categorization than classical techniques. Quantum-inspired information retrieval techniques rely on Hilbert space formalism and fidelity measurements while offering advantages over classical methods by virtue of Grover's search algorithm. The analytical framework includes quantum kernels and probability distributions that extend conventional comparison metrics such as cosine similarity. However, despite theoretical advantages, noisy and resource-limited NISQ devices make practical implementation very challenging. Therefore, hybrid quantum-classical approaches are considered a viable approach. The current study will review the mathematical underpinnings of quantum NLP research, complexity assessments, and comparative insights between quantum and classical approaches to highlight the potential benefits and current challenges in quantum NLP research.

It starts with theoretical notions (described in Figure 1), such as Hilbert spaces and Grover's algorithm, and the review structure progresses from a purely mathematical underpinning to a real-world application. It falls into two main areas: one regarding quantum semantic search by fidelity measures and quantum walks, and another on quantum text classification by means of QSVM and QNN/VQC models. In order to identify complementarities and trade-offs, these branches merge under a comparative study that analyzes mathematical methods and empirical behaviors. Furthermore, this frame-

Proceedings of the QuantumNLP: Integrating Quantum Computing with Natural Language Processing, pages 33–43
November 24, 2025

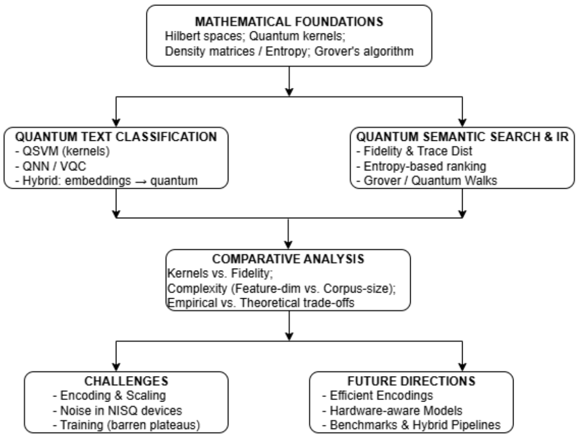


Figure 1: Overview of the analytical review

work indicates problems like noise and encoding and suggests future avenues of research, such as uniform benchmarks and effective encodings.

2 Scope and Review Methodology

With a focus on text classification and semantic search, this article reviews the integration of quantum computing into NLP. It outlines the development from early quantum-inspired frameworks to advanced quantum kernel methods and hybrid models, surveying literature from 2010 to 2025. Among others, IEEE Xplore, ACM Digital Library, and arXiv can be consulted using keywords and phrases such as "quantum NLP" and "quantum semantic search." Only those studies that introduce quantum models for natural language processing (NLP) problems, develop theoretical insights into either text classification or semantic search, and offer analytical contrasts between classical and quantum approaches are reviewed. Works which are purely classical or with no direct relevance to NLP will not be considered. To ensure a systematic progress review in quantum NLP, contributions are grouped into three categories: quantum models for text categorization, quantum approaches for semantic search, and supporting mathematical analyses.

3 Mathematical Foundations

The use of quantum computing in natural language processing (NLP) is mathematically intense. The following section presents the mathematical basics that make up the analytical framework of quantum methods in text classification and semantic search. These are Hilbert spaces, quantum probability, measures of fidelity, kernel methods, and computational complexity.

3.1 Hilbert Spaces and Quantum Text Representation

In quantum mechanics, physical systems are encoded in a complex Hilbert space $\mathcal{H}$, and the system's state is specified as a normalized vector in that space. A pure state is normally written as:

$$|\psi\rangle = \sum_i c_i |i\rangle, \qquad \sum_i |c_i|^2 = 1, \tag{1}$$

where c_i are complex probability amplitudes. $\{|i\rangle\}$ denotes an orthonormal basis of the Hilbert space $\mathcal{H}$.

In NLP, the $|i\rangle$ are linked with a token, word embedding, or latent semantic component. Superposition is built into quantum representations to enable more expressive encoding of semantic relations over classical embeddings.

3.2 Quantum Probability and Density Matrices

Quantum probability is derived from the Born rule. For a state $|\psi\rangle$, the probability of observing basis state $|i\rangle$ is:

$$P(i) = |\langle i|\psi\rangle|^2. \tag{2}$$

For mixed states, a density matrix ρ is defined as:

$$\rho = \sum_k p_k |\psi_k\rangle\langle\psi_k|, \quad \text{with} \quad \mathrm{Tr}(\rho) = 1. \tag{3}$$

This allows ambiguous words to be modeled as probabilistic mixtures of multiple semantic states (Piwowarski et al., 2010). The information content of a state is quantified using the von Neumann entropy:

$$S(\rho) = -\mathrm{Tr}(\rho \log \rho), \tag{4}$$

which generalizes Shannon entropy into quantum systems.

3.3 Similarity and Distance Metrics

Semantic similarity in quantum models is expressed via fidelity:

$$F(\rho, \sigma) = \left(\mathrm{Tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right)^2, \tag{5}$$

where ρ and σ represent query and document states. Fidelity generalizes cosine similarity by embedding comparisons in Hilbert space (van der Meer et al., 2021). Another important measure is the trace distance:

$$D(\rho, \sigma) = \tfrac{1}{2}\mathrm{Tr}\,|\rho - \sigma|, \tag{6}$$

which captures dissimilarity between semantic states.

3.4 Illustrative Comparison: Cosine Similarity vs. Fidelity

To better understand how quantum similarity measures differ from classical ones, consider two simple normalized 2-dimensional vectors representing a query q and document d:

$$q = (1, 0), \qquad d = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right). \tag{7}$$

Cosine Similarity:

$$\begin{aligned}\cos(q, d) &= \frac{q \cdot d}{\|q\|\|d\|} \\ &= \frac{1 \cdot \frac{\sqrt{2}}{2} + 0 \cdot \frac{\sqrt{2}}{2}}{1 \cdot 1} = \frac{\sqrt{2}}{2} \approx 0.707.\end{aligned} \tag{8}$$

Fidelity Measure: When the same vectors are treated as pure quantum states $|q\rangle$ and $|d\rangle$, fidelity is defined as:

$$F(q, d) = |\langle q|d\rangle|^2 = \left(\frac{\sqrt{2}}{2}\right)^2 = 0.5. \tag{9}$$

Interpretation:

- Cosine similarity measures *geometric angle* between classical vectors.
- Fidelity measures *quantum probability overlap* between states.

Although fidelity reduces to the *square* of cosine similarity for pure normalized states, the probabilistic meaning of fidelity is more aligned with quantum measurements. In more complex mixed-state scenarios (e.g., density matrices), fidelity captures richer semantic uncertainty beyond what cosine similarity provides.

3.5 Quantum Kernels and Feature Maps

In classification tasks, quantum kernels extend the classical kernel trick. A quantum feature map $\phi(x)$ encodes data into quantum states, and the kernel function is defined as:

$$k(x, y) = |\langle \phi(x)|\phi(y)\rangle|^2. \tag{10}$$

This helps the model to handle the large feature spaces where fewer resources can be used (Havlíček et al., 2019).

3.6 Complexity Considerations

In theory, quantum algorithms provide great advantages in terms of computational efficiency. Grover's technique reduces unstructured database searches from $O(N)$ to $O($
$sqrtN)$ for large document collections (Grover, 1996). The HHL algorithm can solve linear equations exponentially faster under some conditions. These advances are certainly relevant for NLP tasks involving high-dimensional embeddings and large text corpora; hence, there is the prospect of quantum methods outperforming classical systems for semantic search and classification (Harrow et al., 2009).

3.7 Implications for NLP

This paper emphasizes analytical comparisons to classical models using Hilbert space formalism, quantum probability, and kernel-based feature mapping to realize NLP tasks. In doing so, it is shown that quantum models may decrease processing costs while retaining the semantic information undisturbed. These mathematical frameworks, in essence, serve as the foundation for approaches using quantum machine learning on linguistic data in practical quantum models for text classification, such as QSVMs, QNNs, and hybrid quantum-classical architectures.

4 Quantum Computing for Text Classification

One of the main applications of NLP is text categorization. It is useful for various purposes, like spam filtering and sentiment analysis etc.. While they deliver accurate results, conventional models, including transformer-based models such as BERT and support vector machines, are computationally expensive, particularly in the scenario of high-dimensional feature spaces. A suitable replacement is quantum computing, which enhances classification efficiency through the application of large Hilbert spaces and ideas of superposition and entanglement. The primary focus of this section's coverage of quantum models for text classification is quantum support vector machines, quantum neural networks, and hybrid architectures.

As shown in Table 1, recent developments in quantum computing for text classification are summarized chronologically, covering the period from 2019 to 2025. The table highlights the main approaches, encoding strategies, model types, as well

Table 1: Chronological comparison of quantum (and quantum-like) approaches for text classification.

Year	Approach (citation)	Encoding	Model Type	Advantages	Limitations
2019	QSVM with quantum kernels (Havlíček et al., 2019)	Angle / amplitude	QSVM kernel classifier	Exponential feature mapping; effective separation in high-dim spaces	Sensitive to noise; evaluated only on small datasets
2019	Quantum Convolutional Neural Networks (QCNN) (Cong et al., 2019)	Structured encoding	Convolutional-like QNN	Parameter efficient; locality aware; some robustness to noise	Task-specific design; deeper circuits required
2021	Quantum Neural Networks (QNN) (Abbas et al., 2021)	Angle / amplitude	Variational quantum circuits (VQC)	High expressivity; end-to-end trainable	Barren plateaus (gradient vanishing); noisy hardware limits
2022	Quantum SVM for Text Classification (Li et al., 2022)	Amplitude / angle from embeddings	QSVM + hybrid embedding	Uses word2vec/BERT embeddings; competitive accuracy	Inherits embedding costs; limited to small corpora
2024	Quantum Support Vector Classifier on NISQ hardware (Suzuki, 2024)	Angle / amplitude	QSVC on IBM/IonQ devices	First hardware validation; real device results	Strongly affected by noise; dataset scaling issues
2024	Quantum Self-Attention Neural Networks (QSANN) (Li, 2024)	Classical embeddings $\rightarrow$ quantum attention	Hybrid quantum-classical attention	Captures contextual dependencies; integrates attention with QNN	Only small simulator tests; hardware results pending
2024	Hybrid transfer learning (BERT + QSVM/QNN) (Anonymous, 2024)	Pretrained embeddings $\rightarrow$ quantum classifier	Hybrid pipeline	Combines classical embeddings with quantum classifiers; practical for NISQ	Dependent on pretrained models; added quantum overhead
2025	Quantum-like wave model for semantic classification (Gruzdeva et al., 2025)	Semantic units $\rightarrow$ wave embeddings	Quantum-like wave model	Captures interference in semantics; accuracy gains over baselines	Semi-heuristic; not hardware-based; small datasets only
2025	Hybrid QTL with kernel self-attention (Chen and Lou, 2025)	Classical embeddings + quantum kernel	Hybrid transfer-learning	Improves feature separability; tested on real datasets	Complex architecture; hardware scaling challenges
2025	Single-Qudit QNN (SQ-QNN) (Souza and Portugal, 2025)	Angle encoding into qudits	Qudit-based QNN	Reduces qubit needs by using qudits; supports multi-class tasks	Still theoretical/simulator-level; qudit hardware needed

as their analytical advantages and limitations, providing a structured comparison of progress in this domain.

4.1 Quantum Support Vector Machines (QSVM)

An optimal separating hyperplane within a feature space is defined by the Support Vector Machine, which is basically a supervised learning algorithm. It is used to classify the data. The Quantum Support Vector Machine maps data into high-dimensional Hilbert spaces. Unitary operations are used to map data points into quantum states in QSVM. A quantum kernel defines the similarity of these states. Basically, the architecture allows QSVMs to utilize polynomial resources (Havlíček et al., 2019; Schuld and Petruccione, 2019) on quantum hardware and operate within an exponentially dimensional feature space. With weighted kernel evaluations on training sets, a decision function in a QSVM is analogous to traditional SVMs. QSVMs are an interesting method because quantum kernels are capable of separating classes that polynomial-time classical kernels are not.

4.2 Quantum Neural Networks (QNNs)

Quantum Neural Networks (QNNs) are constructed using Variational Quantum Circuits (VQCs), which consist of parameterized unitary gates optimized with a classical optimizer. For an input x, the quantum state is expressed as:

$$|\psi(\theta, x)\rangle = U(\theta, x)|0\rangle, \tag{11}$$

where θ represents trainable parameters. The probability of assigning class y is given by:

$$P(y|x) = |\langle y|\psi(\theta, x)\rangle|^2. \tag{12}$$

QNNs have been shown to achieve expressive power that surpasses shallow classical neural networks, particularly in handling non-linear and high-dimensional relationships (Abbas et al., 2021). However, they face challenges such as barren plateaus, where gradients vanish exponentially with circuit depth. Recent studies have also demonstrated the practical use of QNN-based classifiers specifically for NLP tasks (Pandey et al., 2024).

4.3 Data Encoding Strategies

Encoding is an essential link between textual data and quantum computation, which requires converting units of language into quantum states with semantic integrity preserved. Effective encoding techniques have impact on representational expressivity and hardware viability in NISQ devices and are crucial, not only for text categorization but also for quantum semantic search and information retrieval.

Different encoding strategies have been discussed here:

- **Basis encoding:** Each token or feature is placed directly into a basic quantum state.

Here,it gives a clear but sparse form in the qubit form.

- **Amplitude encoding:** dense embeddings such as word2vec or BERT are normalized and encoded into the amplitudes of a quantum state:
$$|x\rangle = \frac{1}{\|x\|} \sum_i x_i |i\rangle, \tag{13}$$
providing an exponentially compact form of feature representation.

- **Angle encoding:** numerical features are converted into rotation angles of single-qubit gates, offering low circuit depth at the cost of reduced representational capacity.

Amplitude encoding is particularly appealing for NLP tasks, as it enables dense semantic embeddings to be represented efficiently in Hilbert space while still benefiting from quantum parallelism (Schuld and Petruccione, 2019). As quantum NLP advances, developing encoding techniques that balance compactness, expressiveness, and noise resilience will remain a key challenge.

4.4 Hybrid Quantum-Classical Models

The NISQ era of existing quantum devices faces challenges in quantum model development in NLP due to the low number of qubits and pervasive noise (Suzuki, 2024). Hybrid quantum-classical models leverage classical embeddings generated using word2vec and BERT, which can then be fed into a quantum classifier such as the QNN or QSVM (Devlin et al., 2019; Abbas et al., 2021; Li et al., 2022). The combination leverages quantum capabilities for improved feature separation and the quality of the classification of particularly complex text distributions but draws on classical strengths for embeddings (Havlíček et al., 2019). Somewhat limited in accuracy due to the hardware, QSVMs and quantum kernel classifiers have produced successful results on IBM and IonQ devices (Suzuki, 2024). A completely quantum solution, hybrid models represent an exciting way to incorporate quantum computing into NLP applications (Anonymous, 2024).

4.5 Complexity Analysis

Quantum Support Vector Machines employ polynomial circuit resources to evaluate similarities in exponentially vast spaces. This is done through the use of the quantum kernel trick. Analytical comparisons shows that:

- For the d-dimensional embedding, the classical SVM kernel evaluation mainly requires $O(d)$ operations.

- Quantum kernel evaluation scales as $O(\text{poly}(n))$. Here n is used as the number of qubits, which represents the feature space.

This advantage is significant when processing high-dimensional language embeddings is done. This is frequently used in NLP applications. However, the reliability of kernel estimate on NISQ hardware. It restricted connection, and qubit noise limit realistic speedup.

Complexity of QNN-based Models: Variational Quantum Circuits are a novel computational regime in QNNs, with complexities dependent on the number of qubits n, the circuit depth L and trainable parameters $|\theta|$. At each iteration of the gradient-based training, it is necessary to run the quantum circuit multiple times, incurring a certain cost.

$$O\left(L \cdot \text{poly}(n) \cdot |\theta|\right). \tag{14}$$

For large n or L, QNNs suffer from problems like *barren plateaus* due to vanishing gradients, which make optimization costly. In contrast with QSVMs, that rely on quantum kernel evaluation for computation, QNNs are plagued by scalability issues due to optimization overhead and hardware noise. Because of optimization complexity and coherence restrictions in NISQ devices, increased expressiveness of QNNs can hardly be exploited in practice.

4.6 Applications in NLP

Quantum models, like QSVMs and QNNs, have competed with conventional models in various NLP tasks, including sentiment analysis, spam filtering, and fake news detection (Pandey et al., 2024). Specifically, QSVMs are good at categorizing reviews, while QNNs have high performance in identifying trustworthy news sources and filtering spam using hybrid architectures. Moreover, applying quantum feature spaces to enable data-efficient solutions holds promising advances for multilingual and low-resource languages. This work points toward a path for future research in scalable quantum hardware development, as it emphasizes not only what has been achieved with supervised quantum models in text classification but also the potential for quantum-inspired methods in information retrieval and semantic search (discussed in Section 5).

5 Quantum Computing for Semantic Search and Information Retrieval

Quantum semantic search, by exploiting Hilbert space representations and quantum similarity measures, improves the ranking of documents, thereby outperforming conventional models based on lexical matching, such as TF-IDF and BM25 (Piwowarski et al., 2010; van der Meer et al., 2021). Quantum-inspired information retrieval models leveraged amplitude-encoded quantum states to model documents and queries with the aim of incorporating semantic aspects and uncertainties (Schuld and Petruccione, 2019). While quantum distance measures, such as trace distance, quantify semantic dissimilarity (van der Meer et al., 2021), quantum similarity measures, such as fidelity, augment classical cosine similarity by reflecting probabilistic overlaps (Piwowarski et al., 2010). Hybrid quantum-classical methods combine classical embeddings with quantum techniques to make the most of existing NISQ hardware (Yamada et al., 2024; Devlin et al., 2019), while some quantum algorithms, such as Grover's search, obtain significant savings in search time (Grover, 1996). In this changing approach to IR, with quantum technologies still evolving, the representation of deeper semantic relevance and uncertainty points to a more expressive future for search algorithms (Zhang et al., 2023; Gupta et al., 2025).

5.1 Hilbert Space Representations of Documents

The documents and queries are mainly represented as vectors in a Hilbert space $\mathcal{H}$ at the time of the quantum-inspired retrieval process. This is denoted as:

$$|d\rangle = \sum_i \alpha_i |t_i\rangle, \qquad |q\rangle = \sum_i \beta_i |t_i\rangle, \tag{15}$$

Here $|t_i\rangle$ represents the basis vectors, and α_i, β_i represents the normalized weights.

By using the fidelity, the similarity between a query q and a document d has measured:

$$\mathcal{F}(q, d) = \left|\langle d|q\rangle\right|^2. \tag{16}$$

Compared to traditional cosine similarity, this gives the squared inner product of the two states, which provides a more descriptive similarity metric (Piwowarski et al., 2010; van der Meer et al., 2021).

5.2 Quantum Probability and Entropy Measures

Quantum IR can also use density matrices to encode uncertainty in semantic states. For a document mixture, the density operator is given as:

$$\rho_d = \sum_i p_i |d_i\rangle\langle d_i|, \tag{17}$$

where p_i are probability weights. The similarity between documents can then be computed using trace distance or von Neumann entropy:

$$S(\rho) = -\mathrm{Tr}(\rho \log \rho). \tag{18}$$

Entropy-based ranking allows capturing semantic diversity and ambiguity, beyond what is possible in classical IR frameworks (Zhang et al., 2023).

5.3 Grover's Algorithm for Document Retrieval

Grover's quantum.search algorithm achieves a quadratic speedup for unstructured search problems (Grover, 1996). For a collection of N documents, Grover's algorithm will locate a matching document in $O(\sqrt{N})$ time, as opposed to $O(N)$ classically. Mathematically, successive applications of the Grover operator G increase the likelihood of the target state $|d^*\rangle$:

$$G = (2|s\rangle\langle s| - I) \cdot (I - 2|d^*\rangle\langle d^*|), \tag{19}$$

where $|s\rangle$ is the uniform superposition of all states. This gives theoretical speedups for large-scale IR.

5.4 Hybrid Quantum-Classical IR Models

Recent efforts integrate classical embeddings (e.g., BERT, word2vec) with quantum fidelity-based retrieval. Queries and documents are first embedded in dense vector spaces, then encoded into quantum states for matching. Such hybrid approaches provide practical pathways for deploying quantum IR on NISQ-era hardware (Yamada et al., 2024).

5.5 Chronological Comparison of Approaches

Table 2 summarizes the basic developments in quantum IR approaches from 2010 to 2025. It highlights the encodings, models, advantages, and limitations.

Table 2: Chronological comparison of quantum computing approaches for semantic search and information retrieval.

Year	Approach (citation)	Encoding	Model Type	Advantages	Limitations
2010	Quantum-inspired IR framework (Piwowarski et al., 2010)	Term basis states	Hilbert space retrieval	Introduced fidelity-based query-document similarity; linked IR to quantum probability	Conceptual framework; no hardware implementation
2019	Quantum probability ranking model (Zuccon and Azzopardi, 2019)	Amplitude encoding of terms	Quantum-inspired ranking model	Probabilistic interpretation of ranking; novel use of quantum probability	Early-stage model; tested on small corpora
2021	Quantum algorithms for IR (van der Meer et al., 2021)	Amplitude encoding	Hybrid algorithms for retrieval	Theoretical speedups using Grover's search and quantum walks	Lacks large-scale hardware benchmarks
2023	Entropy-based quantum IR (Zhang et al., 2023)	Density matrices	Entropy ranking framework	Incorporates semantic diversity and ambiguity; entropy-based document ranking	Simulator-based; hardware scaling not addressed
2024	Hybrid embedding + quantum fidelity search (Yamada et al., 2024)	BERT embeddings → quantum states	Hybrid quantum-classical IR	Integrates deep embeddings with quantum fidelity search; suitable for NISQ devices	Dependent on pretrained embeddings; limited qubits
2025	Quantum walk-based semantic retrieval (Gupta et al., 2025)	Amplitude encoding of graph embeddings	Quantum walk retrieval	Explores semantic search using quantum walks over document graphs; potential retrieval efficiency gains	Experimental stage; scalability to large corpora unproven

As shown in Table 2, approaches span from foundational quantum-inspired frameworks in 2010 to recent hybrid and quantum walk-based retrieval models in 2025, demonstrating the evolution from conceptual theory to practical hybrid implementations.

6 Comparative Analytical Insights

In this section, two core NLP tasks are described text categorization in Table 1 and semantic search/information retrieval in Table 2. These are compared using quantum techniques. It mainly highlights each domain's unique issues, their mathematical formulations, and development paths while also pointing out their trade-offs and complementarities.

6.1 Mathematical Underpinnings

Quantum text classification methods are predominantly kernel-based or variational, relying on mappings into exponentially large Hilbert spaces and parameterized quantum circuits. Quantum Kernel's analytical construction is as follows:

$$k(x, y) = |\langle \phi(x) | \phi(y) \rangle|^2, \tag{20}$$

QSVMs and hybrid kernel models are supported by the idea like, (Havlíček et al., 2019; Li et al., 2022). To check the similarity between a particular query and the document, quantum probability, density matrices, and fidelity is used for semantic search methods.

$$F(q, d) = |\langle q|d \rangle|^2, \quad S(\rho) = -\mathrm{Tr}(\rho \log \rho). \tag{21}$$

Quantum kernels for text classification define unique decision limits, making the distinction among categories of data easier. On the other hand, semantic search and information retrieval focus on the relevance of the materials to the queries. They aim to emphasize relevant pages and provide effective representation of semantic meaning through the use of probability-based similarity and entropy measurements.

6.2 Computational Complexity

Whereas QNNs suffer from optimization problems, QSVMs offer implicit embeddings in $O(2^n)$ dimensions using polynomial resources for classification (Abbas et al., 2021). Grover's search reduces the query complexity in information retrieval from $O(N)$ to $O(\sqrt{N})$; quantum walk-based approaches provide better document graph exploration (Gupta et al., 2025). While retrieval puts an emphasis on query scaling and ranking efficiency, classification emphasizes decision boundary complexity.

6.3 Evolution of Approaches (2010–2025)

From 2010 to 2025, research changed from conceptual formulations to hybrid implementations. For example, some classification advances that show progress towards NISQ practicality are the quantum-inspired classifiers of 2019, the hybrid pipelines with BERT embeddings and QSVMs of 2024, and the single-qudit QNNs of 2025. Some recent examples of retrieval advances include the Hilbert space-based IR models of 2010, entropy-driven ranking in 2023, hybrid embedding-fidelity models in 2024, and quantum walk retrieval in document graphs from 2025. This trajectory shows a convergence towards hybrid paradigms, combining quantum-enhanced classifiers and retrieval systems with classical embeddings.

6.4 Analytical Trade-offs

The comparison analysis shows different types of significant trade-offs:

- **Expressivity vs. Stability:** Entropy-based IR techniques provide stability, but they are less expressive in terms of model power. QNNs have large capacity but tey are difficult to optimize.
- **Scalability:** Classification complexity scales with embedding dimensionality, while retrieval scales with corpus size. Both benefit from quantum asymptotic advantages in distinct regimes.
- **Hardware Realization:** Classification methods (e.g., QSVMs) have been experimentally tested on NISQ devices (Suzuki, 2024), whereas IR models remain mostly simulator-bound, with limited demonstrations on hardware.

6.5 Outlook

The two main approaches in quantum NLP, quantum text categorization and quantum semantic search, are complementary rather than competing. Quantum text classification excels in supervised tasks that involve clear-cut decision boundaries, whereas quantum semantic search adopts probabilityand entropy-based measures to capture the meaning of texts and rank documents. In both areas, researchers are moving toward hybrid quantum-classical architectures, and thus classification and retrieval will eventually be part of NLP systems that have traditional components for preprocessing and quantum circuits for semantic reasoning. As quantum devices improve, various applications and practical quantum advantages for NLP might become possible.

7 Experimental Landscape and Benchmarking Status

Theoretically, quantum NLP demonstrates great possibilities, but empirical verification is not possible due to the constraints of existing NISQ hardware. The majority of research uses quantum simulators and short datasets, focusing on practicality before completeness of performance. Benchmarking trends are presented in this section for semantic search and quantum text classification.

7.1 Datasets Used in Current Studies

In some instances, compact datasets have been employed to benchmark quantum-enhanced classifiers: for instance, the SMS Spam Dataset for binary spam filtering with QSVMs (Li et al., 2022), portions of Amazon or IMDb reviews for sentiment analysis based on QNN-based models (Pandey et al., 2024), and TREC-style toy retrieval sets for query relevance assessment (van der Meer et al., 2021). Due to qubit availability constraints, the IR experiments often employ simulated semantic vectors rather than complete corpus representations (Piwowarski et al., 2010). Hybrid BERT-embedded document matching has seen a bit more development, although its application remains limited to very small corpora (Yamada et al., 2024).

7.2 Evaluation Metrics

Performance evaluation typically combines established classical metrics with quantum-specific similarity measures:

- **Accuracy, Precision, Recall, F1-score** for classification (Suzuki, 2024).
- **Entropy-based ranking** to measure semantic diversity (Zhang et al., 2023).
- **Fidelity** as a probabilistic similarity score between query and documents (Piwowarski et al., 2010).

These mixed metrics reflect an ongoing effort to account for both prediction quality and quantum semantic overlap.

7.3 Simulators vs. Hardware Deployments

The lack of standard benchmarks due to differences in dataset size, encoding methodologies, simulator precision, error models, hardware platforms, and circuit depth limits hinders precise performance comparisons among research (van der Meer et al., 2021). Thus, the assertions on quantum advantage in NLP are prima facie tentative and bound by experimental design (Abbas et al., 2021). Medium-scale data sets may be manageable for future information retrieval as qubit counts and noise robustness will likely continue to improve (Suzuki, 2024). Effective data embeddings may allow multilingual and low-resource tasks to benefit, and standard evaluation metrics accounting for accuracy, fidelity, and complexity analysis will be important moving forward (Yamada et al., 2024). Hardware-aware model design and standard benchmarking are necessary before large-scale demonstrations of quantum NLP performance can be realized (Anonymous, 2024).

7.4 Current Limitations in Benchmarking

Current implementations mainly rely on quantum simulators, such as Qiskit and Cirq, because of issues with noise and coherence in real quantum hardware (van der Meer et al., 2021). While hardware-based evaluation is at an early stage of development, problems such as significant losses in the accuracy of QSVMs on IBM and IonQ systems due to qubit decoherence and gate noise persist (Suzuki, 2024). Additionally, variational circuits suffer from empty plateaus, demanding deeper topologies (Abbas et al., 2021). Therefore, most of the studies of QNNs are simulation-limited. Hybrid approaches, offering a good compromise between expressiveness and feasibility, become the most viable approach for current experimentation (Anonymous, 2024).

8 Challenges and Open Problems

Quantum techniques for text classification and semantic search face numerous obstacles regarding mathematical, hardware, algorithmic, and benchmarking factors. This section outlines these restrictions and identifies unresolved issues for further research.

8.1 Encoding Bottlenecks in NLP Data

Encoding high-dimensional textual data into quantum states remains one of the most significant bottlenecks. Given a document embedding $x \in \mathbb{R}^d$, amplitude encoding maps it into a normalized quantum state:

$$|x\rangle = \frac{1}{\|x\|} \sum_{i=1}^{d} x_i |i\rangle. \tag{22}$$

This requires $O(d)$ operations classically, but preparing an arbitrary d-dimensional state on a quantum computer may require $O(d)$ gates, offsetting quantum speedups. Angle encoding reduces cost by mapping each feature into a rotation, but sacrifices representational richness. Open problems include:

- Developing encoding schemes that balance expressivity with circuit depth.
- Exploring qudit-based encodings that reduce qubit requirements (Souza and Portugal, 2025).
- Designing noise-resilient encodings suitable for NISQ hardware.

8.2 Hardware Constraints and Noise Sensitivity

Most reported quantum NLP experiments have been conducted on simulators. Real NISQ devices introduce gate noise, decoherence, and readout errors. For example, QSVM implementations on IBM and IonQ hardware show a drastic drop in accuracy due to noise (Suzuki, 2024). Moreover, current devices limit circuit depth to < 100 gates for reliable execution, restricting model complexity.The open problems in NLP involve developing specific error mitigation techniques, identifying which NLP workloads are inherently noise-tolerant, such as low-rank embeddings, and exploring qudit-based systems which provide higher information density per physical unit.

8.3 Training Challenges in Quantum Neural Networks

Though the Quantum Neural Networks are a promising model, they still suffer from different serious optimization problems. The barren plateau phenomenon leads to gradients vanishing exponentially with the number of qubits or circuit depth:

$$\mathbb{E}[\nabla_\theta L] \sim O\left(\frac{1}{2^n}\right), \tag{23}$$

where n is the number of qubits. This severely limits scalability (Abbas et al., 2021). Hybrid training with classical optimizers introduces additional cost and convergence instability. Open problems include:

- Gradient-free optimization methods for variational circuits.
- Cost functions that mitigate barren plateaus.
- Scalable architectures such as QCNNs or SQ-QNNs (Cong et al., 2019; Souza and Portugal, 2025).

8.4 Scalability of Quantum IR Models

Semantic search requires efficient ranking over massive document collections. Grover's algorithm provides $O(\sqrt{N})$ query complexity, but practical retrieval requires top-k ranking and probabilistic scoring. Quantum walk retrieval models (Gupta et al., 2025) explore graph-based semantics, but remain untested at scale. Open problems include:

- Extending Grover-based search to ranked retrieval.
- Integrating density matrix entropy-based ranking (Zhang et al., 2023) with large document collections.
- Designing quantum IR systems that scale to billions of documents, analogous to web-scale search engines.

8.5 Hybrid Integration and Efficiency Boundaries

Most NISQ-era implementations are hybrid, combining classical embeddings (e.g., BERT, GloVe) with quantum classifiers or retrieval engines (Li et al., 2022; Anonymous, 2024). While effective, this raises fundamental questions:

- What portion of the pipeline truly benefits from quantum speedup?
- How can hybrid systems avoid classical bottlenecks dominating end-to-end runtime?
- What is the theoretical boundary between classical preprocessing and quantum advantage?

Establishing efficiency thresholds for hybrid quantum NLP architectures remains a critical open problem.

8.6 Evaluation and Benchmarking Gaps

There is currently no standardized framework to evaluate quantum NLP models. Classical benchmarks (e.g., GLUE, TREC) are ill-suited for quantum setups due to small dataset constraints. Open problems include:

- Designing quantum-specific NLP benchmarks with fidelity, entropy, and robustness metrics.
- Establishing evaluation protocols that combine accuracy with complexity analysis.
- Developing open-source datasets small enough for NISQ devices yet representative of real tasks.

8.7 Theoretical Uncertainty of Quantum Advantage

Finally, the biggest open problem is the lack of rigorous proof of quantum advantage in NLP. While complexity-theoretic results such as Grover's speedup are well-established, their direct applicability to semantic search and classification remains uncertain. For classification, empirical studies suggest quantum kernels offer improved separability, but no formal guarantee exists. For retrieval, entropy-based models are theoretically elegant but lack evidence of practical superiority. Future directions include:

- Proving formal conditions under which quantum models outperform classical ones.
- Linking quantum kernel theory with generalization bounds in NLP tasks.
- Exploring quantum information-theoretic limits of semantic search.

9 Conclusion and Future Directions

This work reviews quantum computing methods for natural language processing-related tasks, namely, semantic search and text classification (Havlíček et al., 2019; Li et al., 2022). Focusing on techniques such as quantum kernels, variational quantum neural networks (Abbas et al., 2021; Anonymous, 2024), and entropy-driven ranking, the study explores the trajectory from quantum-inspired models to hybrid quantum-classical systems. A comparison is drawn (in Section 6) in which retrieval performs well in both probabilistic and entropy-based models, which currently are both trending toward hybrid paradigms due to NISQ hardware limitations, whereas classification makes good use of quantum kernels. Some other promising future avenues of research involve effective encoding techniques, hardware-aware models, standardized quantum benchmarking, understanding quantum advantage, and integrated quantum NLP pipelines. Quantum computing indeed offers a great future for NLP applications, despite the challenges at present.

References

Amira Abbas, David Sutter, Christa Zoufal, Aurelien Lucchi, Alessio Figalli, and Stefan Woerner. 2021. The power of quantum neural networks. *Nature Computational Science*, 1(6):403–409.

Anonymous. 2024. Hybrid classical-quantum transfer learning for text classification. *arXiv preprint arXiv:2403.19758*.

Xiaoxiao Chen and Xiaoping Lou. 2025. Enhancing text classification through quantum transfer learning: A hybrid quantum-classical approach with complex kernel self-attention networks. *IEEE Access*.

Iris Cong, Soonwon Choi, and Mikhail D. Lukin. 2019. Quantum convolutional neural networks. In *Proceedings of the 36th International Conference on Machine Learning (ICML)*, pages 972–981.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, pages 4171–4186.

Lov K. Grover. 1996. A fast quantum mechanical algorithm for database search. In *Proceedings of the 28th Annual ACM Symposium on Theory of Computing (STOC)*, pages 212–219.

A. S. Gruzdeva, R. N. Iurev, I. A. Bessmertny, A. Y. Khrennikov, and A. P. Alodjants. 2025. A quantum-like approach to semantic text classification. *Entropy*, 27(7):767.

Arjun Gupta, Priya Mehta, and Rahul Singh. 2025. Quantum walk-based semantic retrieval in document graphs. *Entropy*.

Aram W. Harrow, Avinatan Hassidim, and Seth Lloyd. 2009. Quantum algorithm for linear systems of equations. *Physical Review Letters*, 103(15):150502.

Vojtěch Havlíček, Antonio D. Córcoles, Kristan Temme, Aram W. Harrow, Abhinav Kandala, Jerry M. Chow, and Jay M. Gambetta. 2019. Supervised learning with quantum-enhanced feature spaces. *Nature*, 567(7747):209–212.

G. Li. 2024. Quantum self-attention neural networks for text classification. *Science China Information Sciences*.

Wei Li, Rui Zhang, Yuchen Liu, and Haixu Wang. 2022. Quantum svm for text classification. *Quantum Information Processing*, 21(8):1–18.

Shyambabu Pandey, Partha Pakray, and Riyanka Manna. 2024. Quantum classifier for natural language processing applications. *Computación y Sistemas*, 28(2):695–700.

Benjamin Piwowarski, Ingo Frommholz, Mounia Lalmas, and Keith van Rijsbergen. 2010. A quantum-inspired framework for information retrieval. In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management (CIKM)*, pages 59–68.

Maria Schuld and Francesco Petruccione. 2019. *Quantum Machine Learning: An Introduction*. Springer, Cham, Switzerland.

Leandro C. Souza and Renato Portugal. 2025. Single-qudit quantum neural networks for multiclass classification. *arXiv preprint*. ArXiv:2503.09269.

T. Suzuki. 2024. Quantum support vector machines for classification and regression: Performance on nisq hardware. *Quantum Machine Intelligence*.

Rianne van der Meer, Peter Wittek, and Jae Lee. 2021. Quantum algorithms for information retrieval: Theory and applications. *Quantum Information Processing*, 20(5):1–24.

Taro Yamada, Kenji Suzuki, and Hiroshi Tanaka. 2024. Hybrid quantum-classical semantic search with embedding fidelity. *ACM Transactions on Information Systems*.

L. Zhang, H. Chen, and J. Zhao. 2023. Entropy-based quantum information retrieval. *Information Sciences*, 634:81–94.

Guido Zuccon and Leif Azzopardi. 2019. Quantum probability ranking principle for information retrieval. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 1171–1174.

QCNN-MFND: A Novel Quantum CNN Framework for Multimodal Fake News Detection in Social Media

Arya Suneesh
IIIT Kottayam
Kottayam, India
arya21bcs5@iiitkottayam.ac.in

Balasubramanian P
IIIT Kottayam
Kottayam, India
pbala@iiitkottayam.ac.in

Abstract

Fake news on social media platforms poses significant threats to public trust and information integrity. This research explores the application of quantum machine learning (QML) techniques for detecting fake news by leveraging quantum computing's unique capabilities. Our work introduces a hybrid quantum-classical framework that utilizes quantum convolutional neural networks (QCNNs) with angle and amplitude encoding schemes for processing multimodal features from text and images. Experiments conducted on benchmark datasets - GossipCop and Politifact - demonstrate that our quantum-enhanced model achieves superior performance compared to classical approaches, with accuracy rates of 88.52% and 85.58%, and F1 scores of 93.19% and 90.20% respectively. Our findings establish QML as a viable approach for addressing the challenges of fake news detection in the digital era.

1 Introduction

The proliferation of misinformation on social media threatens information integrity and societal welfare. Current machine learning and deep learning models struggle with accurate fake news identification due to insufficient feature extraction. Effective FND models must integrate textual and visual cues to distinguish between real and fake news, but conventional algorithms struggle to capture the subtle complexities of multi-modal data. We explore quantum machine learning as a promising alternative, focusing on quantum convolutional neural networks (QCNNs). Our research aims to develop a novel FND system that leverages QML techniques to enhance precision and robustness in fake news detection while maintaining computational efficiency.

2 Related Work

2.1 Unimodal FND Methods

Unimodal techniques focus either on textual or visual elements to categorize the news into fake or real.

2.1.1 ML-based FND Methods

Various studies have employed machine learning (ML) techniques for FND (Mishra and Sadia, 2023). Verma et al. (2021) utilized Support Vector Machine (SVM) for feature extraction from news articles but lacked deep learning (DL) models. Ozbay and Alatas (2020) adopted Decision Trees but faced accuracy challenges due to reliance on word count-based features. Esteban-Bravo et al. (2024) investigated early prediction of fake news virality using non-parametric models like Random Forest and Support Vector Classifier (SVC).

2.1.2 DL-based FND Methods

Rai et al. (2022) integrated BERT with LSTM, improving FND but suffered from low accuracy attributed to inadequate contextual features. Chen et al. (2024) tackled linguistic differences between Cantonese and Mandarin with a Deep semantic-aware graph convolutional network (SA-GCN) and CantoneseBERT on the Cantonese rumour dataset. Bazmi et al. (2023) emphasized the role of users' socio-cognitive biases and partisan bias with the Multi-View Co-Attention Network (MVCAN) but overlooked the influence of political viewpoints and credibility assessments of users.

2.1.3 QML-based FND Methods

Quantum machine learning (QML) techniques have shown promise in FND. Aishwarya et al. (2023) conducted a comprehensive review of Quantum Machine Learning techniques for FND. Their study revealed that QKNN, when integrated with Genetic and Evolutionary Feature Selection (GEFeS),

Proceedings of the QuantumNLP: Integrating Quantum Computing with Natural Language Processing, pages 44–52
November 24, 2025

achieved an impressive accuracy of 83.8%, surpassing the performance of conventional KNN algorithms. Tian and Baskiyar (2021) showcased the effectiveness of QKNN combined with Genetic and Evolutionary Feature Selection.

2.2 Multimodal FND Methods

Multimodal FND methods integrate both textual and visual features for detection.

2.2.1 DL-based FND Methods

Raja et al. (2024) proposed Dilated Temporal CNNs (DTCN), BiLSTM, and Contextualized Attention Mechanism (CAM), achieving impressive accuracy of 93.97% on the Dravidian_Fake dataset. Singhal et al. (2020) employed SpotFake+ but faced issues with prolonged training time and information loss from VGG-19's pooling layer. Kaliyar et al. (2021) employed Feed Forward Neural Networks with multiple CNN channels for local sequential feature extraction, yet generalization ability remains unexplored. Singh et al. (2023) employed multimodal learning techniques with NasNet Mobile for image analysis and BERT+ELECTRA for text processing, achieving 85% accuracy on the Twitter MediaEval Dataset and Weibo Corpus.

2.2.2 QML-based FND Methods

Qu et al. (2024) proposed QMFND, a quantum multimodal fusion-based model designed specifically for FND on social media platforms. By employing quantum encoding and quantum convolutional neural networks (QCNNs), QMFND achieved notable accuracies of 87.9% and 84.6% on the Gossipcop and Politifact datasets, respectively. However, the performance of QMFND is subject to limitations imposed by current hardware constraints and significant background noise in the operating environment of quantum computers.

3 Preliminaries

3.1 Pre-trained Language Models

Pre-trained language models form the basis for extracting representations from news text, using transformer architectures to capture contextual relationships. **BERT** processes text bidirectionally through masked language modeling, predicting randomly masked tokens from surrounding context. **XLNet** employs permutation-based autoregressive pre-training, capturing bidirectional context without relying on `[MASK]` tokens by considering all sequence permutations. **DistilBERT** is a compressed version of BERT that retains 97% of its language understanding while running 60% faster through knowledge distillation from a larger teacher model.

3.2 Pre-trained Image Models

Pre-trained convolutional neural networks extract visual features from images. These models learn hierarchical representations through successive convolutional layers.

VGG architectures (VGG16 and VGG19) utilize small 3×3 convolutional filters throughout the network. They stack multiple convolutional layers before pooling operations. This design enables learning complex features while maintaining computational efficiency.

ResNet50 introduces residual connections to address vanishing gradient problems. Skip connections allow gradients to flow directly through shortcuts. The architecture consists of 50 layers organized into residual blocks. Each block contains convolutional layers with identity mappings.

EfficientNet applies compound scaling to balance network depth, width, and resolution. It uses mobile inverted bottleneck blocks (MBConv) as building components. Squeeze-and-excitation optimization improves channel interdependencies. This architecture achieves superior accuracy with fewer parameters.

3.3 Understanding Quantum Mechanisms

Quantum computing transcends classical computing principles, offering the potential for unprecedented computational power and efficiency. One fundamental aspect of quantum computing is quantum encoding, a technique that transforms classical information into quantum states, enabling it to be processed and manipulated by quantum algorithms.

Quantum encoding transforms classical data into quantum states, exploiting superposition and entanglement to exponentially increase information density and computational capabilities beyond classical methods.

Several encoding approaches exist, each with distinct advantages LaRose and Coyle (2020):

Angle Encoding represents data through rotational angles of quantum gates. Parameterized rotation operations encode information directly into angular parameters. This method offers simplicity and hardware efficiency for near-term quantum devices.

$$|\psi\rangle = \bigotimes_{i=1}^{n} R_Y(x_i)|0\rangle = \bigotimes_{i=1}^{n} \begin{pmatrix} \cos(x_i/2) \\ \sin(x_i/2) \end{pmatrix} \quad (1)$$

where $R_Y(x_i)$ represents a Y-axis rotation gate parameterized by the data value x_i.

Amplitude Encoding represents data through relative amplitudes of quantum states. This scheme leverages superposition, enabling multiple information pieces to coexist simultaneously within a quantum system. For a normalized classical data vector $\mathbf{x}$, amplitude encoding creates:

$$|\psi\rangle = \frac{1}{\|\mathbf{x}\|} \sum_{i=0}^{2^n-1} x_i |i\rangle \quad (2)$$

where $\|\mathbf{x}\|$ is the normalization factor ensuring $\langle\psi|\psi\rangle = 1$, and $|i\rangle$ represents the computational basis states.

4 Proposed Methodology

4.1 Training and Evaluation Framework for Hybrid Model

Our fake news detection system, described in figure 1 combines classical deep learning with quantum computing. The hybrid architecture processes text and visual content through separate pathways before quantum integration.

In data preparation, news articles pass through a text transformer for contextual embeddings, while images are processed via CNN for visual feature extraction. A MultiHeadCrossAttention mechanism creates connections between text and image modalities, helping identify mismatches that often signal deception.

Training follows an epoch-based approach with early stopping to prevent overfitting. Data batches move through classical pathways, get fused, and pass to the QCNN (Cong et al., 2019) component, which leverages quantum principles like superposition and entanglement. We selected QCNN over Q-RNN or Q-LSTM because CNNs naturally preserve spatial locality in quantum circuits, essential for capturing hierarchical patterns in multimodal data. The convolutional structure aligns with quantum gate locality constraints on NISQ devices. We initially used cross-entropy loss before switching to focal loss due to dataset imbalance, and implemented gradient clipping for stability.

The validation process runs after each epoch, computing accuracy, precision, and recall. An early stopping mechanism halts training after three epochs without improvement in validation loss.

Final evaluation includes standard classification metrics and threshold optimization to identify the optimal decision boundary between real and fake news, producing metrics using both default (0.5) and optimized thresholds.

This pipeline balances classical deep learning's strength in feature extraction with quantum computing's advantages in modeling complex relationships.

5 Experimental Settings

5.1 Setup

All experiments were conducted on the Kaggle platform using an NVIDIA Tesla P100 GPU (16 GB VRAM) with 13GB RAM. The models were implemented using PyTorch 2.0 and trained with CUDA 12 acceleration. For transformer-based language models, we utilized the Hugging Face Transformers library. Image processing was handled with torchvision and quantum circuit simulations were executed using Pennylane with PyTorch interface.

The datasets were preprocessed using standard NLP techniques for textual data, including tokenization, normalization, and sequence padding. For image data, we employed standard preprocessing pipelines with resizing to 224×224 pixels, normalization using ImageNet statistics (mean=[0.485, 0.456, 0.406], std=[0.229, 0.224, 0.225]), and augmentation techniques including random horizontal flips and color jitter during training.

5.2 Dataset Analysis

The experiments were conducted on two benchmark fake news datasets: Gossipcop and Politifact. As described in Table 1, the dataset statistics reveal several notable characteristics. A significant class imbalance exists in both GossipCop and Politifact datasets, with real news consistently outnumbering fake news. The GossipCop (GC) dataset maintains approximately an 80-20 split between real and fake news in both train and test sets. The Politifact (PF) dataset shows a different ratio, with approximately 65-35 split in the training set shifting to 72-28 in the test set.

Text length analysis exposes distinct patterns between the two sources: Politifact articles are generally longer, with mean lengths of 8,919 and 9,494 characters for train and test sets respectively, compared to GossipCop's shorter average of around

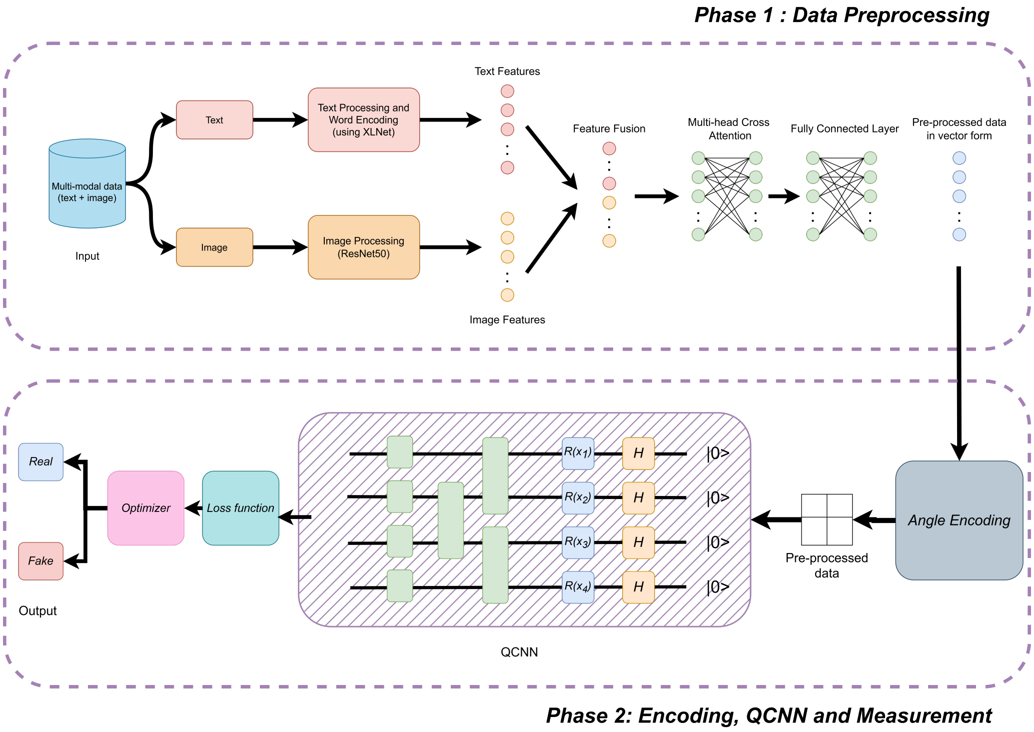


Figure 1: Overview of Training Process of QCNN-MFND

Table 1: Dataset Statistics Comparison

Metric	GC Train	GC Test	PF Train	PF Test
Total samples	10,010	2,830	381	104
Real	7,974 (79.7%)	2,285 (80.7%)	246 (64.6%)	75 (72.1%)
Fake	2,036 (20.3%)	545 (19.3%)	135 (35.4%)	29 (27.9%)
Mean text length	3,427.5	3,460.8	8,919.2	9,494.2
Std dev	5,872.6	6,433.2	17,501.6	18,349.9
Min	34.0	57.0	42.0	45.0
Median	2,072.0	2,046.5	2,511.0	2,966.5
Max	100,096.0	100,055.0	100,155.0	100,077.0

3,400 characters.

5.3 Evaluation Metrics

Due to class imbalance, we employed multiple standard metrics for binary classification problems:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{4}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{5}$$

$$\text{F1 Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{6}$$

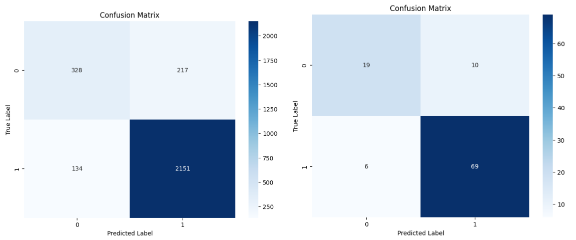
where TP, TN, FP, FN represent true positives, true negatives, false positives, and false negatives respectively.

6 Results and Analysis

6.1 Textual Feature Analysis

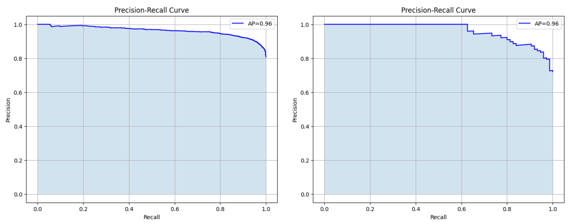
We evaluated multiple transformer-based language models for textual feature extraction. Tables 2 and 3 present the performance metrics across both datasets.

XLNet achieved the highest accuracy (0.876) on the GossipCop dataset, while DistilBERT demonstrated superior performance on Politifact with the highest accuracy (0.9135) and F1 score (0.9379). These results highlight the effectiveness of transformer-based models for fake news detec-



(a) GossipCop dataset (b) Politifact dataset

Figure 2: Confusion matrices for XLNet on both datasets.



(a) GossipCop dataset (b) Politifact dataset

Figure 3: Precision-Recall curves for XLNet on both datasets.

tion, with different architectures exhibiting distinct strengths across different news domains.

Fig. 2 shows the confusion matrices for XLNet performance on both datasets, demonstrating strong classification performance with minimal false negatives. Fig. 3 presents the precision-recall curves, indicating robust performance across different threshold values.

6.2 Visual Feature Analysis

We evaluated six prominent CNN architectures for visual feature extraction. Table 4 shows the performance comparison on the GossipCop dataset.

ResNet50 achieved the highest performance (79.93% accuracy, 0.7802 F1 score) among all CNN models. Modern architectures generally demonstrated better optimization with lower loss values compared to traditional VGG models.

Table 2: Performance of Transformer Models on GossipCop

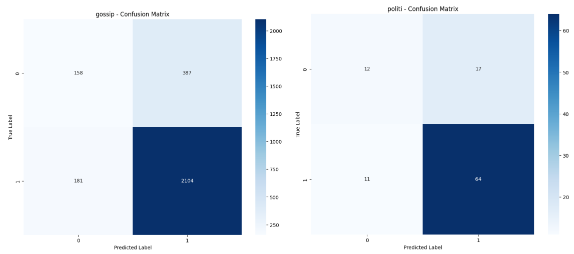
Model	Acc.	Prec.	Rec.	F1	Loss
BERT	0.871	0.920	0.920	0.920	0.747
RoBERTa	0.874	0.913	0.933	0.923	0.643
MPNet	0.872	0.926	0.915	0.920	**0.605**
DistilBERT	0.869	0.913	0.926	0.919	0.691
XLNet	**0.876**	0.908	**0.941**	**0.925**	0.707

Table 3: Performance of Transformer Models on Politifact

Model	Acc.	Prec.	Rec.	F1	Loss
BERT	0.846	0.873	0.920	0.896	**0.388**
RoBERTa	0.875	0.888	**0.947**	0.916	0.857
MPNet	0.837	0.914	0.853	0.883	0.456
DistilBERT	**0.914**	**0.971**	0.907	**0.938**	0.584
XLNet	0.846	0.873	0.920	0.896	0.585

Table 4: Performance of CNN Models on GossipCop

Model	Acc.	Prec.	Rec.	F1	Loss
VGG16	0.765	0.745	0.765	0.753	0.992
VGG19	0.786	0.759	0.786	0.769	**0.658**
ResNet50	**0.799**	**0.772**	**0.799**	**0.780**	0.974
EfficientNet	0.763	0.765	0.763	0.764	0.999
ViT	0.728	0.742	0.728	0.735	1.018
ConvNeXt	0.783	0.761	0.783	0.769	1.110



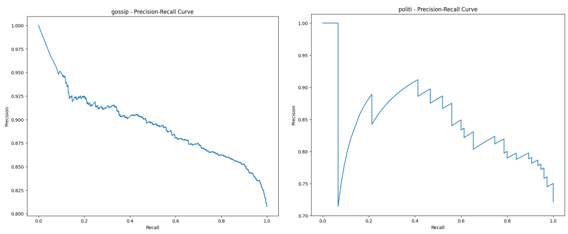
(a) GossipCop dataset (b) Politifact dataset

Figure 4: Confusion matrices for ResNet50 on both datasets.

Fig. 4 shows the confusion matrices for ResNet50 on both datasets, while Fig. 5 displays the corresponding precision-recall curves, demonstrating consistent performance across different news domains.

Table 5: Performance of CNN Models on Politifact

Model	Acc.	Prec.	Rec.	F1	Loss
VGG16	0.721	0.724	0.721	0.723	1.719
VGG19	0.721	0.730	0.721	0.725	1.623
ResNet50	**0.731**	0.715	**0.731**	0.720	0.923
EfficientNet	0.731	0.720	0.731	0.724	**0.744**
ViT	0.731	0.720	0.731	0.724	1.170
ConvNeXt	0.721	**0.744**	0.721	**0.729**	0.791



(a) GossipCop dataset (b) Politifact dataset

Figure 5: Precision-Recall curves for ResNet50 on both datasets.

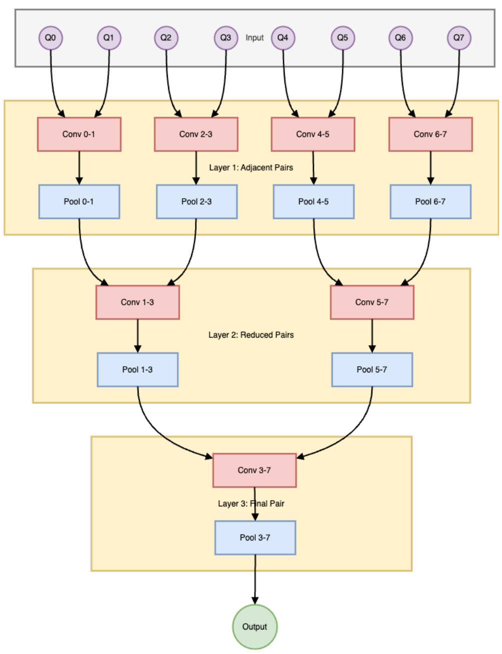


Figure 6: Architecture of the implemented QCNN showing the hierarchical quantum processing structure.

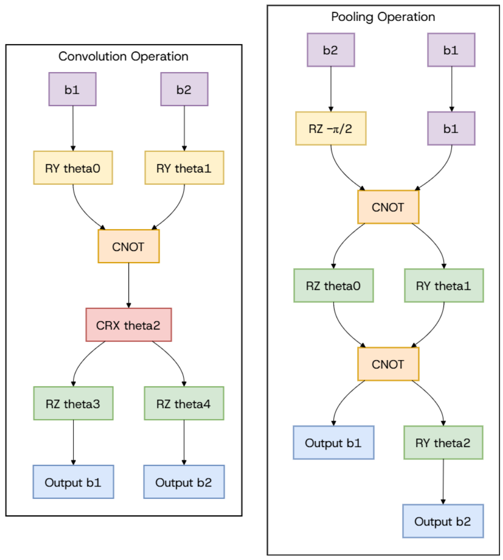


Figure 7: Detailed view of convolution and pooling layers in the QCNN, illustrating quantum gate operations.

6.3 Quantum Convolutional Neural Network Implementation

The implemented QCNN architecture consists of three primary components: quantum convolution layers, quantum pooling layers, and a measurement layer. The network operates on 8 qubits and implements a hierarchical structure with multiple conv-pool operations at different scales. We selected 8 qubits as a balance between expressivity and current NISQ device limitations, aligning with typical quantum hardware availability.

Fig. 6 illustrates our QCNN architecture, while Fig. 7 details the convolution and pooling layer operations, demonstrating the quantum gate operations used for feature extraction and compression.

Each convolution operation implements initial RY rotations, CNOT entanglement, controlled-RX rotation, and final RZ rotations. Each pooling operation uses parameterized rotations and CNOT gates to compress quantum information while preserving relevant features.

6.4 Experimental Setup

6.4.1 Model Architecture

A hybrid quantum-classical model was implemented with five key components. The image pathway uses pretrained CNNs (VGG16, EfficientNet, ResNet50) to extract features, projecting them to lower dimensions ($qbits/2$) for fusion. The text pathway processes input through XLNet, with mean-pooled features projected to $qbits/2$ dimensions. MultiHeadCrossAttention aligns image features with text context. The fusion component concatenates features and compresses them via a linear layer to $qbits$ dimensions. Finally, the QCNN processes the fused features to produce class probabilities.

6.4.2 Training Protocol

Loss Functions: Two loss functions addressed class imbalance:

Cross-Entropy Loss (CE) with class weighting and label smoothing:

$$\mathcal{L}_{\text{CE}} = -\frac{1}{N}\sum_{i=1}^{N} w_{y_i}\Big[y_i \log(p_i) + (1-y_i)\log(1-p_i)\Big] + \lambda\|\theta\|_2^2 \quad (7)$$

where w_{y_i} is class weight (inverse frequency), p_i is predicted probability, and $\lambda = 0.1$.

Focal Loss (FL) down-weights easy examples:

$$\mathcal{L}_{\text{FL}} = -\frac{1}{N}\sum_{i=1}^{N} \alpha_{y_i}(1-p_i)^{\gamma}\Big[y_i \log(p_i) + (1-y_i)\log(1-p_i)\Big] \tag{8}$$

where $\alpha_{y_i} = \frac{\text{\# minority class}}{\text{\# total samples}}$ and $\gamma = 2.0$.

Training Setup: AdamW optimizer (lr=2×10^{-4}), OneCycleLR scheduler, batch size 32, 25 epochs with early stopping (patience=3).

6.5 Performance Across Model Configurations

We evaluated multiple combinations of text feature extractors (XLNet), image feature extractors (VGG16, EfficientNet, ResNet50), and quantum encoding methods (Angle, Amplitude) on both datasets. Table 6 summarizes the performance metrics for each configuration.

On the GossipCop dataset, ResNet50 + XLNet with angle encoding trained with focal loss achieved the highest performance (88.52% accuracy, 93.19% F1 score). Angle encoding consistently outperformed amplitude encoding when paired with the same image feature extractor.

On the Politifact dataset, EfficientNet+XLNet with angle encoding and focal loss achieved the best results (88.46% accuracy, 92.31% F1 score). Focal loss significantly improved performance across both datasets, particularly evident in recall performance.

6.6 Classical versus Quantum Models

Table 7 presents the comparison between traditional classical approach (XLNet+ResNet50) and our proposed quantum model (QCNN-MFND).

The comparative analysis reveals our hybrid model outperforms classical approaches on the GossipCop dataset, with a remarkable 65% reduction in false negatives, critical for minimizing missed fake news instances. While the classical approach performs marginally better on Politifact's smaller dataset, this suggests our quantum model requires larger datasets to fully optimize its parameters.

7 Conclusion

We successfully developed QCNN-MFND, a novel framework leveraging quantum computing principles for fake news detection on social media. By combining QML with deep learning approaches, our architecture integrates XLNet for text analysis, ResNet50 for visual features, and quantum convolutional neural networks for multimodal fusion. The framework achieves impressive results—**88.52% accuracy and 93.19% F1 score on GossipCop**, and **85.58% accuracy with 90.20% F1 score on Politifact** - demonstrating significant advantages in minimizing missed fake news instances. Our experiments reveal that quantum computing offers particular benefits for larger datasets, providing a balanced precision-recall trade-off that prioritizes false negative reduction. This advancement represents an important step toward creating more trustworthy information ecosystems, with potential for greater improvements as quantum computing technology continues to mature.

Future directions include building a web application to enable real-time detection, further QCNN architectural exploration, and explainable quantum models for fake news detection.

Limitations

Several limitations merit consideration. We simulated quantum circuits on classical hardware; real quantum devices introduce noise and hardware constraints not captured in our experiments. Our 8-qubit architecture faces deployment challenges on current NISQ devices.

The datasets present additional constraints. The Politifact dataset's small size (381 training samples) limits model learning capacity. Both datasets focus exclusively on English-language social media news, leaving cross-domain and multilingual generalization untested. The significant class imbalance (80-20 and 65-35 splits) affects detection performance despite focal loss mitigation.

Our evaluation scope remains limited to two social media datasets. Temporal robustness, adversarial testing, and real-time inference performance remain unexplored. Training requires high-end GPU resources, and deployment costs on actual quantum hardware are quite high.

References

C. Aishwarya, Manikanadan Venkatesan, and P. Prabhavathy. 2023. Research oriented reviewing of quantum machine learning. *International Research Journal on Advanced Science Hub*, 5:165–177.

Parisa Bazmi, Masoud Asadpour, and Azadeh Shakery. 2023. Multi-view co-attention network for fake

Table 6: Experiments using proposed model QCNN-MFND with various configurations

Dataset	Text	Image	Encoding	Acc.	Prec.	Rec.	F1
GossipCop	XLNet	VGG16	Angle	0.877	0.902	0.951	0.926
		EfficientNet		0.881	0.906	0.952	0.928
		EfficientNet*		0.882	0.899	0.963	0.930
		ResNet50*		**0.885**	0.894	**0.973**	**0.932**
		VGG16	Amplitude	0.877	0.899	0.955	0.926
		EfficientNet		0.875	**0.916**	0.931	0.923
		ResNet50*		0.884	0.902	0.962	0.931
Politifact	XLNet	VGG16	Angle	0.846	0.883	0.907	0.895
		EfficientNet		0.837	0.837	0.960	0.894
		EfficientNet*		**0.885**	0.889	**0.960**	**0.923**
		ResNet50*		0.856	0.885	0.920	0.902
		VGG16	Amplitude	0.875	0.897	0.933	0.915
		EfficientNet		0.875	**0.908**	0.920	0.914
		ResNet50*		0.769	0.823	0.867	0.844

* Models trained with focal loss criterion.

Table 7: Performance Comparison Between Classical and Proposed Quantum Model

Dataset	Metric	Classical (without QCNN)	Proposed (with QCNN)
GossipCop	Accuracy	87.39%	**88.52%**
	F1 Score	0.922	**0.932**
	Precision	**0.918**	0.894
	Recall	0.926	**0.973**
Politifact	Accuracy	**87.50%**	85.58%
	F1 Score	**0.914**	0.902
	Precision	**0.908**	0.885
	Recall	0.920	0.920

news detection by modeling topic-specific user and news source credibility. *Information Processing & Management*, 60(1):103146.

Xinyu Chen, Yifei Jian, Liang Ke, Yunxiang Qiu, Xingshu Chen, Yunya Song, and Haizhou Wang. 2024. A deep semantic-aware approach for cantonese rumor detection in social networks with graph convolutional network. *Expert Systems with Applications*, 245:123007.

Iris Cong, Soonwon Choi, and Mikhail Lukin. 2019. Quantum convolutional neural networks. *Nature Physics*, 15:1–6.

Mercedes Esteban-Bravo, Lisbeth d. l. M. Jiménez-Rubido, and Jose M. Vidal-Sanz. 2024. Predicting the virality of fake news at the early stage of dissemination. *Expert Systems with Applications*, 248:123390.

Rohit Kumar Kaliyar, Anurag Goswami, and Pratik Narang. 2021. Fakebert: Fake news detection in social media with a bert-based deep learning approach. *Multimedia Tools and Applications*, 80(8):11765–11788.

Ryan LaRose and Brian Coyle. 2020. Robust data encodings for quantum classifiers. *Physical Review A*, 102(3):032420.

Alok Mishra and Halima Sadia. 2023. A comprehensive analysis of fake news detection models: A systematic literature review and current challenges. *Engineering Proceedings*, 59(1):28.

Feyza Altunbey Ozbay and Bilal Alatas. 2020. Fake news detection within online social media using supervised artificial intelligence algorithms. *Physica A: Statistical Mechanics and its Applications*, 540:123174.

Zhiguo Qu, Yunyi Meng, Ghulam Muhammad, and Prayag Tiwari. 2024. Qmfnd: A quantum multimodal fusion-based fake news detection model for social media. *Information Fusion*, 104:102172.

Nishant Rai, Deepika Kumar, Naman Kaushik, Chandan Raj, and Ahad Ali. 2022. Fake news classification using transformer based enhanced lstm and bert. *International Journal of Cognitive Computing in Engineering*, 3:98–105.

Eduri Raja, Badal Soni, Candy Lalrempuii, and Samir Kumar Borgohain. 2024. An adaptive cyclical learning rate based hybrid model for dravidian fake news detection. *Expert Systems with Applications*, 241:122768.

Prabhav Singh, Ridam Srivastava, K.P.S. Rana, and Vineet Kumar. 2023. Semi-fnd: Stacked ensemble based multimodal inferencing framework for faster fake news detection. *Expert Systems with Applications*, 215:119302.

Shivangi Singhal, Anubha Kabra, Mohit Sharma, Rajiv Ratn Shah, Tanmoy Chakraborty, and Ponnurangam Kumaraguru. 2020. Spotfake+: A multimodal framework for fake news detection via transfer learning (student abstract). In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 13915–13916.

Ziyan Tian and Sanjeev Baskiyar. 2021. Fake news detection: an application of quantum k-nearest neighbors. In *2021 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 1–6.

Pawan Kumar Verma, Prateek Agrawal, Ivone Amorim, and Radu Prodan. 2021. Welfake: word embedding over linguistic features for fake news detection. *IEEE Transactions on Computational Social Systems*, 8(4):881–893.

Quantum Natural Language Processing: A Comprehensive Survey of Models, Architectures, and Evaluation Methods

Arpita Vats
Boston University
LinkedIn*

Rahul Raja
Carnegie Mellon University
Stanford University
LinkedIn*

Ashish Kattamuri
Proofpoint*

Abhinav Bohra
Amazon*

Abstract

Quantum Natural Language Processing (QNLP) is an emerging interdisciplinary field at the intersection of quantum computing, natural language understanding, and formal linguistic theory. As advances in quantum hardware and algorithms accelerate, QNLP promises new paradigms for representation learning, semantic modeling, and efficient computation. However, existing literature remains fragmented, with no unified synthesis across modeling, encoding, and evaluation dimensions.In this work, we present the first systematic and taxonomy driven survey of QNLP that holistically organizes research spanning three core dimensions: computational models, encoding paradigms, and evaluation frameworks. First, we analyze foundational approaches that map linguistic structures into quantum formalism, including categorical compositional models, variational quantum circuits, and hybrid quantum classical architectures. Second, we introduce a unified taxonomy of encoding strategies, ranging from quantum tokenization and state preparation to embedding based encodings, highlighting tradeoffs in scalability, noise resilience, and expressiveness. Third, we provide the first comparative synthesis of evaluation methodologies, benchmark datasets, and performance metrics, while identifying reproducibility and standardization gaps.We further contrast quantum inspired NLP methods with fully quantum implemented systems, offering insights into resource efficiency, hardware feasibility, and real world applicability. Finally, we outline open challenges such as integration with LLMs and unified benchmark design, and propose a research agenda for advancing QNLP as a scalable and reliable discipline.

1 Introduction

The intersection of quantum computing and natural language processing (NLP) has given rise to the emerging field of QNLP. Traditional NLP methods rely heavily on classical statistical and neural approaches, which, despite recent breakthroughs in LLMs (Brown et al., 2020), face fundamental challenges in scalability, representation efficiency, and capturing complex compositional semantics (Bender et al., 2021). Quantum computing, with its inherent parallelism and high-dimensional Hilbert space representations, offers a fundamentally new computational paradigm that can potentially overcome some of these limitations (Meichanetzidis et al., 2020; Varmantchaonala et al., 2024). Specifically, quantum models promise exponential speedups in linear algebra operations, richer encoding of linguistic structures, and novel mechanisms for semantic composition grounded in quantum theory. Foundational frameworks such as categorical compositional distributional models (DisCoCat) (Coecke et al., 2010) leverage quantum formalisms to represent grammatical structure, while hybrid quantum classical architectures demonstrate the feasibility of encoding word embeddings and performing sentence classification tasks on near-term quantum hardware (Lorenz et al., 2021b). Recent work further explores quantum algorithms for compositional text processing (Zhang et al., 2024) and surveys near term QNLP applications (Wiebe et al., 2024).

This paper provides a systematic survey of QNLP across three core dimensions: (i) computational models that define how linguistic structure and semantics can be mapped to quantum circuits and algorithms, (ii) encoding paradigms that determine how text tokens, syntactic dependencies, or embeddings are represented in quantum states, and (iii) evaluation frameworks that assess the effectiveness, efficiency, and robustness of QNLP methods. By categorizing and analyzing existing approaches, we highlight key tradeoffs in expressiveness, scalability, and noise resilience. Furthermore, we contrast quantum inspired NLP tech-

*This work does not relate to the authors' positions at LinkedIn, Proofpoint, or Amazon.

Proceedings of the QuantumNLP: Integrating Quantum Computing with Natural Language Processing, pages 53–64
November 24, 2025

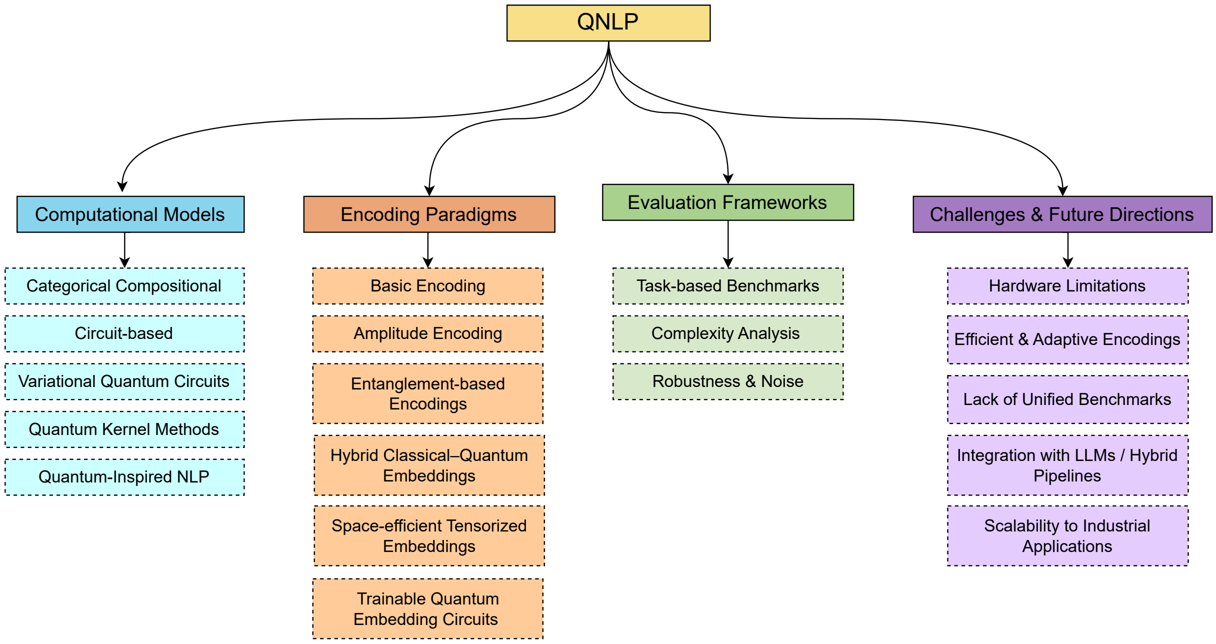


Figure 1: Taxonomy of QNLP highlighting core components computational models, encoding paradigms, evaluation frameworks, and future challenges.

niques, which adapt ideas from quantum mechanics within classical settings, with implementations on actual quantum hardware, thereby clarifying both theoretical promise and current practical limitations. The overall evolution of QNLP approaches from foundational categorical frameworks to hybrid quantum classical architectures is illustrated in Figure 1, which presents the taxonomy of major model families and their interrelations across computational, encoding, and evaluation dimensions.

2 Background

2.1 Quantum Computing Fundamentals

Quantum computing leverages the laws of quantum mechanics to perform computations beyond the reach of classical machines. Its fundamental unit of information, the *qubit*, generalizes the classical bit by existing in a superposition of states. A quantum state $|\psi\rangle$ is a vector in a complex Hilbert space $\mathcal{H}$ (Moretti and Oppio, 2017), where the state of a single qubit can be expressed as:

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle\,, \quad \alpha, \beta \in \mathbb{C}, \quad |\alpha|^2 + |\beta|^2 = 1.$$

Here, α and β are complex amplitudes, and the normalization condition ensures a probabilistic interpretation. Multiple qubits are represented through tensor products, e.g., $|\psi\rangle_{AB} = |\psi\rangle_A \otimes |\psi\rangle_B$. Entanglement arises when such states cannot be decomposed into tensor products, a phenomenon critical to quantum advantage in algorithms.

Quantum computation is driven by unitary operators U acting on states:

$$|\psi'\rangle = U\,|\psi\rangle\,,$$

which ensure reversibility and preserve probability amplitudes. Measurement collapses the superposition into classical outcomes, with probabilities determined by the squared amplitudes of the state vector. Together, superposition, entanglement, unitary evolution, and measurement define the computational paradigm of quantum mechanics.

2.2 Quantum Machine Learning Foundations

QML studies how quantum mechanical principles can enhance or accelerate learning algorithms (Schuld et al., 2015). It leverages the expressive power of quantum states and the computational efficiency of quantum operations to address tasks in classification, regression, clustering, and generative modeling.

A key concept is the *quantum feature map*, which encodes classical data $x \in \mathbb{R}^d$ into a quantum state $|\phi(x)\rangle$ within a high-dimensional Hilbert space $\mathcal{H}$. This encoding induces a kernel:

$$k(x, x') = |\langle\phi(x)|\phi(x')\rangle|^2,$$

allowing quantum models to exploit feature spaces that may be exponentially larger than those accessible classically (Schuld et al., 2015). Quantum kernels have been investigated for support vector machines (SVMs) and nearest-neighbor methods, showing theoretical potential for improved separability.

Another foundational algorithm is the Harrow–Hassidim–Lloyd (HHL) method, which provides exponential speedups for solving linear systems of equations (Harrow et al., 2009). Since solving linear systems underpins many ML tasks (e.g., regression, Gaussian processes), HHL exemplifies how quantum algorithms could drastically reduce complexity from polynomial to logarithmic in the number of variables. In the near term, *variational*

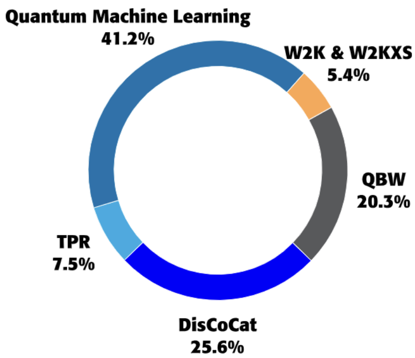


Figure 2: Adoption rates of QNLP models derived from the analyzed papers (Varmantchaonala et al., 2024).

quantum algorithms (VQAs) have become the dominant paradigm for NISQ era devices (Cerezo et al., 2021b). These models use parameterized quantum circuits $U(\theta)$, where θ denotes tunable gate parameters, to transform input states. The circuits are trained by minimizing an objective function:

$$C(\theta) = \langle \psi_0 | U^\dagger(\theta) H U(\theta) | \psi_0 \rangle,$$

with a classical optimizer updating θ based on quantum hardware evaluations. Variational circuits are flexible and have been applied to supervised learning (e.g., quantum classifiers), unsupervised tasks (e.g., clustering), and generative models.
Another critical building block is the *quantum neural network* (QNN), which uses variational circuits as analogues of neural layers. Entanglement plays a role similar to non-linear activation functions by enabling complex correlations between inputs. Hybrid QNNs combine quantum layers with classical networks, demonstrating performance gains in cases such as image and text classification.

From a complexity-theoretic perspective, QML offers potential advantages when classical methods suffer from the curse of dimensionality. Quantum states inhabit exponentially large Hilbert spaces naturally, enabling compact representation of complex data distributions. However, practical challenges remain, including noise, barren plateaus in variational optimization (McClean et al., 2018), and efficient data encoding (also known as the quantum data-loading problem).

For QNLP specifically, QML foundations provide the computational substrate: quantum feature maps offer new embedding paradigms for tokens, variational circuits serve as sequence-processing units, and entanglement provides a mechanism for modeling compositionality and long-range linguistic dependencies. These align with the goals of QNLP frameworks such as DisCoCat and hybrid quantum–classical pipelines, making QML an indispensable component of quantum language understanding.The distribution of QNLP model adoption across surveyed studies is shown in Figure 2, highlighting the dominance of Quantum Machine Learning based frameworks, followed by DisCoCat and Quantum Bag-of-Words models.

2.3 Natural Language Processing

NLP provides the computational basis for representing and interpreting linguistic data. Its core principle, *distributional semantics*, states that words appearing in similar contexts tend to have similar meanings. Early models such as Latent Semantic Analysis (LSA), Word2Vec, and GloVe encoded words as dense vectors $\mathbf{e}_w \in \mathbb{R}^d$, capturing semantic similarity through geometric proximity.

Modern NLP advances this idea through *contextual embeddings* using Transformer architectures such as BERT (Devlin et al., 2019), GPT (Radford et al., 2019), and T5 (Raffel et al., 2020). The self-attention mechanism

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) V$$

enables long-range dependency modeling by relating all tokens within a sequence (Vaswani et al., 2017). Despite their success, Transformers face $O(n^2)$ time and memory complexity with sequence length n, motivating efficient variants such as sparse and linearized attention.

Classical NLP also employs grammatical formalisms context-free grammars (CFGs), dependency parsing, and formal semantics to capture compositionality, yet integrating syntax with distributed semantics at scale remains challenging.

Quantum approaches address this limitation: quantum states in high-dimensional Hilbert spaces can encode inter-token dependencies through *entanglement*. Frameworks such as DisCoCat (Categorical Compositional Distributional Models) (Coecke et al., 2010) unify grammar and semantics via category theory, suggesting that QNLP can yield richer and more efficient representations than classical embeddings.

2.4 Quantum Classical Hybrids

Fully fault-tolerant quantum computers remain a long-term goal, but present-day devices fall into the category of Noisy Intermediate-Scale Quantum (NISQ) systems (Preskill, 2018). These machines contain on the order of 50–500 qubits, which are sufficient for exploring quantum advantage but are limited by decoherence, gate errors, and connectivity constraints. As a result, most practical QML and QNLP approaches rely on hybrid quantum–classical methods. **Variational Circuits** is a central paradigm in the NISQ era is the use of variational quantum circuits (VQCs) as shown in Figure 3. These are parameterized circuits $U(\theta)$ with tunable gates, where parameters θ are optimized iteratively by a classical optimizer. Given an input state $|\psi_0\rangle$ and a Hamiltonian H encoding the objective, the optimization task is defined as:

$$C(\theta) = \langle\psi_0|U^\dagger(\theta)HU(\theta)|\psi_0\rangle.$$

The quantum device computes expectation values, while the classical optimizer updates θ using gradient-based or gradient-free methods (Jäger et al., 2025). This feedback loop exploits quantum representational capacity while avoiding long quantum coherence times, which are difficult to sustain on NISQ devices. In a typical hybrid learning pipeline, classical pre-processing transforms raw data into a form suitable for quantum encoding (e.g., token embeddings or feature normalization). The encoded data are passed to a quantum circuit that performs transformations, such as entangling operations to capture correlations. The measurement outcomes are then post-processed by classical neural layers or decision functions. This integration allows quantum circuits to act as specialized layers within a larger classical deep learning framework.

For natural language tasks, hybrid models provide a practical compromise between expressiveness and feasibility. Classical components handle tasks such as subword tokenization, syntactic parsing, or initial embedding generation, while the quantum layer captures higher-order dependencies using entanglement. For example, a hybrid QNLP pipeline might map token embeddings into quantum states, apply a variational circuit to model contextual interactions, and then use a classical classifier to predict sentiment or semantic similarity. Such approaches combine the scalability of classical preprocessing with the structural advantages of quantum computation.

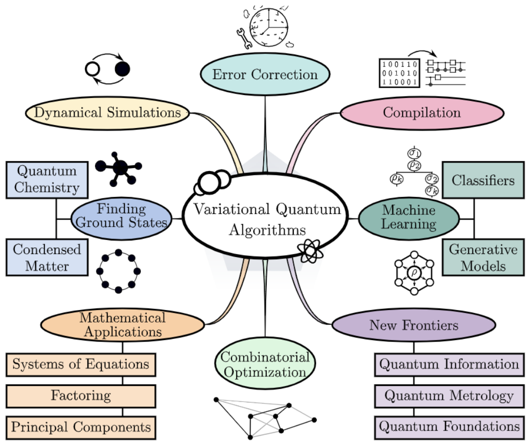


Figure 3: Applications of Variational Quantum Algorithms (VQAs) across optimization, simulation, machine learning, and emerging quantum domains. (Cerezo et al., 2021a).

3 Computational Models for QNLP

Several computational paradigms have been proposed for QNLP, each exploiting different aspects of quantum mechanics to model linguistic structure, meaning, and tasks. This section surveys categorical compositional frameworks, circuit based models, variational approaches, quantum kernel methods, and quantum inspired NLP techniques.

3.1 Categorical Compositional Models

The *categorical compositional distributional model* (DisCoCat) (Wu and Wang, 2019) was one of the first frameworks to unify grammatical structure and distributional semantics in a quantum-compatible setting. It leverages *compact closed categories* to map syntactic derivations to tensor contractions in Hilbert spaces. Each word is represented as a tensor, and sentence meaning arises compositionally through linear maps:

$$\vec{s} = f(\vec{w}_A \otimes \vec{w}_B), \quad f : A \otimes B \to C,$$

with entanglement naturally encoding word dependencies.

Building on this foundation, several extensions have been proposed: DisCoCirc (Chang et al., 2023): introduces discourse-level dynamics by updating word states via variational quantum circuits, e.g., $|w\rangle' = U_c |w\rangle$. Quantum Graph Transformers (QGT) (Xu et al., 2025): integrate dependency graphs with quantum self-attention, where attention weights are computed by parameterized circuits:

$$\alpha_{ij} = \frac{\exp(\langle \phi(x_i)|U(\theta)|\phi(x_j)\rangle)}{\sum_k \exp(\langle \phi(x_i)|U(\theta)|\phi(x_k)\rangle)}.$$

Quantum Context-Sensitive Embeddings (QCSE) (Liu et al., 2025b): generalize contextual embeddings (e.g., BERT) into Hilbert space with $|w, c\rangle = U(C)\,|w\rangle$. Quantum Text Pretraining Networks (QTP-Net) (Zhang et al., 2025): encode word senses as quantum superpositions $|w\rangle = \sum_i \alpha_i\,|s_i\rangle$ aligned with knowledge bases. MultiQ-NLP (Wang et al., 2024): extends composition to multimodal data, using entanglement to model cross-modal dependencies (text–image).

Together, these models have evolved DisCoCat from a purely categorical semantic formalism into dynamic, contextual, pretrained, and multimodal frameworks, demonstrating the adaptability of QNLP across linguistic and hybrid tasks.

3.2 Quantum Circuit-based Models

Quantum circuits map linguistic structure directly into hardware-executable operations. Tokens are encoded into quantum states, syntactic relations are represented by entangling gates, and grammatical reductions correspond to circuit modules (Ge et al., 2024). For example, a dependency relation between two words may be represented as a controlled rotation or CNOT gate applied between their corresponding qubits (Hu and Kais). Sentence meaning then emerges from the full circuit state, with measurements providing semantic outputs (Lan et al., 2024). An example of such a circuit implementation for a simple transitive sentence is shown in Figure 4.

Circuit-based approaches highlight the structural parallel between parse trees and quantum circuit diagrams, making them natural candidates for syntax-sensitive tasks (Liu et al., 2025a).They are particularly attractive for experiments on NISQ devices since circuits can be compiled directly into gate sequences supported by current hardware (Venturelli et al., 2019). However, their scalability depends on efficient encoding schemes and noise-aware compilation, as circuit depth grows with sentence length. Hybrid pipelines that combine shallow circuits with classical post-processing are commonly used to mitigate hardware limitations. A recent circuit-based approach proposes Quantum Parameter Adaptation (QPA), where quantum neural networks are used during training to generate classical model weights. This enables parameter efficient fine tuning of LLMs while keeping inference entirely classical (Liu et al., 2025a).

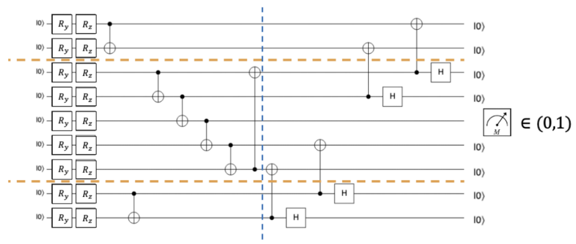


Figure 4: Quantum circuit for a transitive sentence. The circuit based on DisCoCat model, maps a simple sentence into quantum operations. Qubits on the left encode word embeddings via rotation gates, while the right region represents grammatical contractions through entangle gates such as CNOT

3.3 Variational Quantum Models

Variational quantum circuits (VQCs) $U(\theta)$ extend circuit-based models by introducing tunable parameters θ optimized via classical loops (Liu et al., 2024).

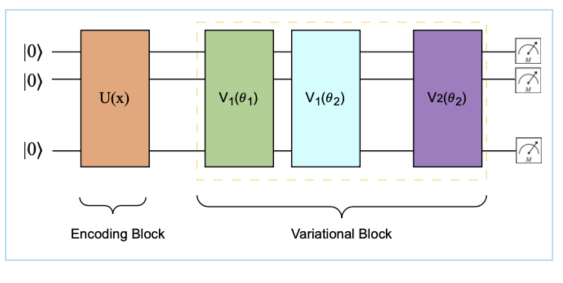


Figure 5: Variational Quantum Circuit (VQC) architecture illustrating how linguistic inputs are encoded into quantum states and processed by parameterized variational layers whose parameters are trained in a classical optimization loop. (Liu et al., 2024).

This paradigm makes VQCs the most widely explored approach in QNLP. Tokens are embedded into quantum states via feature maps, processed through parameterized entangling layers, and measured to produce outputs (Kankeu et al., 2025).

Training minimizes a loss function:

$$C(\theta) = \sum_i \ell(y_i, f_\theta(x_i)),$$

where ℓ is typically cross-entropy or mean squared error.

VQCs have been applied to tasks such as text classification, semantic similarity, and sentiment analysis. They benefit from the expressive capacity of entanglement to capture contextual information, and from their hybrid nature which integrates well with classical neural networks. Key challenges include barren plateaus in optimization, noise-induced instability, and the high cost of quantum state preparation (Novák et al., 2025). Recent work explores hardware efficient ansätze and error aware training to address these limitations (Gujju et al., 2025), making VQCs a practical testbed for QNLP research.

3.4 Quantum Kernel Methods

Quantum kernel methods leverage quantum feature maps $|\phi(x)\rangle$ that embed linguistic data into Hilbert spaces of potentially exponential dimension. The induced kernel is defined as:

$$k(x, x') = |\langle\phi(x)|\phi(x')\rangle|^2,$$

which can be used with classical machine learning models such as support vector machines (SVMs) or Gaussian processes (Wang et al., 2025). These methods are particularly well-suited to similarity-based tasks, including semantic textual similarity (STS), paraphrase detection, and clustering of embeddings (Herbold, 2024). They offer the advantage of being mathematically rigorous, providing provable separability properties in high-dimensional spaces. However, scalability is a major limitation, since evaluating kernels requires repeated state preparation and measurement. Approximate quantum kernel estimation and hybrid quantum–classical kernel learning have been proposed as intermediate solutions.

4 Encoding Paradigms

4.1 Basic Encoding

A recent proposal introduces a *learnable basic encoding layer* that maps each token to a qubit register with minimal parameter overhead (Munikote, 2024). Instead of relying purely on fixed rotation or amplitude maps, the method applies small parameterized gates on basis states, adapting them during training to better reflect token distributions. Concretely, a token index i is first mapped to a basis state $|i\rangle$, and then acted upon by a shallow trainable unitary $E(\phi)$:

$$|\psi_i\rangle = E(\phi)\,|i\rangle\,.$$

Here, $E(\phi)$ is composed of single-qubit rotations and entanglers whose parameters ϕ are learned jointly with the downstream task, offering a flexible compromise between rigid encodings and heavy variational circuits.

The scheme retains the discrete structure of token identities while allowing adaptation to semantic space, enabling gradients to flow directly through the encoder (Baek et al., 2025). Because only a small unitary is applied, the circuit depth overhead remains modest, making it compatible with NISQ devices. Its parameters can absorb differences in token frequency or contextual distributions, positioning this method between static basis encoding and hybrid embeddings. As such, it provides a more expressive and scalable representation for QNLP tasks than one-hot or rotation-only mappings.

4.2 Amplitude Encoding

Embed dense vectors into amplitudes:

$$\mathbf{e} \in \mathbb{R}^d \;\mapsto\; |\phi(\mathbf{e})\rangle = \frac{1}{\|\mathbf{e}\|} \sum_{j=1}^{d} e_j\,|j\rangle\,.$$

This method is highly qubit-efficient ($\log d$) and preserves inner-product geometry, allowing similarity to be computed via inner products in Hilbert space. The main drawback is that state preparation can be computationally expensive, often requiring $\mathcal{O}(d)$ operations, and the resulting states are sensitive to noise. To mitigate this, amplitude encoding is often combined with problem-specific *quantum feature maps*, enabling kernel methods that exploit the high-dimensional Hilbert space structure (Schuld and Killoran, 2019).

Recent advances show that amplitude encoding can deliver exponential data compression in hybrid quantum-classical architectures. For instance, a dataset with $d = 2^n$ features can be represented using only n qubits, whereas angle encoding would require d. Chen et al. (Chen et al., 2025) integrate amplitude encoding into hybrid Quantum Neural Networks (QNNs) for recovery rate prediction, demonstrating superior generalization on small-sample, high-dimensional financial datasets. Embedding amplitude-encoded inputs into Parameterized Quantum Circuits (PQCs) preserves unitarity and avoids costly orthogonality constraints, yielding two key benefits: improved computational efficiency through fewer qubits and parameters, and richer representational capacity compared to angle encoding for tasks requiring high-dimensional embeddings.

4.3 Entanglement-based Encodings

Introduce entanglers (CNOT/CZ) to correlate token subsystems (Schuld et al., 2021):

$$|\psi\rangle = U_{\text{ent}}\big(|w_1\rangle \otimes \cdots \otimes |w_n\rangle\big).$$

This approach explicitly captures syntactic and semantic dependencies by creating correlations between token representations, mirroring categorical contraction in compositional semantics. Entanglement allows local word embeddings to be combined into global sentence states, enriching expressivity beyond independent encodings.

The trade-off is that entanglement substantially increases circuit depth and noise sensitivity, especially on NISQ hardware (González-García et al., 2022). Efficient design therefore requires carefully chosen ansätze and compilation strategies to minimize gate counts and error accumulation. When optimized, entanglement-based encodings provide a direct mechanism for modeling relational structure, but scalability remains a major challenge compared to simpler schemes.

4.4 Hybrid Embedding Strategies

A hybrid approach first uses a classical model (e.g., BERT or Word2Vec) (Devlin et al., 2019) to compute an embedding $\mathbf{e}$, and then applies a feature map $\mathbf{e} \mapsto |\phi(\mathbf{e})\rangle$ followed by a trainable quantum circuit $U(\theta)$ before measurement. This combines the semantic richness of pretrained embeddings with quantum layers that can model higher-order correlations and capture non-linear dependencies in Hilbert space (Döschl and Bohrdt, 2025).

Such strategies represent the most practical and NISQ-friendly pathway, since heavy semantic lifting is done classically and quantum resources are reserved for expressive refinements. By leveraging classical pretraining, hybrid embeddings reduce qubit demands and training cost, while still offering the potential to uncover representational structures inaccessible to purely classical methods. This makes them a dominant design choice for early QNLP systems and applied quantum machine learning pipelines.

4.5 Space-efficient tensorized embeddings.

A line of work factorizes the embedding matrix into low-order tensor products inspired by entanglement, yielding *word2ket*-style embeddings that compress parameters by $10^2\times$ or more with negligible accuracy loss on standard NLP tasks (Panahi et al., 2019). These embeddings can be used purely classically or as quantum-ready parametrizations (tensor factors $\Rightarrow$ shallow preparation circuits). This offers a principled bridge between tensor-network structure and learnable word representations.

4.6 Trainable quantum embedding circuits.

A 2024 study proposes a recurrent quantum embedding neural network (RQENN) with a *trainable encoding* based on parameterized binary indices that learns token embeddings *within* a small quantum circuit cell; the cell is reused across sequence positions to capture context with fewer qubits and measurements than prior QNLP approaches (Varmantchaonala et al., 2025). Reported results show reduced parameter count and bits used, and accuracy gains over earlier QNLP baselines on a text-like vulnerability detection task, highlighting the value of *learned* encoders (vs. fixed maps) under NISQ constraints (Kea et al., 2024).

4.7 Resource Cost Modeling

We characterize encodings by qubits q, depth L, state-prep cost T_{prep}, and shot complexity m. For amplitude encoding,

$$\mathbf{e} \in \mathbb{R}^d \mapsto |\phi(\mathbf{e})\rangle = \frac{1}{\|\mathbf{e}\|} \sum_{j=1}^{d} e_j \, |j\rangle\,, \quad (1)$$

$$q = \lceil \log_2 d \rceil, \qquad T_{\text{prep}} = \Theta(d). \quad (2)$$

with low depth but prep-bound runtime. Angle/rotation encoding yields $q = \Theta(d)$, $T_{\text{prep}} = \Theta(d)$ and often better robustness on NISQ. Entanglement-based composition adds syntax or graph-induced two qubit layers; we report $L = L_0 + E$ where E is the number of entangling edges (Susulovska, 2024). For hybrid embeddings, q is constant (few qubit head) with classical compute absorbing semantics; we report wall clock and device usage alongside accuracy.

5 Evaluation Frameworks

Evaluation in QNLP spans both empirical performance and theoretical efficiency. At the task level, models are assessed on standard NLP objectives such as sentiment classification, semantic similarity, and sequence labeling, with accuracy, F1, or correlation metrics compared against compute-matched classical baselines (Tomal et al., 2025). Because quantum circuits produce probabilistic outputs, metrics are accompanied by confidence intervals derived from measurement shots, and evaluations must also report resource costs including qubit counts q, circuit depth L, gate complexity, state-preparation cost T_{prep}, and shot budgets m, ensuring fairness under NISQ constraints (Ma and Li, 2024). To validate results beyond simulation, a hardware-in-the-loop protocol is followed: device backend, transpilation strategy, calibration snapshot, and shot counts are disclosed, with paired simulator–device runs performed using identical seeds (Nguyen et al., 2017). Robustness is further probed through noise modeling, barren-plateau stress tests, and lightweight error mitigation (readout calibration, zero-noise extrapolation, and gradient-preserving initialization).

Beyond raw task performance, evaluation emphasizes comparability and reproducibility. Canonical ablations such as removing entanglers, swapping amplitude versus angle encodings, reducing data re-uploading depth, or replacing quantum heads with classical ones are standardized to attribute improvements to specific design choices (Aktar et al., 2025). Benchmarking remains challenging due to the lack of large standardized QNLP corpora, so we propose compact, structure-sensitive tasks (compositional classification, semantic similarity, and sequence labeling with nested constituents) with fixed splits and optional precomputed embeddings for hybrid models. Together with artifact release (QASM circuits, seeds, calibration snapshots, and ablation configs) (Li et al., 2022), these practices enable like for like comparisons across models and clarify where QNLP shows unique strengths capturing compositionality, contextual dependencies, and high-dimensional correlations while highlighting the tradeoffs in scalability, noise resilience, and hardware feasibility relative to classical NLP systems (Lhoest et al., 2021).

6 Challenges and Future Directions

Despite encouraging theoretical advances and early experiments, QNLP still faces significant challenges. Current NISQ hardware limits circuit depth, qubit counts, and gate fidelity, restricting scalability and necessitating noise-resilient encodings and carefully designed variational ansätze (Preskill, 2018; McClean et al., 2018). Encoding strategies such as amplitude or entanglement-based mappings offer expressive representational power but suffer from costly state preparation and noise sensitivity, motivating the exploration of adaptive encodings and resource efficient parameterization methods that balance expressivity with hardware feasibility (Chen et al., 2025). At the evaluation level, the absence of standardized QNLP benchmarks makes comparisons across models difficult; task-specific corpora and quantum-compatible evaluation suites are needed to validate theoretical speedups and measure robustness under realistic conditions (Lorenz et al., 2021a).

Looking ahead, hybrid quantum classical pipelines remain the most practical path, though their advantage over strong classical baselines such as transformers is not yet conclusive. Further research into quantum inspired embeddings and hybrid variational architectures may clarify where QNLP offers unique value (Huang et al., 2021; Kartsaklis et al., 2021). Achieving scalability will require moving beyond toy corpora to industrial-scale applications such as semantic search, question answering, and multimodal reasoning. Meeting these goals will demand not only algorithmic innovation but also advances in quantum hardware and close collaboration between NLP researchers and quantum computing specialists, ensuring QNLP matures into a robust framework for structure-sensitive language tasks.

7 Conclusion

QNLP lies at the intersection of quantum computing and natural language processing, introducing new paradigms for compositional semantics, efficient representation, and contextual modeling. This survey reviews foundational models DisCoCat, circuit based, variational, and hybrid architectures alongside encoding strategies, evaluation frameworks, and open challenges. Although still nascent, advances in hybrid embeddings, quantum feature maps, and noise mitigation indicate near-term feasibility. Future progress will hinge on scalable benchmarks, tighter integration with classical NLP, and improved quantum hardware. QNLP thus holds promise to advance beyond proof-of-concept studies and deliver tangible computational gains for structure sensitive language tasks.

References

Shamminuj Aktar, Andreas Bärtschi, Abdel-Hameed A. Badawy, and Stephan Eidenbenz. 2025. Quantum graph transformer for nlp sentiment classification.

Junyeob Baek, Hosung Lee, Christopher Hoang, Mengye Ren, and Sungjin Ahn. 2025. Discrete jepa: Learning discrete token representations without reconstruction.

Emily M Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*.

Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. Language models are few-shot learners. In *Advances in Neural Information Processing Systems*, volume 33.

M. Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C. Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R. McClean, Kosuke Mitarai, Xiao Yuan, Lukasz Cincio, and Patrick J. Coles. 2021a. Variational quantum algorithms. *Nature Reviews Physics*, 3(9):625–644.

Marco Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C Benjamin, Suguru Endo, Keisuke Fujii, et al. 2021b. Variational quantum algorithms. *Nature Reviews Physics*, 3.

Eric Chang, John Smith, and Ling Zhao. 2023. Variational quantum classifiers for natural-language text: A discocirc approach. In *Proceedings of the Quantum NLP Workshop*.

Ying Chen, Paul Griffin, Paolo Recchia, Lei Zhou, and Hongrui Zhang. 2025. Hybrid quantum neural networks with amplitude encoding: Advancing recovery rate predictions.

Bob Coecke, Mehrnoosh Sadrzadeh, and Stephen Clark. 2010. Mathematical foundations for a compositional distributional model of meaning. *Lambek Festschrift*.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*.

Fabian Döschl and Annabelle Bohrdt. 2025. Importance of correlations for neural quantum states.

Yan Ge, Wu Wenjie, Chen Yuheng, Pan Kaisen, Lu Xudong, Zhou Zixiang, Wang Yuhan, Wang Ruocheng, and Yan Junchi. 2024. Quantum circuit synthesis and compilation optimization: Overview and prospects.

Guillermo González-García, Rahul Trivedi, and J. Ignacio Cirac. 2022. Error propagation in nisq devices for solving classical optimization problems. *PRX Quantum*, 3(04):040326. "NISQ-Computer: Quantum entanglement can be a double-edged sword" via MPQ news.

Yaswitha Gujju, Romain Harang, and Tetsuo Shibuya. 2025. Llm-guided ansätze design for quantum circuit born machines in financial generative modeling.

Aram W Harrow, Avinatan Hassidim, and Seth Lloyd. 2009. Quantum algorithm for linear systems of equations. *Physical Review Letters*, 103(15):150502.

Steffen Herbold. 2024. Semantic similarity prediction is better than other semantic similarity measures.

Zheyong Hu and Sabre Kais. Characterizing quantum circuits with qubit functional configurations. *Scientific Reports*.

Hsin-Yuan Huang, Richard Kueng, and John Preskill. 2021. Power of data in quantum machine learning. *Nature Communications*, 12(1):2631.

Jonas Jäger, Thierry Nicolas Kaldenbach, Max Haas, and Erik Schultheis. 2025. Fast gradient-free optimization of excitations in variational quantum eigensolvers.

Ivan Kankeu, Stefan Gerd Fritsch, Gunnar Schönhoff, Elie Mounzer, Paul Lukowicz, and Maximilian Kiefer-Emmanouilidis. 2025. Quantum-inspired embeddings projection and similarity metrics for representation learning.

D. Kartsaklis, R. Lorenz, K. Meichanetzidis, and B. Coecke. 2021. lambeq: An efficient high-level python library for quantum nlp. In *Proceedings of the 16th Conference on Quantum Physics and Logic (QPL)*.

Kimleang Kea, Won-Du Chang, Hee Chul Park, and Youngsun Han. 2024. Enhancing a convolutional autoencoder with a quantum approximate optimization algorithm for image noise reduction.

Michael Lan, Philip Torr, and Fazl Barez. 2024. Towards interpretable sequence continuation: Analyzing shared circuits in large language models.

Quentin Lhoest, Albert Villanova del Moral, Yacine Jernite, Abhishek Thakur, Patrick von Platen, Suraj Patil, Julien Chaumond, Mariama Drame, Julien Plu, Lewis Tunstall, Joe Davison, Mario Šaško, Gunjan Chhablani, Bhavitvya Malik, Simon Brandeis, Teven Le Scao, Victor Sanh, Canwen Xu, Nicolas Patry, Angelina McMillan-Major, Philipp Schmid, Sylvain Gugger, Clément Delangue, Théo Matussière, Lysandre Debut, Stas Bekman, Pierric Cistac, Thibault Goehringer, Victor Mustar, François Lagunas, Alexander M. Rush, and Thomas Wolf. 2021. Datasets: A community library for natural language processing.

Ang Li, Samuel Stein, Sriram Krishnamoorthy, and James Ang. 2022. Qasmbench: A low-level quantum benchmark suite for nisq evaluation and simulation. *ACM Transactions on Quantum Computing*.

Chen-Yu Liu, Samuel Yen-Chi Chen, Kuan-Cheng Chen, Wei-Jia Huang, and Yen-Jui Chang. 2024. Programming variational quantum circuits with quantum-train agent.

Chen-Yu Liu, Chao-Han Huck Yang, Hsi-Sheng Goan, and Min-Hsiu Hsieh. 2025a. A quantum circuit-based compression perspective for parameter-efficient learning. In *The Thirteenth International Conference on Learning Representations*.

Xiaoming Liu, Yifan Zhao, and Ming Chen. 2025b. Qcse: Pretrained quantum context-sensitive word embeddings. *arXiv preprint arXiv:2509.05729*.

R. Lorenz, D. Kartsaklis, K. Meichanetzidis, and B. Coecke. 2021a. Qnlp experiments with lambeq. In *Proceedings of Quantum Physics and Logic (QPL)*.

Robin Lorenz, Daniel Pearson, and et al. 2021b. Qnlp: Quantum natural language processing. *arXiv preprint arXiv:2102.12846*.

Ning Ma and Heng Li. 2024. Understanding and estimating the execution time of quantum programs.

Jarrod R McClean, Sergio Boixo, Vadim N Smelyanskiy, Ryan Babbush, and Hartmut Neven. 2018. Barren plateaus in quantum neural network training landscapes. *Nature Communications*, 9(1):4812.

Konstantinos Meichanetzidis, Abdelghani Toumi, Giacomo De Felice, Bob Coecke, et al. 2020. Quantum natural language processing on near-term quantum computers. *arXiv preprint arXiv:2005.04147*.

Valter Moretti and Marco Oppio. 2017. Quantum theory in real hilbert space: How the complex hilbert space structure emerges from poincaré symmetry. *Reviews in Mathematical Physics*, 29(06):1750021.

Nidhi Munikote. 2024. Comparing quantum encoding techniques.

Van Hoa Nguyen, Yvon Besanger, Quoc Tuan Tran, Tung Lam Nguyen, Cederic Boudinet, Ron Brandl, Frank Marten, Achilleas Markou, Panos Kotsampopoulos, Arjen A. van der Meer, Effren Guillo-Sansano, Georg Lauss, Thomas I. Strasser, and Kai Heussen. 2017. Real-time simulation and hardware-in-the-loop approaches for integrating renewable energy sources into smart grids: Challenges actions.

Vojtěch Novák, Ivan Zelinka, and Václav Snášel. 2025. Optimization strategies for variational quantum algorithms in noisy landscapes.

Aliakbar Panahi, Seyran Saeedi, and Tom Arodz. 2019. word2ket: Space-efficient word embeddings inspired by quantum entanglement. In *International Conference on Learning Representations 2020*.

John Preskill. 2018. Quantum computing in the nisq era and beyond. *Quantum*, 2:79.

Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. 2019. Language models are unsupervised multitask learners. Technical report, OpenAI.

Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, et al. 2020. Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of Machine Learning Research*, 21.

Maria Schuld and Nathan Killoran. 2019. Quantum machine learning in feature hilbert spaces. *Physical Review Letters*, 122(4):040504.

Maria Schuld, Ilya Sinayskiy, and Francesco Petruccione. 2015. An introduction to quantum machine learning. *Contemporary Physics*, 56(2):172–185.

Maria Schuld, Ryan Sweke, and Johannes Jakob Meyer. 2021. Effect of data encoding on the expressive power of variational quantum-machine-learning models. *Physical Review A*, 103(3).

N. A. Susulovska. 2024. Geometric measure of entanglement of quantum graph states prepared with controlled phase shift operators.

S. M. Yousuf Iqbal Tomal, Abdullah Al Shafin, Debojit Bhattacharjee, MD. Khairul Amin, and Rafiad Sadat Shahir. 2025. Quantum-enhanced attention mechanism in nlp: A hybrid classical-quantum approach.

Charles M. Varmantchaonala, Jean Louis K. E. Fendji, Julius Schöning, and Marcellin Atemkeng. 2024. Quantum natural language processing: A comprehensive survey. *IEEE Access*, 12:99578–99598.

Charles M. Varmantchaonala, Niclas GÖtting, Nils-Erik SchÜtte, Jean Louis E. K. Fendji, and Christopher Gies. 2025. Qcse: A pretrained quantum context-sensitive word embedding for natural language processing.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in Neural Information Processing Systems (NeurIPS)*.

Davide Venturelli, Minh Do, Bryan O'Gorman, Jeremy Frank, Eleanor Rieffel, Kyle E. C. Booth, Thanh Nguyen, Parvathi Narayan, and Sasha Nanda. 2019. Quantum circuit compilation: An emerging application for automated reasoning. In *Scheduling and Planning Applications Workshop*.

Leyang Wang, Yilun Gong, and Zongrui Pei. 2025. Quantum and hybrid machine-learning models for materials-science tasks.

Yichen Wang, Zhe Sun, and Xin Hu. 2024. Multiqnlp: Multimodal structure-aware quantum natural language processing. *arXiv preprint arXiv:2411.04242*.

Nathan Wiebe, Bob Coecke, et al. 2024. Near-term advances in quantum natural language processing. *Annals of Mathematics and Artificial Intelligence*, 92(3-4):501–520.

Yanying Wu and Quanlong Wang. 2019. A categorical compositional distributional modelling for the language of life.

Jian Xu, Feng Li, and Hao Chen. 2025. Quantum graph transformers for sentiment classification. *arXiv preprint arXiv:2506.07937*.

Hao Zhang, Jiaqi Wang, and Qiang Li. 2025. Qtpnet: Quantum text pre-training network for nlp tasks. *arXiv preprint arXiv:2506.00321*.

Li Zhang, Wei Chen, et al. 2024. Quantum algorithms for compositional text processing. *arXiv preprint arXiv:2408.06061*.

A Appendix

Task	Method	Design Highlights	Input Data Type	Label Type	Loss
Sentence Classification	DisCoCat (Coecke et al., 2010)	Maps grammatical reductions to tensor contractions in Hilbert space (compact-closed categories); sentence meaning via categorical compositionality with quantum-ready tensors.	Tokenized sentences	Sentiment / Topic	Cross-entropy
	VQC-QNLP (Gujju et al., 2025)	Parameterized quantum circuit $U(\theta)$ on encoded tokens; hybrid loop minimizes expectation; entanglement captures long-range dependencies under NISQ.	Token embeddings	Binary / Multi-class	Weighted cross-entropy
Semantic Similarity	QBW (Lorenz et al., 2021a)	Quantum Bag-of-Words; embeds words as quantum states; measures similarity via state fidelity/overlaps instead of cosine distance.	Sentence pairs	Similarity / Paraphrase	Fidelity or MSE
	Quantum Kernel (QK-NLP) (Schuld and Killoran, 2019; Wang et al., 2025)	Quantum feature map $\lvert\phi(x)\rangle$ induces kernel $k(x, x') = \lvert\langle\phi(x)\lvert\phi(x')\rangle\rvert^2$; classical SVM/GP on quantum kernel matrix.	Sentences / embeddings	STS / Entailment	Hinge loss / GP NLL
Sequence Labeling	DisCoCirc (Chang et al., 2023)	Discourse-aware extension of DisCoCat; circuit evolution updates word states across context; syntax–semantics via variational updates.	Token sequences	POS / NER / chunks	Token-level cross-entropy
	QCSE (Liu et al., 2025b)	Quantum Context-Sensitive Embeddings: context unitary $U(C)\,\lvert w\rangle$ entangles tokens; contextual vectors in Hilbert space for tagging.	Token sequences	Sequence tags	MSE / cross-entropy
Hybrid Embedding Learning	Hybrid-QNN (Chen et al., 2025)	Classical encoder (e.g., BERT) $\rightarrow$ amplitude/angle map $\rightarrow$ shallow PQC refinement; few-qubit head for NISQ robustness.	Pretrained text embeddings	Sentiment / Intent	Cross-entropy (hybrid)
Low-Resource / Multi-Modal	MultiQ-NLP (Wang et al., 2024)	Entangles text–image qubits; cross-modal attention via controlled rotations; improves transfer in few-shot regimes.	Text–image pairs	Match / Tags	Contrastive (InfoNCE)
Sense Modeling / Pretraining	QTP-Net (Zhang et al., 2025)	Encodes word senses as quantum superpositions $\lvert w\rangle = \sum_i \alpha_i \lvert s_i\rangle$; learns sense mixture via measurement-driven objectives.	Large text corpora	Sense / Masked tokens	NLL; superposition reconstruction
Encoding Learning	Trainable Basic Encoding (Munikote, 2024)	Learnable encoder $E(\phi)$ on basis states prior to PQC; low-depth, NISQ-friendly alternative to fixed angle/amplitude maps.	Token indices	Task-specific	Task loss + encoder reg.
Resource-Efficient Embeddings	word2ket / Tensorized (Panahi et al., 2019)	Factorizes embedding matrix into low-order tensor products; quantum-ready prep with shallow circuits; large parameter compression.	Vocabulary embeddings	Task-specific	Task loss; tensor-factor regs

Table 1: **Summary of representative Quantum Natural Language Processing (QNLP) models across core linguistic tasks.** The table aligns prior work by task, model type, and architectural design to illustrate how quantum principles are applied to language understanding. **Task** denotes the linguistic objective (e.g., classification, similarity, or tagging); **Method** names the quantum or hybrid framework; **Design Highlights** summarize each model's encoding scheme (amplitude, angle, entanglement, or hybrid), circuit structure, and optimization strategy. **Input** and **Label Type** describe the data and prediction targets, while **Loss / Objective** lists the corresponding training criterion. Together, these entries show how QNLP architectures integrate formal semantics with quantum computation, balancing expressivity, resource efficiency, and NISQ-era feasibility.

Encoding Paradigm	Core Idea / Map	Qubits q	State-Prep Cost T_{prep}	Strengths	Limitations
Basic / Learnable Encoding	Token index $i \mapsto \lvert i\rangle$ with shallow trainable unitary $E(\phi)\,\lvert i\rangle$	$\Theta(\log V)$ (index map)	Low (shallow $E(\phi)$)	Very low depth; parameter-efficient; preserves discrete identity; NISQ-friendly	Needs downstream entanglers/PQC for expressivity; tuning still task-dependent
Angle / Rotation Encoding	Map features to single-qubit rotations (e.g., $R_y(\cdot)/R_z(\cdot)$) per dimension; supports data re-uploading	$\Theta(d)$	$\Theta(d)$	Simple, robust, transparent geometry; pairs well with re-uploading in VQCs	Linear qubit growth with d; underuses Hilbert space unless combined with entanglement
Amplitude Encoding	$\mathbf{e} \in \mathbb{R}^d \mapsto \lvert\phi(\mathbf{e})\rangle = \frac{1}{\lVert\mathbf{e}\rVert} \sum_j e_j \lvert j\rangle$ (inner-products preserved)	$\lceil \log_2 d \rceil$	$\Theta(d)$ (state loading)	Exponential compression of d; strong for kernel/similarity tasks; unitary-friendly	Expensive loaders; noise-sensitive; benefits from high-fidelity prep
Entanglement-based Composition	Apply U_{ent} (CNOT/CZ) to correlate token subsystems; syntax/relations via entanglers	Task-dependent	Entanglers dominate	Directly captures compositional/relational structure; aligns with categorical semantics	Increases depth and error on NISQ; careful compilation needed
Hybrid Embedding Strategies	Classical embedding $\mathbf{e}$ (e.g., BERT/Word2Vec) $\rightarrow$ quantum feature map $\lvert\phi(\mathbf{e})\rangle \rightarrow$ PQC $U(\theta)$	Few-qubit heads common	Modest; depends on chosen feature map	Best near-term trade-off; leverages pretrained semantics; smaller q / shots	Classical front-end may dominate compute; quantum benefit is task- and map-dependent
Space-efficient Tensorized (word2ket)	Factorize embedding matrix into low-order tensor products; shallow quantum prep from factors	By factorization design	Low (from tensor factors)	$10^2\times$ compression reported; principled bridge to tensor networks; shallow circuits	Quality depends on factorization rank/structure; extra design choices required
Trainable Quantum Embedding Circuits	Small reusable quantum cell learns token/context encoding in-circuit; reused across positions	Few (cell reused)	Low–moderate (per-cell)	Parameter-efficient; context-aware; fewer qubits/shots than naïve per-token circuits	Requires careful training/stability on NISQ; generalization may be dataset-dependent

Table 2: Encoding paradigms discussed in this survey. V: vocabulary size; d: feature dimension. For fair NISQ comparisons, report q, circuit depth L, state-prep cost T_{prep}, and shot budgets m alongside task metrics.

A Survey of Quantum Natural Language Processing: From Compositional Models to NISQ-Era Empiricism

Arpan Phukan and Asif Ekbal
Department of Computer Science and Engineering, IIT Patna
arpan_2121cs33@iitp.ac.in, asif@iitp.ac.in

Abstract

Quantum Natural Language Processing (QNLP) has emerged as a novel paradigm that leverages the principles of quantum mechanics to address fundamental challenges in language modeling, particularly in capturing compositional meaning. This survey charts the evolution of QNLP, from its theoretical foundations in the Distributional Compositional Categorical (DisCoCat) framework to its modern implementation on Noisy Intermediate-Scale Quantum (NISQ) hardware. We review the primary architectural approaches, including variational quantum circuits and tensor networks, and summarize the growing body of empirical work in tasks such as text classification, sentence similarity, and question answering. A recurring finding is the potential for QNLP models to achieve competitive performance with significantly fewer parameters than their classical counterparts. However, the field is critically constrained by the limitations of NISQ-era hardware. We conclude by discussing these challenges and outlining the future trajectory towards achieving a demonstrable quantum advantage and building more interpretable, efficient language models.

1 Introduction

Quantum Natural Language Processing (QNLP) is an integrative and rapidly developing field that applies the principles of quantum computing to the challenges of natural language processing (Pallavi and Prasanna Kumar, 2025). It is motivated by a foundational hypothesis that extends beyond the simple pursuit of computational speedup: the idea that language is "quantum native" (Widdows et al., 2024). This proposition suggests that the mathematical formalism of quantum mechanics, particularly the compositional structure of Hilbert spaces, provides a natural and perhaps ideal framework for modeling the compositional nature of linguistic meaning. By grounding language in a physical computational model, QNLP seeks a paradigm shift from the purely statistical and often opaque methods of classical NLP to a more structured and interpretable approach (Phukan et al., 2024).

This pursuit is driven by the persistent limitations of classical models. Even state-of-the-art Large Language Models (LLMs) struggle to robustly handle the principle of compositionality, the process by which the meanings of individual words combine according to grammatical rules to form the meaning of a sentence (Song et al., 2025). Many classical architectures effectively treat sentences as a "bag of words" or a flat sequence of tokens, failing to capture the deep, hierarchical relationships encoded in syntax (Chen et al., 2024b). Furthermore, natural language is inherently ambiguous. A phrase such as "The bank was crowded" presents a challenge that classical models resolve through statistical inference (Wu et al., 2021). Quantum mechanics, with its principles of superposition and entanglement, offers a potentially more efficient solution, allowing for the simultaneous representation and processing of multiple meanings within a single quantum state (Schuld and Killoran, 2019; Phukan et al., 2025). This quantum representation can then "collapse" to a definite meaning as more context becomes available, a process that arguably mirrors human cognitive processing of ambiguity (Phukan and Ekbal, 2023).

Finally, the exponential growth in the parameter counts and energy consumption of classical LLMs has created an urgent need for more efficient and scalable learning paradigms (Ji and Jiang, 2026). QNLP presents a potential path toward models that are not only more powerful but also more resource-efficient (Phukan et al., 2024; Phukan and Ekbal, 2023).

This survey provides a comprehensive overview of the QNLP landscape. Section 2 details the foundational Distributional Compositional Categorical

Proceedings of the QuantumNLP: Integrating Quantum Computing with Natural Language Processing, pages 65–75
November 24, 2025

(DisCoCat) (Coecke et al., 2010) framework. Section 3 reviews the primary architectures used for implementing QNLP models on near-term quantum devices. Section 4 summarizes the empirical progress across key NLP tasks. Section 5 offers a critical discussion of the field's current challenges and future outlook.

2 Foundational Framework: Compositionality via DisCoCat

The theoretical cornerstone of modern QNLP is the Distributional Compositional Categorical (DisCoCat) framework (Coecke et al., 2010), which provides a mathematically rigorous unification of two central pillars of linguistic theory: the distributional hypothesis (a word's meaning is defined by its context) and the principle of compositionality (the meaning of a whole is a function of the meaning of its parts and how they are combined).

The DisCoCat model operates through a formal mapping between grammar and meaning. On the grammatical side, it employs a categorial grammar, typically a pregroup grammar, where words are assigned abstract grammatical types. For instance, a noun might be assigned type n, while a transitive verb that takes a noun as its object and a noun as its subject to form a sentence would have the type $n^r s n^l$, where s is the type for a sentence and the superscripts r and l denote right and left adjoints, respectively (Refer Figure 1). A sequence of words is considered grammatical if its sequence of types can be reduced to the sentence type s through a series of predefined rules (Yeung and Kartsaklis, 2021).

On the semantic side, word meanings are represented as vectors (or more generally, tensors) in a high-dimensional Hilbert space, following standard distributional semantics. The central innovation of DisCoCat is the use of category theory to define a structure-preserving function that maps the category of grammar to the category of vector spaces (i.e., semantics). This ensures that the reduction of grammatical types corresponds directly to a specific mathematical operation on the meaning vectors, namely, tensor contraction (Sadrzadeh et al., 2018). This entire compositional process can be visualized and reasoned about using string diagrams, an intuitive graphical calculus where boxes represent word meanings (tensors) and wires represent their grammatical types (tensor indices).

The profound insight that catalyzed the field of QNLP was the observation that the mathematical structure of pregroup grammars and the structure of quantum processes both form a rigid monoidal category. This shared structure allows for a direct and systematic translation from a sentence's grammatical string diagram to a quantum circuit. In this mapping, words become quantum states or operations, and the grammatical rules dictating their composition become entangling gates or measurements (Correia et al., 2022). This correspondence makes quantum computers the "native environment" for executing DisCoCat models. The framework can thus be conceptualized as a "compiler" for language: it takes a high-level linguistic structure (a sentence) as input, parses it according to a formal grammar, and outputs a low-level, executable representation (a quantum circuit) (Peral-García et al., 2024; Laakkonen et al.). Toolkits such as `lambeq`[1] and `DisCoPy`[2] have been developed to automate this compilation pipeline, providing a principled method for generating quantum algorithms for NLP tasks.

3 Architectures for Quantum Language Models

The implementation of QNLP models on present-day hardware has led to a variety of architectural approaches (Refer Table 1). These can be seen as existing on a spectrum, reflecting a fundamental trade-off between adherence to the pure, linguistically-grounded theory of DisCoCat and the pragmatic need to achieve robust performance on noisy, resource-constrained quantum devices.

3.1 Variational Quantum Circuits (VQCs)

Variational Quantum Circuits[3] (VQC) (Qi et al., 2023) are the dominant paradigm for executing machine learning tasks on Noisy Intermediate-Scale Quantum (NISQ) hardware (Phukan et al., 2024). A VQC is a quantum circuit that includes gates with adjustable parameters (e.g., rotation angles). It operates within a hybrid quantum-classical loop: the circuit is executed on a quantum processing unit (QPU), the output is measured to compute a classical loss function, and a classical optimizer updates the circuit parameters to minimize this loss, analogous to training a classical neural network(Bashiri and Naderi, 2024; Hong and Lopez,

[1] https://github.com/CQCL/lambeq
[2] https://discopy.org/
[3] https://pennylane.ai/qml/glossary/variational_circuit

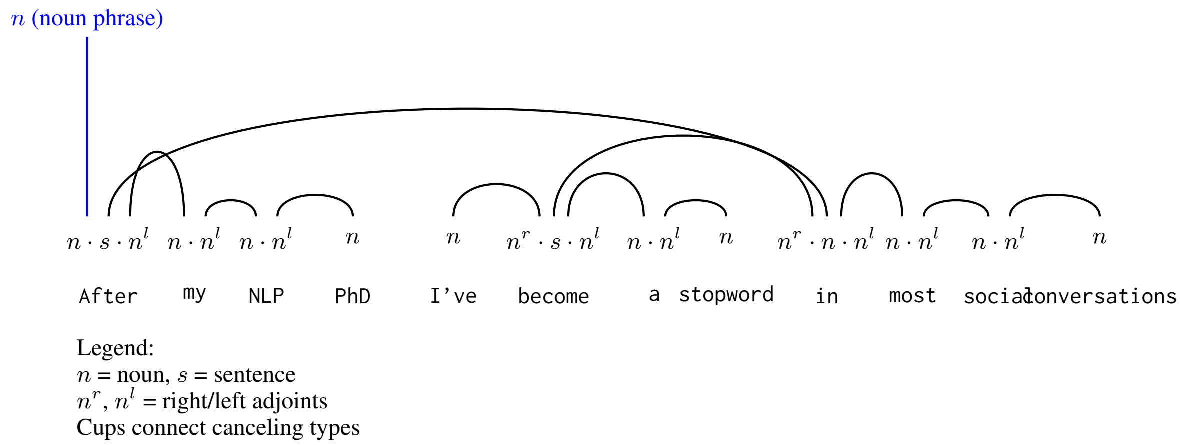


Figure 1: DisCoCat compositional diagram for "After my NLP PhD, I've become a stopword in most social conversations."

2025). Instead of loading pre-computed word vectors, the meanings of words are learned directly as the parameters of the quantum circuits (Zeng and Coecke, 2016). The DisCoCat framework provides the "quantum circuit skeleton" based on a sentence's grammar, and the free parameters within this structure are then optimized end-to-end for a specific downstream task, such as text classification.

3.2 Tensor Network (TN) Representations

Tensor Networks are a set of techniques originating from many-body quantum physics for efficiently representing and manipulating large, high-dimensional tensors (Christandl et al., 2024). This framework is deeply connected to QNLP, as quantum circuits themselves can be formally described as a specific class of tensor network (Rieser et al., 2023; Zhang et al., 2019). TNs, particularly one-dimensional structures like Matrix Product States (MPS), are naturally suited for modeling sequential data like language, as they are designed to efficiently capture local correlations (Berezutskii et al., 2025; Zhang et al., 2019; Teixeira et al.). They serve a dual role in QNLP: as a powerful tool for classically simulating quantum language models and as a class of machine learning models in their own right, offering a structured approach that lies between classical recurrent models and full quantum implementations (Berezutskii et al., 2025).

3.3 Hybrid Quantum-Classical Models

Representing the most pragmatic end of the architectural spectrum, hybrid models seek to enhance proven classical architectures by replacing specific components with quantum counterparts (Pandey et al., 2022, 2023; Phukan and Ekbal, 2023; Pandey and Pakray, 2023; Phukan et al., 2025). This approach aims to leverage quantum effects for computationally challenging subroutines while retaining the overall power and stability of classical frameworks.

A prominent example is the Quantum Transformer, which replaces classical modules like the self-attention mechanism with a VQC-based implementation (Concepcion, 2025; Kerenidis et al., 2024). The goal is to harness quantum properties like entanglement to capture complex contextual relationships between tokens more efficiently than is possible classically (Chen et al., 2024a). A specific instantiation of this idea is the Quantum Self-Attention Neural Network (QSANN) (Li et al., 2024a), which introduces a quantum version of self-attention designed to be scalable and implementable on NISQ devices. Notably, this model bypasses the need for rigid syntactic pre-processing required by pure DisCoCat models, making it more readily applicable to larger, real-world dataset.

These distinct architectural philosophies highlight the field's dynamic search for an optimal balance between theoretical elegance and empirical viability in the NISQ era.

4 Empirical Progress in QNLP Tasks

Despite the constraints of NISQ hardware, a growing body of empirical work has begun to explore the capabilities of QNLP models on a range of standard NLP tasks (Refer Table 2). A consistent theme emerging from these experiments is not necessarily a quantum speedup in terms of wall-clock time, but a significant advantage in terms of parameter and

Approach	Core Principle	Strengths	Limitations/Challenges
Compositional (DisCoCat)	Maps grammatical structure directly to quantum processes via category theory.	Theoretically grounded, highly interpretable, "quantum-native" foundation.	Relies on rigid grammatical parsing, can be brittle; may be inefficient without variational training.
Variational QNLP (VQC-based)	Uses grammar-informed circuit skeletons with parameters learned via a hybrid quantum-classical loop.	Enables training on NISQ hardware, learns word meanings from data, avoids need for QRAM.	Susceptible to barren plateaus, optimization is challenging, performance is limited by hardware noise.
Hybrid (e.g., Quantum Transformer)	Replaces specific components of classical architectures (e.g., attention) with quantum circuits.	Leverages power of established classical models, targets specific computational bottlenecks.	Less theoretically pure, potential for quantum advantage is localized to a single component.

Table 1: Comparison of Various Approaches

data efficiency.

4.1 Text Classification

Text classification is a foundational NLP task that involves assigning predefined labels to text data, with applications ranging from spam detection to topic analysis (Taha et al., 2024; Sun et al., 2023; Li et al., 2024b). It is the most widely explored task in experimental QNLP, serving as a primary benchmark for new models (Peral-García et al., 2024).

4.1.1 Sentiment Classification

Sentiment classification, a task focused on identifying the emotional tone (e.g., positive, negative, neutral) of a text, has been a key testbed for QNLP (Jim et al., 2024). Early proof-of-concept experiments on IBM's quantum computers successfully demonstrated that VQC-based models derived from DisCoCat could be trained to solve simple classification problems (Ganguly et al., 2022). These studies also showed that on carefully constructed datasets, syntax-aware models provided a performance advantage over syntax-agnostic "bag-of-words" baselines, validating the core premise of the compositional approach (Kando et al., 2022).

More recent work has scaled these experiments to real-world data. The QSANN model (Li et al., 2024a), for example, was evaluated on sentiment classification tasks using subsets of the Yelp [4], IMDb (Maas et al., 2011), and Amazon datasets (McAuley and Leskovec, 2013). It was shown to outperform a comparable classical self-attention baseline while using dramatically fewer trainable parameters, for instance, 49 quantum parameters versus 785 classical parameters on the Yelp dataset. Other architectures, such as Quantum Recurrent Neural Networks (QRNNs) (Pandey et al., 2024), have also been proposed for classification, though current implementations remain limited to smaller datasets. While comparative studies acknowledge that classical methods generally still achieve higher accuracy on standard benchmarks, the potential for QNLP models to learn effectively with greater model efficiency is a key finding (Zhang et al., 2024).

4.1.2 Aspect-Based Sentiment Analysis (ABSA)

Aspect-Based Sentiment Analysis (ABSA)[5] is a fine-grained sentiment analysis task that aims to identify the sentiment expressed towards specific aspects or features within a text (Hua et al., 2024). For example, in the sentence "The phone has a great camera but a poor battery life," ABSA would identify a positive sentiment towards the "camera" (aspect) and a negative sentiment towards the "battery life" (aspect) (Zhang et al., 2022). This level of detail is crucial for applications like customer feedback analysis. While ABSA is a significant area of research in classical NLP, its exploration using dedicated QNLP models is still an emerging field (Kayal et al., 2025).

4.2 Sentence Similarity

The task of measuring the semantic similarity between two sentences classically involves encoding them into vector embeddings and calculating a distance metric, such as cosine similarity (Gao et al., 2025). Quantum approaches have explored this task using several methods, such as semantic matching frameworks that leverage the density matrix formalism in Hilbert space (Zhang et al., 2025).

4.2.1 Semantic Textual Similarity

One prominent technique for semantic textual similarity is Quantum Kernel Methods (Schuld et al., 2015; Schuld and Killoran, 2019), where sentences are mapped to quantum states via a quantum feature map. The similarity between two sentences is then calculated as the inner product (or fidelity)

[4] `http://www.yelp.com/dataset_challenge`

[5] `https://www.sciencedirect.com/topics/computer-science/aspect-based-sentiment-analysis`

of their corresponding quantum states, $f(x_i, x_j) = | < \psi(x_i)|\psi(x_j) > |^2$. This value serves as a kernel that can be fed into a classical support vector machine for classification or similarity tasks (Egginger et al., 2024).

Another innovative, quantum-inspired approach involves using a quantum circuit-based architecture (simulated on classical hardware) as a projection head to compress high-dimensional sentence embeddings from a classical model like BERT (Kankeu et al., 2025). The similarity between the resulting compressed representations is measured using a metric based on the fidelity of quantum states. This method achieved performance competitive with classical techniques but with 32 times fewer parameters, and it demonstrated particularly strong performance in low-data settings.

4.2.2 Paraphrase Identification

Paraphrase identification is a specific binary classification task within sentence similarity that determines whether two sentences convey the same meaning, even if they use different wording (Alvi et al., 2021). This task is critical for applications such as plagiarism detection, text summarization, and improving question answering systems (Palivela, 2021). While a core challenge in classical NLP, the development of specialized QNLP models for paraphrase identification is an active area of research.

4.2.3 Information Retrieval (IR)

Information Retrieval is the task of finding relevant documents or information from a large collection in response to a user's query. The connection to sentence similarity is foundational; vector space models, which represent both queries and documents as vectors, measure relevance based on the similarity between these vectors (Sordoni and Nie, 2013). This vector-based paradigm is well-suited for quantum implementation. Mathematical models for language explicitly motivated by quantum theory have been successfully applied to IR (Fan et al., 2024), suggesting that quantum computers could offer a natural and efficient environment for these tasks. This link is formally demonstrated by frameworks that generalize the probability framework of quantum physics for interactive retrieval tasks (Piwowarski and Lalmas, 2009).

4.3 Question Generation (QG)

Question Generation is the task of automatically generating natural language questions from a given input, such as a text passage or a knowledge base (Mulla and Gharpure, 2023). QG has important applications in creating educational materials, augmenting datasets for QA systems, and enhancing conversational agents (Duan et al., 2017; Dey and Rudrapal, 2024). While QNLP has been successfully applied to the dual task of question answering, direct applications to the generative task of QG are still in the early stages of exploration, with current research still focused on addressing core challenges in classical settings (Dey and Rudrapal, 2024), and thus represent a key direction for future research.

4.4 Question Answering (QA)

Question answering remains a nascent but promising application area for QNLP, with research exploring how quantum algorithms can provide advantages like potential speedups (Correia et al., 2022). While classical QA systems involve complex pipelines for information retrieval and answer generation, quantum-inspired techniques are beginning to emerge. One such innovation is the Quantum Fusion Module (QFM) (Duan et al., 2024), designed for open-domain QA. This module applies principles from quantum theory to fuse the token embeddings from a question and a candidate passage into a single representation analogous to a "quantum mixed state." This fused representation allows a classifier to more accurately predict whether the passage contains the answer, thereby improving the performance of a larger, classical T5-based model (Raffel et al., 2020).

4.5 Processing Multimodal Data

A significant new field for QNLP is the processing of multimodal data, which involves integrating information from different formats such as text, images/video, and audio (Li et al., 2021; Phukan and Ekbal, 2023; Phukan et al., 2024, 2025). The MultiQ-NLP framework (Hawashin and Sadrzadeh, 2024) has been developed specifically for this purpose, extending the compositional, structure-aware models of QNLP to handle both text and images. In this framework, both linguistic components and image features are translated into quantum circuits. This allows the model to create a unified representation that captures the interactions between modalities. When tested on an image classification task that requires understanding the relationship between a subject, verb, and object (SVO-Probes), the structure-aware quantum model performed on par with state-of-the-art classical models, demonstrating the viability of QNLP for complex multimodal reasoning. Phukan and Ekbal (2023) introduced QeMMA, a framework for multimodal sentiment analysis. On the CMU-MOSEI dataset (Zadeh et al., 2018), their model, implemented with Qiskit[6] and executed on quantum hardware, outperformed classical baselines by 3.52% in accuracy and 10.14% in F1-score, providing concrete evidence of the empirical advantage of quantum approaches.

Further advancing this line of work, Phukan et al. (2024) proposed a hybrid quantum-classical architecture designed to jointly tackle the interrelated tasks of sarcasm, emotion, and sentiment detection. Their model integrates a Variational Quantum Circuit (VQC) with a classical deep neural network, hypothesizing that quantum entanglement and state space properties can more effectively model the nuanced cross-modal interactions and task correlations present in such complex inference problems.

4.5.1 Emotion Recognition

Emotion recognition is a more granular task than sentiment analysis, aiming to identify specific emotions like joy, anger, or guilt from data (Rasool et al., 2025). While much of the work in this area uses classical NLP models on text, QNLP research has begun to explore this task, often by processing multimodal data. Quantum-enhanced models have been applied to recognize emotions from physiological signals using quantum-enhanced Support Vector Machines (SVMs) (Bayro and Jeong, 2025), and from facial expressions using Quantum Neural Networks (QNNs) (Alsubai et al., 2024). Furthermore, hybrid quantum-classical models have been developed for multimodal data, including speech and text (Li et al., 2025). Such frameworks are increasingly designed to analyze sentiment, emotion, and other related states simultaneously, leveraging quantum properties to model the complex correlations between them (Li et al., 2025).

4.5.2 Sarcasm Detection

Sarcasm detection is a particularly challenging NLP task due to the inherent incongruity between the literal and intended meaning of an utterance (Xi et al., 2025; Lu et al., 2025). Quantum computing has been applied to this problem to better model such complex linguistic phenomena, for instance, through Quantum Neural Networks (QNNs) for analyzing sentiment, emotion, and sarcasm (Singh et al., 2025; Phukan et al., 2024). Other quantum-based models designed directly for this task include the Quantum Fuzzy Neural Network (QFNN) for multimodal sarcasm detection (analyzing text, audio, and visuals) (Tiwari et al., 2024) and frameworks that use the mathematics of quantum probability to jointly model sarcasm, sentiment, and emotion (Singh et al., 2025; Phukan et al., 2025).

[6] https://www.ibm.com/quantum/qiskit

5 Challenges and Future Outlook

The trajectory of QNLP is shaped by a dynamic interplay between the ambitious theoretical promises of quantum computation and the stark realities of current hardware. Progress in the field can be understood as advancing on two parallel but interdependent fronts: a hardware-aware effort to extract value from today's imperfect machines, and a theory-forward exploration of what will be possible with the fault-tolerant computers of the future.

5.1 The NISQ Bottleneck

The primary constraint on QNLP research is the Noisy Intermediate-Scale Quantum (NISQ) era of hardware (Rai et al., 2022). Today's quantum processors are limited by several fundamental factors:

- Low Qubit Counts: Systems with 50-100 qubits are typical, which severely restricts the size and complexity of the language models that can be implemented (Riel, 2021; Pan et al., 2025).

- High Error Rates: Qubits are extremely sensitive to their environment, suffering from decoherence that introduces noise and leads to high error rates (typically 0.1%-1%) in gate operations and measurements (Rai et al., 2022). This noise accumulates rapidly, corrupting the results of all but the simplest computations (Khan et al., 2024).

- Limited Circuit Depth: Because noise accumulates with every operation, there is a strict limit on the number of gates (circuit depth, typically 20-100 layers) that can be applied before the quantum signal is overwhelmed by noise (Pan et al., 2025).

Task	Model/Approach	Dataset(s)	Key Finding / Reported Advantage
Text Classification	DisCoCat + VQC (Ganguly et al., 2022)	Synthetic (Ganguly et al., 2022)	Syntax-aware model outperforms bag-of-words.
Text Classification	QSANN (Li et al., 2024a)	Yelp[4], IMDb (Maas et al., 2011), and Amazon datasets (McAuley and Leskovec, 2013)	Outperforms classical baseline with >90% fewer parameters.
Sentence Similarity	Quantum-Inspired Projection Head (Kankeu et al., 2025)	TREC 2019 DL (Voorhees et al., 2020) and TREC 2020 DL (Craswell et al., 2021)	Competitive performance with 32x fewer parameters; excels in low-data regimes.
Question Answering	Quantum-Inspired Fusion-in-Decoder (QFiD) (Duan et al., 2024)	Natural Questions (Kwiatkowski et al., 2019) and TriviaQA (Joshi et al., 2017)	Quantum-inspired fusion improves relationship prediction over baseline.
Multimodal Classification	MultiQ-NLP (Hawashin and Sadrzadeh, 2024)	SVO-Probes (Hendricks and Nematzadeh, 2021)	Performs on par with classical models while offering better interpretability.
Multimodal Classification	HQNN (VQC+NN) (Phukan et al., 2024)	Extended MUStARD(Chauhan et al., 2020)	Outperforms classical baselines for multitask sarcasm, sentiment, and emotion detection.

Table 2: QNLP Tasks

These hardware limitations collectively mean that current QNLP experiments are necessarily confined to small-scale datasets and simplified "toy" problems. Scaling these models to handle the complexity and volume of data processed by modern LLMs is, for now, impossible.

5.2 Algorithmic Hurdles

Beyond hardware, significant algorithmic challenges must be overcome. A primary obstacle in training VQCs is the phenomenon of barren plateaus, where the gradient of the loss function vanishes exponentially as the number of qubits increases, rendering the optimization process intractable for larger models (Pande, 2023). Furthermore, the task of data encoding, efficiently translating classical text data into quantum states, is a non-trivial problem that can itself become a computational bottleneck (Ranga et al., 2024). Active research on the hardware-aware front is focused on developing mitigation strategies, including noise-aware training protocols (Rahman and Zhuang, 2025), quantum error mitigation techniques (Atban et al., 2025), and the design of specialized loss functions and optimizers (Pande, 2023).

6 Conclusion

The long-term vision for QNLP is to achieve a demonstrable quantum advantage, where a quantum computer solves a practical NLP problem more efficiently, more accurately, or with fewer resources than is possible with any classical machine. While theoretical results have shown that quantum language models are, in principle, more expressive than their classical counterparts (i.e., the problem is BQP-complete), a practical advantage has not yet been realized.

Perhaps one of the most compelling promises of the theory-forward front is the potential for QNLP to deliver more interpretable and trustworthy AI. By building models on the transparent, compositional structure of DisCoCat, QNLP offers a path away from the "black box" nature of many LLMs, toward systems whose reasoning can be traced and understood.

Ultimately, the future of QNLP depends on the co-evolution of hardware and theory. As quantum technology matures beyond the NISQ era and toward fault-tolerance, the full potential of these linguistically-grounded, quantum-native models can be explored. The goal is not merely to replicate classical NLP on a different substrate, but to fundamentally reimagine language modeling by leveraging the unique computational capabilities of the quantum world, potentially leading to a new generation of AI that is more powerful, efficient, and reliable.

7 Limitations

This survey provides a broad overview of the QNLP landscape, but it is subject to several inherent limitations reflecting the nascent and rapidly evolving nature of the field.

First, the scope of this review is primarily centered on the theoretical lineage from the DisCoCat framework to its empirical implementation on NISQ-era hardware. While we touch upon quantum-inspired classical models and tensor networks, an exhaustive analysis of all classical algorithms that leverage principles from quantum theory is beyond the scope of this paper.

Second, the field of QNLP is advancing at an exceptional pace. This survey represents a snapshot of the research landscape at the time of writing. New hardware developments, algorithmic breakthroughs, and empirical findings are published frequently, and some emerging work may not be captured here.

References

Shtwai Alsubai, Abdullah Alqahtani, Abed Alanazi, Mohemmed Sha, and Abdu Gumaei. 2024. Facial emotion recognition using deep quantum and advanced transfer learning mechanism. *Frontiers in Computational Neuroscience*, 18:1435956.

Faisal Alvi, Mark Stevenson, and Paul Clough. 2021. Paraphrase type identification for plagiarism detection using contexts and word embeddings. *International Journal of Educational Technology in Higher Education*, 18(1):42.

Furkan Atban, Muhammed Yusuf Küçükkara, and Cüneyt Bayılmış. 2025. Enhancing variational quantum classifier performance with meta-heuristic feature selection for credit card fraud detection. *The European Physical Journal Special Topics*, pages 1–14.

Hadis Bashiri and Hassan Naderi. 2024. Comprehensive review and comparative analysis of transformer models in sentiment analysis. *Knowledge and Information Systems*, 66(12):7305–7361.

Allison Bayro and Heejin Jeong. 2025. Enhancing emotional response detection in virtual reality with quantum support vector machine learning. *Computers & Graphics*, 128:104196.

Aleksandr Berezutskii, Minzhao Liu, Atithi Acharya, Roman Ellerbrock, Johnnie Gray, Reza Haghshenas, Zichang He, Abid Khan, Viacheslav Kuzmin, Dmitry Lyakh, and 1 others. 2025. Tensor networks for quantum computing. *Nature Reviews Physics*, pages 1–13.

Dushyant Singh Chauhan, SR Dhanush, Asif Ekbal, and Pushpak Bhattacharyya. 2020. Sentiment and emotion help sarcasm? a multi-task learning framework for multi-modal sarcasm, sentiment and emotion analysis. In *Proceedings of the 58th annual meeting of the association for computational linguistics*, pages 4351–4360.

I-Chi Chen, Harshdeep Singh, VL Anukruti, Brian Quanz, and Kavitha Yogaraj. 2024a. A survey of classical and quantum sequence models. In *2024 16th International Conference on COMmunication Systems & NETworkS (COMSNETS)*, pages 1006–1011. IEEE.

Jiaao Chen, Xiaoman Pan, Dian Yu, Kaiqiang Song, Xiaoyang Wang, Dong Yu, and Jianshu Chen. 2024b. Skills-in-context: Unlocking compositionality in large language models. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 13838–13890.

Matthias Christandl, Vladimir Lysikov, Vincent Steffan, Albert H Werner, and Freek Witteveen. 2024. The resource theory of tensor networks. *Quantum*, 8:1560.

Bob Coecke, Mehrnoosh Sadrzadeh, and Stephen Clark. 2010. Mathematical foundations for a compositional distributional model of meaning. *arXiv preprint arXiv:1003.4394*.

Wesley Concepcion. 2025. Hybrid quantum transformer for natural language processing.

Adriana D Correia, Michael Moortgat, and Henk TC Stoof. 2022. Quantum computations for disambiguation and question answering. *Quantum Information Processing*, 21(4):126.

Nick Craswell, Bhaskar Mitra, Emine Yilmaz, and Daniel Campos. 2021. Overview of the TREC 2020 deep learning track. *CoRR*, abs/2102.07662.

Debopam Dey and Dwijen Rudrapal. 2024. An analysis of automatic question generation research progress and challenges. In *International Conference on Data Science and Network Engineering*, pages 247–258. Springer.

Nan Duan, Duyu Tang, Peng Chen, and Ming Zhou. 2017. Question generation for question answering. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 866–874, Copenhagen, Denmark. Association for Computational Linguistics.

Ruixue Duan, Xin Liu, Zhigang Ding, and Yangsen Zhang. 2024. Quantum-inspired fusion for open-domain question answering. *Electronics*, 13(20):4135.

Sebastian Egginger, Alona Sakhnenko, and Jeanette Miriam Lorenz. 2024. A hyperparameter study for quantum kernel methods. *Quantum Machine Intelligence*, 6(2):44.

Zipeng Fan, Jing Zhang, Peng Zhang, Qianxi Lin, Yizhe Li, and Yuhua Qian. 2024. Quantum-inspired language models based on unitary transformation. *Information Processing & Management*, 61(4):103741.

Srinjoy Ganguly, Sai Nandan Morapakula, and Luis Miguel Pozo Coronado. 2022. Quantum natural language processing based sentiment analysis using lambeq toolkit. In *2022 Second International Conference on Power, Control and Computing Technologies (ICPC2T)*, pages 1–6. IEEE.

Hui Gao, Peng Zhang, Jing Zhang, and Chang Yang. 2025. Qsim: a quantum-inspired hierarchical semantic interaction model for text classification. *Neurocomputing*, 611:128658.

Hala Hawashin and Mehrnoosh Sadrzadeh. 2024. Multimodal structure-aware quantum data processing. *arXiv preprint arXiv:2411.04242*.

Lisa Anne Hendricks and Aida Nematzadeh. 2021. Probing image-language transformers for verb understanding. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 3635–3644.

Ying-Yi Hong and Dylan Josh Domingo Lopez. 2025. A review on quantum machine learning in applied systems and engineering. *IEEE Access*.

Yan Cathy Hua, Paul Denny, Jörg Wicker, and Katerina Taskova. 2024. A systematic review of aspect-based sentiment analysis: domains, methods, and trends. *Artificial Intelligence Review*, 57(11):296.

Zhenya Ji and Ming Jiang. 2026. A systematic review of electricity demand for large language models: evaluations, challenges, and solutions. *Renewable and Sustainable Energy Reviews*, 225:116159.

Jamin Rahman Jim, Md Apon Riaz Talukder, Partha Malakar, Md Mohsin Kabir, Kamruddin Nur, and Mohammed Firoz Mridha. 2024. Recent advancements and challenges of nlp-based sentiment analysis: A state-of-the-art review. *Natural Language Processing Journal*, 6:100059.

Mandar Joshi, Eunsol Choi, Daniel Weld, and Luke Zettlemoyer. 2017. TriviaQA: A large scale distantly supervised challenge dataset for reading comprehension. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1601–1611, Vancouver, Canada. Association for Computational Linguistics.

Shunsuke Kando, Hiroshi Noji, and Yusuke Miyao. 2022. Multilingual syntax-aware language modeling through dependency tree conversion. In *Proceedings of the Sixth Workshop on Structured Prediction for NLP*, pages 1–10, Dublin, Ireland. Association for Computational Linguistics.

Ivan Kankeu, Stefan Gerd Fritsch, Gunnar Schönhoff, Elie Mounzer, Paul Lukowicz, and Maximilian Kiefer-Emmanouilidis. 2025. Quantum-inspired embeddings projection and similarity metrics for representation learning. *arXiv preprint arXiv:2501.04591*.

Mrinmoy Kayal, Mohinikanta Sahoo, Jayadeep Pati, and Ranjan Kumar Behera. 2025. Quantum-inspired aspect-based sentiment analysis using natural language processing. In *Advances in Quantum Inspired Artificial Intelligence: Techniques and Applications*, pages 151–169. Springer.

Iordanis Kerenidis, Natansh Mathur, Jonas Landman, Martin Strahm, Yun Yvonna Li, and 1 others. 2024. Quantum vision transformers. *Quantum*, 8:1265.

Misha Urooj Khan, Muhammad Ahmad Kamran, Wajiha Rahim Khan, Malik Muhammad Ibrahim, Muhammad Umair Ali, and Seung Won Lee. 2024. Error mitigation in the nisq era: Applying measurement error mitigation techniques to enhance quantum circuit performance. *Mathematics*, 12(14):2235.

Tom Kwiatkowski, Jennimaria Palomaki, Olivia Redfield, Michael Collins, Ankur Parikh, Chris Alberti, Danielle Epstein, Illia Polosukhin, Jacob Devlin, Kenton Lee, and 1 others. 2019. Natural questions: a benchmark for question answering research. *Transactions of the Association for Computational Linguistics*, 7:453–466.

Tuomas Laakkonen, Konstantinos Meichanetzidis, and Bob Coecke. Quantum algorithms for compositional text processing. *EPTCS 406*, page 162.

Guangxi Li, Xuanqiang Zhao, and Xin Wang. 2024a. Quantum self-attention neural networks for text classification. *Science China Information Sciences*, 67(4):142501.

Qiuchi Li, Dimitris Gkoumas, Christina Lioma, and Massimo Melucci. 2021. Quantum-inspired multimodal fusion for video sentiment analysis. *Information Fusion*, 65:58–71.

Xintong Li, Jinya Jiang, Ria Dharmani, Jayanth Srinivasa, Gaowen Liu, and Jingbo Shang. 2024b. Open-world multi-label text classification with extremely weak supervision. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 15084–15096, Miami, Florida, USA. Association for Computational Linguistics.

YaoChong Li, Yi Qu, Ri-Gui Zhou, and Jing Zhang. 2025. Qmlsc: A quantum multimodal learning model for sentiment classification. *Information Fusion*, 120:103049.

Ming Lu, Zhiqiang Dong, Ziming Guo, Xiaoming Zhang, Xinxi Lu, Tianbo Wang, and Litian Zhang. 2025. A multi-modal sarcasm detection model based on cue learning. *Scientific Reports*, 15(1):10261.

Andrew L. Maas, Raymond E. Daly, Peter T. Pham, Dan Huang, Andrew Y. Ng, and Christopher Potts. 2011. Learning word vectors for sentiment analysis. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 142–150, Portland, Oregon, USA. Association for Computational Linguistics.

Julian McAuley and Jure Leskovec. 2013. Hidden factors and hidden topics: understanding rating dimensions with review text. In *Proceedings of the 7th ACM conference on Recommender systems*, pages 165–172.

Nikahat Mulla and Prachi Gharpure. 2023. Automatic question generation: a review of methodologies, datasets, evaluation metrics, and applications. *Progress in Artificial Intelligence*, 12(1):1–32.

Hemant Palivela. 2021. Optimization of paraphrase generation and identification using language models in natural language processing. *International Journal of Information Management Data Insights*, 1(2):100025.

Gundala Pallavi and Rangarajan Prasanna Kumar. 2025. Quantum natural language processing and its applications in bioinformatics: a comprehensive review of methodologies, concepts, and future directions. *Frontiers in Computer Science*, 7:1464122.

Yi Pan, Hanqi Jiang, Junhao Chen, Yiwei Li, Huaqin Zhao, Lin Zhao, Yohannes Abate, Yingfeng Wang, and Tianming Liu. 2025. Bridging classical and quantum computing for next-generation language models. *arXiv preprint arXiv:2508.07026*.

Mandaar B Pande. 2023. Review of optimization techniques and barren plateaus in training of quantum machine learning problems. In *International Conference on Computing and Network Communications*, pages 3–15. Springer.

Shyambabu Pandey, Nihar Jyoti Basisth, Tushar Sachan, Neha Kumari, and Partha Pakray. 2023. Quantum machine learning for natural language processing application. *Physica A: Statistical Mechanics and its Applications*, 627:129123.

Shyambabu Pandey, Pankaj Dadure, Morrel VL Nunsanga, and Partha Pakray. 2022. Parts of speech tagging towards classical to quantum computing. In *2022 IEEE Silchar Subsection Conference (SILCON)*, pages 1–6. IEEE.

Shyambabu Pandey and Partha Pakray. 2023. Bi-quantum long short-term memory for part-of-speech tagging. In *Proceedings of the 20th International Conference on Natural Language Processing (ICON)*, pages 301–307.

Shyambabu Pandey, Partha Pakray, and Riyanka Manna. 2024. Quantum classifier for natural language processing applications. *Computación y Sistemas*, 28(2):695–700.

David Peral-García, Juan Cruz-Benito, and Francisco José García-Peñalvo. 2024. Comparing natural language processing and quantum natural processing approaches in text classification tasks. *Expert Systems with Applications*, 254:124427.

Arpan Phukan and Asif Ekbal. 2023. Qemma: Quantum-enhanced multi-modal sentiment analysis. In *Proceedings of the 20th International Conference on Natural Language Processing (ICON)*, pages 815–821.

Arpan Phukan, Anas Anwarul Haq Khan, and Asif Ekbal. 2025. Qumin: quantum multi-modal data fusion for humor detection. *Multimedia Tools and Applications*, 84(18):18855–18872.

Arpan Phukan, Santanu Pal, and Asif Ekbal. 2024. Hybrid quantum-classical neural network for multimodal multitask sarcasm, emotion, and sentiment analysis. *IEEE Transactions on Computational Social Systems*, 11(5):5740–5750.

Benjamin Piwowarski and Mounia Lalmas. 2009. A quantum-based model for interactive information retrieval. In *Conference on the Theory of Information Retrieval*, pages 224–231. Springer.

Jun Qi, Chao-Han Yang, and Pin-Yu Chen. 2023. Qtn-vqc: An end-to-end learning framework for quantum neural networks. *Physica Scripta*, 99(1):015111.

Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. 2020. Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of Machine Learning Research*, 21(140):1–67.

Maqsudur Rahman and Jun Zhuang. 2025. Nqnn: Noise-aware quantum neural networks for medical image classification. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 433–442. Springer.

Anshuman Rai, Amey Talekar, Javed Khan, and Ahlam Ansari. 2022. A review article on quantum natural language processing. *Int J Res Appl Sci Eng Technol*, 10(1).

Deepak Ranga, Aryan Rana, Sunil Prajapat, Pankaj Kumar, Kranti Kumar, and Athanasios V Vasilakos. 2024. Quantum machine learning: Exploring the role of data encoding techniques, challenges, and future directions. *Mathematics*, 12(21):3318.

Abdur Rasool, Saba Aslam, Naeem Hussain, Sharjeel Imtiaz, and Waqar Riaz. 2025. nbert: Harnessing nlp for emotion recognition in psychotherapy to transform mental health care. *Information*, 16(4):301.

Heike Riel. 2021. Quantum computing technology. In *2021 IEEE international Electron devices meeting (IEDM)*, pages 1–3. IEEE.

Hans-Martin Rieser, Frank Köster, and Arne Peter Raulf. 2023. Tensor networks for quantum machine learning. *Proceedings of the Royal Society A*, 479(2275):20230218.

Mehrnoosh Sadrzadeh, Dimitri Kartsaklis, and Esma Balkır. 2018. Sentence entailment in compositional distributional semantics. *Annals of Mathematics and Artificial Intelligence*, 82(4):189–218.

Maria Schuld and Nathan Killoran. 2019. Quantum machine learning in feature hilbert spaces. *Physical review letters*, 122(4):040504.

Maria Schuld, Ilya Sinayskiy, and Francesco Petruccione. 2015. An introduction to quantum machine learning. *Contemporary Physics*, 56(2):172–185.

Jaiteg Singh, Kamalpreet Singh Bhangu, Abdulrhman Alkhanifer, Ahmad Ali Alzubi, and Farman Ali. 2025. Quantum neural networks for multimodal sentiment, emotion, and sarcasm analysis. *Alexandria Engineering Journal*, 124:170–187.

Jiajun Song, Zhuoyan Xu, and Yiqiao Zhong. 2025. Out-of-distribution generalization via composition: a lens through induction heads in transformers. *Proceedings of the National Academy of Sciences*, 122(6):e2417182122.

Alessandro Sordoni and Jian-Yun Nie. 2013. Looking at vector space and language models for ir using density matrices. In *International Symposium on Quantum Interaction*, pages 147–159. Springer.

Xiaofei Sun, Xiaoya Li, Jiwei Li, Fei Wu, Shangwei Guo, Tianwei Zhang, and Guoyin Wang. 2023. Text classification via large language models. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 8990–9005, Singapore. Association for Computational Linguistics.

Kamal Taha, Paul D Yoo, Chan Yeun, Dirar Homouz, and Aya Taha. 2024. A comprehensive survey of text classification techniques and their research applications: Observational and experimental insights. *Computer Science Review*, 54:100664.

Erico Souza Teixeira, Yara Rodrigues Inácio, and Pamela TL Bezerra. Applying quantum tensor networks in machine learning: A systematic.

Prayag Tiwari, Lailei Zhang, Zhiguo Qu, and Ghulam Muhammad. 2024. Quantum fuzzy neural network for multimodal sentiment and sarcasm detection. *Information Fusion*, 103:102085.

Ellen M Voorhees, Nick Craswell, Bhaskar Mitra, Daniel Campos, and Emine Yilmaz. 2020. Overview of the trec 2019 deep learning track.

Dominic Widdows, Willie Aboumrad, Dohun Kim, Sayonee Ray, and Jonathan Mei. 2024. Quantum natural language processing. *KI-Künstliche Intelligenz*, 38(4):293–310.

Sixuan Wu, Jian Li, Peng Zhang, and Yue Zhang. 2021. Natural language processing meets quantum physics: A survey and categorization. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 3172–3182.

Zhonghao Xi, Bengong Yu, and Haoyu Wang. 2025. Multimodal sarcasm detection based on sentiment-clue inconsistency global detection fusion network. *Expert Systems with Applications*, 275:127020.

Richie Yeung and Dimitri Kartsaklis. 2021. A CCG-based version of the DisCoCat framework. In *Proceedings of the 2021 Workshop on Semantic Spaces at the Intersection of NLP, Physics, and Cognitive Science (SemSpace)*, pages 20–31, Groningen, The Netherlands. Association for Computational Linguistics.

AmirAli Bagher Zadeh, Paul Pu Liang, Soujanya Poria, Erik Cambria, and Louis-Philippe Morency. 2018. Multimodal language analysis in the wild: Cmu-mosei dataset and interpretable dynamic fusion graph. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2236–2246.

William Zeng and Bob Coecke. 2016. Quantum algorithms for compositional natural language processing. *EPTCS*, 221(arXiv: 1608.01406):67–75.

Chenchen Zhang, Qiuchi Li, Dawei Song, and Prayag Tiwari. 2025. Quantum-inspired semantic matching based on neural networks with the duality of density matrices. *Engineering Applications of Artificial Intelligence*, 140:109667.

Hui Zhang, Qinglin Zhao, and Chuangtao Chen. 2024. A light-weight quantum self-attention model for classical data classification. *Applied Intelligence*, 54(4):3077–3091.

Lipeng Zhang, Peng Zhang, Xindian Ma, Shuqin Gu, Zhan Su, and Dawei Song. 2019. A generalized language model in tensor space. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 7450–7458.

Wenxuan Zhang, Xin Li, Yang Deng, Lidong Bing, and Wai Lam. 2022. A survey on aspect-based sentiment analysis: Tasks, methods, and challenges. *IEEE Transactions on Knowledge and Data Engineering*, 35(11):11019–11038.

Author Index