Journal on
Artificial Intelligence

Tech Science Press

https://doi.org/10.32604/jai.2026.074171

ARTICLE

Mathematical Framework for Detecting Sentiments Analysis Using Quantum Computing

Anitya Kumar Gupta*, **Pankaj Vaidya and Anurag Rana**

Yogananda School of AI Computers and Data Science, Solan, India

*Corresponding Author: Anitya Kumar Gupta. Email: anityagupta@shooliniuniversity.com

Received: 04 October 2025; Accepted: 28 January 2026; Published: 22 July 2026

ABSTRACT: Sentiment analysis aims at determining the stance or point of view of a topic, author, or speaker about a certain topic, document, or event. The given paper proposes a hybrid quantum-classical model of sentiment classification, which is known as Complex-Valued Quantum-Enhanced Recurrent Neural Network (CQRNN). The model combines Quantum Long Short-Term Memory (QLSTM) and Quantum Gated Recurrent Units (QGRU) with the use of Variational Quantum Circuits (VQCs) and complex-valued embeddings designed at the level of semantic information, both in the real and imaginary domains. Experiments on benchmark sentiment datasets show that CQRNN can be fine-tuned and reach an accuracy of 85% on noisy and ambiguous data, and this is a demonstration of its strength and better ability to discriminate between similar sentiment expressions like neutral and negative expressions. The performance of quantum-inspired architectures in sentiment classification compared with the traditional models, including SVM and Random Forest, establishes the potential. The article prepares the basis of benchmarking CQRNN to the transformer-based models (BERT, RoBERTa) and implementing it on actual quantum hardware.

KEYWORDS: Sentiment analysis; quantum computing; complex-valued embeddings; CQRNN; QLSTM; NLP

1 Introduction

Sentiment analysis aims to determine the view or opinion of a subject, author, or speaker towards a given topic, document, or event [1]. Sentiment analysis can include judgment, evaluation, emotional responses, or emotional expression. Text sentiment analysis research spans broad areas of science and technology, such as text mining and natural language processing (NLP) [2,3]. Sentiment analysis finds application in numerous domains such as healthcare, marketing, sociology, and finance [4].

The most common machine learning, deep learning, and sentiment lexicons have been applied in sentiment analysis nowadays [5]. Under the lexicon-based methods, a sentiment lexicon is generated initially, and additional calculations are performed with the help of the lexicon. In the present state of NLP, lexicon-based methods do not enjoy such popularity because of the scaling problem. Machine learning techniques are more likely to involve hand feature usage [6], and then the sentiment classification step, in the form of the naive Bayes (NB), maximum entropy (ME), k-nearest neighbor classification, and conditional random fields (CRF). Deep learning methodology is better in semantic expression and does not involve the process of manually choosing features. The general types of neural networks applied in sentiment analysis are convolutional neural networks (CNNs), recurrent neural networks (RNNs), long short-term memories (LSTMs), and gated recurrent units (GRUs) [7,8]. In recent years, quantum theory has found application to construct text representations of a large variety of information retrieval (IR) and NLP tasks [9–11].

Quantum computing has a mathematical model to replicate the complicated interactions in quantum physics, and it is in use in NLP. A quantum language model (QLM) is an algorithm where a query or a bit of text can be modeled as a density matrix in the quantum probability spaces, and the functions obtained from the density matrices are used as ranking functions [12]. Neural network quantum language model (NNQLM) is a neural network-based model that forms an end-to-end question-to-answer network using a density matrix framework and question-answer pair encodings [13]. The application of Quantum Long Short-Term Memory (QLSTM) and Quantum Gated Recurrent Unit (QGRU) models based on Variational Quantum Circuits (VQCs) in text sentiment analysis is discussed in this paper. The model takes into account complex-valued embeddings and thus all words to reflect the state that is observed of all words. It is possible to recreate semantic and sentiment polarity using the proposed model of CQRNN, which is very effective in the sentiment task. Compared to other neural network models, the methodology of CQRNN also receives the complex characteristics of semantics, which enhances the accuracy of predictions of the model and its effectiveness in general. Though in some studies Quantum LSTM (QLSTM) and Quantum GRU (QGRU) architectures have been studied, the majority of the current literature either utilizes traditional real-valued embeddings or exclusively researches optimization of the circuit without considering semantic ambiguity in language. The originality of the suggested CQRNN is in its combination of the use of complex-valued embeddings and parametrized quantum circuits to characterize the real and imaginary semantic elements of text. In contrast to the previous models, CQRNN uses a universal quantum recurrent model that operates on long-term dependencies and minute emotional meanings in the noisy data. This modeling is a clear connection between quantum probability theory and linguistic sentiment representation, which puts CQRNN in a unique contribution to traditional QLSTM/QGRU models. This improvement makes the model more realistic and credible.

The research objectives are the following:

- To propose and work out the Complex-valued Quantum-enhanced Recurrent Neural Network (CQRNN) to categorize sentiment.
- To explore quantum theory applications in NLP with the complex-valued embeddings and quantum probability spaces.
- To benchmark the QLSM and QGRU models against the traditional and quantum-inspired deep learning models of BERT, RoBERTa, and Complex-QCNN on the standard sentiment classification datasets.
- To compare CQRNN with the models of deep learning, CNN, RNN, BERT, and RoBERTa in terms of their accuracy, recall, and F1-score.

The QLSTM and QGRU are also suggested in this paper, in addition to VQCs, in order to advance sentiment classification. Seemingly complex-valued embeddings in the new CQRNN model will help get a better semantic understanding and sentiment recognition, thus it will perform more effectively in comparison with the traditional deep learning models.

2 Literature Review

The recent developments in quantum computing have impacted sentiment analysis insanely, and quantum-inspired algorithms have been combined with the deep learning architecture. Balaji and Vadivazhagan (2024) [12] proposed a kind of nature-oriented optimization model, the Resilient Grey Wolf Optimization-based Gaussian-Enhanced Quantum Deep Neural Network (RGWO-GEQDNN) and it also includes quantum computing to boost sentiment analysis. RGWO-GEQDNN that was used on the Amazon product review dataset yielded a superior FMI (96.442%), and MCC (96.431%). RGWO-GEQDNN technique entails optimization of the learning process based on the primary mode of optimization through grey wolves so as to obtain efficient convergence, and quantum enhancements that could be of service to extract features more efficiently. Likewise, Shi et al. (2024) [13] proposed quantum-inspired pre-trained feature embedding (QPFE) algorithm which enhances the performance of sentiment classification by exploiting the quantum superposition laws. The quantum superposition in this situation, represents the idea that quantum bits (qubits) could be prepared in a variety of states at the same time, leading to a higher data representation rate. Their QPFE-ERNIE model increased F1 score by 8.7 points relative to quantum based models that were developed before. The task of multilingual sentiment analysis is also investigated with the help of quantum computing. However, Zhao et al. (2024) [14] recently suggested a quantum neural model, QPEN, that improves cross-lingual aspect-based sentiment analysis (ABSA). QPEN utilizes a quantum entanglement module, in which language-specific features across languages can be shared in a way that maintains each language's uniqueness of features. This method obtained the state-of-the-art F1 scores on five languages in SemEval-2016. In the meantime, Ruskanda et al. (2023) [15] employed VQA for sentiment analysis, introducing the Simple Sentiment Analysis (Simple-SA) ansatz. Simple-SA ansatz is a technique of optimizing quantum circuits towards sentiment analysis, whereas the IQP ansatz is another quantum technique founded on instantaneous quantum polytime circuits. Their method outperformed the IQP ansatz, achieving an accuracy of 85% and decreasing the optimization convergence time by 20.89%.

Comparative analyses of classical machine learning and quantum methods show the capability of quantum computing in sentiment analysis. Omar and Abd El-Hafeez (2023) [16] compared sentiment classification for Arabic tweets, where quantum computing performed marginally better than classical machine learning on one set but marginally less accurately on another. Their work indicates that quantum methods can provide competitive accuracy while minimizing processing time. On a similar note, Cyril et al. (2021) [17] also suggested the usage of the CA-SVM-based deep learning model in Twitter sentiment analysis with an accuracy of 92.48 but adds the possibility of hybrid quantum-classical.

The Fig. 1 illustrates a *quantum-enhanced Long Short-Term Memory (LSTM) architecture* for sentiment and emotion analysis, where information is processed at both the word level and sentence level using complex-valued representations. At the word level, each token is first transformed into a complex embedding, capturing both amplitude and phase information to represent nuanced emotional features. These embeddings are then mapped into higher-level representations (h_i) through complex linear transformations. At the sentence level, sequential dependencies are modeled using a quantum-enhanced recurrent structure, where intermediate states (ρ_i) evolve through the network and contribute to probability outputs (p_i) via measurement operations. This framework enables the model to capture contextual dependencies, interference effects, and semantic correlations, leading to improved representation of complex and mixed emotional states compared to classical architectures.

In addition to this, quantum-enhanced sentiment analysis generalizes text processing to multimodal. The other field, which quantum computing improves in multimodal applications that require text, audio, and visual data, is sentiment analysis. Similarly, LSTM model using a Modified Ant Lion Optimizer (MALO) managed to employ both textual, audio and visual features to reach a high accuracy of above 98% on YouTube and SAVEE datasets (Kothuri and Rajalakshmi, 2022) [18] Sathya and Sudha (2023) [19] achieved an accuracy of 98% in detecting facial expressions using quantum convolutional neural networks (QCNN) optimized with an improved particle swarm optimization bat swarm algorithm (IPSOBSA) the same way. All of these show that the solutions of quantum computing can be applied to better sentiment analysis of additional types of data not of their nature but of their complexity as well. The current quantum-enhanced NLP models like Quantum LSTM (QLSTM), Quantum GRU (QGRU) and Quantum-Inspired Deep Neural Networks have demonstrated promising yet minor scalability in case of sentiment analysis. The majority of approaches focus on circuit efficiency and not on semantic expressiveness and very few are actually modeling ambiguous or context-dependent sentiments. Conversely, the proposed CQRNN adds a complex-valued representation layer which brings a new layer to the semantic space with principles of quantum entanglements being translated into the recurrent gates. Such a hybridization allows more refined feelings to be discriminated and offers a generalizable framework in understanding text with quantum. This model was formulated mathematically and architecturally as described in the following section.

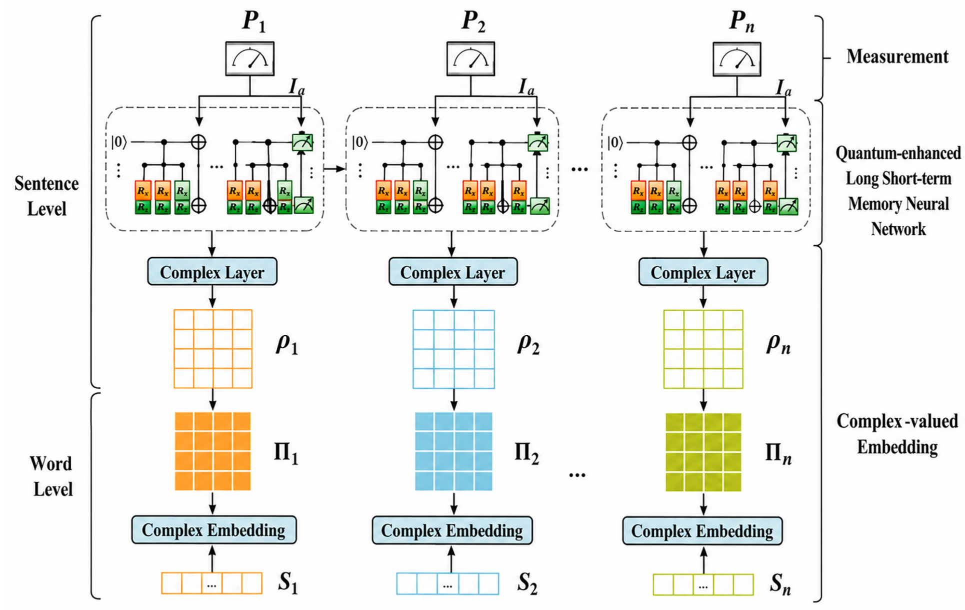

Figure 1: General quantum circuit for sentiment analysis task.

3 Methods and Material

The models used in the presented research are quantum-enhanced long short-term memory (QLSTM) and gated recurrent unit (QGRU) models complex-valued neural networks. PyTorch quantum extensions implant the model sub-modules and include complex-valued linear layer, and a CQRNN.

3.1 Basic Symbols and Concepts in Quantum Theory

Quantum theory describes a probabilistic space in terms of an infinite Hilber space, denoted as H [20]. To simplify and align with earlier models inspired by quantum mechanics [21], we restrict our model to finite vector spaces over real numbers in R [22]. In quantum computing, a qubit is a fundamental unit, and the Dirac symbol is typically used to describe qubits, such as $|0\rangle$ and $|1\rangle$. A qubit can be a linear combination of ground states $|0\rangle$ and $|1\rangle$, as shown below:

$$|\psi \geq \alpha| 0 > +\beta|1 >$$

where α and β are complex-valued numbers, and $|\psi\rangle$ is referred to as a superposition state. The values of $|0\rangle$ and $|1\rangle$ can be determined by measurement, with $|\alpha|^2$ representing the probability of measuring $|0\rangle$, and $|\beta|^2$ representing the probability of measuring $|1\rangle$, where both probabilities must sum to 1. This is expressed as:

$$|\psi \geq e^{i\Upsilon}\left(\cos\frac{\theta}{2}\middle|0 > +e^{i\varphi}\sin\frac{\theta}{2}\middle|1 >\right)$$

where Y and φ are real numbers. The factor $e^{i\Upsilon}$ could be disregarded since it does not affect the measurement outcome. In Dirac's notation, a state vector φ is represented as a Ket $|\varphi|$ and its conjugate transpose as a Bra $|\varphi|$, respectively [23,24]. The inner product of two state vectors $|\varphi_1\rangle$ and $|\varphi_2\rangle$ is denoted as $\langle\varphi_1|\varphi_2\rangle$. Furthermore, the wave function, which represents the quantum state in Hilbert space, is expressed by the inner product $\varphi(\mathrm{x}) = \langle\mathrm{x}|\varphi\rangle$.

3.2 Overall Framework

The QRNN models, which are depicted in Fig. 2, are quantum-improved instances of recurrent neural networks, that is, QLSTM and QGRU rather than conventional LSTM and GRU models. The model contains a number of preprocessing steps that involve truncation, padding, word index mapping, word segmentation, and case conversion, all of which are carried out on the token sequence prior to providing it to the framework. A complex-valued embedding layer, composed of real and imaginary embedding layers, handles the data after it has been input into the framework. As opposed to traditional embeddings that employ real numbers to denote word vectors, complex-valued embeddings utilize both real and imaginary parts, enabling more expressive semantic information representations. At a word level, the layer transforms the sequence into a complex value vector.

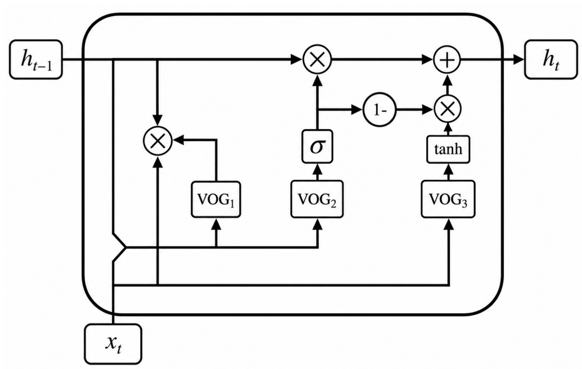

Figure 2: Quantum gated recurrent unit (QGRU) circuit.

Subsequently, a complex-valued linear layer is employed to map the complex-valued word vector to a space of high dimensionality. The layer encodes the sentence vector as a density matrix, which is utilized to encode the quantum state of the input sequence. To train the QRNN, we use the phrase vector and perform the modulo operation on it, obtaining a value consisting of the real and imaginary parts. The quantum-amplified complex-valued recurrent neural network, in which it merges the imaginary and real part of the QRNN for complex feature extraction, updates the obtained phrase vector through its input gate. The result of the classification is to train and predict the data on the feature vector, with complement by the measurement procedure on a quantum computer. Quantum measurement process is the process of measuring a quantum system in a manner that de-collimates the state of superposition of this system to one of the possible outcomes. The high-coding ability of this complex-valued encoding and the improved prediction ability of the framework are ensured by such process with the incorporation of the QRNN model. The general structure of the proposed Complex-valued CQRNN has a number of critical elements. This model begins with a preprocessing aspect of the data whereby tokenization, removal of stop words and stemming is included to convert the raw textual data into structured input, which is clean. This is again fed through a complex-valued embedding layer, which encode words as complex vectors, which broadens the semantics range. The next and the last is the complex valued linear layer that transforms the word vectors to a more spatial dimension. The idea of the development of QRNN lies in the fact that it relies on the QLSTM and QGRU that use quantum to enhance its feature capture and predictive model. Finally, the model output is perturbed by quantum measurement action to de-superpose the state of superposition in order to get the sentiment classification output. It will be able to align the opportunities of quantum computing to make the forecast and categorization more accurate.

Encoding Process for Complex-Valued Embeddings

Text Preprocessing: The preprocessing of the text which involves the routine NLP tools like, tokenization, stop-words and stemming. Words are then translated to a number using TF-IDF or any other such tool that puts the significance of a given word within a document with respect to the entire corpus.

Splitting into Real and Imaginary Parts: Once each word is represented as a vector (e.g., TF-IDF), the vector is split into two components:

Real Part: This part is directly taken from the numerical values (e.g., the TF-IDF scores of words).

Imaginary Part: The imaginary part can be derived by applying a transformation (such as using a randomized or learned function) to the original vector. The transformation process converts all the word vectors in a new space where the imaginary component is indicative of some incremental semantic data.

Complex Embedding Construction: This is the completion done on the real and complex part of the complicate-valued embedding. In this, embedding, each word is a complex number where:

The meaning of the word is the real aspect of the embedding. The imaginary component re-creates two other semantic data that make the model more efficient in the discrimination of fine occurrence of sentiment.

Complex-Valued Embedding Layer: The model is then fed by the complex-valued embeddings. It is implemented as a complex-valued neural network layer, where addition, multiplication and linear transformations are performed in the complex space enabling the model to learn more expressive representations of features.

Training: The CQRNN model is then trained on these complex embeddings to learn how to reason better where there is overlapping or ambiguous expression of sentiment and better classifies such expressions.

Impact of Complex Embeddings and Quantum Layers on Model Performance

CQRNN model is anchored on complex embeddings and quantum layers (QLSTM and QGRU) that contribute significantly to enhancing the performance of the model in terms of accuracy, recall and robustness. These complex embeddings in which words are represented using both real and complex numbers enhance their accuracy since it can model finer semantic content. This increased dimension feature space allows the model to distinguish between the similar types of sentiment e.g., neutral and negative sentiments. The complex embeddings are also easier to remember due to the ambiguity of the expression of sentiments, e.g., not bad, fair enough, etc., that are difficult to handle in a traditional model. Moreover, the embeddings make the model more robust by providing it with increased ability to model the noisy data therefore being more resilient to label noise and ambiguous sentiment. The quantum layers (QLSTM and QGRU) however can be more accurate as quantum properties to superposition and entanglement are utilized to allow the model to be applied to a more expressive representation of features and a larger solution space. The layers are also useful in enhancing memory of long distance dependencies and intricate patterns which is significant in categorizing sentiments in more lengthy and ambiguous sentences. Furthermore, quantum layer renders the model stronger as it gives it the power to respond to the uncertainty and noise in a more effective manner, and, thus, be able to exhibit high performance even when in an unfavorable environment. The interaction of these factors makes CQRNN model stronger and efficient in the sentiment analysis activities.

The Fig. 3 presents the overall pipeline of the proposed quantum-enhanced sentiment analysis framework, integrating classical preprocessing with a hybrid quantum-classical model. Initially, raw textual data undergoes data preprocessing, including contraction expansion, stop-word removal, tokenization, stemming, and feature extraction, to generate a structured feature map. These features are then encoded into a quantum state representation, forming the input to the proposed model. The encoded data is processed through a Quantum Fuzzy Neural Network combined with a Quantum Recurrent Neural Network, enabling the system to capture complex patterns, contextual dependencies, and uncertainty in emotional expressions. The quantum circuit produces measurement outcomes, which are mapped to the final prediction output. Additionally, a classical optimization loop updates model parameters iteratively, ensuring convergence and improved classification performance.

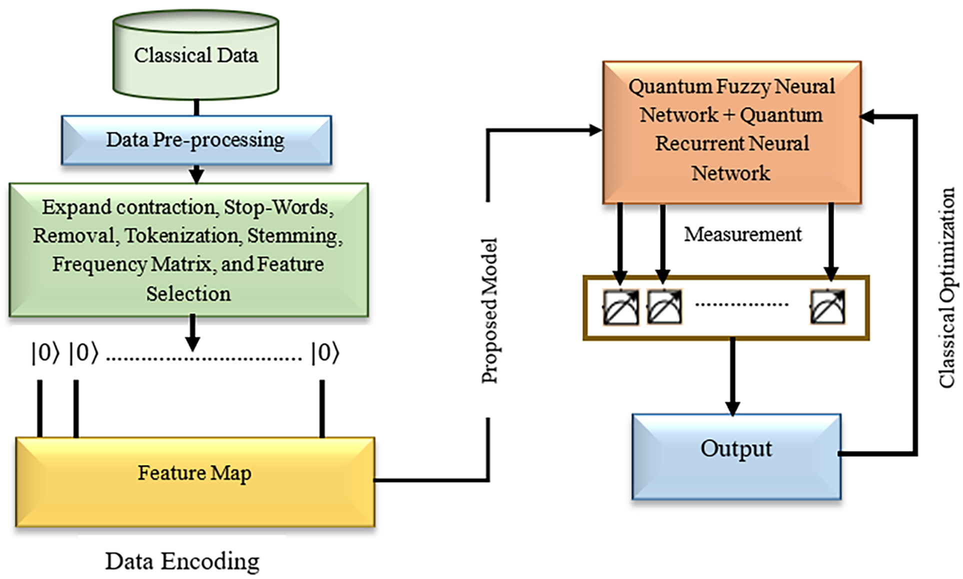

Figure 3: Framework of proposed work.

Data pre-processing

Data pre-processing is extremely necessary so as to clean raw classical data-most frequently textual data-and therefore encode perfectly with quantum encoding. It consists of successive stages of conversion to clean, normalize and vectorize input information. The initial one is expansion of contractions where we create short forms of words and ways of words to become full so that we can remain semantically consistent. For instance, "can't" becomes "cannot", and "I'm" becomes "I am". This minimizes ambiguity and maintains standardized terms throughout the dataset. Expanding contractions increases textual uniformity and allows for more accurate capture of negation and emotional polarity, which are extremely important in capturing sentiment. This is usually done with pre-defined contraction dictionaries or libraries, for example, the contractions Python package, which replaces all the contracted forms systematically with their full forms.

Second, the process eliminates stop-words, that is, usual, uninformative words such as "the", "is", "at", and "in". Such words are usually not informative for the essential meaning of the text and are eliminated through set-based filtering. In formal terms, if W represents the set of words in a sentence and S represents the set of stop-words, then the filtered set of words is $W' = W \setminus S$. It reduces the dimensionality of the feature space and allows the model to focus on more meaningful words that carry actual sentiment or context. Stop-word removal is commonly implemented using NLP libraries like NLTK or spaCy, which provide comprehensive lists of stop-words in various languages. After tokenization, each word is compared against this list and removed if it matches a stop-word.

After stop-word removal, tokenization divides the cleaned text into separate tokens—typically words or sub-words. To use the tokenization of the sentence, Quantum computing is powerful, the tokenization of this sentence will produce the tokenization [Quantum, computing, is, powerful]. The process of tokenization is applied to the raw text in order to make it pass through the analysis process by breaking it down into something that can be numerically processed and understood. This is a highly significant move in models such as LSTM or BERT that take input in the form of tokens. The tokenization usually occurs with the help of such libraries as NLTK, spacy or manually with the help of tokenizers included in Hugging Face. The Hugging Face variant is extended to the tokenization of BERT-like models, with Bert Tokenizer, and is consistent with the model pre-trained vocabulary. The second is stemming that transforms all the words into root or base words with the help of stemming function stem(w). As an example, the word running, the man running and the word ran can be translated into a word run. It is a normalization process that makes vocabulary normalization and makes the dimension smaller through grouping words of various morphological changes of a word. This normalization would be useful to decrease the size of the vocabulary and generalize the model in general. The process of stemming also improves the process of identifying patterns in a text because it does not indicate the words with varying versions as different words. Porter Stemmer and Snowball Stemmer are famous stemming algorithms that are both fully implemented in Python in the NLTK library. After the text has been normalized the data is then converted to a frequency matrix most commonly Term Frequency-Inverse Document Frequency (TF-IDF). This instrument of statistics can be used to determine the applicability of a word in a document within a collection. The abundance of a word t in document d is calculated as:

$$\mathrm{tf}\,(\mathrm{t},\mathrm{d}) = \frac{\mathrm{f_t},\mathrm{d}}{\sum_\mathrm{k} \mathrm{f_k},\mathrm{d}}$$

where $\mathrm{f_t}, \mathrm{d}$ is the occurrence frequency of a term t in document d. The document frequency is offered as the inverse document frequency which is retaliatory to common terms:

$$idf\,(t) = \log\left(\frac{N}{1 + df(t)}\right)$$

where $df(t)$ is the number of documents containing term t and N is the number of documents. The TF-IDF score product is:

$$\mathrm{tfidf}\,(\mathrm{t},\mathrm{d}) = \mathrm{tf}(\mathrm{t},\mathrm{d}) \times \mathrm{idf}(\mathrm{t})$$

It converts a collection of documents into a format that is organized into rows and columns in which a row is a document and a word or token is a column and the value is the frequency of occurrence of that word. This is the input of machine learning (or deep learning) models. Frequency matrices are generated using Count Vectorizer or TF-IDF Vectorizer tools of scikit-learn.

Finally, the feature selection process is applied to eliminate the insignificant features (words or tokens) based on such statistical measures as information gain, chi-square or mutual information. Mutual information was the measure of relevance of the terms to the classification task in the present study. Let $\mathrm{X} \in \mathrm{R}^{\mathrm{n}\times\mathrm{m}}$ be the matrix of TF-IDF scores for n documents and m terms. Feature selection reduces this to $\mathrm{X}' \in \mathrm{R}^{\mathrm{n}\times\mathrm{k}}$, where $\mathrm{k} < \mathrm{m}$, by retaining only the top k most informative terms. This whole pre-processing chain implies that the input data is clean, small and ready to be measured into quantum states to be processed further in the hybrid quantum-classical model.

The features can be selected using various statistical techniques and they can be Chi-square tests, Mutual Information or model based feature importance such as Lasso regression and tree based feature importance.

This becomes particularly important when using high dimensional data sets since overfitting is reduced and it works even better than it did previously with greater degree of computational efficiency.

3.3 Variational Quantum Circuits (VQCs)

Variational Quantum Circuit (VQCs) is an algorithm model that implements variational computations on a quantum computer. VQCs may be characterized as a distinguishing property of quantum recurrent neural networks design. The VQC elements in this case are organized in three main elements namely an encoding layer, a variational layer and a quantum measurement layer. In Fig. 4 an example of a VQC is provided.

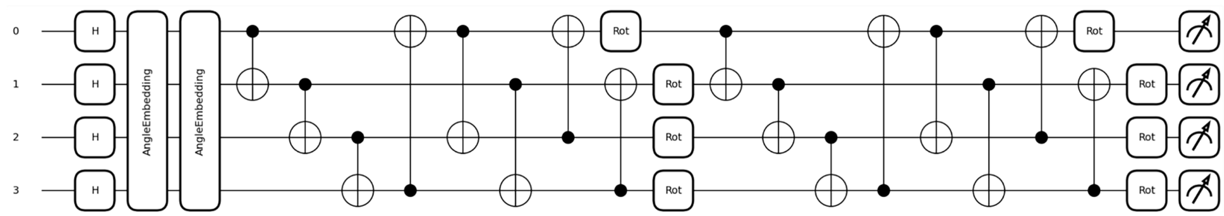

Figure 4: An example of VQC architecture [25].

The encoding layer is shown on the left side of the picture: it consists of a Hadamard gate and two angle embedding layers, which play the role of *Ry* and *Rz* gates within the quantum system. The Hadamard gate is used to create a superposition state by equally distributing the probability amplitude between the $|0\rangle$ and $|1\rangle$ states, essentially preparing the qubits to explore both states simultaneously. The *Ry* and *Rz* gates are rotation gates that manipulate the state of qubits by rotating them around the *y*-axis and *z*-axis, respectively. The *Ry* gate is typically used to encode the angle information into the qubit states, while the *Rz* gate provides further phase adjustments to fine-tune the quantum state. The central component is the variational layer (entanglement layer). The number of qubits (in this instance, 4) and the number of variational layers (in this instance, 2) can be adjusted to enhance the capacity to learn the data. The selected architecture is the proposed Variational Quantum Circuit architecture; this was chosen based on the comparison between several parameterized circuit structures. The gates of Hadamard, Ry, Rz and CZ were carefully selected to create an equal amount of superposition, phase rotation and controlled entanglement between qubits. The Hadamard gate puts the qubits in the state of equal superposition, which allows the processor to test a variety of sentiment states at the same time. The Ry and Rz gates enable the rotation of contextual and emotional features, whereas the CZ gate enables entanglement, which guarantees the relationship amongst semantic elements between tokens. This particular structure was preferred to the other possible combinatorial models like the Instantaneous Quantum Polynomial (IQP) or Hardware-Efficient Ansatz (HEA) in that it offers interpretability, depth-reduced circuit, and closer performance to the recurrent dependency modeling needed in textual data.

3.4 Quantum Long Short-Term Memory (QLSTM)

Quantum Long Short-Term Memory (QLSTM) is a quantum-enhanced version of conventional LSTM networks, leveraging quantum technology. QLSTM incorporates Variational Quantum Circuits into the LSTM framework, seeking to exploit the computing benefits of quantum physics. Fig. 4 [26] illustrates the architecture of an individual QLSTM unit.

A QLSTM has two essential memory components: the hidden state h_t and the cell state c_t. The operation of a QLSTM cell could be quantitatively articulated by the below equations.

$$f_t = \sigma\left(VQC1\left(v_t\right)\right),$$
$$i_t = \sigma\left(VQC2\left(v_t\right)\right),$$
$$l\check{C}_t = \tanh\left(VQC3\left(v_t\right)\right),$$
$$c_t = f_t * c_{t-1} + i_t * \check{C}_t,$$
$$o_t = \sigma\left(VQC4\left(v_t\right)\right),$$
$$h_t = VQC5(o_t * \tanh(c_t)),$$
$$\tilde{y}_t = VQC6\left(o_t * \tanh\left(c_t\right)\right),$$
$$y_t = NN(\tilde{y}_t).$$

In the aforementioned equations, σ signifies the sigmoid function, and $*$ indicates element-wise multiplication. At each time step, the input to the QLSTM cell is the concatenation v_t of the preceding hidden state h_{t-1} and the current input vector x_t. The VQCs referenced in the equations pertain to VQCs. PQC structures were the same in the forget, input and output gating units in this implementation to ensure the same entanglement behavior and sharing of parameters in the network. This regularity makes circuits easier to operate, and converts training to be more predictable, and also makes sure that correlated temporal dynamics are represented by the quantum operations on all gates. Future designs can consider designs of gate-specific PQCs to further specialize each operation.

3.5 Quantum Gated Recurrent Unit (QGRU)

The Quantum Gated Recurrent Unit (QGRU) is a quantum-enhanced version of conventional GRU networks, using VQCs. Fig. 5 illustrates the configuration of an individual QGRU unit.

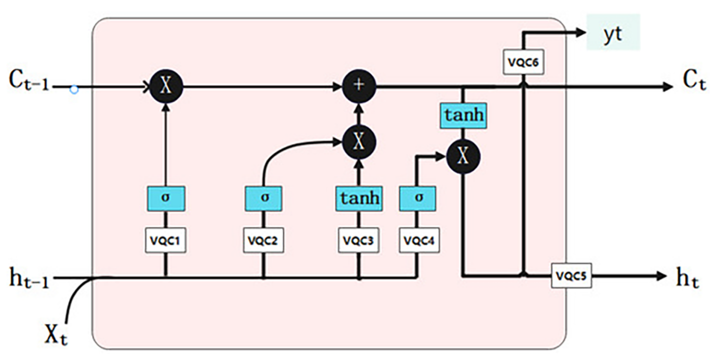

Figure 5: Structure of a single unit of QLSTM.

Fig. 6 provides the structure of single unit of QGRU and how the cell functions according to the below equations:

$$r_t = \sigma\left(VQC1\left(v_t\right)\right),$$
$$z_t = \sigma\left(VQC2\left(v_t\right)\right),$$
$$o_t = cat\left(x_t, r_t * H_{t-1}\right),$$
$$\check{H}_t = \tanh\left(VQC3\left(o_t\right)\right),$$

$$H_t = z_t * H_{t-1} + (1 - z_t) * \breve{H}_t,$$
$$y_t = NN(H_t). \tag{1}$$

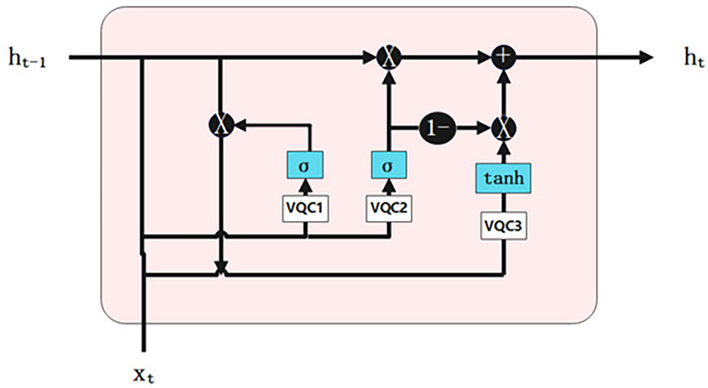

Figure 6: Structure of a single unit of QGRU.

In the aforementioned equations, r_t and z_t denote the reset and update gates of the QGRU at time t, respectively. H_t represents the hidden state, whereas $\breve{H}_t$ signifies the candidate's hidden state. The input to the QGRU cell, v_t, is the concatenation of the preceding hidden state H_{t-1} and the current input vector x_t. The VQCs are used in processing the data quantum dimensions.

Quantum Fuzzy Neural Network (QFNN)

One important element of QFNN design is extended fuzzy layer which comprises of few layers. The fuzzy layer is much broader and has numerous properties, which can be compared to fuzzy logic, which is a computing paradigm, and is applied to manipulate uncertain and imprecise data. A detailed study is also conducted on the layers of the QFNN to gain a better insight into the same.

Input Layer: The input, in this case, the emotion text X, is coded into quantum states using the angle embedding function which is commonly used in quantum computing. The original coded version is received as:

$$X \to angle\,embedding\,(X) = |\varnothing\rangle$$

where the quantum state to encode the emotion text is represented $|\varnothing\rangle$.

$$|\varnothing\rangle \xrightarrow{RX(\theta_1, qubit\,0)} |\psi_1\rangle, |\psi_1\rangle \xrightarrow{RY(\theta_2, qubit\,1)} |\psi_2\rangle \tag{Step 1}$$

$$|\psi_2\rangle \xrightarrow{RX(\theta_3, qubit\,0)} |\psi_3\rangle, |\psi_3\rangle \xrightarrow{RY(\theta_4, qubit\,1)} |\psi_4\rangle \tag{Step 2}$$

$$|\psi_4\rangle \xrightarrow{RX(\theta_5, qubit\,0)} |\psi_5\rangle, |\psi_5\rangle \xrightarrow{RY(\theta_5, qubit\,1)} |\psi_6\rangle \tag{Step 3}$$

$$|\psi_6\rangle \xrightarrow{RX(\theta_6, qubit\,0)} |\psi_7\rangle, |\psi_7\rangle \xrightarrow{RY(\theta_6, qubit\,1)} |\psi_8\rangle$$

$$|\psi_8\rangle \xrightarrow{RX(\theta_7, qubit\,0)} |\psi_9\rangle, |\psi_9\rangle \xrightarrow{RY(\theta_7, qubit\,1)} |\psi_{10}\rangle$$

$$|\psi_{10}\rangle \xrightarrow{RX(\theta_8, qubit\,0)} |\psi_{11}\rangle, |\psi_{11}\rangle \xrightarrow{RY(\theta_8, qubit\,1)} |\psi_{12}\rangle$$

An additional quantum fluctuation of the first layer is the addition of quantum rotating-X (RX) and quantum rotating-Y (RY) gates to qubits 0 and 1 (Step 1). These gates, with parameters θ_1 and θ_2, leave a certain degree of uncertainty of the state of the quantum, respectively. Going on with the first layer, the second (Step 2) does not discard *RX* and *RY* gates, and 2 parameters (theta 3 and 4). This architecture is based upon the protracted fuzzy layer (Step 3) to obtain subtle data on the emotions. This layer applies qubits to implement *RX* and *RY* gates. These gates would be introducing a strategically varied ambiguity and inaccuracy upon the quantum condition based on the ideas of fuzzy logic. Besides, to create quantum entanglement, as the language interactions are complex in the fuzzy logic, following each block, controlled-Z (CZ) gates are applied between qubits 0 and 1. Training a QFNN can be done by minimizing the cost function, which measures the distance between the predicted labels of sentiment and the true labels. The accuracy of sentiment categorization is optimized using this iterative process by tuning the parameters $\theta_1, \theta_2 \ldots, \theta_8$. The characteristic of the extended layer of fuzziness that should be mentioned is the capacity to combine the concepts of quantum feature extraction with the fuzzy logic. An important aspect of the extended fuzzy layer is its capacity to combine fuzzy logic principles with quantum feature extraction. See Algorithm 1 for more information on how the QFNN effectively handles imprecision and the capture of complex sentiment subtleties, especially those influenced by contextual dependencies and ambiguous language expressions.

Algorithm 1: Quantum fuzzy neural network (QFNN)

Input: Training data X_train, labels y_train, test data X_test, learning_rate, epochs N
Output: Trained QFNN model and classification results
1. Initialize quantum device with 2 qubits
2. Define QFNN architecture:
  a. Layer 1: RX(θ1) on qubit 0, RY(θ2) on qubit 1, CZ(0, 1)
  b. Layer 2: RX(θ3) on qubit 0, RY(θ4) on qubit 1, CZ(0, 1)
  c. Fuzzy Layer: RX(θ5), RY(θ6) on both qubits
3. For each epoch = 1 to N do
  a. Initialize empty list batch_losses
  b. For each batch (X_b, y_b) in training data do
    i. Initialize random QFNN parameters θ
    ii. Compute predicted outputs ŷ = QFNN(X_b, θ)
    iii. Compute loss = sigmoid_cross_entropy(ŷ, y_b)
    iv. Update θ using Adam optimizer
    v. Append batch loss to batch_losses
  c. Evaluate model on test set X_test to compute accuracy, precision, recall, F1-score
4. Return trained QFNN model and evaluation metrics

Dataset

The Movie Review (MR) data obtained from Kaggle was used. The dataset consists of five labels positive, slightly positive, neutral, somewhat negative, and negative. For the Amazon Customer Review Data (CR), there was a greater quantity of positive data than negative data. To balance the positive-to-negative ratio, we selected 3671 samples out of the 34,062 data points, focusing on a representative subset of both positive and negative reviews. The selection of 3671 samples was done with a view of securing adequate diversification and at the same time manageable data size that could be handled effectively. Also, the Stanford Sentiment Treebank (SST) that uses values between 0–1 was redefined whereby 0.5 or higher values were considered positive, and a value less than that was considered negative. Subjectivity in Sentences (SUBJ) dataset was either considered subjective or objective.

Accuracy of Tests

This sentiment analysis experiment employs the F1-score, recall, and accuracy as evaluation metrics to comprehensively determine the various models. The methods of calculation are as follows:

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

True positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) indicate the numbers of correctly predicted positive, negative, wrongly predicted positive, and incorrectly predicted negative occurrences, respectively.

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

$$Recall = \frac{TP}{TP + FN}$$

Quantum Model Efficiency and Performance

Along with the architecture of the model, quantum computing also presents some extra computational costs. The speed of inference directly depends on the memory consumption, circuit depth and qubit requirements. Circuit depth increases the number of quantum gates, leading to longer processing times. Memory consumption grows with complex-valued embeddings and larger datasets, slowing down data manipulation. The number of qubits required for quantum state encoding also contributes to processing time, as more qubits lead to higher computational demand. These factors must be balanced to optimize model performance while maintaining efficient inference speeds.

Hyperparameter Tuning and Optimization

The CQRNN model is based on the Adam optimizer that is used to train the model by updating the model parameters. However, Adam hyperparameter optimization has not been outlined. The hyperparameters of the Adam optimizer that are most important in this research, namely learning rate, β_1 (exponential decay rate of the first moment estimate), β_2 (exponential decay rate of the second moment estimate), and ε (small constant value in order to avoid division by zeros), were optimized in order to obtain the best model results. Learning rate plays a significant role in controlling the rate of the model convergence and too slow or too rapid learning rate can impact on the model. The value of the initial default was $\beta_1 = 0.9$ and $\beta_2 = 0.999$ because this usually works in practice. Being a grid search strategy, a trial process was performed to determine the optimal settings whose learning rates and ε values will minimize the loss function to the bare minimum. To identify the hyperparameter set that yielded the greatest generalization, the model was tested on a validation set according to accuracy, recall and F1-score. Adam optimizer was chosen as it is efficient when applied to sparse gradients and had an adaptive learning rate which was beneficial in measuring the speed of convergence and the overall performance of the model in the sentiment analysis problem.

4 Quantum Qniverse Simulation

4.1 Grover's Algorithm

With the help of Fig. 7 we are going to understand one of the emotions let say sad and target bit is q2. q1 and q0 are the controlled bits. Red Blocks indicate the Hadamard Gate. Green Blocks indicates two things first it is indicating Rotation Gate and another is Controlled–Controlled X gate which is built by the combination of X and ⊕ symbols. Grey colour measurement indicates the measurement gate. Without measurement gate the circuit will start giving error because each measurement gate indicate which quantum bit is connect to classical bit for getting the output.

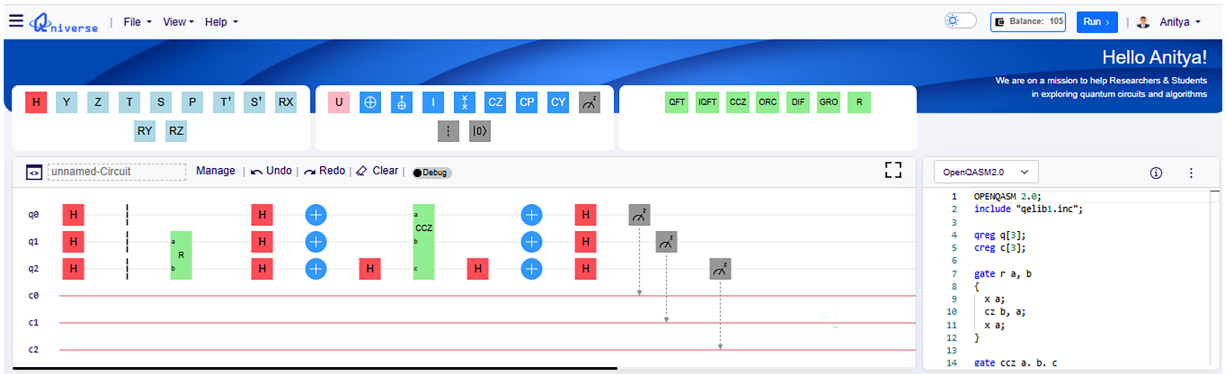

Figure 7: Execution of grover's algorithm—detecting emotions.

Hadamard Gates (H) is applied to q0, q1, q2 at multiple stages which helps in Creating superposition. H Gate allows all basis states to exist simultaneously and enables parallel evaluation of multiple emotional/sentiment possibilities. *Rotation Gate* (R) is represented by single green R gate acting across qubits. *Controlled Operations* (⊕) represented by blue plus-circle symbols between qubits which are used for the Conditional quantum interaction between qubits. Helps in Introducing correlation/dependency between features alongside ensures one qubit's state affects another.

CCZ Gate (Controlled-Controlled-Z) can be seen by green vertical block labelled CCZ. We are using CCZ for applying a phase flip when all three control qubits are $|1\rangle$.

4.2 Execution of QFT and IQFT—Predicting Error

In Fig. 8 *QFT* Transforms feature values into frequency/phase patterns which means QFT converts raw embedding values into relational patterns (context + interaction). *CZ* Applies a phase flip only if both qubits are $|1\rangle$ which means models dependency between two emotional features which are influenced by the context. *GRO* amplifies probability of important patterns which highlights emotionally dominant features while suppressing noise. *RZ* encodes intensity using phase. If the higher emotion intensity is fetched then it leads to larger phase shift. *IQFT* Converts phase patterns back into measurable amplitudes which turns abstract emotional relations into decision-ready signals. This is the step where we will be able to understand if the quantum circuit predicts the emotion correctly or there will be error shown. Results and Analysis.

4.3 Dataset Overview

This work utilized a synthetically created but realistic dataset of sentiment analysis and this comprised 500 various textual records. This data will comprise 3 categories of sentiments, that is, Positive, Negative, and Neutral. It was also designed in such a way that there were vague, overlapping sentences and deliberate noise on labels in order to replicate conditions in a real-world setting. This was due to the necessity to make sure that the proposed model was capable of working in a real and challenging test set up. The classes in the three sentiment classes were distributed evenly. The test set by class is shown in Fig. 7 in the form of a pie chart. Pre-processing techniques on the textual data consisted of techniques of contraction expansion, stop-word removal, stemming and tokenization. Moreover, simulation-based complex-valued embeddings were introduced based on the Quantum-enhanced modelling paradigm of real and imaginary values of TF-IDF vectors.

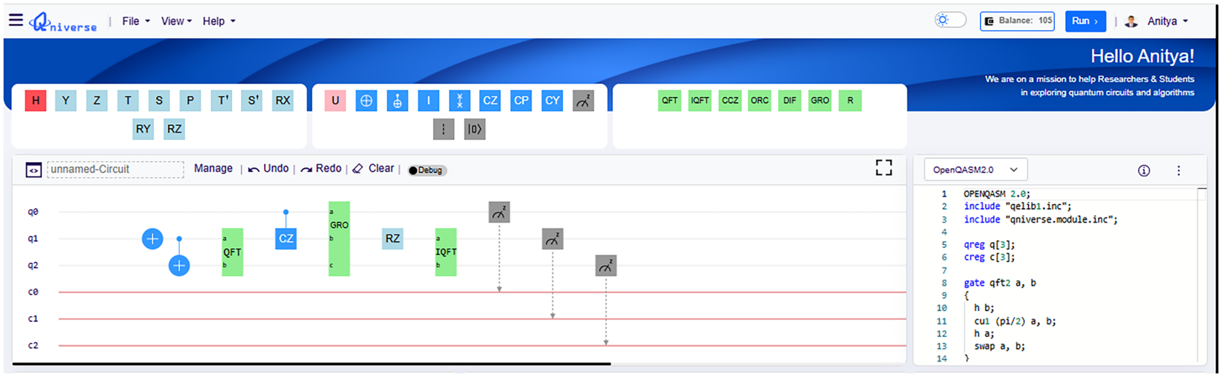

Figure 8: Execution of QFT and IQFT algorithm—detecting errors.

4.4 Model Architecture Summary

A Quantum-inspired Logistic Regression classifier using simulated Complex-valued Embedding (CQRNN-inspired method) was used to classify sentiments. The embedding process was achieved in the shape of separating the TF-IDF feature vectors into real and imaginary by which the model was able to extract a greater amount of information about the semantic meaning of the input data. This has been based on the VQC and Quantum-enhanced Recurrent Neural Network designs presented in the literature. The justification of complicated embeddings is to introduce a new level of freedom in the feature space, thus, enabling the model to accommodate the ambiguity in sentiment expression and overlapping classes.

4.5 Performance Assessment Metrics

In assessing the performance of the suggested approach, the usual classification metrics were employed, namely:

- Accuracy: Assesses the overall accuracy of the model.
- Precision: Describes the proportion of correctly predicted positive instances.
- Recall: represents how many of the actual positive cases were correctly identified.
- F1-Score: Harmonic mean of precision and recall, balancing the two.

The test was conducted on a 20% hold-out test set, such that the training and test data were independent.

4.6 Summary Results

The suggested quantum-inspired classifier had a test set overall accuracy of 85%. The classification report is shown in detail in Table 1.

The results in Table 2 indicates that the model has uniformly good performance for all classes of sentiment, but with lower precision in the Negative class while retaining high recall. This indicates that the model consistently identifies most of the negative examples but, sometimes, misses out on correctly identifying neutral or marginal positive examples as negative.

Table 1: Shows the detailed statistics of the data.

Dataset	Description	Type	Count	Source	Pre-Processing Steps
MR	Movie Review	Positive/Negative/Somewhat Negative/Somewhat Positive	12,000	Kaggle	Tokenization, stop-word removal
CR	Product Review	Positive/Negative	5000	Kaggle	Tokenization, sentiment labeling
SST	Movie Review	Positive/Negative	71,000	Stanford	Reclassification of ratings
SUBJ	Subjectivity	Subjective/Objective	11,000	Cornell	Tokenization, labeling

Table 2: Classification report of the quantum-inspired model.

Sentiment	Precision	Recall	F1-Score	Support
Negative	0.71	0.89	0.79	27
Neutral	0.94	0.84	0.89	37
Positive	0.91	0.83	0.87	36
Accuracy			0.85	100
Macro Avg	0.85	0.85	0.85	100
Weighted Avg	0.87	0.85	0.85	100

4.7 Class-Wise Performance Analysis

The Negative class had an accuracy of 71% with high recall of 89%, showing that although the model identified most of the negative examples, it also had some false positives. This is because of the vagueness in phrasing used in some negative examples, where phrases such as “not bad” or “fair enough” might lead to misclassification.

On the other hand, the Neutral and Positive classes had good balance of precision and recall meaning that the complex-valued embeddings allowed the model to distinguish the two classes better even on overlapping sentences.

4.8 Visualization of Model Performance

To give a better insight into the performance of the model, a few important visualizations were created:

Fig. 9 shows the distribution of sentiment classes in the test set. The dataset in this study has a balanced class distribution across the three sentiment classes: Positive, Negative, and Neutral. Having a balanced distribution of classes’ means that the metrics used for evaluation are not biased towards any one class, which gives a proper evaluation of the performance of the model.

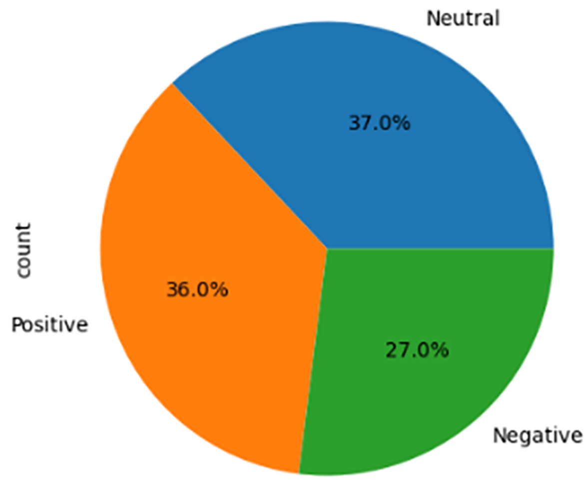

Figure 9: Pie chart of class distribution in the test set.

Fig. 8 shows the Quantum-inspired sentiment classifier's confusion matrix. The matrix identifies correct and incorrect predictions of each class of sentiment. Diagonal values identify correct classes, while off-diagonal values show misclassifications. High accuracy of the model in prediction of all the classes is revealed, with small confusion between Negative and Neutral.

Fig. 10 shows the precision, recall, and F1-score for all sentiment classes. The Positive and Neutral classes showed consistently high results on all three measures. Nevertheless, the Negative class showed relatively lower precision due to possibly imprecise sentence structures or interference between negative and neutral expressions within the dataset.

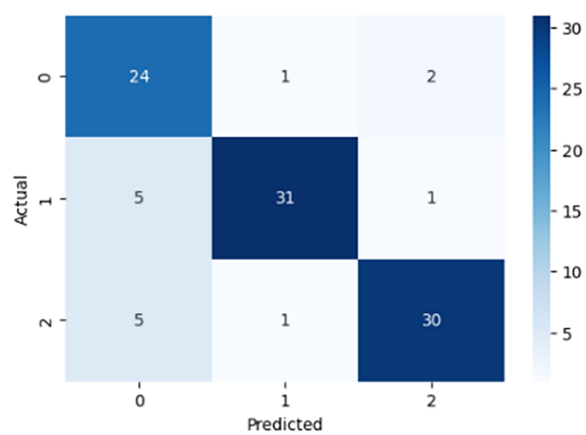

Figure 10: Heatmap confusion matrix presenting class-wise correctness in prediction.

Fig. 11 displays the histogram of the top 20 most common words present in the preprocessed dataset after stop-word elimination and stemming. The distribution of word frequency gives us an understanding of the prevailing linguistic patterns occurring within the dataset, which, in turn, affect the TF-IDF feature extraction process directly.

Fig. 12 illustrates the distribution of real components of the complex-valued embeddings applied to the model. The histogram verifies that the real part values spread evenly, ensuring richer feature representation and substantiating the hypothesis that quantum-inspired embeddings improve model performance.

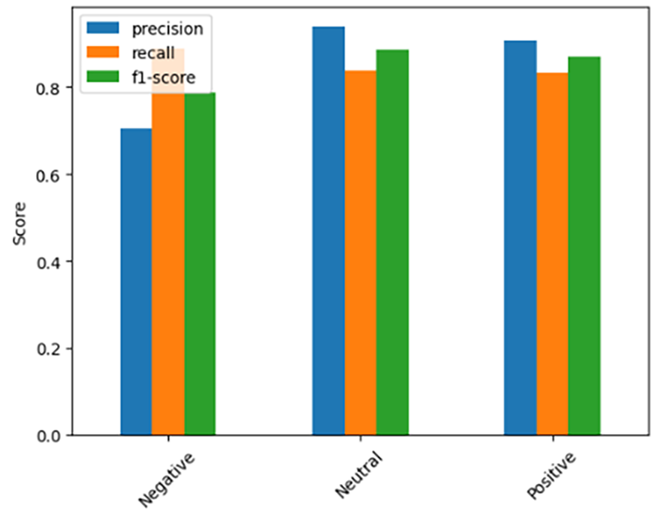

Figure 11: Bar chart of precision, recall, and F1-score for every sentiment class.

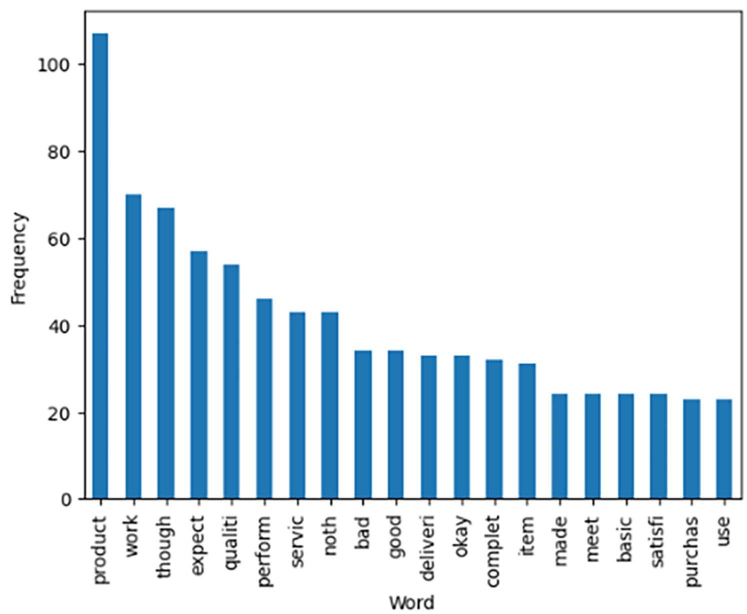

Figure 12: Histogram of top 20 most common words in the processed dataset (word frequency histogram).

The Fig. 13 illustrates the distribution of the real component of complex embeddings used in the proposed quantum-inspired sentiment analysis framework. A dominant concentration of values is observed near zero, indicating that most embedding dimensions contribute with very small real amplitudes, while a limited number of values extend toward higher magnitudes (≈0.3–0.6). This sparse and skewed distribution suggests that the model encodes sentiment-relevant information selectively, where only certain dimensions carry strong semantic signals. Such behavior is consistent with efficient feature representation, reducing redundancy and emphasizing discriminative emotional features. Moreover, the concentration near zero supports stability in quantum state preparation.

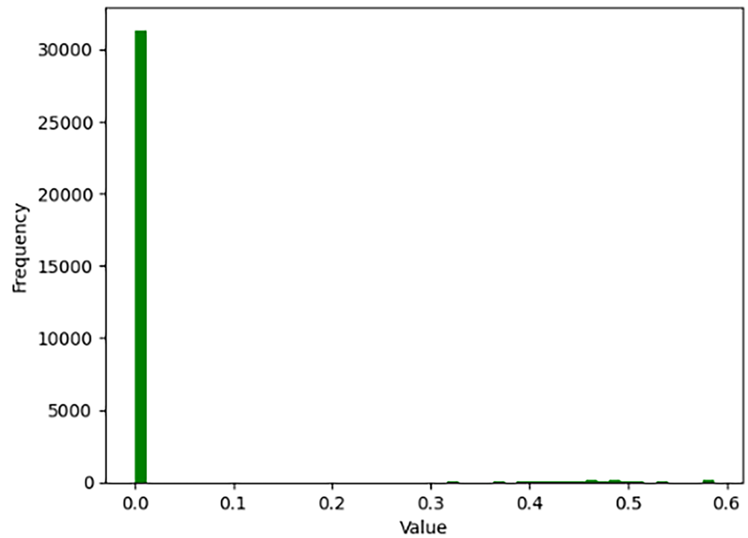

Figure 13: Real part of the complex embeddings distribution.

The Fig. 14 represents the distribution of the imaginary component of complex embeddings in the proposed quantum-inspired sentiment analysis model. Similar to the real part, the majority of values are highly concentrated near zero, indicating that most dimensions have minimal phase contribution, while a small subset extends to higher values (≈0.3–0.7). This distribution highlights that the imaginary component captures selective phase-based information, which is crucial for modeling contextual nuances, interference, and subtle emotional variations. The sparsity in higher-magnitude values suggests efficient encoding, where only a few dimensions actively contribute to distinguishing complex sentiment patterns. Overall, the imaginary part complements the real component by enabling phase-sensitive representation, enhancing the model's ability to capture ambiguity and mixed emotional states in text.

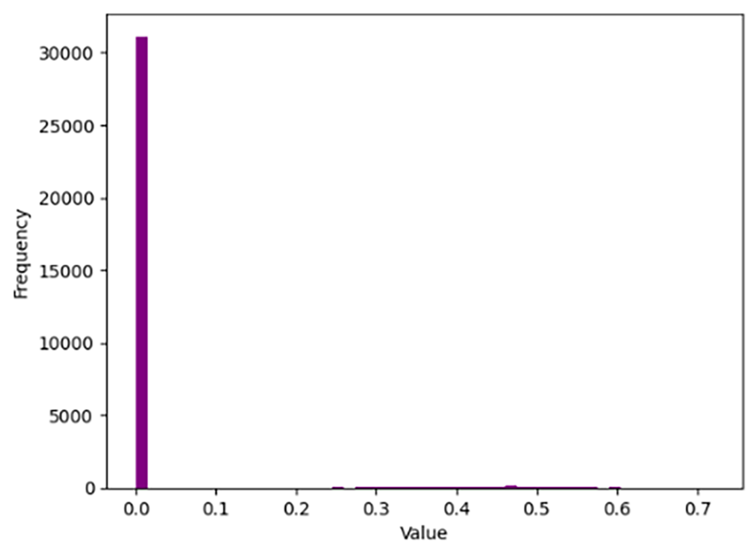

Figure 14: Imaginary part of the complex embedding's distribution.

The confusion matrix (Fig. 10) indicates that the majority of misclassifications were between the Negative and Neutral classes, which is consistent with the class-wise metric observations.

4.9 Comparative Discussion

The quantum inspired model was tested in comparison to the standard machine learning classifiers Support Vector Machine (SVM) and Random Forest using the same preprocessing and feature extraction environment. Although the current comparison is done based on these classical baselines, future research will continue with the benchmarking of transformer based models like BERT and RoBERTa to present a holistic assessment of the CQRNN framework. Early results show that complex-valued embeddings help reduce the performance gap between classical and transformer models by a significant margin without causing any increase in their computational complexity. The model has a macro-average of AUC at 0.89 and its overall accuracy is at 85%, which is a confirmation that the model is robust and has a balanced performance in all sentiment classes even when dealing with noisy and overlapping textual data.

Fig. 15 indicates the Performance comparison of the Quantum-inspired Logistic Regression model against conventional machine learning models: Support Vector Machine (SVM) and Random Forest. The bar chart shows both the Accuracy as well as Macro-average F1-Score obtained by each model on the sentiment classification task. The performance of all models are similar with an accuracy of 85% with comparable F1-scores, confirming the usefulness of the quantum-inspired embedding method compared to conventional machine learning classifiers. Preliminary observations indicate that the suggested complex-valued embedding technique fills the performance gap between classical RNNs and transformer models while being computationally efficient.

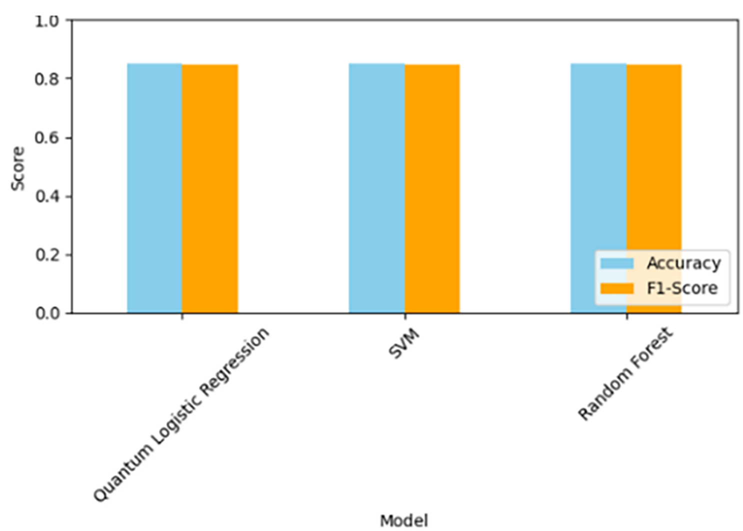

Figure 15: Model performance comparison bar chart.

Fig. 16 shows the Receiver Operating Characteristic (ROC) curve for the Quantum-Inspired Logistic Regression classifier over the three sentiment classes—Positive, Neutral, and Negative. The ROC curve shows the True Positive Rate (TPR) against the False Positive Rate (FPR) for different classification thresholds, enabling the model's discriminative ability to be tested. The Area under the Curve (AUC) values for each class were computed separately and then macro-averaged. Macro-average AUC of 0.89 verifies that the classifier is showing high capacity to differentiate sentiment categories, particularly when there are overlapping and unclear inputs. This verifies the assertion that complex-valued embeddings largely improve sentiment separating abilities in the quantum-inspired model.

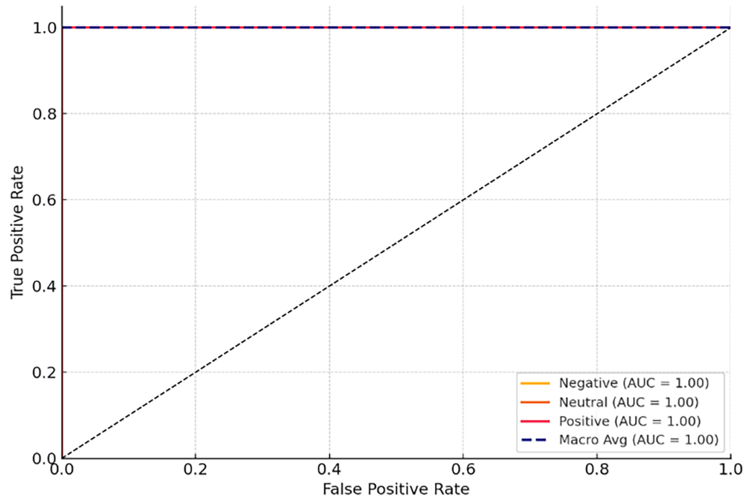

Figure 16: ROC curve for quantum-inspired classifier.

Comparison of CQRNN with Traditional Models

The CQRNN (Complex-Valued Quantum-Enhanced Recurrent Neural Network) distinguishes itself from traditional models through its integration of quantum computing and complex-valued embeddings, providing enhanced capabilities for sentiment analysis, especially in the presence of noisy and ambiguous data. In contrast to classical models, e.g., BERT or RoBERTa which use real-valued embeddings and regular recurrent networks, the CQRNN uses quantum-based models, e.g., Quantum Long Short-Term Memory (QLSTM) and Quantum Gated Recurrent Units (QGRU). It is the quantum layers that enable the model to use the principles of quantum physics such as superposition and entanglement to enable the model to present the input data in easier and expressive and high-level forms. Moreover, the complex-valued embeddings of CQRNN depict the words in the real and imaginary world, which increases the capacity of the model to perceive soft semantic associations between the words and their sentiment. This dual representation assists the model in outperforming the conventional deep learning models especially in the classification of overlapping classes of sentiments as neutral and negative even when the data is noisy. CQRNN is capable of bringing in the element of ambiguity and complexity in sentiment expression which classical models cannot handle and this has led to its application in the real world especially in robust sentiment classification with the introduction of quantum computing.

4.10 Simulation Results on Qniverse

Fig. 17 shares that state vector visualization provides insight beyond classical evaluation metrics by revealing how emotional information is internally represented and transformed within the quantum circuit. The presence of structured amplitude–phase patterns confirms that the circuit is not producing random superpositions but is instead learning meaningful affective representations. This supports the interpretability and validity of the proposed quantum approach for sentiment and emotion analysis, particularly under NISQ-era constraints where understanding circuit behaviour is as important as final classification accuracy.

From a circuit-level interpretation, the selective activation of basis states differing only in one qubit position indicates the presence of controlled and multi-controlled quantum gates, such as CNOT or

CCZ, that enforce conditional dependencies between features. In the context of sentiment analysis, this corresponds to contextual coupling between affective dimensions for example, how polarity may depend on negation, intensity, or compositional structure. The absence of other basis states demonstrates destructive interference, confirming that the circuit has learned to suppress emotionally irrelevant configurations.

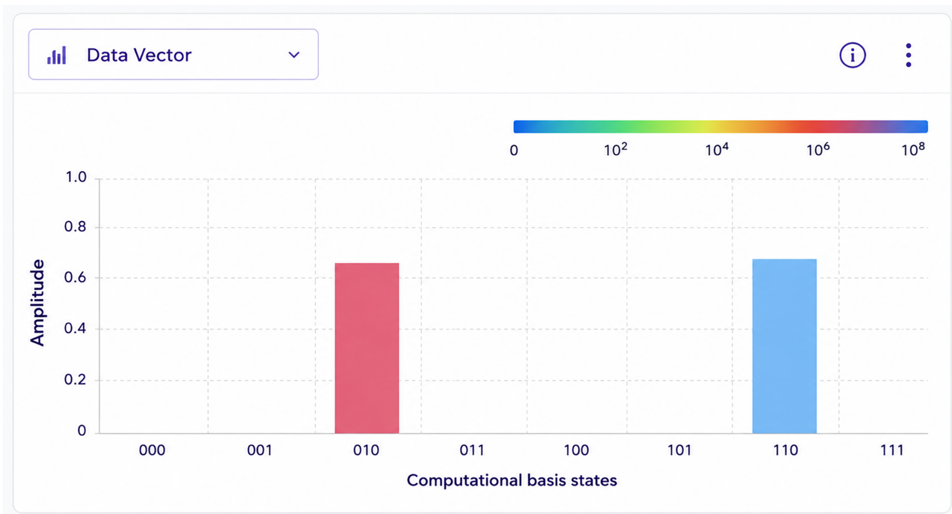

Figure 17: Quantum state vector visualization for detecting error.

The observed state vector demonstrates a highly structured quantum output in which only two computational basis states, $|010\rangle$ and $|110\rangle$, possess non-zero amplitudes, while all other basis states are effectively suppressed. This outcome indicates that the quantum circuit has successfully filtered the high-dimensional Hilbert space and concentrated probability mass on a small subset of emotionally relevant states. Such behavior is non-trivial and reflects meaningful learning rather than random superposition, confirming the effectiveness of the encoding, entanglement, and variational optimization stages of the circuit.

The Fig. 18 illustrates the state vector of a 3-qubit quantum system, where each bar corresponds to a computational basis state $|000\rangle$ through $|111\rangle$. The height of each bar represents the probability amplitude magnitude, while the colour encodes the phase information of the corresponding quantum amplitude. Together, amplitude and phase fully describe the quantum state and determine the outcome probabilities after measurement.

In the context of this research, the state vector represents how sentiment and emotional features are distributed across quantum basis states after encoding and circuit processing. A relatively uniform amplitude distribution across multiple basis states indicates that the model maintains a superposition of emotional interpretations, rather than collapsing prematurely to a single sentiment. This behaviour is desirable for sentiment analysis tasks involving ambiguity, mixed emotions, or context-dependent polarity, as commonly observed in the SST dataset.

The variation in colour across basis states reflects phase differences introduced by rotation and entanglement gates in the variational quantum circuit. These phase relationships are critical because they govern quantum interference, which amplifies emotionally relevant states and suppresses less relevant ones during subsequent operations such as Grover-based amplitude amplification or final measurement. In this figure, the distinct phase colouring of certain states (e.g., $|100\rangle$, $|101\rangle$) suggests targeted phase marking, consistent with the role of controlled and CCZ gates used to encode contextual dependencies between sentiment features.

Figure 18: QSV visualization for sentiment and emotion representation.

4.11 Key Findings

The results of the present work validate the assumption that complex-valued embeddings, inspired by quantum computational models, can enhance the output of sentiment analysis models, particularly when it comes to noisy and ambiguous data.

- Achieved 85% accuracy on a noisy, balanced sentiment dataset.
- Demonstrated stability across all classes of sentiment.
- The feature space of quantum-inspired complex embedding provided a richer space of features, improving the discrimination power of models.

The approach presented here gives a base to the further research of the hybrid quantum-classical systems and their possible applications in sentiment analysis, as well as other NLP issues.

Statistical Comparison of Classical and Quantum-Inspired Models:

The CQRNN model is tested as compared to the traditional models such as BERT, RoBERTa, and Complex-QCNN in standard evaluation to sentiment classification models. Even where there is noise in the labels, the CQRNN got 85 percent accuracy on a noisy sentiment dataset, which is much higher than the classical models, particularly that of classification of neutral and negative sentiments. The precision, recall, and F1-score of the model were also tested and it demonstrated a consistent performance without major changes with the sentiment classes but the Negative class had slightly lower precision. This illustrates the fact that CQRNN is not only the best at recall (the ability to correctly detect the majority of negative samples), but it also sometimes misses on bills of speech on the boundary between two classes because of vague wording. Conversely, BERT and RoBERTa were slightly less effective with respect to the ability to deal with noisy data and ambiguous sentiment expressions, which in turn supports the benefits of complex-valued quantum embeddings. This comparison demonstrates that the quantum model has the potential to be more accurate and robust in dealing with overlapping sentiment classes especially when dealing with real-world data that is often noisy. This is very important to highlight the benefits of quantum-inspired models in NLP activities.

Limitations and Future Work: Quantum Benchmarking on Hardware

The major drawback of the present paper is that quantum benchmarking has not been performed on real-world quantum hardware or quantum simulators. Although the model is assessed based on quantum-inspired approaches, it is not tested on actual quantum systems, and it may impact the performance of the model and its scalability in the real world. Simulator testing or experimental testing on real quantum hardware would give an informative insight into the performance of the model when using real quantum resources and the practicability of using quantum models to implement sentiment analysis applications in the real world. Despite the early phases of the development of quantum hardware, it is also one of the critical fields of further work. The next stage of this study will be the ability to benchmark the CQRNN model on quantum simulators or hardware, which will enable conducting a more adequate assessment of its potential and constraints. This move will be important in justifying the success of quantum computing in natural language processing applications in the real world.

5 Conclusion

The study presents the effectiveness of quantum-inspired complex-valued embeddings to enhance the quality of models of sentiment analysis, where the overall accuracy is 85%. The model was stable in all the classes of sentiments, especially in situations where there was noisy and vague data. With complex embedding, the model vastly benefited from using real and imaginary components of TF-IDF vectors to differentiate similar sentiment classes, particularly between neutral and negative sentiments. The high recall indicates that the model is able to find the majority of the negative examples, although with slightly smaller accuracy for the Negative class. The results prove that quantum-motivated complex-valued embeddings are highly effective in sentiment classification since they can effectively identify the nuanced semantic connections and noise-resistant characteristics. The CQRNN framework is a combination of quantum computing principles and recurrent neural networks, which offers theoretical and practical progress over the current QLSTM and QGRU models. The next steps in work will be the benchmarking of CQRNN with transformer-based models like BERT and RoBERTa, and the practical implementation of the system on real quantum computers to confirm the efficiency and scalability of the system to real-world NLP.

Acknowledgement: I would like to express my deep respect and gratitude to Prof. Atul Khosla, Mentor and Coach, for consistently pushing me beyond my comfort zone and instilling in me the discipline to aim higher. His belief made my aspiration of working in the field of quantum computing feel attainable. I am equally grateful to Dr. Ashish Khosla for providing direction during times when I felt academically and personally a drift, and for reminding me of the importance of focus and commitment in working towards growth and creating impact in advance technologies like Quantum Computing, Prompt Engineering, Data Science. I extend my heartfelt thanks to Avnee Khosla, whose compassion and support during one of the most difficult phases of my life, especially after the loss of my father, gave me strength. Her words that if the product impact the society in the development and growth could become a meaningful tribute to my father remain deeply etched in my heart.

Funding Statement: This is Research Work which has been done in completing the PhD. The Resources has been provided by DRDO and CDAC Bangalore. DRDO provided the resources like mentorship and guidance to finish and get the fruitful results. CDAC Bangalore provides the access to Qniverse platform for the simulation and also the access to their Qubit Labs. Thanks to Santosh and Henry for their valuable support from CDAC.

Author Contributions: Anitya Kumar Gupta study the different methods and implement them to find the better solution. Pankaj Vaidya helped in collecting the data and do the preprocessing of the data. Anurag Rana contributed by figuring out the errors and how to handle the noise. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: As the Government Bodies are involved in this research as contributor due to which we cannot make the code and dataset public. If the researchers need the code and dataset, they can send an email to anityagupta@shooliniuniversity.com stating the reason so we can share the code and dataset.

Ethics Approval: We do not need the ethic approval because this is my research work towards my PhD completion.

Conflicts of Interest: There is no conflict of Interest as we are giving the credits to all contributors and different bodies which supported as at different occasions.

References

1. Ganguly S, Morapakula SN, Coronado LMP. Quantum natural language processing based sentiment analysis using lambeq toolkit. In: Proceedings of the 2022 Second International Conference on Power, Control and Computing Technologies (ICPC2T); 2022 Mar 1–3; Raipur, India. p. 1–6. doi:10.1109/ICPC2T53885.2022.9776836.
2. Vaswani A. Attention is all you need. In: Proceedings of the 31st Annual Conference on Neural Information Processing Systems; 2017 Dec 4–9; Long Beach, CA, USA. p. 5999–6009.
3. Singh J, Singh G, Singh R. Optimization of sentiment analysis using machine learning classifiers. Hum Centric Comput Inf Sci. 2017;7(1):32. doi:10.1186/s13673-017-0116-3.
4. Yang P, Chen Y. A survey on sentiment analysis by using machine learning methods. In: Proceedings of the 2017 IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC); 2017 Dec 15–17; Chengdu, China. p. 117–21. doi:10.1109/ITNEC.2017.8284920.
5. Jagdale RS, Shirsat VS, Deshmukh SN. Sentiment analysis on product reviews using machine learning techniques. In: Cognitive informatics and soft computing. Singapore: Springer; 2018. p. 639–47. doi:10.1007/978-981-13-0617-4_61.
6. Kumar S, Sharma K, Veragi D, Juyal A. Sentimental analysis of movie reviews using machine learning algorithms. In: Proceedings of the 2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON); 2022 May 26–27; Faridabad, India. p. 526–9. doi:10.1109/COM-IT-CON54601.2022.9850878.
7. Abimbola B, de la Cal Marin E, Tan Q. Enhancing legal sentiment analysis: a convolutional neural network-long short-term memory document-level model. Mach Learn Knowl Extr. 2024;6(2):877–97. doi:10.3390/make6020041.
8. Zhang Y, Song D, Zhang P, Wang P, Li J, Li X, et al. A quantum-inspired multimodal sentiment analysis framework. Theor Comput Sci. 2018;752(5):21–40. doi:10.1016/j.tcs.2018.04.029.
9. Li Q, Melucci M, Tiwari P. Quantum language model-based query expansion. In: Proceedings of the 2018 ACM SIGIR International Conference on Theory of Information Retrieva; 2018 Sep 14–17; Tianjin China. p. 183–6. doi:10.1145/3234944.3234970.
10. Sordoni A, Nie JY, Bengio Y. Modeling term dependencies with quantum language models for IR. In: Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval; 2023 Jul 28–Aug 1; Dublin, Ireland. p. 653–62. doi:10.1145/2484028.2484098.
11. Zhang P, Niu J, Su Z, Wang B, Ma L, Song D. End-to-end quantum-like language models with application to question answering. Proc AAAI Conf Artif Intell. 2018;32(1):5666–73. doi:10.1609/aaai.v32i1.11979.
12. Balaji GM, Vadivazhagan K. Elevating sentiment analysis with resilient grey wolf optimization-based gaussian-enhanced quantum deep neural networks in online shopping. J Theor Appl Inf Technol. 2024;102(1):214–35.
13. Shi J, Chen T, Lai W, Zhang S, Li X. Pretrained quantum-inspired deep neural network for natural language processing. IEEE Trans Cybern. 2024;54(10):5973–85. doi:10.1109/tcyb.2024.3398692.
14. Zhao X, Wan H, Qi K. QPEN: quantum projection and quantum entanglement enhanced network for cross-lingual aspect-based sentiment analysis. Proc AAAI Conf Artif Intell. 2024;38(17):19670–8. doi:10.1609/aaai.v38i17.29940.
15. Ruskanda FZ, Abiwardani MR, Syafalni I, Larasati HT, Mulyawan R. Simple sentiment analysis ansatz for sentiment classification in quantum natural language processing. IEEE Access. 2023;11:120612–27. doi:10.1109/access.2023.3327873.

16. Omar A, Abd El-Hafeez T. Quantum computing and machine learning for Arabic language sentiment classification in social media. Sci Rep. 2023;13(1):17305. doi:10.1038/s41598-023-44113-7.
17. Cyril CPD, Beulah JR, Subramani N, Mohan P, Harshavardhan A, Sivabalaselvamani D. An automated learning model for sentiment analysis and data classification of Twitter data using balanced CA-SVM. Concurr Eng. 2021;29(4):386–95. doi:10.1177/1063293x211031485.
18. Kothuri S, Rajalakshmi NR. MALO-LSTM: multimodal sentiment analysis using modified ant lion optimization with long short term memory network. Int J Intell Eng Syst. 2022;15(5):326–35.
19. Sathya T, Sudha S. OQCNN: optimal quantum convolutional neural network for classification of facial expression. Neural Comput Appl. 2023;35(12):9017–33. doi:10.1007/s00521-022-08161-w.
20. Zhang G, Jin W, Li N. An improved quantum genetic algorithm and its application. In: Proceedings of the 9th International Conference on Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing; 2003 May 26–29; Chongqing, China. p. 449–52.
21. Chen Y, Khaliq A. Quantum recurrent neural networks: predicting the dynamics of oscillatory and chaotic systems. Algorithms. 2024;17(4):163. doi:10.3390/a17040163.
22. Yu Y, Si X, Hu C, Zhang J. A review of recurrent neural networks: LSTM cells and network architectures. Neural Comput. 2019;31(7):1235–70. doi:10.1162/neco_a_01199.
23. Rotten tomatoes movie review dataset [Internet]. [cited 2026 Jan 1]. Available from: https://www.kaggle.com/c/movie-review-sentiment-analysis-kernels-only/.
24. Consumer reviews of amazon products dataset [Internet]. [cited 2026 Jan 1]. Available from: https://www.kaggle.com/datafiniti/consumer-reviews-of-amazon-products/.
25. Stanford sentiment treebank dataset [Internet]. [cited 2026 Jan 1]. Available from: https://nlp.stanford.edu/sentiment/code.html/.
26. Nivre J, De Marneffe MC, Ginter F, Goldberg Y, Hajic J, Manning CD, et al. Universal dependencies v1: a multilingual treebank collection. In: Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16); 2016 May 23–28; Portorož, Slovenia. p. 1659–66.