




## Article

# An Effective Strategy for Sentiment Analysis Based on Complex-Valued Embedding and Quantum Long Short-Term Memory Neural Network

Zhulu Chu <sup>1</sup>, Xihan Wang <sup>1</sup>, Meilin Jin <sup>1</sup>, Ning Zhang <sup>2</sup>, Quanli Gao <sup>1,\*</sup> and Lianhe Shao <sup>1,\*</sup>

<sup>1</sup> School of Computer Science, Xi'an Polytechnic University, Xi'an 710600, China;

210721073@stu.xpu.edu.cn (Z.C.); xihan\_wang@xpu.edu.cn (X.W.); kaelyn@stu.xpu.edu.cn (M.J.)

<sup>2</sup> Library Information and Digital Library, National Library of China, Beijing 100081, China; zning@nlc.cn

\* Correspondence: gaoquanli@nwu.edu.cn (Q.G.); shaolianhe@xpu.edu.cn (L.S.)

**Abstract:** Sentiment analysis aims to study, analyse and identify the sentiment polarity contained in subjective documents. In the realm of natural language processing (NLP), the study of sentiment analysis and its subtask research is a hot topic, which has very important significance. The existing sentiment analysis methods based on sentiment lexicon and machine learning take into account contextual semantic information, but these methods still lack the ability to utilize context information, so they cannot effectively encode context information. Inspired by the concept of density matrix in quantum mechanics, we propose a sentiment analysis method, named Complex-valued Quantum-enhanced Long Short-term Memory Neural Network (CQLSTM). It leverages complex-valued embedding to incorporate more semantic information and utilizes the Complex-valued Quantum-enhanced Long Short-term Memory Neural Network for feature extraction. Specifically, a complex-valued neural network based on density matrix is used to capture interactions between words (i.e., the correlation between words). Additionally, the Complex-valued Quantum-enhanced Long Short-term Memory Neural Network, which is inspired by the quantum measurement theory and quantum long short-term memory neural network, is developed to learn interactions between sentences (i.e., contextual semantic information). This approach effectively encodes semantic dependencies, which reflects the dispersion of words in the embedded space of sentences and comprehensively captures interactive information and long-term dependencies among the emotional features between words. Comparative experiments were performed on four sentiment analysis datasets using five traditional models, showcasing the effectiveness of the CQLSTM model.

**Keywords:** sentiment analysis; density matrix; complex-valued neural network; quantum-enhanced long short-term memory neural network

**MSC:** 68T07; 68T50



**Citation:** Chu, Z.; Wang, X.; Jin, M.; Zhang, N.; Gao, Q.; Shao, L. An Effective Strategy for Sentiment Analysis Based on Complex-Valued Embedding and Quantum Long Short-Term Memory Neural Network. *Axioms* **2024**, *13*, 207. <https://doi.org/10.3390/axioms13030207>

Academic Editors: Florin Felix Nichita and Oscar Humberto Ross

Received: 3 February 2024

Revised: 6 March 2024

Accepted: 19 March 2024

Published: 21 March 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Sentiment analysis aims to determine the speaker's, author's or another subject's viewpoint or attitude towards a particular topic, document or event. It may involve judgment, evaluation, emotional states or emotional expressions. Currently, the research in text sentiment analysis spans various prominent scientific and technological directions, particularly in the applications of natural language processing and text mining technologies. Additionally, sentiment analysis finds applications across diverse domains, such as sociology, finance, marketing and healthcare services.

Currently, sentiment analysis methods primarily include those based on sentiment lexicons, machine learning and deep learning. Lexicon-based methods involve the prior establishment of a sentiment lexicon, followed by calculations based on this lexicon. The methods based on machine learning require manual feature extraction as a preliminary

step. Subsequently, machine learning algorithms, such as naïve Bayes (NB), maximum entropy (ME), k-nearest neighbor classification (KNN) and conditional random field (CRF), are employed to conduct sentiment classification on the text. The method based on deep learning does not require artificial feature extraction and has strong semantic expression ability. Commonly used neural network models are convolutional neural network (CNN), recurrent neural network (RNN), LSTM (long-term short memory), GRU (gated recurrent unit), etc.

In recent years, quantum theory, as a mathematical framework for simulating complex interactions and dynamics in quantum physics, has been applied to construct text representations for various information retrieval (IR) and natural language processing (NLP) tasks [1–4]. For instance, the quantum language model (QLM) [1] represents queries or documents as density matrices on quantum probability spaces, using metrics based on density matrices as ranking functions. The neural network quantum language model (NNQLM) [5] based on neural network builds an end-to-end question–answering network using a density matrix representation and modeling question–answer pairs. Zhang et al. [6] utilized an enhanced version of QLM for sentiment analysis on Twitter. These quantum theory-based models can be considered as an extension of classical methods, as they can capture the inherent complexity in interactions. This prompts us to explore the use of quantum theory as a theoretical basis to capture utterance interactions. Recently, a quantum long short-term memory neural network (QLSTM) [7,8] has been proposed. Will the combination of a complex-valued neural network and QLSTM have positive effects on sentiment analysis?

In this paper, we propose the Complex-valued Quantum-enhanced Long Short-term Memory Neural Network (CQLSTM), which combines complex-valued embedding with a quantum-enhanced long short-term memory neural network to jointly capture the interactions between words and sentences for text sentiment analysis. The model employs complex-valued embedding to represent each word, where each embedding vector serves as the observed state of the word. Furthermore, a sentence corresponds to a mixed state represented by the density matrix. Compared to traditional word embedding approaches, the combination of complex-valued embedding and density matrices allows for the encoding of more semantic dependencies and feature interactions between sentiment words. Inspired by quantum measurement theory, the model employs a complex-valued neural network to extract the interactions between words. This interaction is integrated into two input gates of QLSTM, taking the density matrix as input. The hidden states obtained twice are combined and fed back to the measurement function, ultimately determining the emotional state of the sentence.

The proposed CQLSTM model effectively models both semantic and sentiment polarity, achieving notable performance in sentiment analysis tasks. In comparison to other neural network models, the CQLSTM approach excels in extracting advanced semantic features, thereby enhancing the model’s predictive accuracy and effectiveness. This improvement contributes to the practicality and reliability of the model.

Our contributions are as follows:

- Grounded in quantum computation theory and complex-valued neural networks, we introduce the Complex-valued Quantum-enhanced Long Short-term Memory Neural Network;
- We are integrating complex-valued embedding with a quantum-enhanced long short-term memory neural network;
- The Complex-valued Quantum-enhanced Long Short-term Memory Neural Network is developed and used for sentiment analysis.

The remainder of the paper is structured as follows: Section 2 conducts a comprehensive review of the literature and provides a summary of the related work. In Section 3, we meticulously detail the materials and methods employed in this study. Section 4 engages in a comprehensive discussion and analysis of our experimental results. Finally, Section 5 encapsulates our findings and draws definitive conclusions.

## 2. Related Work

The related work primarily presents the findings and research in sentiment analysis, encompassing lexicon-based sentiment analysis methods and neural network frameworks. The key discoveries, methodologies and limitations from these studies are synthesized. Considering these challenges, the rationale for introducing CQLSTM is outlined and the advantages of CQLSTM are elucidated.

### 2.1. Previous Research Results

The lexicon-based method requires the initial creation of a sentiment lexicon. This involves manually annotating sentiment direction and intensity within the established lexicon. The annotated lexicon is employed for the text sentiment analysis [9–11]. This approach heavily relies on the quality of the sentiment lexicon and the manually constructed sentiment rules. This approach will also test the lexicon's quality, the applicability to specific domain and the researchers' level of expertise. The advantage of this method is that it generally does not necessitate storing vast corpora or designing complex training algorithms; however, its drawback lies in its low efficiency.

Compared to lexicon-based methods, machine learning network models exhibit greater adaptability and generalization capabilities, as well as higher predictive accuracy and computational efficiency. The advantages inherent in machine learning network models position them as a highly promising and widely applicable technology in sentiment analysis. This has emerged as a focal point of intense research activity in the current academic landscape. The sentiment analysis methods based on machine learning utilize statistical knowledge [12] and machine learning algorithms. They train convergent learning models based on annotated sentiment labels and effective text features to accomplish the classification of subjective documents. A myriad of techniques has been explored in sentiment analysis, with notable examples including support vector machine (SVM), random forest (RF), convolutional neural networks (CNN) [13] and recurrent neural network (RNN). This approach treats sentiment analysis as a supervised classification task, typically involving large-scale annotated datasets and training classification models.

While the functionality of machine learning models is relatively straightforward, the process of feature extraction in machine learning is typically complex. The results of machine learning classification are influenced by the quality of the extracted features. In addition, there is a common issue of low generalization ability.

In recent years, deep learning has driven a new wave of artificial intelligence, playing an irreplaceable role in fields such as natural language processing. Its main concept involves constructing a neural network model for text sentiment classification. Sentiment analysis based on deep learning has attracted an increasing number of researchers due to its powerful performance advantages. Kim [14] was the first to apply convolutional neural networks (CNN) to sentence classification tasks. Despite having only one convolutional layer and minimal tuning, the neural network achieved significant results on datasets, such as movie reviews and consumer comments, demonstrating the robust learning capability of deep convolutional networks. With the rise of attention mechanisms in the image domain, scholars have gradually introduced them into the field of natural language processing [14–17]. Meanwhile, some researchers are dedicated to combining convolutional neural networks with long short-term neural networks. This hybrid approach has achieved promising results compared to individual methods [18,19].

While deep learning has achieved significant success in the field of sentiment classification, scholars have identified challenges with data-driven models, including poor interpretability and the requirement for extensive training data.

### 2.2. Motivation for Complex-Valued Quantum-Enhanced Long Short-Term Neural Network

The quantum mechanics theory serves as a formal mathematical framework for studying the interactions and evolution of microscopic particles and their constituent systems. In recent years, quantum mechanics has found applications in the field of natural language

processing. Numerous studies have revealed an intrinsic unity between the mathematical framework of quantum mechanics and the foundational principles of natural language processing. The initial effort to integrate quantum mechanics and natural language processing was pioneered by van Rijsbergen [20], who introduced an unified framework for information retrieval. This framework consolidated traditional information retrieval models within Hilbert space, enhancing the generality of information retrieval models. Building on van Rijsbergen's work, researchers embarked on exploring the intersection of natural language processing and quantum mechanics. This research falls into the following two main categories: the first involves investigating the semantics and syntax of natural language to develop a comprehensive understanding from an abstract level. While the second delves into exploring quantum-like phenomena within natural language. The quantum-like phenomena encompass non-classical probabilistic events, interference phenomena during information retrieval and quantum entanglement phenomena within the dependency relationships between words. Researchers aim to incorporate these quantum-like phenomena into models and algorithms for natural language processing. Typically, such quantum-like algorithms or quantum-inspired algorithms do not necessitate deployment on quantum computing platforms. Relevant natural language processing tasks can be efficiently executed using classical computing platforms. However, some algorithms are constructed using tensor network architectures, offering the potential for deployment on quantum platforms. Stein [21] applied quantum natural language processing (QNLP) to sentiment analysis in the financial field, aiming to explore how to use quantum computing methods to improve the effectiveness of emotion analysis in the financial field.

The density matrix is a fundamental concept in quantum statistical physics. In quantum mechanics, when a system is in its initial state, its description relies on a wave function or a state vector. However, when the system is in a mixed state, its representation shifts to a density matrix. Recently, density matrix-based approaches have been employed in natural language processing tasks, including question and answer (QA) [5], information retrieval (IR) [1,22] and sentiment analysis (SA) [6]. Sordoni, Nei and Bengio [1] successfully incorporated quantum probability into language modeling (LM), introducing a quantum language model (QLM). The model represents a query or document as a density matrix within a quantum probabilistic space and computes density matrix-based metrics to function as a ranking mechanism. In this process, the accurate estimation of the density matrix is crucial to its operational efficiency. QLM is rooted in the mathematical framework of quantum computation theory, aiming to simulate natural language by drawing inspiration from the analogies between natural language and quantum systems. These analogies include word polysemy and particle superposition, semantics in specific scenarios and particle collapse to specific ground states after measurement, semantic changes over time and space, continuous evolution of particles over time, etc. [23]. Quantum language models are more closely integrated with natural language than classical neural networks, offering enhanced interpretability and the potential for exponential acceleration [24]. Moreover, these quantum-inspired language models operate within the mathematical framework of Hilbert space. This distinctive characteristic imparts a heightened level of expressive capabilities, contributing to their robust performance in various linguistic tasks.

The neural network-based quantum language model (NNQLM) [5] establishes an end-to-end network for question answering (QA), aiming to collectively model a question-answer pair through density matrix representations. This involves the development of a neural network-based quantum language model tailored for QA tasks. The model designs a density matrix representation based on word embedding for individual sentences. Similarity is then computed through a joint representation of two sentences, facilitating the matching of question and answer functions. Zhang [6] applied a modified version of the quantum language model (QLM) for Twitter sentiment analysis. Additionally, in [25] the authors employed density matrices in the click-through rate (CTR) prediction task. They achieved favorable results on the Criteo and Avazu datasets using the density matrix-based convolutional neural network (CNN) model. These endeavors collectively showcase the

capability of density matrices to effectively capture implicit interactions. Quantum theory-based models, such as those utilizing density matrices, can be viewed as a generalization of classical approaches, offering the capacity to comprehend the inherent complexity of interactions. This motivates further exploration into leveraging quantum theory as a foundational framework for capturing intra-utterance and inter-utterance interactions, aiming to extract more comprehensive feature interaction information.

Quantum neural networks face the challenge of gradient vanishing when trained on large datasets. The CQLSTM framework employed in this paper effectively circumvents the issues of gradient explosion and gradient vanishing commonly encountered in general recurrent network models. In the model's input section, complex-valued encoding is utilized to enable a quantum bit to carry and transmit multiple classical quantum bits (i.e., a quantum bit contains  $|0\rangle$  and  $|1\rangle$  [26,27] or linear combinations of states). Transmitting one quantum bit can convey two classical pieces of information, significantly expanding the channel capacity to twice that of classical transmission channels. With the incorporation of multi-layer activation functions, diverse data are mapped to distinct amplitudes, enhancing the network's ambiguity and improving the accuracy and certainty of network pattern recognition.

### 3. Materials and Methods

In this paper, complex-valued neural network are jointly used with quantum-enhanced long short-term memory neural network. The model sub-modules complex-valued linear layer and QLSTM are constructed using Python's pytorch. In this section, we describe the proposed CQLSTM model in detail.

#### 3.1. Basic Symbols and Concepts in Quantum Theory

In quantum theory, the probabilistic space is represented by an infinite Hilbert space denoted as  $H$  [28]. For simplicity and consistency with previous quantum-inspired models [2,5], we confine our model to finite vector spaces over real numbers in  $R$ . In quantum computing, a qubit is the basic unit, and the Dirac symbol is usually used to describe a qubit, such as  $|0\rangle$  and  $|1\rangle$ . A qubit can be a linear combination of ground states  $|0\rangle$  and  $|1\rangle$ , such as follows:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \quad (1)$$

where  $|\psi\rangle$  is often called a superposition state, and both  $\alpha$  and  $\beta$  are complex-valued numbers.  $|0\rangle$  and  $|1\rangle$  can be obtained from measuring, where  $|\alpha|^2$  is the probability of obtaining  $|0\rangle$ ,  $|\beta|^2$  is the probability of obtaining  $|1\rangle$ , and satisfy  $|\alpha|^2 + |\beta|^2 = 1$ . According to the following Euler's formula:  $e^{ix} = \cos x + i \sin x$ , the equation can also be expressed as follows:

$$|\psi\rangle = e^{i\gamma} \left( \cos \frac{\theta}{2} |0\rangle + e^{i\varphi} \sin \frac{\theta}{2} |1\rangle \right), \quad (2)$$

where  $\gamma, \varphi$  are all real numbers. Additionally,  $e^{i\gamma}$  can be omitted since it does not have any observable effect. In the framework of Dirac's notation, a state vector  $\varphi$  and its transpose are denoted as a Ket  $|\varphi\rangle$  and a Bra  $\langle\varphi|$ , respectively [26,27]. The inner product between two state vectors  $|\varphi_1\rangle$  and  $|\varphi_2\rangle$ , is represented as  $\langle\varphi_1|\varphi_2\rangle$ . Furthermore, the wave function's representation, serving as a mathematical description of the quantum state in Hilbert space, is articulated through the inner product  $\varphi(x) = \langle x|\varphi\rangle$ .

Quantum probability (QP) serves as a broader framework encompassing classical probability theory. In QP, an event is conceptualized as a subspace of Hilbert space, denoted by an orthogonal projector  $\Pi$ . Assuming  $|u\rangle$  represents a unit vector ( $\|u\|^2 = 1$ ), the projector  $\Pi$  in the direction  $u$  is expressed as  $|u\rangle\langle u|$ . A density matrix  $\rho = \sum_i p_i |u_i\rangle\langle u_i|$  effectively captures the quantum state. This density matrix  $\rho$  is characterized by symmetry ( $\rho^T = \rho$ ), positive semi-definite ( $\rho \geq 0$ ) and a trace of 1. The quantum probability measure is intricately linked to the density matrix, adhering to the following two key conditions: (1) for each projector  $|u\rangle\langle u|$ ,  $\mu(|u\rangle\langle u|) \in [0, 1]$ , (2) for any orthonormal basis



$\{|e_i\rangle\}$ ,  $\sum_{i=1}^n \mu(|e_i\rangle\langle e_i|) = 1$ . Gleason's theorem rigorously establishes the existence of a mapping function, revealing  $\mu(|u\rangle\langle u|) = \text{tr}(\rho |u\rangle\langle u|)$  for any  $|u\rangle$ .

Events are redefined not merely as subsets but as subspaces, specifically as projection sets on subspaces. Establishing a bijection between the stator space and the projection, it can be accurately asserted that  $|u\rangle$  itself constitutes a fundamental event. In general, any  $\text{Ket } |v\rangle = \sum_i v_i |u_i\rangle$  is termed a superposition of  $\{|v_i\rangle\}$ , where  $\{|v_1\rangle, \dots, |v_n\rangle\}$  forms an orthonormal basis.

The quantum probability measure  $\mu$  serves as a broader concept than the classical probability measure. Hence, it adheres to the condition  $\sum_i \mu(|u_i\rangle\langle u_i|) = 1$ . Gleason's theorem establishes that in any real vector space with a dimension greater than 2, there exists a one-to-one correspondence between the quantum probability measure  $\mu$  and the density matrix  $\rho$ . This correspondence is expressed as follows:

$$\mu_\rho(|v\rangle\langle v|) = \text{tr}(\rho |v\rangle\langle v|), \quad (3)$$

Gleason's theorem posits that the density matrix can be interpreted as a suitable quantum extension of classical probability distribution. It assigns a quantum probability to each infinite two-tuple.

When given a density matrix  $\rho$ , two-tuples can be identified through the decomposition of its eigenvectors, establishing a two-tuple for each eigenvector. This decomposition is expressed as  $\rho = R \Lambda R^T = \sum_{i=1}^n \lambda_i |r_i\rangle\langle r_i|$ , where  $|r_i\rangle$  signifies the eigenvector, and  $\lambda_i$  represents the corresponding eigenvalue. It's noteworthy that a diagonal density matrix can effectively represent a traditional probability distribution.

### 3.2. Overall Framework

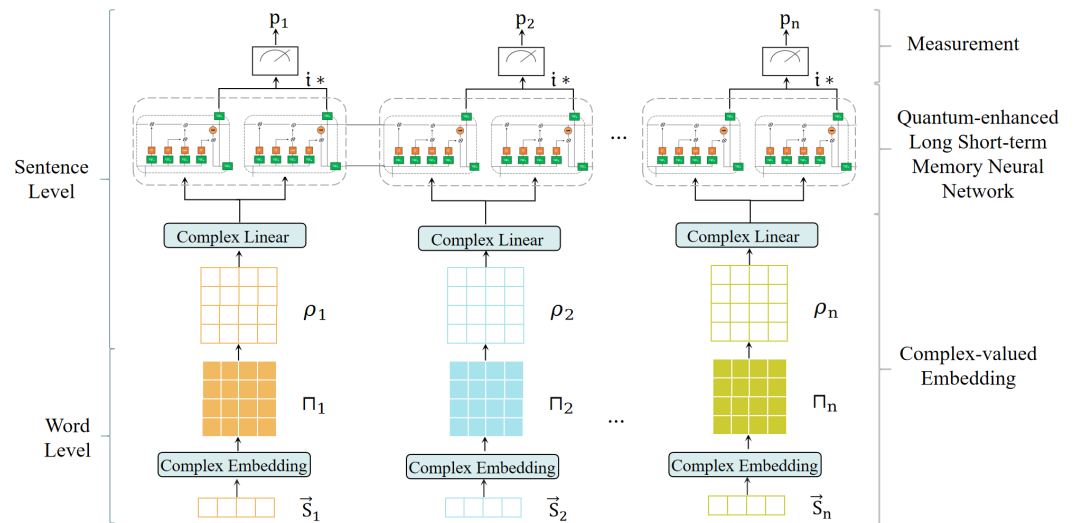
Figure 1 illustrates the architecture of the CQLSTM model. The model takes a pre-processed token sequence as input, incorporating operations such as truncation, padding, word index mapping, word segmentation and case conversion. Once the data are input into the framework, it undergoes processing through a complex-valued embedding layer, consisting of a real embedding layer and an imaginary embedding layer. This layer transforms the sequence into a complex-valued word vector at the word level. Subsequently, the complex-valued word vector is projected into a high-dimensional space by a complex-valued linear layer, representing the sentence vector through the density matrix. In order to input the sentence vector into QLSTM for training, we take the sentence vector modulo operation. The value after modulo operation also includes real parts and imaginary parts. The resulting sentence vector is fed into the input gate of the Complex-valued Quantum-enhanced Long Short-term Memory Neural Network, which is composed of both real and imaginary QLSTM components for advanced feature extraction. The feature vector obtained is combined and the quantum measurement process is employed to fit and predict the data, ultimately yielding the final classification result. This approach retains the robust coding performance of complex-valued encoding while enhancing the framework's predictive capabilities through the integration of the QLSTM model.

#### 3.2.1. Complex-Valued Embedding Module

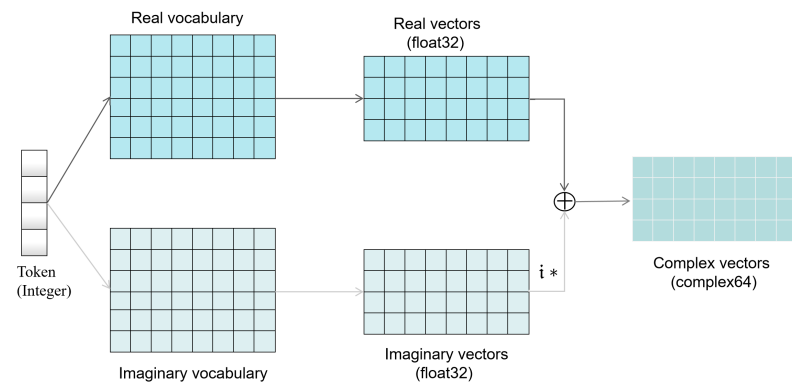
In this module, the input comprises a token sequence derived from pre-processed text, presented as an integer vector. The complex-valued embedding module encompasses a real embedding layer and an imaginary embedding layer. As depicted in Figure 2, the complex-valued embedding process unfolds by sequentially passing the text token sequences through these two embedding layers. Ultimately, the computation of the complex word vector representation for each token is achieved using the following equation:

$$[w_1, w_2, \dots, w_n] = [r_1, r_2, \dots, r_n] + i \times [i_1, i_2, \dots, i_n], \quad (4)$$

The imaginary part is denoted by  $i$ .



**Figure 1.** The architecture of CQLSTM. For explanation see text.



**Figure 2.** The process of complex-valued embedding.

The complex-valued embedding layer converts data into quantum bits. In contrast to employing one-hot [29] vectors for word embedding, our approach involves the utilization of complex-valued embedding. Specifically, RoBERTa [30] is employed as the real part of the complex embedding, while the imaginary part is represented by a self-trained word embedding. The real embedding and imaginary embedding capture different types of textual features, fully leveraging the heterogeneous nature of complex-valued neural networks. This approach utilizes a high-dimensional space to express more global information. Specifically, complex-valued embedding maps token IDs corresponding to words into a complex vector space, mirroring the process of constructing quantum states in quantum computing. Each word undergoes a mapping from a discrete space to a high-dimensional Hilbert space. This transition culminates in the creation of a complex-valued column vector, symbolizing the rich and nuanced representation of the respective word in the quantum-inspired framework.

In this module, the word embedding information obtained from the complex-valued embedding module needs to be translated into the complex-valued space. In the preceding layer, the complex-valued embedding module has already mapped the words into a high-dimensional complex-valued vector space. The computation of the density matrix representation for the sentence unfolds as a pivotal process in this intricate procedure as follows:

$$\rho = \sum_i p_i |u_i\rangle\langle u_i|, \quad (5)$$

where the weights of the words  $p_i$  are trainable through the attentional mechanism,  $|u_i\rangle$  is a pure state vector with probability  $p_i$ . By default, each word is assigned an equal weight.

The density matrix, which leverages quantum bits for quantum computation, has the ability to encode more semantic dependencies. It represents diverse potential combinations of 0 and 1, concatenating them to offer increased computational power compared to an equivalent number of binary files.

### 3.2.2. Complex-Valued Quantum-Enhanced Long Short-Term Memory Neural Network Module

Quantum state evolution elucidates the transition of a quantum system from one state to another. In the realm of quantum computing theory, this transformation of quantum states and density matrices is accomplished through the application of quantum gates. Each quantum gate is associated with a unitary matrix whose dimensions align with the number of quantum bits ( $n = 2^m$ ), upon which it operates. In this study, we emulate the dynamics of quantum systems employing complex-valued linear and quantum-enhanced long short-term memory (QLSTM) networks. We configure the evolutionary modules to have inputs and outputs conforming to  $n \times n$ . Consequently, the original dimensionality of the quantum system remains unchanged post feature learning within the sentence.

The core component of a complex-valued neural network is the complex-valued linear layer, often denoted as the fully connected layer. Within this layer, each neuron forms connections with all neurons in the preceding layer, creating a comprehensive network of interconnections. The computational process is described as follows:

$$f = WX + b, \quad (6)$$

where  $W$  is the weight matrix in the network, and  $b$  signifies the bias in the network layer. Trabelsi et al. [31] introduced an implementation of the complex-valued linear layer using the pytorch library. This layer employs two real-valued linear layers to calculate the real and imaginary parts individually. It derives the updated real and imaginary components based on the principles of complex-valued computation. The specific computation process is outlined as follows:

$$f_r(X) = W_r X + b_r, \quad (7)$$

$$f_i(X) = W_i X + b_i, \quad (8)$$

$$f_c(X) = f_r(X_r) - f_i(X_i) + i \times [f_r(X_i) - f_i(X_r)], \quad (9)$$

$$= W_r X_r - W_i X_i + b_r - b_i + i \times [W_r X_i + W_i X_r + b_r + b_i], \quad (10)$$

where  $i$  is the imaginary part.

In the QLSTM model, the substitution of the classical neural network in the LSTM unit with variational quantum classifiers (VQCs) extends the LSTM into the quantum domain. VQCs are quantum circuits equipped with quantum parameters capable of accommodating the inherent noise present in quantum computations. These parameters can be iteratively optimized using classical gradient descent [32]. The VQC plays the roles of both feature extraction and data compression. Figure 3a shows a schematic diagram of the QLSTM architecture and the components of the VQC used in the QLSTM are shown in Figure 3b. The VQC consists of the following three parts: data encoding, a variational stratification layer and a quantum measurement. The data encoding layer converts the output of the complex-valued linear layer into quantum states. The variational layer updates the circuit parameters by a gradient descent algorithm.

The CQLSTM consists of two QLSTMs, one of which is the virtual part. The results are linearly combined to obtain the final eigenvector and fed into the measurement function. To satisfy the relevant requirements, our CQLSTM model uses a 2-layer circuit with each VQC gate composed of 4 quantum bits with a word embedding size of 64 and vocab size as the output size. The QLSTM mathematical construction is given as follows:

$$f_t = \sigma(VQC_1(v_t)), \quad (11)$$



$$i_t = \sigma(VQC_2(v_t)), \quad (12)$$

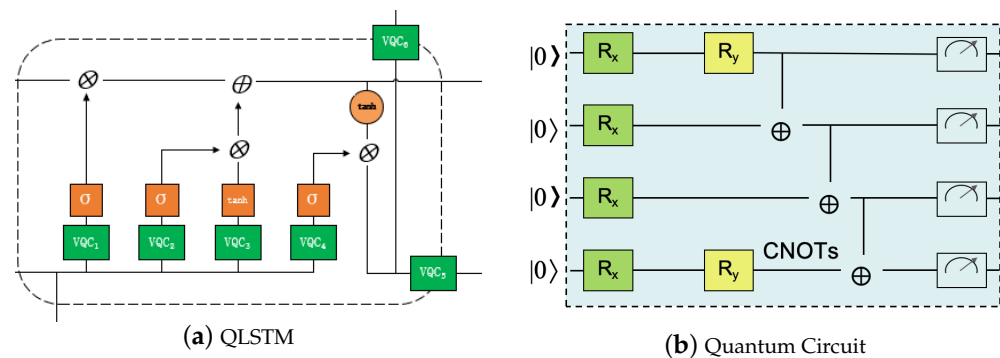
$$\tilde{C} = \tanh(VQC_3(v_t)), \quad (13)$$

$$c_t = f_t * c_{t-1} + i_t * \tilde{C}, \quad (14)$$

$$o_t = \sigma(VQC_4(v_t)), \quad (15)$$

$$h_t = VQC_5(o_t * \tanh(c_t)), \quad (16)$$

$$y_t = VQC_6(o_t * \tanh(c_t)), \quad (17)$$



**Figure 3.** Structure of QLSTM (a) and its component VQC (b).

### 3.2.3. Measurement Module

We use the measurement, which is based on quantum-computing to predict the output of the model. The measurement module utilizes the high-level features acquired from preceding modules as input to predict the classification result. The determination of the number of measurement basis vectors aligns with the count of classification labels. The ultimate classification outcome is determined by selecting the output category with the highest predicted value [23]. Additionally, considering that the measurement output is a real value, we construct a prediction vector by extracting the probability value for each label. This vector is subsequently concatenated with the prediction results from other models, fostering an integrated model fusion approach that contributes to superior experimental outcomes.

### 3.3. Dataset

We used the following four sentiment-analysis datasets: the Customer Review (CR), the Opinion polarity dataset (MPQA), the Movie Review (MR) and the Sentence Subjectivity (SUBJ). The specifics of these datasets are provided in Table 1.

**Table 1.** Introduction to the four sentiment analysis datasets.

Dateset	Description	Type	Count
CR	Product reviews	pos/neg	4k
MPQA	Opinions	pos/neg	11k
MR	Movie reviews	pos/neg	11k
SUBJ	Subjectivity	subj/obj	10k

### 3.4. Accuracy of Tests

In the sentiment analysis experiment, the evaluation metrics of accuracy and F1-score are employed to comprehensively assess the performance of various models. The calculation methods are outlined below.

- Accuracy: Accuracy is defined as the ratio of correctly predicted items to the total number of items as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \quad (18)$$

True positives ( $TP$ ), true negatives ( $TN$ ), false positives ( $FP$ ) and false negatives ( $FN$ ) represent the counts of correctly predicted positive instances, correctly predicted negative instances, incorrectly predicted positive instances and incorrectly predicted negative instances, respectively.

- F1-score: The F1-score represents the harmonic mean of precision and recall.

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall}, \quad (19)$$

Precision and recall are calculated as follows:

$$Precision = \frac{TP}{TP + FP}, Recall = \frac{TP}{TP + FN}, \quad (20)$$

In our experiment, the utilized loss function is the cross-entropy loss, defined as follows:

$$Loss = -\frac{1}{n} \sum_i [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)], \quad (21)$$

where  $n$  denotes the number of samples,  $y_i$  is the label for the  $i$ -sample and  $p_i$  is the predicted probability value for the positive class.

## 4. Results

We used PennyLane [33] version 0.13.0 and its plugin to perform simulations of quantum processes with the Qulacs library [34]. It is a VQC simulator that runs on a GPU. In the experiment, we set several hyperparameters, with the learning rate configured at 0.001. This choice was made to ensure model stability during training and prevent potential instability resulting from excessively large parameter updates. The batch size was 32 and the embedding dim was set to 768. All biases are set to 0, and the dropout rate was set to 0.5. These settings were chosen to ensure ample iterations over the dataset during training, facilitating robust learning of data features and enhancing overall model performance.

### 4.1. Comparison Models

We employed five classical models—TextCNN, LSTM, GRU, ELMo (Embeddings from Language Models) and BERT (Bidirectional Encoder Representations from Transformers)—as comparative benchmarks for our experiments. The descriptions of these models are provided below.

TextCNN [14] is evolved on the basis of neural networks. The essence of CNN is a multi-layer perceptron, which uses local connection and weight sharing. On the one hand, it reduces the number of weights and makes the network accessible to optimize. On the other hand, it reduces the risk of network overfitting.

LSTM [35] is a special kind of RNNs. Compared with traditional RNNs, an LSTM can effectively capture important long time dependencies in sequences and can solve the gradient problem.

GRU [36] belongs to the family of recurrent neural networks (RNNs). Similar to the long short-term memory (LSTM) architecture, GRU is specifically crafted to address challenges related to long-term memory and gradient-related issues. However, GRU exhibits a faster computational speed compared to LSTM.

ELMos [37] comprises bidirectional LSTMs as its fundamental components. Trained with a large corpus using the language model as the objective, ELMo obtains a shared semantic representation. This representation is subsequently applied to downstream

NLP tasks, leading to a notable enhancement in model performance. ELMo excels in providing word-level semantic representations and has proven effective across various downstream tasks.

BERT [38] is a language pretraining model (PLM). For instance, ELMo and GPT are models that are autoregressive (AR) and focus solely on one-directional information, where they predict the succeeding word based on the preceding words. BERT, on the other hand, uses context information to reconstruct original data from noisy data, making it an autoencoding model. The pretraining of BERT involved the following two tasks: next sentence prediction (NSP) and masked language model (MLM). BERT delivers a 768-dimensional vector for every token and includes a unique token ([CLS]) as part of its output.

#### 4.2. Experimental Results and Analysis

In the sentiment analysis study, several traditional NLP models such as long short-term memory (LSTM), GRU, TextCNN, Bidirectional Encoder Representations from Transformers (BERT), embeddings from language models (ELMo) were selected for comparison. The experimental outcomes for four sentiment analysis datasets are presented in Tables 2 and 3. Table 2 displays the results based on accuracy, whereas Table 3 showcases the results based on the F1-score metric. Additionally, Figure 4 illustrates a graphical representation of the accuracy metric results from the experiment, followed by an in-depth analysis of the findings. The source code is publicly released at Github <https://github.com/czlquantum/CQLSTM>, accessed on 4 March 2024.

**Table 2.** Accuracy for different experiments. Best results have been indicated in bold. The best-performing model is indicated in bold.

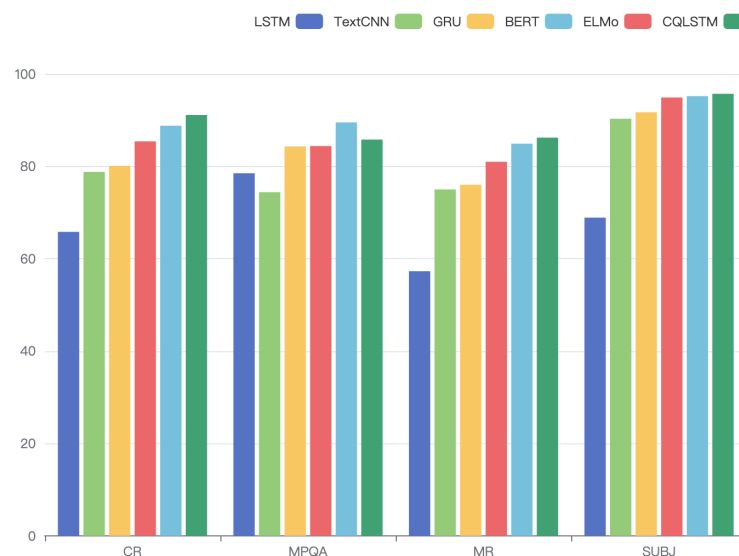
MODEL	CR	MPQA	MR	SUBJ
LSTM	65.8	78.5	57.3	68.9
TextCNN	78.8	74.4	75	90.3
GRU	80.1	84.3	76	91.7
ELMo	85.4	84.4	81	94.9
BERT	88.8	<b>89.5</b>	84.9	95.2
CQLSTM (Ours)	<b>91.1</b>	85.8	<b>86.2</b>	<b>95.7</b>

**Table 3.** F1-score for different experiments. The best-performing model is indicated in bold.

MODEL	CR	MPQA	MR	SUBJ
LSTM	56	71.4	57.3	68.8
TextCNN	71.2	75.5	75.9	90.1
GRU	71	73.3	75.5	91.8
ELMo	79.9	74.2	82.3	94.9
BERT	85.2	82.3	84.6	95.2
CQLSTM (Ours)	<b>90.5</b>	<b>82.4</b>	<b>86.2</b>	<b>95.7</b>

1. In comparison to traditional LSTM, TextCNN and GRU models, as well as popular pretraining models like ELMo and BERT, our model has demonstrated substantial advantages. Specifically, the enhancement achieved by CQLSTM is highlighted through the result comparisons. The performance disparities between CQLSTM and LSTM across various datasets are as follows: CR (+25.3%), MPQA (+7.3%), MR (+28.9%) and SUBJ (+26.8%), with an average improvement of 22.08% in accuracy metrics. When considering the F1-score metrics, the improvements seen with CQLSTM compared to LSTM are as follows: CR (+34.51%), MPQA (+11%), MR (+28.9%) and SUBJ (+26.86%), averaging a 25.32% enhancement. These findings underscore the notable performance leap of CQLSTM over LSTM, attributed to its foundation in complex-valued neural networks and quantum-computing theory, enabling the learning of intricate textual features;

2. In comparison to the pretraining models, the CQLSTM outperforms both ELMo and BERT in terms of experimental outcomes. Compared to BERT, CQLSTM has better results (under the F1-score metrics) in CR, MPQA, MR and SUBJ as follows: CR (+5.3%), MR (+1.6%) and SUBJ (+0.5%); on the MPQA dataset, the accuracy of CQLSTM is slightly lower than BERT (+3.7%), but the CQLSTM has better experimental results under the F1-score metric (+0.1%). From the overall results, CQLSTM performs slightly better than BERT on the MPQA dataset, which may be due to the fact that the MPQA dataset is all short sentences and words. CQLSTM is not as accurate as BERT in analyzing short sentences and words and is more adept at dealing with long texts; CQLSTM is not as accurate as BERT in analyzing short sentences and words and is more adept at dealing with long texts;
3. From Figure 4, it is evident that CQLSTM consistently outperforms, maintaining the highest level across all four sentiment analysis datasets. This observation underscores the notable performance advantage that CQLSTM exhibits in this context.



**Figure 4.** The accuracy scores of four sentiment analysis datasets evaluated by six different models.

#### 4.3. Ablation Study

In this subsection, we design a series of sub-models to study the impact of different components of the CQLSTM model as follows:

- The QLSTM model does not include a complex-valued embedding module;
- The model, which extracts features using the Complex-valued Quantum-enhanced Long Short-term Memory Neural Network module, is called QIM.

From Tables 4 and 5 we observe that CQLSTM achieves the best performance. The experimental results reveal that CQLSTM outperforms the QLSTM model, showcasing significant improvements. Specifically, when comparing the accuracy metrics across the four datasets, CQLSTM exhibits an average improvement of 14.7%, with individual improvements of CR (+18.1%), MPQA (+3.8%), MR (+16.2%) and SUBJ (+20.7%). In contrast, the comparison based on the F1-score metrics between CQLSTM and QLSTM results in an average improvement of 15.55%, with individual improvements of CR (+21.9%), MPQA (+2.4%), MR (+19.2%) and SUBJ (+18.7%).

Compared with QIM model, CQLSTM experiment results were significantly improved. The results of the accuracy measurement are as follows: CR (+1.8%), MR (+0.3%), SUBJ (+0.5%); the comparison results under the F1-score measure were as follows: CR (+1.9%), MR (+0.3%) and SUBJ (+0.6%). QIM model performs better on the dataset MPQA, because MPQA is composed of short sentences and words, which can extract less context information.

The results show that feature extraction between words and context makes a positive contribution to judge the emotional polarity of sentences. Compared with QLSTM, QIM

performs better. QIM with mixed complex-valued can extract more features than single QLSTM. Compared with QIM, CQLSTM has better performance. This shows that our proposed complex-valued embedding based on density matrix can efficiently encode semantic dependencies and their probability distributions. At the same time, the accuracy of QIM outperforms the classical model, which proves that modeling interdiscourse interaction is beneficial to sentiment analysis.

These findings underscore the superior performance of CQLSTM over QLSTM and QIM. The complex-valued embedding used by CQLSTM, which when coupled with the incorporation of complex-valued encoding and density matrix, enables the expression of more concealed information compared to traditional vector encoding methods. Moreover, the synergistic effect of complex-valued embedding and a quantum-enhanced long short-term memory neural network enhances feature extraction, allowing for the better capture of advanced semantic features, particularly interactions between words and sentences.

**Table 4.** Accuracy results of ablation experiments. The best-performing model is indicated in bold.

MODEL	CR	MPQA	MR	SUBJ
QLSTM	73	82	70	75
QIM	89.3	<b>89.1</b>	85.9	95.2
CQLSTM (Ours)	<b>91.1</b>	85.8	<b>86.2</b>	<b>95.7</b>

**Table 5.** F1-score Results of ablation experiments. The best-performing model is indicated in bold.

MODEL	CR	MPQA	MR	SUBJ
QLSTM	68.6	80.0	67	77
QIM	88.6	<b>87.2</b>	85.9	95.1
CQLSTM (Ours)	<b>90.5</b>	82.4	<b>86.2</b>	<b>95.7</b>

## 5. Conclusions and Discussion

This paper proposes an effective strategy for sentiment analysis based on complex-valued embedding and a quantum long short-term memory neural network, which can efficiently encode hybrid semantic subspaces and extract high-level semantic features. The main idea is to combine quantum theory with complex-valued neural networks to deal with emotion analysis tasks. In this study, words are depicted as quantum states, sentences are depicted as density matrices, interactions between words as quantum state evolution in sentences and the associated sentence labels as a quantum state measurement with a subsequent collapse to the ground state. The incorporation of this physical interpretation into the model enhances comprehension. Additionally, certain aspects of natural language can be elucidated through quantum effects. For example, the multiple meanings of words can be effectively captured through quantum entanglement, thereby improving the model's interpretability. CQLSTM extends quantum applications to the field of natural language processing, and the neural network layer of CQLSTM is based on complex-valued neural networks, which can be easily transplanted to future quantum computers. We look forward to implementing our proposed algorithm on a true quantum computer in the future.

The CQLSTM operates within the Hilbert space, where the word vector is depicted as a complex value, expanding the representation space beyond that of real values. The semantic information of the context is conveyed through the real part, while the imaginary part captures non-semantic details like word positioning, sentiment and ambiguity. The complex number space offers deep-learning algorithms broader representation options compared to real-number spaces, thereby enhancing model development prospects. Moreover, this study opens avenues for further exploration in diverse sentiment analysis. Various research endeavors have applied sentiment analysis in numerous domains. For instance, Daniel et al. [39] investigated sentiment among Twitter users in response to events involving companies like Microsoft, Walmart and Apple. Hasselgren et al. [40] introduced a method for recommending stocks based on sentiment trends extracted from



Twitter. Wu et al. [41] emphasized the effectiveness of sentiment in predicting stock prices for small-sized companies. In Wang et al. [42] conducted sentiment analysis on the Guba platform to anticipate bear markets. In [43], sentiment analysis was employed to forecast fundraising outcomes from online comments. Kauffmann et al. [44] enhanced product recommendations by incorporating sentiment classification from Amazon reviews.

As quantum hardware advances, there is an opportunity to enhance and scale quantum models for tackling larger and more intricate problems. The integration of classical and quantum models can capitalize on their respective strengths, leading to increased accuracy and computational efficiency. Future efforts will explore the direct development of quantum convolutional neural networks for multiple regression, aiming to achieve affective polarity classification for short sentences and words. Additionally, in terms of adapting and optimizing model structure, we will explore the integration of various quantum neural network layers, including quantum attention mechanisms, to enhance the model's expressive capacity and generalization capabilities. To tackle the potential challenges related to long-term dependencies encountered by QLSTM when processing extended sequences, we aim to explore the integration of other recurrent neural networks with quantum theory. Examples include the gated recurrent unit (GRU) or transformer, which could augment the model's capability to effectively capture and understand prolonged dependencies.

**Author Contributions:** Conceptualization, Z.C., L.S. and Q.G.; methodology, Z.C., L.S. and X.W.; software, Z.C., M.J. and X.W.; validation, L.S., Q.G., M.J. and X.W.; formal analysis, Z.C., L.S. and Q.G.; investigation, Z.C., L.S. and Q.G.; resources, Z.C., L.S. and N.Z.; data curation, Z.C., L.S. and M.J.; writing—original draft preparation, Z.C.; writing—review and editing, L.S., Q.G., N.Z., M.J. and X.W.; visualization, Z.C.; supervision, L.S., N.Z. and X.W.; project administration, Z.C., L.S. and Q.G.; funding acquisition, L.S., Q.G. and X.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by the Natural Science Foundation of China (No.62072362, 12101479) and the Shaanxi Provincial Key Industry Innovation Chain Program (No.2020ZDLGY07-05), Natural Science Basis Research Plan in Shaanxi Province of China (No.2021JQ-660), Xi'an Major Scientific and Technological Achievements Transformation Industrialization Project (No.23CGZHCHYH0008).

**Data Availability Statement:** We have used four datasets including the following: Customer Review (CR) [45], Opinion polarity dataset (MPQA) [46], Movie Review (MR) [47] and Sentence Subjectivity (SUBJ) [47].

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Sordoni, A.; Nie, J.-Y.; 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, Dublin, Ireland, 28 July–1 August 2013; pp. 653–662.
2. Wang, P.; Wang, T.; Hou, Y.; Song, D. Modeling relevance judgement inspired by quantum weak measurement. In *Advances in Information Retrieval: 40th European Conference on IR Research, ECIR 2018, Grenoble, France, 26–29 March 2018, Proceedings 40*; Springer: Cham, Switzerland, 2018; pp. 424–436.
3. 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 Retrieval, Tianjin, China, 14–17 September 2018; pp. 183–186.
4. Zhang, Y.; Song, D.; Zhang, P.; Wang, P.; Li, J.; Li, X.; Wang, B. A quantum-inspired multimodal sentiment analysis framework. *Theor. Comput. Sci.* **2018**, *752*, 21–40. [CrossRef]
5. Zhang, P.; Niu, J.; Su, Z.; Wang, B.; Ma, L.; Song, D. End-to-end quantum-like language models with application to question answering. In Proceedings of the AAAI Conference on Artificial Intelligence, New Orleans, LA, USA, 2–7 February 2018; Volume 32.
6. Zhang, Y.; Song, D.; Li, X.; Zhang, P. Unsupervised sentiment analysis of twitter posts using density matrix representation. In *Advances in Information Retrieval: 40th European Conference on IR Research, ECIR 2018, Grenoble, France, 26–29 March 2018, Proceedings 40*; Springer: Cham, Switzerland, 2018; pp. 316–329.
7. Chen, S.Y.-C.; Yoo, S.; Fang, Y.-L.L. Quantum long short-term memory. In Proceedings of the ICASSP 2022—2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Singapore, 23–27 May 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 8622–8626.

8. Di Sipio, R.; Huang, J.H.; Chen, S.Y.; Mangini, S.; Worring, M. The dawn of quantum natural language processing. In Proceedings of the ICASSP 2022—2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Singapore, 23–27 May 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 8612–8616.
9. Asghar, M.Z.; Khan, A.; Ahmad, S.; Qasim, M.; Khan, I.A. Lexicon-enhanced sentiment analysis framework using rule-based classification scheme. *PLoS ONE* **2017**, *12*, e0171649. [[CrossRef](#)] [[PubMed](#)]
10. Baroni, M.; Zamparelli, R. Nouns are vectors, adjectives are matrices: Representing adjective-noun constructions in semantic space. In Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing, Cambridge, MA, USA, 9–11 October 2010; pp. 1183–1193.
11. Polanyi, L.; Zaenen, A. Contextual valence shifters. In *Computing Attitude and Affect in Text: Theory and Applications*; Springer: Dordrecht, The Netherlands, 2006; pp. 1–10.
12. Lin, F.; Yu, Y. Chinese text sentiment classification based on extreme learning machine. In Proceedings of the ELM-2016: International Conference on Extreme Learning Machine 2016, Singapore, 13–15 December 2016; Springer: Cham, Switzerland, 2017; pp. 171–181.
13. Yin, W.; Schütze, H. Convolutional neural network for paraphrase identification. In Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Denver, CO, USA, 31 May–5 June 2015; pp. 901–911.
14. Kim, Y. Convolutional neural networks for sentence classification. *arXiv* **2014**, arXiv:1408.5882.
15. Su, Y.J.; Chen, C.H.; Chen, T.Y.; Cheng, C.C. Chinese microblog sentiment analysis by adding emoticons to attention-based CNN. *J. Internet Technol.* **2020**, *21*, 821–829.
16. Wang, Y.; Huang, M.; Zhu, X.; Zhao, L. Attention-based lstm for aspect-level sentiment classification. In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, Austin, TX, USA, 1–5 November 2016; pp. 606–615.
17. Ma, D.; Li, S.; Zhang, X.; Wang, H. Interactive attention networks for aspect-level sentiment classification. *arXiv* **2017**, arXiv:1709.00893.
18. Sosa, P.M. Twitter sentiment analysis using combined lstm-cnn models. *Eprint Arxiv* **2017**, 2017, 1–9.
19. Wang, J.; Yu, L.C.; Lai, K.R.; Zhang, X. Dimensional sentiment analysis using a regional cnn-lstm model. In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, Berlin, Germany, 7–12 August 2016; Volume 2, pp. 225–230.
20. Van Rijsbergen, C.J. *The Geometry of Information Retrieval*; Cambridge University Press: Cambridge, UK, 2004.
21. Stein, J.; Christ, I.; Kraus, N.; Mansky, M.B.; Müller, R.; Linnhoff-Popien, C. Applying qnlp to sentiment analysis in finance. In Proceedings of the 2023 IEEE International Conference on Quantum Computing and Engineering (QCE), Bellevue, WA, USA, 17–22 September 2023; IEEE: Piscataway, NJ, USA, 2023; Volume 2, pp. 20–25.
22. Li, Q.; Li, J.; Zhang, P.; Song, D. Modeling multi-query retrieval tasks using density matrix transformation. In Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, Santiago, Chile, 9–13 August 2015; pp. 871–874.
23. Lai, W.; Shi, J.; Chang, Y. Quantum-inspired fully complex-valued neutral network for sentiment analysis. *Axioms* **2023**, *12*, 308. [[CrossRef](#)]
24. Wu, S.; Li, J.; Zhang, P.; Zhang, Y. Natural language processing meets quantum physics: A survey and categorization. In Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, Online and Punta Cana, Dominican Republic, 7–11 November 2021; pp. 3172–3182.
25. Niu, T.; Hou, Y. Density matrix based convolutional neural network for click-through rate prediction. In Proceedings of the 2020 3rd International Conference on Artificial Intelligence and Big Data (ICAIBD), Chengdu, China, 28–31 May 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 46–50.
26. Nielsen, M.A.; Chuang, I.L. *Quantum Computation and Quantum Information*; Cambridge University Press: Cambridge, UK, 2010.
27. Zhang, G.; Jin, W.; Li, N. An improved quantum genetic algorithm and its application. In *Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing: 9th International Conference, RSFDGrC 2003, Chongqing, China, 26–29 May 2003 Proceedings 9*; Springer: Berlin/Heidelberg, Germany; Cambridge University Press: Cambridge, UK, 2003; pp. 449–452.
28. Yanofsky, N.S.; Mannucci, M.A. *Quantum Computing for Computer Scientists*; Cambridge University Press: Cambridge, UK, 2008.
29. Liang, J.; Chen, J.; Zhang, X.; Zhou, Y.; Lin, J. One-hot encoding and convolutional neural network based anomaly detection. *J. Tsinghua Univ. (Sci. Technol.)* **2019**, *59*, 523–529.
30. Liu, Y.; Ott, M.; Goyal, N.; Du, J.; Joshi, M.; Chen, D.; Levy, O.; Lewis, M.; Zettlemoyer, L.; Stoyanov, V. Roberta: A robustly optimized bert pretraining approach. *arXiv* **2019**, arXiv:1907.11692.
31. Trabelsi, C.; Bilaniuk, O.; Zhang, Y.; Serdyuk, D.; Subramanian, S.; Santos, J.F.; Mehri, S.; Rostamzadeh, N.; Bengio, Y.; Pal, C.J. Deep complex networks. *arXiv* **2017**, arXiv:1705.09792.
32. Griol-Barres, I.; Milla, S.; Cebrián, A.; Mansoori, Y.; Millet, J. Variational quantum circuits for machine learning. An application for the detection of weak signals. *Appl. Sci.* **2021**, *11*, 6427. [[CrossRef](#)]
33. Bergholm, V.; Izaac, J.; Schuld, M.; Gogolin, C.; Ahmed, S.; Ajith, V.; Alam, M.S.; Alonso-Linaje, G.; AkashNarayanan, B.; Asadi, A.; et al. PennyLane: Automatic differentiation of hybrid quantum-classical computations. *arXiv* **2018**, arXiv:1811.04968.
34. Suzuki, Y.; Kawase, Y.; Masumura, Y.; Hiraga, Y.; Nakadai, M.; Chen, J.; Nakanishi, K.M.; Mitarai, K.; Imai, R.; Tamiya, S.; et al. Qulacs: A fast and versatile quantum circuit simulator for research purpose. *Quantum* **2021**, *5*, 559. [[CrossRef](#)]

35. Sundermeyer, M.; Schlüter, R.; Ney, H. Lstm neural networks for language modeling. In Proceedings of the Thirteenth Annual Conference of the International Speech Communication Association, Portland, OR, USA, 9–13 September 2012.
36. Dey, R.; Salem, F.M. Gate-variants of gated recurrent unit (gru) neural networks. In Proceedings of the 2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS), Boston, MA, USA, 6–9 August 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1597–1600.
37. Peters, M.E.; Ammar, W.; Bhagavatula, C.; Power, R. Semi-supervised sequence tagging with bidirectional language models. *arXiv* **2017**, arXiv:1705.00108.
38. Devlin, J.; Chang, M.-W.; Lee, K.; Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv* **2018**, arXiv:1810.04805.
39. Daniel, M.; Neves, R.F.; Horta, N. Company event popularity for financial markets using twitter and sentiment analysis. *Expert Syst. Appl.* **2017**, *71*, 111–124. [[CrossRef](#)]
40. Hasselgren, B.; Chrysoulas, C.; Pitropakis, N.; Buchanan, W.J. Using social media & sentiment analysis to make investment decisions. *Future Internet* **2022**, *15*, 5.
41. Wu, D.D.; Zheng, L.; Olson, D.L. A decision support approach for online stock forum sentiment analysis. *IEEE Trans. Syst. Man Cybern. Syst.* **2014**, *44*, 1077–1087. [[CrossRef](#)]
42. Sun, Y.; Fang, M.; Wang, X. A novel stock recommendation system using guba sentiment analysis. *Pers. Ubiquitous Comput.* **2018**, *22*, 575–587. [[CrossRef](#)]
43. Wang, W.; Guo, L.; Wu, Y.J. The merits of a sentiment analysis of antecedent comments for the prediction of online fundraising outcomes. *Technol. Forecast. Soc. Change* **2022**, *174*, 121070. [[CrossRef](#)]
44. Kauffmann, E.; Peral, J.; Gil, D.; Ferrández, A.; Sellers, R.; Mora, H. Managing marketing decision-making with sentiment analysis: An evaluation of the main product features using text data mining. *Sustainability* **2019**, *11*, 4235. [[CrossRef](#)]
45. Hu, M.; Liu, B. Mining and summarizing customer reviews. In Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Seattle, WA, USA, 22–25 August 2004; pp. 168–177.
46. Wiebe, J.; Wilson, T.; Cardie, C. Annotating expressions of opinions and emotions in language. *Lang. Resour. Eval.* **2005**, *39*, 165–210. [[CrossRef](#)]
47. Nivre, J.; De Marneffe, M.C.; Ginter, F.; Goldberg, Y.; Hajic, J.; Manning, C.D.; McDonald, R.; Petrov, S.; Pyysalo, S.; Silveira, N.; et al. Universal dependencies v1: A multilingual treebank collection. In Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16), Istanbul, Turkey, 23–25 May 2016; pp. 1659–1666.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.