machine learning & knowledge extraction

MDPI

Article

A Quantum-Probability-Inspired Complex-Valued Model for Multilingual Stance Detection

Muhammad Ebrahim Ahmadi and Monireh Hosseini *

IT Department, Faculty of Industrial Engineering, K. N. Toosi University of Technology, Tehran 19991-43344, Iran; ahmadi@email.kntu.ac.ir

* Correspondence: hosseini@kntu.ac.ir

Abstract

In this study, we propose a quantum-probability-inspired complex-valued model for multilingual stance detection. The model brings together ideas from granular computing and quantum theory to better capture semantic meaning across different languages. The proposed model combines contextual embeddings, graph convolutional networks, and a quantum-inspired feature interaction module (QFIM) to capture complex, high-order, and non-linear relationships in multilingual data. The QFIM models quantum amplitude-like interactions to represent overlapping semantic patterns in the latent space. This design helps the system better distinguish subtle differences in stance expressions. To strengthen the representation, a granulation mechanism based on contextual embedding similarity is employed to extract semantically coherent text granules. These granules expand the feature space and help align different elements of the stance structure more accurately. Experimental results on standard benchmark datasets show that the proposed model consistently achieves better performance than existing state-of-the-art methods.

Keywords: stance detection; quantum; granular computing

1. Introduction

Stance detection refers to identifying whether a text expresses a favorable, opposing, or neutral position toward a specific target [1]. With the rapid growth of digital communication, this task has gained considerable attention. It plays an important role in related applications such as fake-news detection, sentiment analysis, argument evaluation, and policy recommendation systems [2,3]. As information is increasingly shared across social media platforms, news outlets, and online forums in multiple languages, accurate cross-lingual stance detection has become especially important. However, most existing models are developed for monolingual settings and do not generalize effectively to multilingual contexts. Multilingual data introduces additional semantic variation, syntactic differences, and cultural nuances that make the problem more challenging. For example, prior studies [4–6] focus primarily on monolingual stance detection.

Existing approaches to multilingual stance detection can generally be grouped into two main categories: (1) language-agnostic feature extraction through machine translation followed by monolingual classification [7], and (2) multilingual transformer-based models such as XLM-R and mBERT that encode texts from different languages using shared parameters [8]. Although these methods have contributed to progress in the field, several limitations remain. Translation-based pipelines may introduce inaccuracies that distort subtle stance expressions. Meanwhile, pretrained multilingual transformers, despite

Check for updates

Academic Editor: Jaroslaw Krzywanski

Received: 3 January 2026
Revised: 27 March 2026
Accepted: 1 April 2026
Published: 18 May 2026

Copyright: © 2026 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.

their strong representational capacity, often struggle to capture complex language-specific semantic dependencies and may rely on relatively shallow cross-lingual alignment.

This study introduces a quantum-probability-inspired complex-valued deep learning model for multilingual stance detection. The model draws on two complementary theoretical perspectives: granular computing and quantum theory. In natural language processing, granular computing (GrC) emphasizes representing linguistic information at multiple levels of granularity, including words, phrases, sentences, and broader semantic units. Instead of treating text as a flat sequence of tokens, GrC groups information into semantically coherent granules. This structure supports more organized and interpretable learning. Within NLP, granular computing helps address ambiguity, preserve contextual relevance, and represent hierarchical linguistic patterns that are often overlooked in token-level approaches. Organizing text into meaningful granules enables the model to focus on the most relevant parts of the input, which can improve performance in tasks such as stance detection, sentiment analysis, and information retrieval. Granules may be fine-grained, capturing detailed linguistic elements, or coarse-grained, reflecting broader conceptual groupings [9]. This flexibility allows the model to adjust its level of abstraction depending on task requirements, reflecting aspects of human language processing.

Quantum theory, by contrast, provides a mathematical foundation for describing phenomena such as superposition, interference, and entanglement. These principles offer a useful perspective for modeling overlapping and context-dependent meanings in language. Integrating quantum-probability-inspired complex-valued concepts into a deep learning framework makes it possible to address semantic ambiguity and cross-lingual variability more effectively in multilingual settings.

The central component of the model is the quantum-inspired feature interaction module (QFIM). It builds a complex-valued latent space that mimics quantum amplitude interactions between feature vectors. This design captures non-linear relationships and overlapping semantic patterns that assist in identifying subtle stance signals. In contrast to true quantum computing systems, the architecture runs on classical hardware and relies on mathematical formulations inspired by quantum principles, such as phase modulation and amplitude interference, to enhance neural network representations.

A context-aware granulation mechanism is also introduced to divide each comment into smaller semantic units based on contextual embedding similarity with the associated query. Each granule represents a coherent text fragment that preserves the argumentative or affective structure of the input. The input representation is then expanded by incorporating the top-k granules most relevant to the stance target. This additional coarse-grained information provides broader contextual support for more accurate stance detection. The key contributions of this paper are summarized as follows:

- A quantum-probability-inspired complex-valued stance detection framework is proposed, incorporating amplitude-based feature interactions to capture complex, high-order semantic dependencies in multilingual textual data.
- A granulation-based feature augmentation algorithm is presented, enabling the extraction of semantically coherent substructures from textual data.
- An end-to-end architecture is designed that integrates graph-based contextual reasoning, quantum-inspired projection, and granular feature augmentation, providing a robust and flexible solution for multilingual stance detection.

The following research questions guide this study:

1. To what extent does incorporating the quantum-inspired feature interaction module improve multilingual stance detection performance compared with conventional transformer-based models?

2. How does the context-aware granulation mechanism affect the model's ability to capture complex stance patterns across different languages?

Theoretical and Practical Implications

This study offers both theoretical and practical contributions to multilingual stance detection and natural language processing. From a theoretical perspective, the framework connects quantum-probability-inspired complex-valued concepts with NLP through a feature interaction module that represents semantic relationships using amplitude-based interactions, superposition, and interference. At the same time, granular computing improves distinguishability by structuring text into semantically coherent units. By modeling higher-order dependencies and language-specific nuances, the framework strengthens multilingual stance detection and responds to ongoing challenges in cross-lingual NLP.

In practical terms, the model achieves improved performance on benchmark datasets and demonstrates potential for applications such as fake-news detection and policy recommendation. The quantum-inspired and granular mechanisms can also be extended to related NLP tasks, including sentiment analysis and hate speech detection. Therefore, this approach is an effective framework for stance detection, with relevance for both research and real-world NLP systems.

The remainder of the paper is structured as follows. Section 2 reviews related work. Section 3 presents the proposed model. Section 4 describes the experimental evaluation. Section 5 concludes the study and outlines directions for future research.

2. Background

Recent developments at the intersection of natural language processing (NLP) and quantum computing have led to the emerging field of quantum natural language processing (QNLP). This interdisciplinary area applies principles from quantum mechanics to address complex language processing tasks with novel methodologies [10]. For example, Abbaszadeh et al. [11] examined how quantum NLP can improve machine translation. Pandey et al. [12] reviewed the progression from quantum computing to quantum machine learning, highlighting applications in NLP where quantum-based models outperform classical methods in tasks such as parts-of-speech tagging for code-mixed social media text. Shi et al. [13] proposed a pretrained quantum-inspired deep neural network that combines quantum principles with classical text embeddings, improving sentiment analysis and word sense disambiguation compared with both classical and earlier quantum-inspired approaches. Tomal et al. [14] introduced a hybrid classical–quantum transformer with a quantum-enhanced attention mechanism, showing improved efficiency and accuracy in NLP tasks like text classification on the IMDb dataset. Ramya et al. [15] presented a hybrid framework for cyber threat detection, demonstrating faster data processing and higher accuracy using quantum algorithms together with advanced NLP models such as BERT. Cavar et al. [16] investigated the compression of classical word embeddings, including fastText, GloVe, and BERT, into quantum states using amplitude encoding, preserving semantic relationships and enabling computation of word similarities in a quantum environment. Li et al. [17] proposed a classical–quantum transfer learning approach by pretraining sentence representations with a complex-valued BERT-like model, enhancing the performance of quantum NLP models in sentence classification tasks. In another work, Yao et al. [18] introduced a self-supervised pretraining method to improve the representation capacity of quantum NLP models, achieving better performance in text classification than purely end-to-end quantum approaches. Pei et al. [19] designed a quantum-based model for one-to-many image generation, employing Pauli-X and reconfigurable beam splitter gates to increase flexibility. A block-based strategy was used to enhance performance on the

MNIST and Fashion-MNIST datasets. Gong et al. [20] introduced a dual-discriminator quantum generative adversarial network. This model integrates a quantum convolutional neural network as a parallel discriminator to capture both local and global image features, which helps improve generation quality and reduce mode collapse on certain real-world datasets. Lou et al. [21] presented a privacy-preserving bidirectional data transmission protocol for smart grids based on semi-quantum computation. The protocol combines mutual identity verification and message authentication while splitting quantum states to enable secure two-way communication, and its security was evaluated through IBM quantum cloud simulations.

Stance detection has been extensively studied in previous work. The task is challenging because textual content often does not explicitly reveal the underlying stance. Various approaches have been proposed to handle these complexities, particularly in multilingual contexts. Vasilakes et al. [22] investigated advanced Vision-Language Models for stance detection in seven languages and observed that these models generally rely more on textual inputs than visual information. Xu et al. [23] developed a multitask learning framework that separates causal and non-causal features to enhance stance representations. Blanco et al. [24] applied deep learning to classify Spanish-language tweets about COVID-19 vaccination, comparing multiple models to evaluate their suitability for public health monitoring. Fan et al. [25] improved zero-shot stance detection using a multitask fine-tuning approach that incorporated debate datasets and external knowledge sources. Similarly, Pu et al. [26] proposed a multitask architecture that combines emotional cues with stance detection mechanisms, resulting in more reliable predictions for social media security. Kumar et al. [27] introduced a hybrid system that combines BERT embeddings with deep convolutional neural networks to detect negative stances in low-resource multilingual social media datasets. Their approach includes data cleaning, BERT fine-tuning, and the use of a deep CNN to extract key linguistic features.

Lan et al. [28] introduced COLA, a collaborative framework using role-infused large language model agents for stance detection. The framework employs multidimensional analysis, role-based reasoning, and stance aggregation, functioning without additional annotations. Mets et al. [29] explored the use of large language models, including GPT-3.5, for automated stance detection in resource-limited environments, achieving performance comparable to supervised approaches. To address biases in large language models, Li et al. [30] proposed a counterfactual data augmentation and calibration strategy to improve generalization and accuracy. Ding et al. [31] presented a prompt-tuning method that incorporates target-focused analysis and natural language rationales, enhancing cross-target stance detection.

Wang et al. [32] developed DEEM, an expert-modeling approach that uses generative LLM capabilities to reduce bias and increase accuracy. Ding et al. [33] introduced PSDCOT, which combines chain-of-thought knowledge with multi-prompt learning to integrate information across tasks effectively. Zhao et al. [34] proposed a zero-shot stance detection approach using contrastive learning and attention-based fusion of multi-perspective features, enabling robust results on unseen topics. Mahmoudi et al. [35] designed DyMoAdapt, a framework combining few-shot learning and GPT-3 for contextual data generation, and also introduced the MGT-VAST dataset to support this work.

Hashemian et al. [36] highlighted the persistent challenges of sentiment analysis on social media. Choi et al. [37] developed a zero-shot stance detection framework using dual-phase learning to improve generalization without relying on labeled data. Jiang et al. [38] addressed the problem of limited target information by applying multi-perspective contrastive learning to distinguish stance features more effectively. Gomez et al. [39] proposed a two-phase classification framework based on topic modeling to extract interpretable

stance explanations from tweets, incorporating features such as sentiment, mental state cues, and embeddings. Alshahrani et al. [40] described a three-step stance classification pipeline consisting of preprocessing, feature extraction, and classification. Lai et al. [41] investigated multilingual stance detection in political discourse using social media data, while Kawintiranon and Singh [42] proposed knowledge-enhanced model fine-tuning. A summary of key stance detection methods and their datasets is provided in Table 1.

Table 1. Summary of prominent stance detection methods and their associated datasets.

Best F1 Score	Dataset	Model	Features
59.21	SE16-T6	SVM [6]	N-gram, SE, Sentiment words
61.04	SE16-T6	Maximum entropy [5]	N-gram, Sentiment, Subject
67.33	SE16-T6	Voting CNN [43]	SE
77.11	SE16-T6	LibSVM [4]	N-gram, POS, Vocabulary, Feelings
70.30	SE16-T6	SVM [44]	N-gram, POS, Vocabulary, Feelings, SE
74.44	SE16-T6	SVM [45]	N-gram, Text, Words, Emotions, Syntactic structure
67.40	SE16-T6	BiGRU-CNN [46]	SE, Modeling objective
61	SE16-T6	LSTM Attention [47]	SE, Vocabulary, Feelings, Dependency parser
71.04	SE16-T6	BiGRU Attention Network [48]	SE, Modeling goal
76.45	SE16-T6	Weighted KNN [40]	TF-IDF, Sentiment words
70.46	SE16-T6	Random Forest, MLP, CNN [49]	SE, POS, Structural, Statistical
73.77	SE16-T6	Roberta, BiLSTM [50]	CE, N-gram
71.24	SE16-T6	DNRFAF, DSRFE, DECCV [50]	N-gram, Vocabulary, Feelings, Structural, SE
68.79	SE16-T6	RNN for LSTM Attention [51]	SE, Modeling objective
64.51	IberEval2017	SVM, LR, CNN for LSTM, BiLSTM [41]	N-gram, BOW, Structural, Vocabulary of feelings

The Transformer architecture [52] has been foundational in deep learning, particularly in natural language processing. By removing recurrence and convolution, Transformers rely on attention mechanisms to process sequences in parallel, improving both efficiency and performance. The architecture consists of encoders and decoders built from repeated blocks that include self-attention and feedforward layers.

BERT [53] introduced bidirectional pre-training of Transformer models, enabling more effective contextual representations. RoBERTa [54] further refined this approach by adjusting training strategies, such as using larger batch sizes and omitting Next Sentence Prediction (NSP). Transformer variants have been applied across a range of tasks, from fake-news detection to genomic data analysis [55–57].

Granular computing contributes to textual understanding by organizing information at multiple levels of abstraction. Liang et al. [58] and Yang et al. [59] applied this approach to question-answering systems and sentiment classification, respectively. Behzadidoost et al. [9,60] extended granular methods to deep learning tasks, including cyberbullying detection and argument analysis.

Despite these developments, multilingual stance detection remains a challenging area. The current study introduces a model that combines quantum-probability-inspired complex-valued representation learning with granular computing to improve stance discrimination.

3. Proposed Model

The proposed model is a multilingual classification framework that combines contextualized embeddings, graph-based relational reasoning, and quantum-inspired feature

interaction. The quantum-inspired feature interaction module is applied during training to capture complex, non-linear, and higher-order feature relationships that traditional linear or shallow layers may overlook. By simulating quantum amplitudes, QFIM represents overlapping patterns in the feature space, improving the modeling of subtle semantic distinctions across multiple languages and supporting reasoning over multiple sentences. This mechanism enhances the network's representational capacity while remaining fully compatible with standard deep learning architectures. Although it does not require quantum hardware, QFIM uses mathematical constructs inspired by quantum probability and complex-valued representations to improve training efficiency and generalization, particularly for natural language tasks with detailed, language-specific dependencies. Another component of the model is a granulation algorithm based on granular computing. This algorithm generates coarse-grained textual granules from individual samples, providing structured sub-units that can enrich feature representations.

Let S be a textual instance consisting of n words, represented as:

$$S = (W_1, W_2, \ldots, W_n)$$

From each instance S, a corresponding text granule G is derived as:

$$G = (g_1, g_2, \ldots, g_m)$$

where typically $m \leq n$, and each g_i is a granulated representation of the corresponding W_i. Formally, the mapping satisfies:

$$|G| \leq |S|$$

The union of all such granules, denoted by $\mathcal{G} = \bigcup_{i=1}^{N} G^{(i)}$, where N is the number of samples, is integrated into the feature space alongside the original dataset samples. This enriched feature space enhances the model's capacity to capture structural and semantic relationships.

3.1. Granule Extraction via Contextual Embedding Similarity for Feature Augmentation

Granule extraction is performed using contextual embedding similarity to identify meaningful semantic substructures within the text. These granules are used to augment the feature space with additional semantic information, supporting more effective multilingual classification.

Let $\mathcal{X}$ be the space of multilingual input texts, and let M denote a pretrained contextual language model. For any input sequence $x \in \mathcal{X}$, let the tokenizer τ produce a token sequence $\tau(x) = \{w_1, \ldots, w_n\}$ with associated token indices $X \in \mathbb{N}^n$ and attention mask $\mathbf{A} \in \{0,1\}^n$. The contextual embeddings $H = M(X, A) \in \mathbb{R}^{n \times d}$ encode token-level semantic representations, where d is the embedding dimension.

To derive a compact representation for x, we apply mean pooling over valid tokens as determined by the attention mask:

$$e_x = \frac{\sum_{i=1}^{n} A_i \cdot H_i}{\sum_{i=1}^{n} A_i},$$

Given an input instance $x = (q, c)$, where q is a question and c is a comment, we define a granular segmentation function that decomposes c into a sequence of semantically coherent subunits, referred to as granules. Let m denote the number of extracted granules. Then,

$$c \mapsto (g_1, g_2, \ldots, g_m), \quad \text{where } g_j \in \mathcal{X},\ 1 \leq j \leq m.$$

Each granule g_j captures a distinct semantic unit within the original comment c, facilitating localized alignment with the task context.

Each granule g_j represents a partial knowledge structure embedded in the comment c. For each granule g_j, we compute the corresponding mean-pooled embedding e_{g_j}. To quantify the alignment between the question and each granule, we compute their semantic proximity using cosine similarity:

$$s_j = \cos(e_q, e_{g_j}) = \frac{e_q \cdot e_{g_j}}{\|e_q\| \cdot \|e_{g_j}\|}.$$

These similarity scores induce a granularity-based ranking over the comment:

$$\mathcal{S} = \{(g_j, s_j)\}_{j=1}^{m}.$$

The top-k granules with the highest relevance to the question are selected as follows:

$$\text{Top}_k(q, c) = \{g_j \mid s_j \text{ ranks among the top } k\}.$$

These granules serve as semantically enriched units of information and are used to augment the feature space for downstream classification. Each data instance is then represented as a tuple:

$$(q, c, \text{Top}_k(q, c), y),$$

where y is the binary label. Algorithm 1 shows the proposed granulation algorithm. This algorithm describes the construction of coarse-grained granules by selecting the most query-relevant context fragments. Lines 1–3 define the input multilingual dataset, tokenizer, encoder, and top-k parameter; initialize the output augmented dataset; and prepare the algorithm state. Lines 4–5 iterate over each instance and partition the context c into syntactically or discourse-motivated fragments. Lines 6–8 tokenize and encode the query q, then compute a pooled query embedding via attention-weighted averaging. Lines 9–15 process each context fragment, where the fragment is tokenized and encoded, and a pooled fragment embedding is computed. Cosine similarity between the query and fragment embeddings is then calculated, and the resulting similarity score is stored. Lines 16–17 rank all fragments by similarity and select the top-k most relevant fragments to form the coarse-grained granule F_{top}, which is then paired with the original label and added to the augmented dataset. Finally, Line 18 returns the constructed dataset E, containing query-aligned, relevance-filtered context representations.

Multilingual stance detection involves variability due to morphological differences, syntactic ordering, and language-specific lexical expressions. Directly modeling full raw sequences can amplify this variability and obscure signals relevant to stance. The granulation algorithm addresses this challenge by extracting semantically coherent fragments, enriching the feature space with additional informative patterns. During training, the model uses these patterns to better understand the relationships between textual elements, improving its ability to distinguish subtle differences between stances in the feature space. This granulation process operates across languages and is designed to support any language, enabling the model to consistently extract semantically coherent fragments regardless of linguistic structure.

Algorithm 1 Constructing coarse-grained granules

Input: Multilingual dataset $\mathcal{D} = \{(q_i, c_i, y_i)\}_{i=1}^{N}$, tokenizer $\mathcal{T}$, encoder $\mathcal{M}$, top-k parameter $k \in \mathbb{N}$

Output: Augmented dataset $\mathcal{E} = \{(\mathcal{F}_{\text{top}}, y)\}$
Initialize $\mathcal{E} \leftarrow \emptyset$
for each instance $(q, c, y) \in \mathcal{D}$ **do**
  Partition c into fragments $\mathcal{F}_c = \{f_1, \ldots, f_M\}$ using syntactic/discourse cues
  Tokenize $q \rightarrow (X_q, A_q) = \mathcal{T}(q)$
  Encode $X_q \rightarrow H_q = \mathcal{M}(X_q, A_q)$
  Compute pooled query embedding:

$$v_q = \frac{\sum_{j=1}^{|X_q|} A_q^{(j)} \cdot H_q^{(j)}}{\sum_{j=1}^{|X_q|} A_q^{(j)}}$$

  Initialize similarity list $S \leftarrow \emptyset$
  for each $f \in \mathcal{F}_c$ **do**
    Tokenize $f \rightarrow (X_f, A_f) = \mathcal{T}(f)$
    Encode $X_f \rightarrow H_f = \mathcal{M}(X_f, A_f)$
    Compute pooled fragment embedding:

$$v_f = \frac{\sum_{j=1}^{|X_f|} A_f^{(j)} \cdot H_f^{(j)}}{\sum_{j=1}^{|X_f|} A_f^{(j)}}$$

    Compute cosine similarity:

$$s_f = \frac{\langle v_q, v_f \rangle}{\|v_q\| \cdot \|v_f\|}$$

    Add (f, s_f) to S
  end for
  Sort S by descending s_f and select top-k fragments to form $\mathcal{F}_{\text{top}}$
  Append $(\mathcal{F}_{\text{top}}, y)$ to $\mathcal{E}$
end for
return $\mathcal{E}$

3.2. Graph Convolutional Network

Conventional stance detection approaches primarily rely on sequential modeling, which may overlook certain token-level interactions beyond standard positional encoding. To model local structural relationships more explicitly, a graph is constructed for each input sentence, and a graph convolutional network is applied over this structure.

The adjacency matrix follows a shifted identity pattern, creating connections between each token and its immediate neighboring token. This design encodes short-range contextual dependencies within the sentence, allowing adjacent tokens to directly exchange information through graph convolution operations. Rather than modeling full syntactic structures, the GCN focuses on enhancing local token interactions, reinforcing short-distance linguistic patterns that may be important for stance representation. By integrating these locally structured representations, the model refines contextual features before passing them to the quantum-inspired feature interaction module.

3.3. Quantum-Inspired Feature Extraction

Stance detection involves more than identifying isolated lexical cues; it requires capturing how expressions position themselves with respect to a target in context. Support and opposition are often conveyed through subtle semantic shifts, implicit evaluations, or the interaction between different components of an instance, such as the relationship between a question and its associated comment. A quantum-probability-inspired complex-valued representation encodes textual features in a form where amplitude reflects the strength

of a semantic signal, while phase represents how that signal relates to other features in the same context. This makes it possible to account for interaction patterns such as reinforcement or contrast. Such modeling is especially relevant in sentences containing negation, contrast, or indirect expressions of support. Rather than treating features as independent numerical values, this formulation captures their interdependence, producing representations that more accurately reflect argumentative orientation than conventional real-valued embeddings.

To capture high-order dependencies and complex-valued interactions among token representations, a component of the proposed model presents a quantum-probability-inspired complex-valued mechanism that operates over the output of a graph convolutional layer. By using principles from quantum computation, such as complex vector representations and phase-based modulation, the model constructs a richer, more expressive feature space that enhances its ability to distinguish subtle semantic patterns. The notation and conceptual framework are inspired by the quantum-inspired approaches introduced by Shi et al. [61] and Lai et al. [62]. Unlike purely real-valued transformations, complex-valued representations allow information to be encoded jointly in amplitude and phase. This double encoding mechanism brings useful patterns to the feature space. The latter feature is conceptually related to the idea of superposition and interference. It is worth noting that quantum-inspired does not signify the utilization of the hardware of quantum mechanics. Instead, it signifies the mathematical structure of complex Hilbert space. The goal is to embed quantum probabilistic geometry into a classical deep learning framework.

Let $H \in \mathbb{R}^{B\times T\times d}$ denote the batch-wise token-level output of the transformer-based model, where B is the batch size, T is the sequence length, and d is the embedding dimension.

We define a batch-wise adjacency matrix $A \in \mathbb{R}^{B\times T\times T}$, where each $A^{(b)}$ is a shifted identity matrix that captures local connectivity:

$$A^{(b)} = \text{Roll}(I_T, 1), \quad \text{for } b = 1, \ldots, B$$

where $I_T \in \mathbb{R}^{T\times T}$ is the identity matrix and Roll shifts each row downward by one.

The graph convolution operation is defined as:

$$H' = \text{ReLU}(AHW + b)$$

where $W \in \mathbb{R}^{d\times d'}$ is a learnable projection matrix, $b \in \mathbb{R}^{d'}$ is a learnable bias vector (broadcasted across batches and time steps), and $H' \in \mathbb{R}^{B\times T\times d'}$ is the updated representation.

The graph structure is incorporated prior to the quantum-inspired mapping in the graph-based transformation. This allows local dependencies and phase-aware feature representations to be captured simultaneously within a single framework.

To model richer feature interactions, the GCN output $H' \in \mathbb{R}^{B\times T\times d'}$ is projected into a quantum-inspired complex-valued space using a learnable projection matrix $W_{\text{mag}} \in \mathbb{R}^{d'\times d'}$:

$$M = H'W_{\text{mag}} \in \mathbb{R}^{B\times T\times d'}.$$

Here, M signifies the magnitude component of the complex representation.

A learnable tensor of phase angles $\theta \in \mathbb{R}^{1\times T\times d'}$ is introduced to simulate quantum-phase encoding. The real and imaginary parts of the complex-valued representation are computed as:

$$\Re(M) = \cos(\theta) \odot M, \quad \Im(M) = \sin(\theta) \odot M,$$

where $\odot$ denotes element-wise multiplication. The above formulation is equivalent to the formula for representing each characteristic in polar form, which directly maps into a

complex vector space. The resulting quantum-inspired activation is computed by summing the squared magnitudes of the real and imaginary components:

$$Q = \Re(M)^2 + \Im(M)^2 \in \mathbb{R}^{B \times T \times d'}.$$

This computation corresponds to the squared magnitude of a complex vector, analogous to measuring probability amplitudes in quantum mechanics. Although the operation is performed on classical hardware, the magnitude-based aggregation maintains the relative phase angles in the complex-valued representations. In this context, the phases correspond to the learnable parameters that modulate the real and imaginary components of each feature, effectively encoding contextual relationships between token representations.

To aggregate this feature matrix, an attention-based pooling mechanism is applied. A learnable context vector $u \in \mathbb{R}^{d'}$ is used to compute attention weights over the token dimension:

$$\alpha_i = \frac{\exp(u^\top Q_i)}{\sum_{j=1}^{T} \exp(u^\top Q_j)} \quad \text{for } i = 1, \ldots, T,$$

where $Q_i \in \mathbb{R}^{d'}$ denotes the representation of the i-th token. The aggregated vector $z \in \mathbb{R}^{B \times d'}$ is then obtained via a weighted sum:

$$z = \sum_{i=1}^{T} \alpha_i Q_i.$$

The step is equivalent to an operator that extracts the most important quantum-inspired information based on the learned weights. This pooled representation is passed through a two-layer feedforward network for classification:

$$h_1 = \text{ReLU}(z), \quad h_2 = \text{Dropout}(h_1),$$

$$\hat{y} = \text{softmax}().$$

The output $\hat{y}$ denotes the predicted class probabilities. The model is trained using sparse categorical cross-entropy loss and optimized via the Adam optimizer.

In this paper, quantum theory is used as a mathematical inspiration rather than as a physical computational model. In the proposed QFIM, semantic interactions are represented through amplitude-based feature transformations, where the magnitude of each component reflects the strength of a semantic signal. This formulation allows multiple contextual cues to be represented jointly, enabling the model to capture overlapping semantic influences within a text segment instead of forcing them into a single linear representation. In addition, relationships between contextual elements are modeled through the phase (alignment) parameters within the QFIM layer, which regulate how features interact with one another and provide a mechanism for capturing complex dependencies across components.

A key distinction between classical attention mechanisms and the QFIM-based approach lies in the way information is aggregated. Classical attention combines features through weighted linear summation, which primarily reflects relative importance. In contrast, QFIM models interactions through amplitude-driven feature composition, where relationships between concepts are encoded through their interaction strength rather than only through additive weighting. This interaction-based formulation facilitates the representation of hierarchical and interdependent semantic structures, supporting more effective modeling of complex linguistic dependencies.

3.4. Model Variants

To enhance the model's performance, several variants of the proposed model were constructed and combined using an ensemble strategy based on the maximum voting method. The considered variants are as follows:

- Quantum-based deep learning model, GCN, and granulation with Global Max Pooling.
- Only the Quantum-based deep learning model, excluding GCN and granulation.
- Quantum-based deep learning model and GCN, without granulation.
- Full architecture with quantum-based deep learning model, GCN, and granulation.
- Quantum-based deep learning model and granulation, without GCN.

The architecture of our proposed quantum-probability-inspired complex-valued deep learning model for multilingual stance detection is illustrated in Figure 1. The model processes multilingual input data through a series of stages, integrating granulation, graph convolution, and QFIM to produce a stance prediction.

The process begins with the Multilingual Input ($\mathcal{X} = (q, c)$) (Node 1), where q represents the question and c denotes the comment. This input is processed by a Pretrained Contextual Language Model (M) (Node 2), such as XLM-RoBERTa, to generate contextual embeddings, denoted as H. These embeddings serve as the foundation for subsequent processing, splitting into two parallel paths.

The first path, labeled "Granulation," focuses on extracting fine-grained features from the comments and questions. This branch comprises four steps:

- Partition Comment ($c \mapsto \{f_1, \ldots, f_M\}$) (Node 3): The context c is divided into a set of feature segments $\{f_1, \ldots, f_M\}$.
- Pooled Embeddings (v_q, v_f) (Node 4): Embeddings are pooled to produce vectors v_q for the query and v_f for each feature segment.
- Cosine Similarity ($s_f = \cos(v_q, v_f)$) (Node 5): The similarity between v_q and v_f is computed using cosine similarity, yielding a score s_f.
- Top-k Granules ($\text{Top}_k(q, c)$) (Node 6): The top-k most relevant granules are selected based on the similarity scores.

The granulation branch finishes up at the Augmented Feature Space ($\mathcal{F}_{\text{top}}$) (Node 7), where it gathers the top-k granules and turns them into a beefed-up feature set for the next steps. This output is then fed into the main processing pipeline via a black arrow, connecting to the Graph Convolutional Layer.

The contextual embeddings H are directly processed by the Graph Convolutional Layer ($H' = \text{ReLU}(AHW + b)$) (Node 8). This layer applies a graph convolution operation. The ReLU activation function produces the updated feature matrix H', which is passed to the next stage, labeled H'.

The subsequent stage involves quantum-inspired feature processing:

- Project to Complex Space ($M = H'W_{\text{mag}}$) (Node 9): The features H' are projected into a complex space using a magnitude weight matrix W_{mag}, yielding M.
- Phase Modulation ($\Re(M) = \cos(\theta) \odot M$) (Node 10): The real part of M is modulated by element-wise multiplication with $\cos(\theta)$, where θ is a phase angle.
- Quantum Activation ($Q = \Re(M)^2 + \Im(M)^2$) (Node 11): A quantum-inspired activation function computes Q as the sum of the squares of the real ($\Re(M)$) and imaginary ($\Im(M)$) parts of M.
- Attention Pooling ($z = \sum \alpha_i Q_i$) (Node 12): An attention mechanism pools the features into a vector z, using attention weights α_i to combine elements Q_i.

The output z from the QFIM stage is fed into a Feedforward Network + Softmax ($\hat{y} = \text{softmax}(\cdot)$) (Node 13), which performs classification to produce the predicted prob-

abilities $\hat{y}$. Finally, the Stance Prediction ($\hat{y}$) (Node 14) represents the model's output, indicating the predicted stance (favor or against).

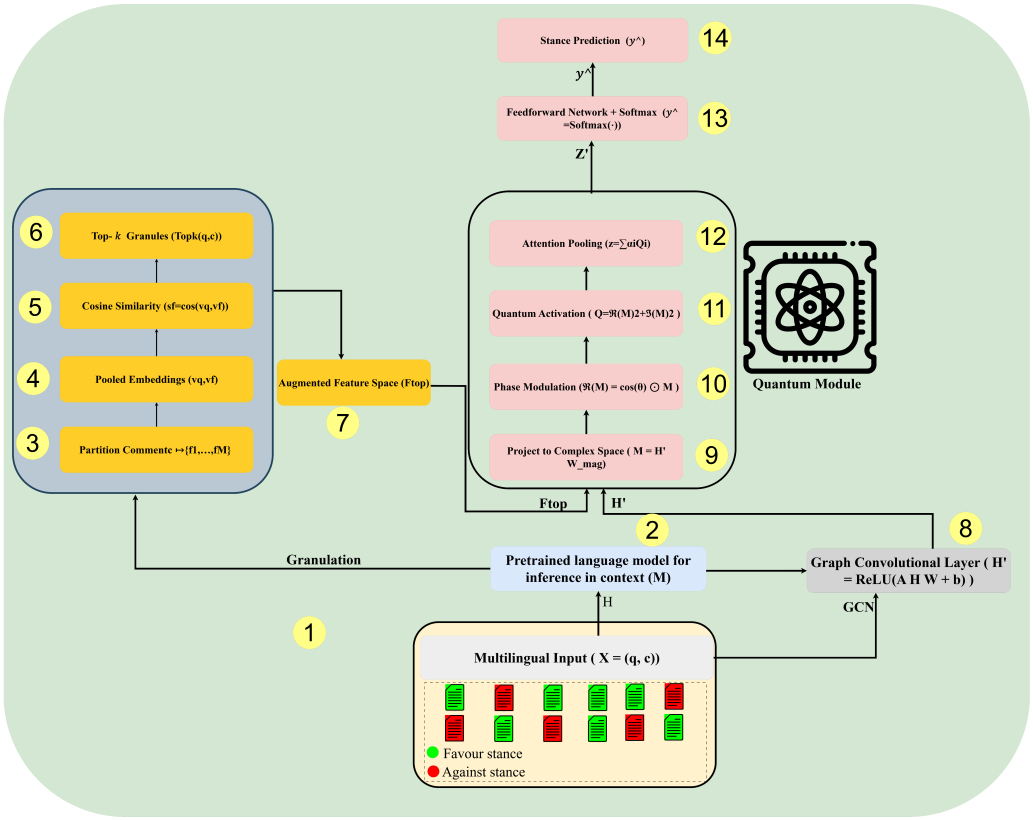


Figure 1. Architecture of the proposed quantum-probability-inspired complex-valued deep learning model for multilingual stance detection.

4. Experimental Results

In this section, the results and data obtained from the experimental evaluation of the proposed model are presented and analyzed.

- Experimental Setup

The model was trained for 10 epochs with a batch size of 64, using the Adam optimizer and a learning rate of 2×10^{-5}. Input sequences were either padded or truncated to a fixed length of 128 tokens. Contextual token embeddings were generated using the pre-trained XLM-RoBERTa-large-XNLI model. These embeddings were processed through a GCN layer with a hidden dimension of 256 to capture local structural information. A QFIM with a hidden dimension of 128 was subsequently applied to enrich feature representations. Token-level features were aggregated using an attention pooling mechanism, and the resulting vector was passed through a dense layer with 64 units and ReLU activation, followed by a dropout layer with a rate of 0.3. The final predictions were obtained from a softmax classification layer. All experiments were executed in TensorFlow within the Google Colab environment, with GPU acceleration to optimize training efficiency.

The dataset used in this study [8] was divided into three subsets: 8000 instances for training, 3926 for validation, and 15,000 for testing. The dataset contains data in German, French, and Italian.

4.1. Evaluation Metrics

Model performance was evaluated using standard classification metrics to enable comparisons with baseline approaches, as defined in Equations (1)–(6). These metrics

rely on four possible classification outcomes. True Positives (TP) are instances correctly predicted as positive, True Negatives (TN) are instances correctly predicted as negative, False Positives (FP) are negative instances incorrectly predicted as positive, and False Negatives (FN) are positive instances incorrectly predicted as negative. Based on these outcomes, the following metrics were computed:

- Precision (Positive Class): the proportion of predicted positives that are correct.
- Recall (Positive Class): the proportion of actual positives correctly identified.
- Precision (Negative Class): the proportion of predicted negatives that are correct.
- Recall (Negative Class): the proportion of actual negatives correctly identified.
- F1 Score: the harmonic mean of precision and recall.
- Accuracy: the proportion of all instances correctly classified.

$$\text{Precision (Positive Class)} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{1}$$

$$\text{Recall (Positive Class)} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{2}$$

$$\text{Precision (Negative Class)} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{3}$$

$$\text{Recall (Negative Class)} = \frac{\text{TN}}{\text{TN} + \text{FN}} \tag{4}$$

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{5}$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{6}$$

All evaluations were conducted using both validation and test datasets to ensure robust and generalizable results.

- Ablation study

The ablation study is used as a systematic approach to evaluate the individual contribution of key components within the proposed model by incrementally removing or altering them and observing the resulting changes in performance. The following model configurations were examined:

1. The pretrained large language model XLM-RoBERTa-large-XNLI was solely utilized without any architectural modifications.
2. A configuration was used in which the quantum model, GCN layer, and granulation mechanism were integrated, with the attention pooling layer replaced by a Global Max Pooling operation.
3. A configuration was evaluated where only the quantum model was used, and both the GCN layer and the granulation mechanism were excluded.
4. A setup was tested in which the quantum model and GCN layer were combined, while the granulation mechanism was omitted.
5. The complete proposed architecture was implemented by integrating the quantum model, GCN layer, and granulation mechanism.
6. A configuration was applied in which the quantum model and granulation mechanism were included, but the GCN layer was removed.
7. An ensemble approach was employed, aggregating predictions from the five aforementioned configurations (steps 2 to 5) to evaluate the effect of model diversity.

Table 2 illustrates the contribution of different architectural components to overall model performance. Using only the pretrained language model (PLM) produced a reasonably strong baseline. Adding the quantum model, GCN layer, and Global Max Pooling

provided modest improvements, particularly in the validation F1 scores for class 0. When granulation was excluded while retaining the quantum model and GCN, test performance decreased slightly, highlighting the usefulness of the granulation mechanism. Removing both the GCN and granulation layers still yielded competitive results, indicating that the quantum model alone provides substantial representational capacity. The complete model, integrating the quantum model, GCN, and granulation, achieved the best overall performance with balanced F1 scores across classes. An ensemble of quantum-based models further improved results on both validation and test sets, achieving the highest observed accuracy (79.57% on validation, 79.25% on test) and balanced F1 scores (79.64% and 79.50% on validation; 78.48% and 79.96% on test), confirming the additive value of each component and the benefits of ensemble integration.

Table 2. Performance comparison of model configurations on validation and test datasets.

Method	Split	Accuracy	Prec 0	Pre 1	Rec 0	Rec 1	F1_0	F1_1
PLM [1] only	Val	78.58	79.76	77.48	76.72	80.45	78.21	78.94
	Test	76.19	76.18	76.19	73.37	78.79	74.75	77.47
QFIM + GCN + MaxPool	Val	78.71	76.71	81.04	82.56	74.83	79.53	77.81
	Test	77.07	74.04	80.39	80.50	73.91	77.13	77.01
Quantum [2]	Val	77.25	77.29	77.22	77.33	77.18	77.31	77.20
	Test	77.21	76.45	77.90	75.95	78.37	76.20	78.14
QFIM + GCN [3]	Val	78.25	78.91	77.61	77.22	79.28	78.06	78.43
	Test	76.13	75.04	77.15	75.38	76.83	75.21	76.99
Full model [4]	Val	78.37	78.96	77.81	77.48	79.28	78.21	78.53
	Test	77.64	77.47	77.79	75.36	79.74	76.40	78.75
QFIM + Gran − noGCN [5]	Val	76.85	77.18	76.52	76.36	77.34	76.77	76.92
	Test	76.43	75.19	77.59	76	76.82	75.59	77.20
QFIM Ensemble [6]	Val	79.57	79.52	79.62	79.77	79.38	79.64	79.50
	Test	79.25	78.18	80.25	78.78	79.68	78.48	79.96

[1] Xlm-roberta-large-xnli; [2] (no GCN, Granulation); [3] without Granulation; [4] (QFIM + GCN + Granulation); [5] QFIM + Granulation without GCN; [6] Ensemble of all QFIM variations.

The radar chart in Figure 2 compares validation and test accuracy along with class-wise F1 scores, providing a clear visual representation across models. The ensemble of QFIM variations achieved the highest scores, reflecting both strong overall accuracy and balanced performance across classes on unseen test data. The full model (QFIM + GCN + Granulation) followed closely behind, significantly improving F1 for Class 1 on the test set, indicating enhanced sensitivity to class-specific patterns. Models that excluded either GCN or granulation exhibited noticeable reductions in test accuracy and class-wise F1 scores.

- Analysis of Validation Results

Table 3 shows how different models performed on the validation dataset. The Proposed Model stood out by achieving the highest overall accuracy at 79.57%. More importantly, it maintained a well-balanced performance across both classes, with F1-scores of 79.64 for class 0 and 79.50 for class 1. This suggests it generalizes well and handles both classes equally, without favoring one over the other. Among the baseline models, XLM-RoBERTa and Multilingual-e5-large came closest, reaching an accuracy of 78.58% and 75.34% showing similarly balanced F1-scores. This highlights its strength as a multilingual pretrained model and confirms its reliability for tasks involving multiple languages. Other pretrained language models, such as LaBSE and mBART, exhibited relatively modest performance. Traditional models like Logistic Regression and SGD Classifier also performed well, with accuracies between 69% and 71%. Although they do not match the transformer-based

models in terms of accuracy, they still provide consistent results and are much easier to interpret, making them good options for simpler or resource-limited settings. Tree-based models performed moderately well too. Random Forest achieved an accuracy of 70.78%, showing that it could still compete thanks to its ensemble structure. Gradient Boosting, on the other hand, had a stronger recall for class 0 (72.69) but struggled with class 1 (recall of 57.50). This imbalance might be due to its sensitivity to noise or overfitting to certain patterns in the training data. Other traditional models such as Naive Bayes, Decision Tree, and K-Nearest Neighbors (KNN) showed moderate performance, with accuracies ranging from 62% to 66%. These models generally found it more difficult to capture the complexity of class 1, which affected their overall effectiveness. Neural models like CNN, CNN-BiLSTM, and Transformer did not perform as well, with accuracies falling below 60%. Their weaker results may be due to a lack of deep contextual understanding or insufficient training on multilingual data. The Hierarchical Attention Network (HAN) showed a pronounced class imbalance, achieving perfect recall for class 0 but failing entirely on class 1, leading to an overall accuracy of 49.90%. This indicates difficulty in capturing features distinguishing class 1.

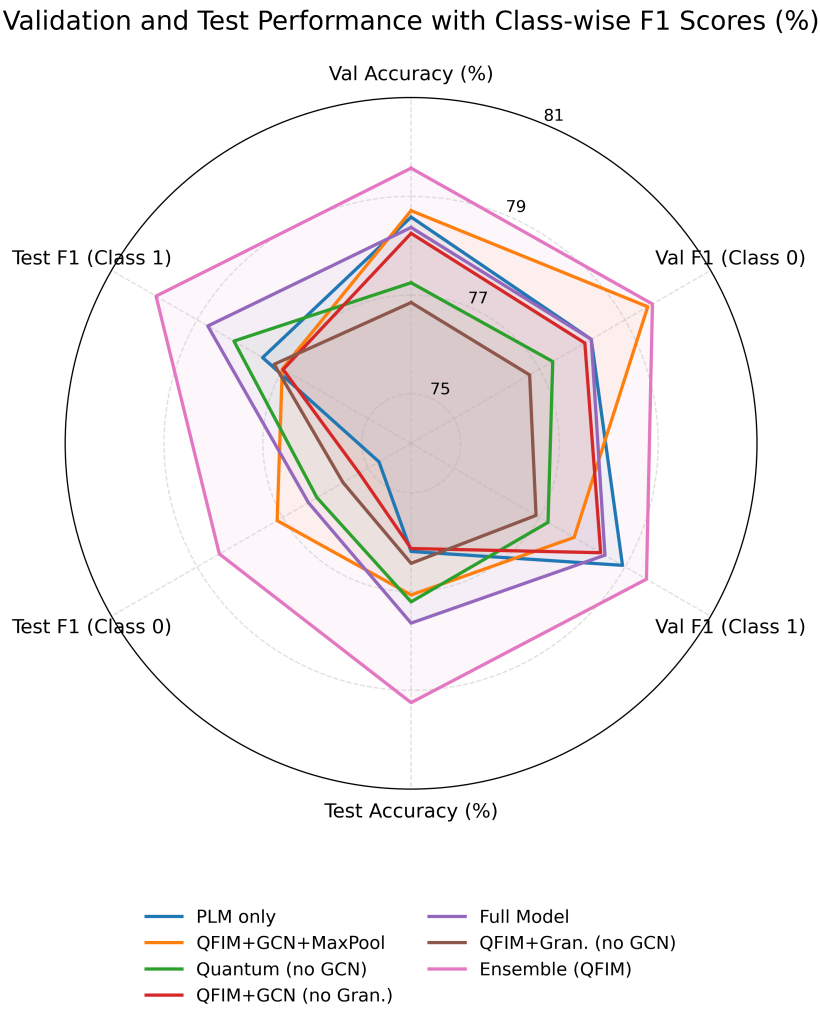


Figure 2. Radar chart comparing validation and test performance across model configurations.

In conclusion, these results clearly highlight the strength of pretrained transformer models, particularly multilingual ones. The proposed model goes a step further by not only achieving the highest accuracy but also maintaining strong balance across both classes.

This demonstrates a clear improvement over all baselines and confirms its effectiveness on the validation set.

Table 3. Comparison of performance of deep Learning and machine learning models against the proposed model in multilingual stance detection under validation data.

Model	Prec 0	Rec 0	F1 0	Prec 1	Rec 1	F1 1	Accuracy
Logistic Regression	71.39	71.82	71.60	71.76	71.33	71.55	71.57
Naive Bayes	65.85	68.81	67.30	67.48	64.46	65.94	66.63
Decision Tree	61.69	63.04	62.36	62.37	61.01	61.68	62.02
KNN	66.29	62.94	64.57	64.86	68.12	66.45	65.54
Random Forest	69.80	73.05	71.39	71.86	68.53	70.15	70.78
Gradient Boosting	63.01	72.69	67.50	67.89	57.50	62.26	65.08
SGD	70.24	68.81	69.52	69.56	70.97	70.26	69.89
Multilingual BERT [63]	68.45	59.47	63.64	64.30	72.70	68.24	66.10
XLM-RoBERTa [64]	79.76	76.72	78.21	77.48	80.45	78.94	78.58
CNN-BiLSTM [65]	59.37	49.31	53.88	56.81	66.40	61.23	57.87
Transformer [52]	56.79	70.24	62.80	61.21	46.77	53.03	58.48
CNN [66]	57.22	73.20	64.23	63.03	45.50	52.85	59.32
HAN [67]	49.90	100	66.58	0	0	0	49.90
Multilingual-e5 [68]	74.76	76.37	75.56	75.95	74.33	75.13	75.34
MBART [69]	67.07	61.56	64.20	64.61	69.90	67.16	65.74
LaBSE [70]	73.11	66.77	69.80	69.54	75.55	72.42	71.17
Proposed Model	79.52	79.77	79.64	79.62	79.38	79.50	79.57

- Analysis of Test Results

Table 4 shows the evaluation of all models on the unseen test dataset, providing an assessment of their ability to generalize to new data. The proposed model achieved the highest accuracy at 79.25%, while maintaining a balanced performance across classes, with F1-scores of 78.48 for class 0 and 79.96 for class 1. This balance indicates that the model performs consistently across categories. This is an important consideration for real-world applications, where misclassification in either class could have significant consequences.

XLM-RoBERTa and Multilingual-e5-large performed well, achieving test accuracies of 76.19% and 74.39%, respectively, with relatively balanced F1-scores. Their consistent performance across validation and test sets demonstrates the effectiveness of multilingual pretrained transformers for general-purpose classification tasks. Not all pretrained models performed equally well; for example, mBART and LaBSE achieved only modest results on the test data. The proposed model outperformed these baselines, attaining higher overall accuracy and maintaining a more balanced performance across both classes. This suggests that it can overcome some limitations of existing pretrained models, particularly in challenging classification scenarios.

Traditional machine learning models, such as Logistic Regression and SGD, showed relatively low performance, with accuracies between 63% and 65%. These models remain valuable due to their simplicity and low computational requirements, making them useful as baseline comparisons. However, their performance varied across classes, suggesting difficulty in capturing patterns for certain categories.

Table 4. Comparison of performance of deep learning and machine learning models against the proposed model in multilingual stance detection under test data.

Model	Prec 0	Rec 0	F1 0	Prec 1	Rec 1	F1 1	Accuracy
Logistic Regression	66.87	62.40	64.56	62.37	66.83	64.52	64.54
Naive Bayes	62.24	60.20	61.20	58.75	60.82	59.77	60.50
Decision Tree	57.23	54.87	56.03	53.64	56.02	54.80	55.43
KNN	58.73	60.19	59.45	56.12	54.62	55.36	57.50
Random Forest	62.48	69.02	65.59	62.57	55.55	58.85	62.52
Gradient Boosting	59.42	71.55	64.92	60.93	47.59	53.44	59.99
SGD	66.51	58.90	62.48	60.73	68.19	64.25	63.38
Multilingual BERT [63]	67.60	58.43	62.68	60.78	69.70	64.94	63.85
XLM-RoBERTa [64]	76.18	73.37	74.75	76.19	78.79	77.47	76.19
CNN-BiLSTM [65]	57.24	48.61	52.57	52.20	60.71	56.13	54.42
Transformer [52]	54.42	67.25	60.16	52.44	39.07	44.78	53.71
CNN [66]	57.09	72.78	63.99	58.09	40.82	47.95	57.43
HAN [67]	51.75	100	68.21	0	0	0	51.75
Multilingual-e5 [68]	74.24	77.67	75.91	74.57	70.84	72.65	74.39
MBART [69]	66.63	59.94	63.11	60.90	67.52	64.04	63.58
LaBSE [70]	72.25	68.80	70.48	67.90	71.41	69.61	70.05
Proposed Model	78.18	78.78	78.48	80.25	79.68	79.96	79.25

Ensemble-based models such as Random Forest and Gradient Boosting produced mixed results. Random Forest achieved a test accuracy of 62.52% and demonstrated strong recall for class 0 at 69.02%, but performance on class 1 was lower. Gradient Boosting similarly showed high recall for class 0 at 71.55%, while recall for class 1 dropped to 47.59, reducing its overall accuracy to 59.99%.

Multilingual BERT experienced a slight drop compared to its validation performance, reaching 63.85% on the test set. This decrease may reflect mild overfitting or subtle domain differences between training and test data. The relatively weak performance emphasizes the need for careful fine-tuning or regularization strategies.

In contrast, models such as CNN, CNN-BiLSTM, and Transformer performed notably worse than their pretrained counterparts, with test accuracies between 53% and 57%. This suggests that, without pretrained knowledge, these models struggled to extract meaningful features from the available training data. The results underscore the importance of leveraging pretrained models, which provide a strong foundation for handling complex language patterns in text classification tasks.

The HAN model encountered significant difficulties on the test data. Although it achieved perfect recall for class 0, it failed to identify any instances of class 1, resulting in a class 1 recall of 0% and an overall accuracy of 51.75%. This behavior indicates that the model defaulted to predicting a single class, likely due to extreme class imbalance, limited training capacity, or insufficient regularization. Without major adjustments, its applicability remains limited.

Overall, these test results confirm the robustness of the proposed model. It achieves the highest overall accuracy while maintaining balanced performance across both classes, a level of fairness that other models did not achieve. Traditional and tree-based models pro-

vide reasonable performance, and pretrained transformers such as XLM-RoBERTa remain strong models; however, the proposed model clearly outperforms all of these approaches.

The strong performance of the proposed model can be attributed to the integration of the quantum-inspired feature interaction module and the granulation mechanism. These components jointly enhance the model's ability to capture subtle and structured patterns in multilingual text, supporting accurate stance classification. The QFIM models complex, non-linear, and high-order semantic relationships through amplitude-based interactions, enabling the representation of dependencies that conventional architectures may not fully capture.

The granulation algorithm further refines the feature space by constructing semantically coherent textual units based on contextual embedding similarity. This process strengthens alignment between related text segments and provides a richer representation of nuanced linguistic structures. The combination of these mechanisms supports stable generalization and balanced performance across classes, leading to consistent improvements over baseline approaches and demonstrating the effectiveness of integrating quantum-probability-inspired complex-valued modeling with granular computing in multilingual stance detection.

Figure 3 presents the generalization performance of the evaluated models, highlighting the differences in accuracy between validation and test phases. The proposed quantum-probability-inspired complex-valued deep learning model records the highest accuracy in both settings, indicating stable performance across varying data distributions. The following observations can be drawn from Figure 3:

1. Generalization in the proposed model: Traditional models such as Logistic Regression and SGD achieve reasonable validation accuracy but experience noticeable declines on the test set (e.g., approximately 7% for Logistic Regression). This pattern indicates sensitivity to validation-specific characteristics and limited adaptability to unseen data. Random Forest and Gradient Boosting show similar trends, with performance decreasing from validation to test (for example, Gradient Boosting declines from 65.08% to 59.99%). In comparison, the proposed model maintains stable performance across both distributions.
2. Challenges for non-pretrained transformer models: Models such as Transformer and CNN-BiLSTM, which lack extensive pretraining, exhibit weak performance and noticeable variation between validation and test phases. This pattern indicates the importance of large-scale pretraining and careful fine-tuning for multilingual stance detection tasks.
3. HAN performance under class imbalance: The HAN achieves perfect recall for class 0 but fails to correctly identify instances of class 1, resulting in a test accuracy of 51.75%. This outcome suggests strong bias toward a single class and limited robustness under imbalanced conditions.
4. XLM-RoBERTa performance: XLM-RoBERTa maintains stable accuracy across validation (78.58%) and test (76.19%) sets, outperforming all baseline models except the proposed approach. This consistency reflects the benefits of large-scale multilingual pretraining.
5. Limitations of CNN-based models: CNN, CNN-BiLSTM, and non-pretrained Transformer architectures consistently show lower performance. These results suggest that convolutional features and standard attention mechanisms may not sufficiently capture complex cross-sentence dependencies in multilingual stance detection.

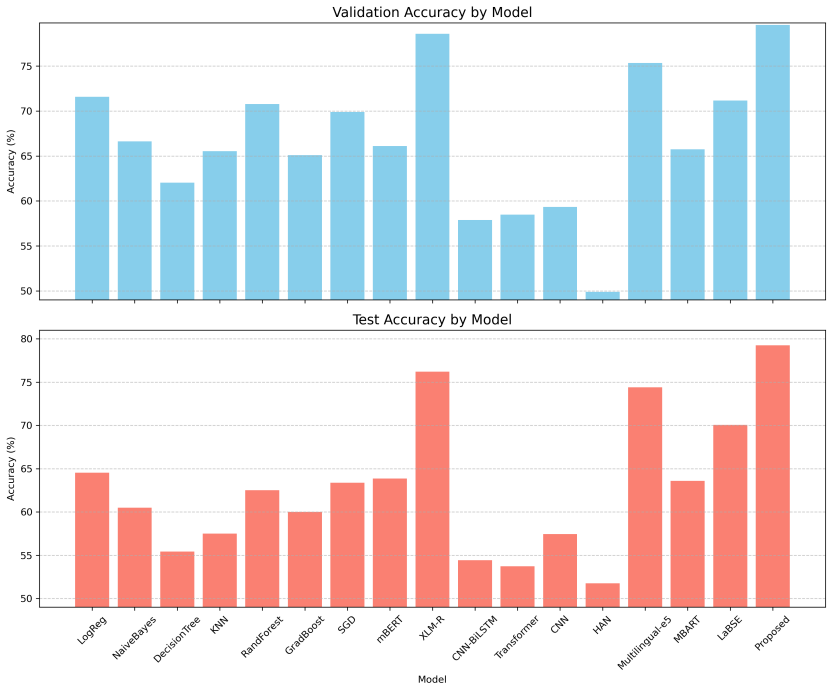


Figure 3. Comparison of validation and test accuracy for various machine learning and deep learning models on the multilingual stance detection task.

Figure 4 presents the F1 scores for classes 0 and 1 across both validation and test sets, providing a detailed view of model behavior in the multilingual stance detection task. The following observations provide insight into class-specific performance patterns:

1. Class-balanced stability of the proposed model: The proposed model achieves nearly identical F1 scores for both classes on the validation set (79.64 for class 0, 79.50 for class 1) and the test set (78.48 for class 0, 79.96 for class 1). The small maximum deviation of approximately 1.5 points indicates that the combination of the quantum-inspired feature interaction module and the granulation mechanism effectively captures detailed semantic patterns for both classes, maintaining consistent performance across linguistic and distributional variations.
2. Class-specific performance disparities: F1 scores indicate that most models, with the exception of the proposed model and XLM-RoBERTa, show significant differences in performance between classes 0 and 1, especially on the test set. For example, Gradient Boosting achieves a test F1 score of 64.92 for class 0 but only 53.44 for class 1, a gap of approximately 11.5 points. This pattern suggests that class 1, which may represent more complex linguistic structures, is more challenging, potentially due to the diversity present in the multilingual dataset.
3. Model complexity versus performance: Figure 4 indicates that simpler models, such as Logistic Regression and Naive Bayes, sometimes outperform more complex non-pretrained neural models, including CNN and Transformer, in terms of F1 scores, particularly for class 1. For instance, Logistic Regression achieves a test F1 of 64.52 for class 1, compared with 47.95 for CNN. This observation suggests that increased architectural complexity alone does not ensure improved performance in multilingual stance detection, especially in the absence of extensive pretraining or effective feature representation.

In response to Research Question 1, the results indicate that the proposed quantum-inspired model with the feature interaction module outperforms traditional transformer-based models, such as XLM-RoBERTa, in multilingual stance detection. For Research

Question 2, the context-aware granulation mechanism enhances the model's capacity to identify complex stance structures across languages by generating semantically coherent textual granules.

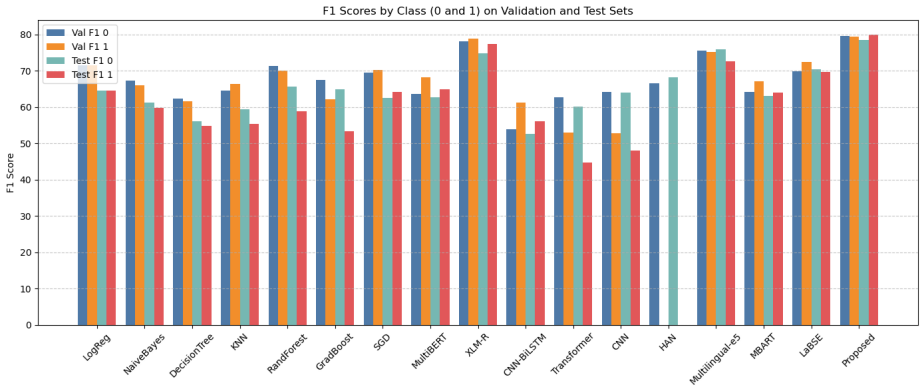


Figure 4. Comparison of F1 scores for deep learning and machine learning models versus the proposed model in multilingual stance detection.

Although the primary focus of this study is not on optimizing computational time, and despite the general consensus that deep learning models tend to have higher time complexity, the combined training and prediction times for all models were measured and compared with their accuracy in Table 5. Models such as Multilingual BERT exhibit lower computational cost but also achieve lower accuracy. In contrast, the proposed model attains the highest accuracy among all models while maintaining a moderate computational cost. For instance, compared with the Multilingual-e5-large model, the proposed model shows higher accuracy with reduced computational time.

Table 5. Training and Prediction Time and Accuracy of Different Models.

Model	Training + Prediction Time	Accuracy (%)
Multilingual-e5-large	94 min 55.94 s	74.39
Multilingual BERT	20 min 01.46 s	63.85
XLM-RoBERTa	64 min 20.86 s	76.19
LaBSE	56 min 16.09 s	70.05
Proposed Model	82 min 45.06 s	77.64

The key findings of this paper can be summarized as follows:

1. The proposed quantum-probability-inspired complex-valued deep learning model improves multilingual stance detection by capturing complex semantic relationships using the proposed granulation algorithm and quantum-inspired feature extraction.
2. A new context-aware granulation mechanism, based on contextual embedding, augments the feature space with semantically coherent textual granules, improving question–comment alignment and enhancing the model's robustness in handling detailed stance detection across languages.
3. Extensive experiments on benchmark datasets demonstrate that the proposed model outperforms state-of-the-art methods, achieving the highest accuracy and balanced F1 scores, emphasizing the additive value of combining quantum-inspired and granular computing techniques to address semantic ambiguity and cross-lingual variability.

4.2. Error Analysis

In this subsection, we present a detailed error analysis of the proposed model (due to the computational intensity of the experiments, the model was run for four epochs).

Figure 5 depicts the model's performance from an error analysis perspective, examining how it behaves across different stance categories, text lengths, and languages. Clear variation appears at the stance level, where the AGAINST class achieves higher accuracy. This performance gap indicates that the model is more effective at capturing AGAINST stances.

In contrast, the performance across different text lengths is much more stable. Short and medium texts obtain very similar Accuracy and Macro-F1 scores, suggesting that the model can reliably handle moderate differences in input length. A similar trend is observed across languages, where the results remain consistent for German, French, and Italian, with French showing a slight advantage. The close alignment between Accuracy and Macro-F1 across languages further indicates balanced predictions and strong cross-lingual generalization. Therefore, these results suggest that the model's main challenge lies in distinguishing between stance categories rather than dealing with variations in text length or language.

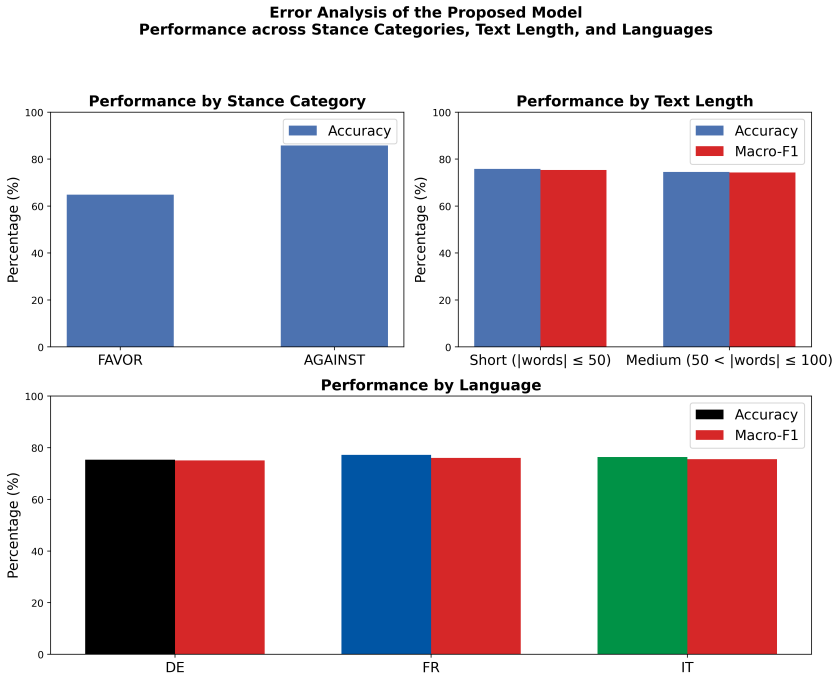


Figure 5. Error analysis of the proposed model showing performance variations across stance categories, text length groups, and languages.

The granulation parameter k plays an important role in controlling the level of semantic detail retained by the proposed context-aware filtering mechanism. Specifically, k determines how many query-relevant fragments are selected from the original comment. As such, lower k values enforce stricter selection of contextual information, producing a more compact and focused representation of the original comment. Conversely, higher k values preserve more contextual content, which may introduce noisy or weakly discriminative granules.

The sensitivity of class-wise F1 scores to varying values of k is illustrated in Figure 6. The plot shows that when $k = 2$, the model achieves the highest and most balanced class-wise F1 scores, with $F1_0$ = 78.48% and $F1_1$ = 79.96%. Afterwards, as k is increased, F1 values for Class 0 continue to decrease, while F1 values for Class 1 remain nearly constant. These results indicate that as larger granularity sizes are incorporated, semantically weaker information is reintroduced. Therefore, the granulation algorithm directly affects overall performance when the size of the granules is changed.

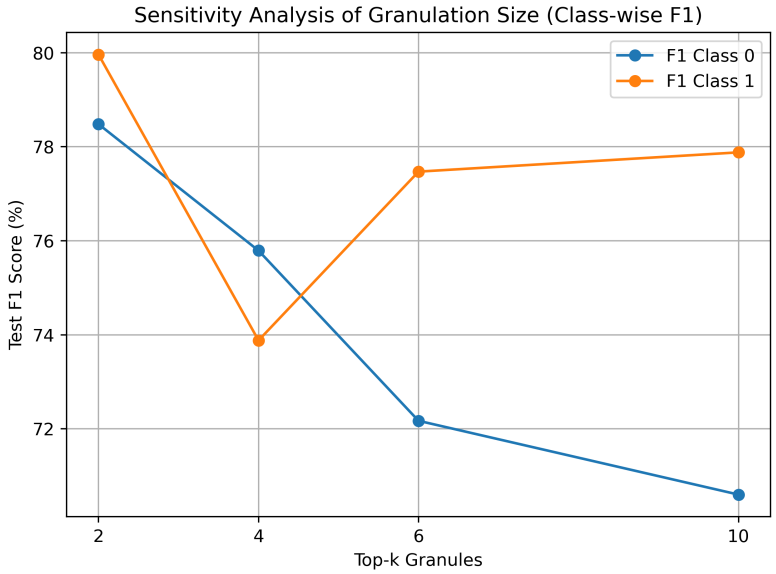


Figure 6. Sensitivity analysis for granulation parameter k on class-wise F1 scores (Test set). The obtained results confirm that smaller k values provide a more balanced and discriminative approach, with optimal values achieved at $k = 2$.

To provide a comprehensive evaluation of the proposed approach, its key advantages and potential limitations are summarized as follows.

Strengths:

- Quantum-probability-inspired complex-valued design: The model employs a quantum-inspired feature interaction module to produce enriched representations, enabling it to capture non-linear and high-order dependencies that are often missed by traditional models.
- Multilingual robustness: Through the integration of context-aware granulation and graph-based reasoning, the model effectively addresses cross-lingual variability and semantic ambiguity in multilingual stance detection.
- Component cooperation: The combination of granular computing, quantum-inspired modules, and graph convolutional networks contributes to improved performance, as demonstrated by ablation studies.
- Empirical performance: Benchmark evaluations indicate that the proposed model surpasses state-of-the-art baselines, including multilingual transformers such as XLM-RoBERTa and Multilingual-e5-large.

Limitations:

- Computational complexity: The inclusion of textual granules increases training overhead, which may require more computational resources than simpler architectures. However, this cost is fixed and does not substantially affect overall training performance.
- Interpretability challenges: The proposed model does not include built-in interpretability features. Understanding the reasoning behind model predictions is important for building trust, diagnosing errors, and ensuring responsible use.
- Limited testing: Evaluation was conducted only on a benchmark dataset. Applying the proposed model to large-scale multilingual data with noisy, informal, or highly varied content could present additional challenges.
- Lack of experimentation on physical quantum hardware: Simulations were not performed on actual quantum devices due to the current limitations of noisy intermediate-scale quantum (NISQ) systems, which have restricted qubit counts and non-negligible

error rates. These constraints can affect computational outcomes and hinder a thorough empirical evaluation of the theoretical contributions. Moreover, the limited availability of quantum hardware remains an additional practical constraint.

5. Conclusions

This study presented a novel quantum-probability-inspired complex-valued deep learning model for multilingual stance detection, integrating principles from granular computing and quantum theory to address semantic ambiguity and cross-lingual variability. The model incorporated a quantum-inspired feature interaction module that captured complex, high-order, and non-linear semantic dependencies through amplitude-based interactions, emulating quantum phenomena such as superposition and interference. A context-aware granulation mechanism, based on contextual embedding similarity, extracted semantically coherent text granules, enhancing question–comment alignment and enriching the feature space with interpretable substructures. Combined with graph convolutional networks, these components enabled effective modeling of nuanced stance cues across diverse languages. Experiments on benchmark datasets demonstrated that the model consistently outperformed state-of-the-art models, achieving the highest accuracy. The ablation study confirmed the additive value of each component, with the ensemble of quantum-inspired variants further enhancing performance. Compared to baseline models, including pretrained transformers like XLM-RoBERTa and traditional machine learning approaches, the model showed superior generalization and robustness, particularly in maintaining class-balanced performance across multilingual data. From a theoretical perspective, the study integrates quantum-inspired principles and granular computing with natural language processing, providing a novel approach to modeling complex semantic relationships. From a practical perspective, the model demonstrates improved performance, making it a useful tool for applications such as fake-news detection, sentiment analysis, and policy recommendation systems, where accurate multilingual stance detection is essential. While the proposed model achieved significant advancements in multilingual stance detection, several directions for future research could further enhance its capabilities. Future research could explore the performance of the model on larger and more diverse multilingual datasets to assess its scalability and robustness across different languages and domains. The quantum-inspired feature interaction module could be adapted for other NLP tasks to evaluate its general applicability. Alternative granulation techniques, such as those based on discourse analysis or syntactic structures, may further improve the extraction of semantically coherent granules, enhancing alignment and interpretability. Expanding the model to include additional languages would broaden its practical utility. Finally, developing methods to visualize and interpret quantum-inspired feature interactions and granule contributions could offer deeper insights into the model's decision-making process, increasing transparency and trust for real-world applications.

Author Contributions: M.H.: supervision, validation, writing—review and editing, project administration. M.E.A.: writing, editing, formal analysis, investigation, implementation, data collection. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

References

1. Mohammad, S.; Kiritchenko, S.; Sobhani, P.; Zhu, X.; Cherry, C. SemEval-2016 Task 6: Detecting Stance in Tweets. In *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016), San Diego, CA, USA, 16–17 June 2016*; Bethard, S., Carpuat, M., Cer, D., Jurgens, D., Nakov, P., Zesch, T., Eds.; Association for Computational Linguistics: Stroudsburg, PA, USA, 2016; pp. 31–41. [CrossRef]
2. Hardalov, M.; Arora, A.; Nakov, P.; Augenstein, I. A Survey on Stance Detection for Mis- and Disinformation Identification. *arXiv* **2021**. [CrossRef]
3. Lauscher, A.; Wachsmuth, H.; Gurevych, I.; Glavaš, G. Scientia potentia Est—On the role of knowledge in computational argumentation. *Trans. Assoc. Comput. Linguist.* **2022**, *10*, 1392–1422. [CrossRef]
4. Pamungkas, E.W.; Basile, V.; Patti, V. Stance classification for rumour analysis in twitter: Exploiting affective information and conversation structure. *arXiv* **2019**, arXiv:1901.01911. [CrossRef]
5. Ebrahimi, J.; Dou, D.; Lowd, D. A joint sentiment-target-stance model for stance classification in tweets. In Proceedings of the COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, Osaka, Japan, 11–16 December 2016; pp. 2656–2665.
6. Sobhani, P.; Mohammad, S.; Kiritchenko, S. Detecting stance in tweets and analyzing its interaction with sentiment. In Proceedings of the Fifth Joint Conference on Lexical and Computational Semantics, Berlin, Germany, 11–12 August 2016; pp. 159–169.
7. Bharathi, A.; Zubiaga, A. Zero-shot Cross-lingual Stance Detection via Adversarial Language Adaptation. *arXiv* **2024**, arXiv:2404.14339.
8. Vamvas, J.; Sennrich, R. X-stance: A multilingual multi-target dataset for stance detection. *arXiv* **2020**, arXiv:2003.08385.
9. Behzadidoost, R.; Mahan, F.; Izadkhah, H. Granular computing-based deep learning for text classification. *Inf. Sci.* **2024**, *652*, 119746. [CrossRef]
10. Guarasci, R.; De Pietro, G.; Esposito, M. Quantum natural language processing: Challenges and opportunities. *Appl. Sci.* **2022**, *12*, 5651. [CrossRef]
11. Abbaszade, M.; Salari, V.; Mousavi, S.S.; Zomorodi, M.; Zhou, X. Application of quantum natural language processing for language translation. *IEEE Access* **2021**, *9*, 130434–130448. [CrossRef]
12. Pandey, S.; Basisth, N.J.; Sachan, T.; Kumari, N.; Pakray, P. Quantum machine learning for natural language processing application. *Phys. A Stat. Mech. Its Appl.* **2023**, *627*, 129123. [CrossRef]
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*, 5973–5985. [CrossRef]
14. Tomal, S.; Shafin, A.A.; Bhattacharjee, D.; Amin, M.; Shahir, R.S. Quantum-Enhanced Attention Mechanism in NLP: A Hybrid Classical-Quantum Approach. *arXiv* **2025**, arXiv:2501.15630.
15. Ramya, P.; Anitha, R.; Rajalakshmi, J.; Dineshkumar, R. Integrating Quantum Computing and NLP for Advanced Cyber Threat Detection. *J. Cybersecur. Inf. Manag.* **2024**, *14*, 186–197. [CrossRef]
16. Cavar, D.; Zhang, C. Semantic Similarities Using Classical Embeddings in Quantum NLP. In Proceedings of the 2024 IEEE International Conference on Quantum Computing and Engineering (QCE), Montreal, QC, Canada, 15–20 September 2024; Volume 2, pp. 450–451.
17. Li, Q.; Wang, B.; Zhu, Y.; Lioma, C.; Liu, Q. Adapting pre-trained language models for quantum natural language processing. *arXiv* **2023**, arXiv:2302.13812. [CrossRef]
18. Yao, B.; Tiwari, P.; Li, Q. Self-supervised pre-trained neural network for quantum natural language processing. *Neural Netw.* **2025**, *184*, 107004. [CrossRef]
19. Pei, J.J.; Gong, L.H.; Qin, L.G.; Zhou, N.R. One-to-many image generation model based on parameterized quantum circuits. *Digit. Signal Process.* **2025**, *165*, 105340. [CrossRef]
20. Gong, L.H.; Chen, Y.Q.; Zhou, S.; Zeng, Q.W. Dual discriminators quantum generation adversarial network based on quantum convolutional neural network. *Adv. Quantum Technol.* **2025**, *8*, e2500224. [CrossRef]
21. Lou, X.; Zan, H.; Wang, Z.; Shi, J.; Zhang, S. Privacy-Preserving Bidirectional Data Transmission of Smart Grid via Semi-Quantum Computation: On Mutual Identity and Message Authentication. *IEEE Trans. Dependable Secur. Comput.* **2025**, *22*, 5537–5548. [CrossRef]
22. Vasilakes, J.; Scarton, C.; Zhao, Z. Exploring Vision Language Models for Multimodal and Multilingual Stance Detection. *arXiv* **2025**, arXiv:2501.17654. [CrossRef]
23. Xu, J.; Liu, B.; Xiao, Y. A Multitask Causality-Inspired Feature Enhancement Method for Stance Detection. *IEEE Trans. Audio Speech Lang. Process.* **2025**, *33*, 773–784. [CrossRef]
24. Blanco, G.; Yáñez Martínez, R.; Lourenço, A. Leveraging deep learning to detect stance in Spanish tweets on COVID-19 vaccination. *JAMIA Open* **2025**, *8*, ooaf007. [CrossRef]
25. Fan, Q.; Lu, J.; Sun, Y.; Pi, Q.; Shang, S. Enhancing zero-shot stance detection via multi-task fine-tuning with debate data and knowledge augmentation. *Complex Intell. Syst.* **2025**, *11*, 151. [CrossRef]

26. Pu, Q.; Huang, F.; Li, F.; Wei, J.; Jiang, S. Integrating Emotional Features for Stance Detection Aimed at Social Network Security: A Multi-Task Learning Approach. *Electronics* **2025**, *14*, 186. [CrossRef]
27. Kumar, S. Negative Stances Detection from Multilingual Data Streams in Low-Resource Languages on Social Media Using BERT and CNN-Based Transfer Learning Model. *ACM Trans. Asian Low-Resour. Lang. Inf. Process.* **2024**, *23*, 1–18. [CrossRef]
28. Lan, X.; Gao, C.; Jin, D.; Li, Y. Stance detection with collaborative role-infused llm-based agents. In Proceedings of the International AAAI Conference on Web and Social Media, Buffalo, NY, USA, 3–6 June 2024; Volume 18, pp. 891–903.
29. Mets, M.; Karjus, A.; Ibrus, I.; Schich, M. Automated stance detection in complex topics and small languages: The challenging case of immigration in polarizing news media. *PLoS ONE* **2024**, *19*, e0302380. [CrossRef]
30. Li, A.; Zhao, J.; Liang, B.; Gui, L.; Wang, H.; Zeng, X.; Wong, K.F.; Xu, R. Mitigating Biases of Large Language Models in Stance Detection with Calibration. *arXiv* **2024**, arXiv:2402.14296.
31. Ding, D.; Chen, R.; Jing, L.; Zhang, B.; Huang, X.; Dong, L.; Zhao, X.; Song, G. Cross-Target Stance Detection by Exploiting Target Analytical Perspectives. In Proceedings of the ICASSP 2024—2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Seoul, Republic of Korea, 14–19 April 2024; pp. 10651–10655.
32. Wang, X.; Wang, Y.; Cheng, S.; Li, P.; Liu, Y. DEEM: Dynamic Experienced Expert Modeling for Stance Detection. *arXiv* **2024**, arXiv:2402.15264. [CrossRef]
33. Ding, D.; Fu, X.; Peng, X.; Fan, X.; Huang, H.; Zhang, B. Leveraging Chain-of-Thought to Enhance Stance Detection with Prompt-Tuning. *Mathematics* **2024**, *12*, 568. [CrossRef]
34. Zhao, X.; Zou, J.; Miao, J.; Tian, L.; Gao, L.; Zhou, B.; Pang, S. Zero-shot stance detection based on multi-perspective transferable feature fusion. *Inf. Fusion* **2024**, *108*, 102386. [CrossRef]
35. Mahmoudi, G.; Behkamkia, B.; Eetemadi, S. Zero-Shot Stance Detection using Contextual Data Generation with LLMs. *arXiv* **2024**, arXiv:2405.11637. [CrossRef]
36. Hashemian, M.; Rahimi, S.K.; Dastjerdi, A.B. An influences-adapted two-phase approach to stance detection in the diachronic perspective. *Expert Syst. Appl.* **2023**, *231*, 120773. [CrossRef]
37. Choi, K.; Ko, Y. Meta-learning with topic-agnostic representations for zero-shot stance detection. *Pattern Recognit. Lett.* **2023**, *171*, 15–20. [CrossRef]
38. Jiang, Y.; Gao, J.; Shen, H.; Cheng, X. Zero-shot stance detection via multi-perspective contrastive learning with unlabeled data. *Inf. Process. Manag.* **2023**, *60*, 103361. [CrossRef]
39. Gómez-Suta, M.; Echeverry-Correa, J.; Soto-Mejía, J.A. Stance detection in tweets: A topic modeling approach supporting explainability. *Expert Syst. Appl.* **2023**, *214*, 119046. [CrossRef]
40. Al-Ghadir, A.I.; Azmi, A.M.; Hussain, A. A novel approach to stance detection in social media tweets by fusing ranked lists and sentiments. *Inf. Fusion* **2021**, *67*, 29–40. [CrossRef]
41. Lai, M.; Cignarella, A.T.; Farías, D.I.H.; Bosco, C.; Patti, V.; Rosso, P. Multilingual stance detection in social media political debates. *Comput. Speech Lang.* **2020**, *63*, 101075. [CrossRef]
42. Kawintiranon, K.; Singh, L. Knowledge enhanced masked language model for stance detection. In Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Online, 6–11 June 2021; pp. 4725–4735.
43. Chen, W.F.; Ku, L.W. Utcnn: A deep learning model of stance classificationon on social media text. *arXiv* **2016**, arXiv:1611.03599. [CrossRef]
44. Xuan, K.; Xia, R. Rumor stance classification via machine learning with text, user and propagation features. In Proceedings of the 2019 International Conference on Data Mining Workshops (ICDMW), Beijing, China, 8–11 November 2019; pp. 560–566.
45. Dey, K.; Shrivastava, R.; Kaushik, S. Twitter stance detection—A subjectivity and sentiment polarity inspired two-phase approach. In Proceedings of the 2017 IEEE International Conference on Data Mining Workshops (ICDMW), New Orleans, LA, USA, 18–21 November 2017; pp. 365–372.
46. Zhu, L.; He, Y.; Zhou, D. Neural opinion dynamics model for the prediction of user-level stance dynamics. *Inf. Process. Manag.* **2020**, *57*, 102031. [CrossRef]
47. Sun, Q.; Wang, Z.; Zhu, Q.; Zhou, G. Stance detection with hierarchical attention network. In Proceedings of the 27th International Conference on Computational Linguistics, Santa Fe, NM, USA, 20–26 August 2018; pp. 2399–2409.
48. Wei, P.; Mao, W.; Zeng, D. A target-guided neural memory model for stance detection in twitter. In *Proceedings of the 2018 International Joint Conference on Neural Networks (IJCNN)*; IEEE: New York, NY, USA, 2018; pp. 1–8.
49. Ahmed, M.; Chy, A.N.; Chowdhury, N.K. Incorporating hand-crafted features in a neural network model for stance detection on microblog. In Proceedings of the 6th International Conference on Communication and Information Processing, Tokyo, Japan, 27–29 November 2020; pp. 57–64.
50. Chen, P.; Ye, K.; Cui, X. Integrating n-gram features into pre-trained model: A novel ensemble model for multi-target stance detection. In *Proceedings of the International Conference on Artificial Neural Networks*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 269–279.

51. Du, J.; Xu, R.; He, Y.; Gui, L. Stance classification with target-specific neural attention networks. In Proceedings of the 26th International Joint Conference on Artificial Intelligence, IJCAI 2017, Melbourne, Australia, 19–25 August 2017; pp. 3988–3994.
52. Vaswani, A. Attention is all you need. *arXiv* **2017**, arXiv:1706.03762.
53. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv* **2018**, arXiv:1810.04805.
54. Liu, Z.; Lin, W.; Shi, Y.; Zhao, J. A robustly optimized BERT pre-training approach with post-training. In *Chinese Computational Linguistics*; Springer: Cham, Switzerland, 2021; pp. 471–484.
55. Alghamdi, J.; Lin, Y.; Luo, S. Towards COVID-19 fake news detection using transformer-based models. *Knowl.-Based Syst.* **2023**, *274*, 110642. [CrossRef]
56. Le, N.Q.K.; Ho, Q.T.; Nguyen, T.T.D.; Ou, Y.Y. A transformer architecture based on BERT and 2D convolutional neural network to identify DNA enhancers from sequence information. *Brief. Bioinform.* **2021**, *22*, bbab005. [CrossRef]
57. Tan, K.L.; Lee, C.P.; Anbananthen, K.S.M.; Lim, K.M. RoBERTa-LSTM: A hybrid model for sentiment analysis with transformer and recurrent neural network. *IEEE Access* **2022**, *10*, 21517–21525.
58. Liang, D.; Wu, Y.; Duan, W. Multiple granularity user intention fairness recognition of intelligent government Q & A system via three-way decision. *Inf. Sci.* **2023**, *631*, 305–326. [CrossRef]
59. Yang, X.; Li, Y.; Meng, D.; Yang, Y.; Liu, D.; Li, T. Three-way multi-granularity learning towards open topic classification. *Inf. Sci.* **2022**, *585*, 41–57. [CrossRef]
60. Behzadidoost, R.; Izadkhah, H.; Mahan, F. A combined multi-margin contrastive learning with granulated data for warrant identification in computational argumentation. *Inf. Sci.* **2024**, *699*, 121756. [CrossRef]
61. Shi, S.; Wang, Z.; Cui, G.; Wang, S.; Shang, R.; Li, W.; Wei, Z.; Gu, Y. Quantum-inspired complex convolutional neural networks. *Appl. Intell.* **2022**, *52*, 17912–17921. [CrossRef]
62. Lai, W.; Shi, J.; Chang, Y. Quantum-Inspired Fully Complex-Valued Neutral Network for Sentiment Analysis. *Axioms* **2023**, *12*, 308. [CrossRef]
63. Pires, T.; Schlinger, E.; Garrette, D. How multilingual is multilingual BERT? *arXiv* **2019**, arXiv:1906.01502. [CrossRef]
64. Conneau, A.; Khandelwal, K.; Goyal, N.; Chaudhary, V.; Wenzek, G.; Guzmán, F.; Grave, E.; Ott, M.; Zettlemoyer, L.; Stoyanov, V. Unsupervised cross-lingual representation learning at scale. *arXiv* **2019**, arXiv:1911.02116.
65. Behzadidoost, R.; Izadkhah, H. An ensembled text-based multi-stacked-cnn-bilstm for rumor detection on social media. In Proceedings of the 2022 9th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS), Bam, Iran, 2–4 March 2022; pp. 1–6.
66. Kalchbrenner, N.; Grefenstette, E.; Blunsom, P. A convolutional neural network for modelling sentences. *arXiv* **2014**, arXiv:1404.2188. [CrossRef]
67. Yang, Z.; Yang, D.; Dyer, C.; He, X.; Smola, A.; Hovy, E. Hierarchical attention networks for document classification. In Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, San Diego, CA, USA, 12–17 June 2016; pp. 1480–1489.
68. Wang, L.; Yang, N.; Huang, X.; Yang, L.; Majumder, R.; Wei, F. Multilingual e5 text embeddings: A technical report. *arXiv* **2024**, arXiv:2402.05672. [CrossRef]
69. Liu, Y.; Gu, J.; Goyal, N.; Li, X.; Edunov, S.; Ghazvininejad, M.; Lewis, M.; Zettlemoyer, L. Multilingual denoising pre-training for neural machine translation. *Trans. Assoc. Comput. Linguist.* **2020**, *8*, 726–742. [CrossRef]
70. Feng, F.; Yang, Y.; Cer, D.; Arivazhagan, N.; Wang, W. Language-agnostic BERT Sentence Embedding. *arXiv* **2020**, arXiv:2007.01852.

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.