

# TGCN-HCA: Temporal Graph Convolution Networks with Hierarchical Co-Attention Mechanisms For Fake News Detection

Muhammad Abdullah<sup>1†</sup>, Zan Hongying<sup>1\*†</sup>, Arifa Javed<sup>1†</sup>,  
Junaid Abdul Wahid<sup>1†</sup>, Kanwal Ahmed<sup>2†</sup>, Hassan Eshkiki<sup>3†</sup>,  
Fabio Caraffini<sup>3\*†</sup>

<sup>1</sup>School of Computing and Artificial Intelligence, Zhengzhou University,  
Zhengzhou, 450001, Henan, China.

<sup>2</sup>School of Software, Henan University, Kaifeng, 470001, Henan, China.

<sup>3</sup>Department of Computer Science, Swansea University, Swansea, SA1  
8EN, Wales, UK.

\*Corresponding author(s). E-mail(s): [iehyzan@zzu.edu.cn](mailto:iehyzan@zzu.edu.cn);  
[fabio.caraffini@swansea.ac.uk](mailto:fabio.caraffini@swansea.ac.uk);

Contributing authors: [abdullah@gs.zzu.edu.cn](mailto:abdullah@gs.zzu.edu.cn);  
[arifa.javed@gs.zzu.edu.cn](mailto:arifa.javed@gs.zzu.edu.cn); [junaid\\_wahid@zzu.edu.cn](mailto:junaid_wahid@zzu.edu.cn);  
[Kanwal@henu.edu.cn](mailto:Kanwal@henu.edu.cn); [h.g.eshkiki@swansea.ac.uk](mailto:h.g.eshkiki@swansea.ac.uk);

†These authors contributed equally to this work.

## Abstract

Social media has become a primary source of information, but its rapid growth has facilitated the spread of misinformation. Existing methods often rely on static graph neural networks (GNNs) to analyze news propagation patterns, overlooking temporal dynamics and semantic nuances. To address this, we propose a Temporal Graph Convolution Network with a Hierarchical Co-Attention (TGCN-HCA) model for effective fake news detection. TGCN-HCA leverages a dynamic graph representation to capture the evolving nature of news propagation. By incorporating temporal information, we can accurately model the sequential spread of news and identify patterns associated with fake news. To enhance semantic understanding, we employ GPT-2 to extract deep contextual features from news content. These features, combined with structural information from the graph, provide a rich representation of news propagation. A hierarchical co-attention mechanism further refines feature representation by jointly considering multiple perspectives:

the importance of nodes within the graph and the relevance of semantic features to the task of fake news detection. By integrating these components, TGCN-HCA effectively captures the complex interplay between temporal dynamics, structural information, and semantic content. Extensive experiments on Twitter15, Twitter16, and Weibo datasets demonstrate the superior performance of TGCN-HCA over state-of-the-art methods. Our model achieves significant accuracy gains of 5.37%, 4.08%, and 1.03% on these datasets, respectively.

**Keywords:** Graph Neural Networks, Deep learning, GPT-2, Fake news, Semantic, NLP

## 1 Introduction

Social media has transformed how information is shared, allowing users to post, forward, and comment on real-time information across various platforms [1]. However, this rapid and widespread adoption has also enabled the swift dissemination of false information, which poses significant risks to public opinion and social stability [2]. The problem is further exacerbated by the ease of creating and sharing fake news using advanced language models. The detrimental effects of fake news are evident in various real-world scenarios. For instance, during the COVID-19 pandemic, misinformation spread rapidly, instilling widespread fear and promoting dangerous health advice, ultimately harming public health [3]. Similarly, false stories influenced voter perceptions during the 2016 U.S. presidential election and may have potentially impacted the election outcome. Moreover, conspiracy theories linking 5G technology to COVID-19 sparked arson attacks in Europe, causing property damage and endangering lives [4]. Additionally, myths surrounding vaccinations have led to decreased vaccination rates, resulting in outbreaks of preventable diseases like measles and whooping cough, posing severe public health risks [5]. These examples underscore the profound impact of fake news on global public safety, health, and political stability.

To counter the spread of misinformation, expert-based fact-checking and machine-learning approaches have been widely employed [6]. Traditional machine learning methods rely on manual feature engineering, incorporating user characteristics, textual content, and propagation patterns to train supervised classifiers [7]. However, these methods are often constrained by labour-intensive feature extraction processes and their inability to capture rumour propagation's complex and evolving nature. Fake news spans diverse topics, styles, and platforms and involves multiple factors, including the content itself, its creators and spreaders, and the broader social context in which it is shared. While deep learning (DL) algorithms, such as convolutional neural networks (CNNs), are effective at identifying local patterns and linguistic features, they struggle to capture broader context and long-range dependencies [8]. Similarly, recurrent neural networks (RNNs) and long short-term memory networks (LSTMs) capture sequential dependencies and contextual information but face computational challenges when applied to large-scale data [9]. Transformers like BERT and GPT, with their attention mechanisms, enhance contextual understanding [10]. Still, their

resource-intensive nature and reliance on large datasets make them difficult to interpret and apply effectively [11]. Furthermore, rather than addressing them holistically, these methods typically focus on one aspect of fake news detection, such as content, propagation, or social context. However, fake news often forms a distinctive propagation structure over time, and existing models struggle to capture the complex relationships inherent in social networks.

Graph convolutional networks (GCNs) offer a promising solution by applying deep learning algorithms to graph-structured data [12]. A GCN-based approach can effectively capture the structural patterns of fake news propagation by modeling the relationships and interactions between nodes, significantly improving detection accuracy [13]. For example, the Graph-aware Co-Attention Networks (GCAN) model has been proposed for explainable fake news detection, focusing on short-text tweets and sequences of retweets without user comments [2]. While it offers explainability, it has data requirements and scalability limitations. The Graph Convolution Network and BERT (GBCA) model combines graph topologies with BERT to train GCNs and enhance fake news detection accuracy [14]. Meanwhile, Hierarchical Aggregation using GNN (HAGNN) addresses issues like data sparsity and user behaviour in rumour detection, although it faces challenges with scalability on large platforms [15]. Another approach, Graph Attention Mechanisms (GANM), manages relationship dependencies while preserving crucial information over long sequences. However, it is constrained by data quality and high computational demands [12]. Finally, the Edge Bayesian GCN model reduces uncertainty in the propagation structure of fake news by reinforcing graph edges, further enhancing its reliability and effectiveness [16].

Overall, existing models encounter significant challenges. They tend to oversimplify the dynamics of news propagation and often neglect temporal dependencies. Semantic features within the propagation structure are overlooked, particularly emotionally salient keywords that can indicate fake news. Therefore, inspired by GBCA [14] and GANM [12] models, a Temporal Graph Convolution Networks with Hierarchical Co-Attention mechanisms (TGCN-HCA) model is introduced. TGCN aims to improve fake news detection by effectively capturing both the temporal dynamics and the semantic features of news content. This model adopts the strengths of temporal graph convolutional and GPT-2, integrated through a hierarchical co-attention mechanism to enhance feature representation. Temporal dependency is the relationships between events or information that change over time. It helps us to understand how things evolve and interact across different periods. Unusual engagement patterns, such as sudden spikes in likes and shares, can also indicate fake news. Temporal graph topologies represent the evolving structure of news dissemination, where each snapshot captures the propagation structure at a specific time step. Edges are defined based on temporal relationships and user interactions (reply or forwarding relationships). The Hierarchical Co-attention mechanism [17] integrates temporal dependencies and semantic features. The model can selectively concentrate on relevant graph regions and news aspects. This approach enhances feature representation and detection accuracy. It effectively tackles the complex relationships between users and news articles. The primary contributions are as follows.

- The proposed model fuses Temporal Graph Convolutional Networks (TGCNs) and GPT-2 model to detect fake news. TGCNs capture temporal dependencies in the propagation structure and build a graph. GPT-2 extracts semantic features from the original fake news text.
- The primary strength of the hierarchical co-attention mechanism lies in its ability to integrate temporal and semantic information. The temporal perspective focuses on the sequence and timing of data, which is crucial for understanding trends and patterns over time. On the other hand, the semantic perspective involves understanding the meaning and context of the data, which is essential for accurate interpretation and decision-making.
- The TGCN-HCA model constructs a graph based on reply and forwarding interactions, effectively mapping the spread of fake news. By integrating features from adjacent nodes through these connections, the model enhances its ability to identify fake news accurately.

The TGCN-HCA model addresses the limitations of existing GNN-based models by incorporating temporal dynamics and semantic features through GPT-2 and a hierarchical co-attention mechanism. This comprehensive approach enhances the model’s ability to detect fake news accurately, effectively handling complex relationships and scaling to large-scale social media platforms. The remainder of this paper is organized as follows: Section 2 presents a literature review on fake news. Section 3 describes the materials and outlines the proposed method. Section 4 discusses the experiments. Section 5 presents the results and discussions. Section 6 provides the concluding remarks.

## 2 Related Work

Fake news detection aims to identify false information using various approaches and abundant data sources, such as text content, user comments, and posting patterns on social media [18]. Recent methods for fake news detection can be grouped into three major categories: classical models, deep learning and transformer-based models, and graph neural network (GNN) models. The impact of fake news can be severe, including harm to individuals and businesses [19], risks to public health [3], and political instability [20]. Social media exacerbates this issue by facilitating the rapid dissemination of false information [2], making the development of effective detection approaches increasingly critical. To address this problem, techniques such as fact-checking [6] and machine learning algorithms [7] have been proposed.

### 2.1 Tradational Approaches

Classical machine learning models, such as Support Vector Machines (SVM), Naive Bayes, and Decision Trees, have been widely employed in fake news detection [21]. For instance, machine learning techniques have been applied to detect misinformation on platforms like Twitter by analyzing the credibility of text and user-specific features, such as posting patterns and account characteristics [22]. Similarly, logistic regression has been used to identify fake news by leveraging features like user behavior and

content-based attributes, such as linguistic patterns [23]. These supervised learning methods have demonstrated effectiveness in distinguishing between credible and non-credible information, particularly when clear feature-based distinctions are available.

Despite their successes, classical machine learning models face significant limitations in fake news detection. These approaches heavily rely on manually engineered features, such as word frequency, syntax, and semantics, to train classifiers [7]. This dependence restricts their ability to capture the complexity and evolving nature of fake news, which often involves subtle manipulations or context-dependent misinformation. Furthermore, these models struggle with challenges such as handling imbalanced datasets, where credible content vastly outnumbers fake content, detecting satire or biased narratives, and comprehending contextual nuances that require deeper semantic understanding.

## 2.2 Deep Learning Approaches

Deep learning models, such as Long Short-Term Memory (LSTM) networks and Word2Vec embeddings, have been introduced to address these limitations. Word2Vec generates context-independent feature vectors, which LSTM layers process to extract key properties [24]. Bi-directional LSTM-RNN models outperform unidirectional models in fake news detection, with attention mechanisms further enhancing their performance by capturing long-range dependencies [25]. Hybrid models combining Convolutional Neural Networks (CNNs) with RNNs have also improved detection accuracy by leveraging spatial and temporal features [26]. In line with recent efforts to achieve adaptive deep learning models in dynamic and data-rich environments [27], such as the design of AI agents for tackling security issues [28], our TGCN-HCA model addresses the ‘spatio-temporal’ challenge by integrating temporal graph convolutional networks (TGCNs) with GPT-2 embeddings, thus enabling real-time adjustments to fake news propagation patterns while capturing emotionally salient semantic features. Our approach continuously refines the model parameters to respond to emerging misinformation trends, while the integration of event detection in social networks [29] complements our work by focusing on temporal and contextual evolution, which TGCN-HCA effectively adapts to by refining both semantic and graph features during training. Moreover, transformer-based models, which include multi-head attention mechanisms, have demonstrated exceptional effectiveness in capturing long-distance tweet dependencies for fake news detection [30][31]. For example, BERT-based models, such as FakeBERT, have advanced the field by incorporating task-specific layers for enhanced semantic understanding [11]. However, despite their success, deep learning models often struggle with multi-source, multi-modal data and fail to capture the complex, non-linear relationships inherent in fake news propagation.

## 2.3 Graphs Based Approaches

Graph Neural Networks (GNNs) have emerged as an effective tool for fake news detection by modeling relationships and interactions in graph-structured representations [15]. GNNs excel at capturing the structural dynamics of fake news propagation and its intricate patterns [32]. Advanced variants, such as Graph Attention Networks (GAN)

and Graph Convolutional Networks (GCN), have further enhanced detection accuracy by improving the modeling of user interactions [33].

For example, Edge Enhanced Bayesian GCN (EBGCN) addresses uncertain relationships between graph edges by incorporating probabilistic reasoning, which improves the model’s robustness, making it more effective at handling noisy or incomplete data in fake news propagation [16]. Similarly, GACN (Graph Attention Convolutional Network) focuses on dynamic user interactions and leverages deep-level node attributes, capturing evolving patterns in misinformation spread and improving the model’s ability to detect emerging fake news trends [34]. Claim-Guided Hierarchical Graph Attention Networks (CHGA) represent conversation threads as undirected interaction graphs, enabling the model to capture complex, multi-layered relationships within news discussions and better understand how misinformation spreads in social networks [35].

Graph Convolution Network and BERT combined with Co-Attention (GBCA) integrates GCNs with BERT to enhance fake news detection by leveraging both the structural dynamics of news propagation and semantic understanding from text data [14]. However, it struggles with temporal dynamics, as it relies on static graph structures and pre-trained BERT embeddings without adapting them in real-time, limiting its ability to capture the evolving nature of fake news propagation and detect context-dependent misinformation over time.

Topic-Event Hierarchical Graph Convolutional Networks (TEH-GCN) detects fake news by aggregating features across predefined topic-event intervals, capturing the relationships between events and topics [17]. However, it faces challenges due to its rigid temporal segmentation, which limits its ability to capture the fluid nature of news dissemination and adapt to real-time changes. Moreover, the focus on predefined event-level interactions restricts its ability to model complex, non-linear relationships across topics, diminishing its flexibility in detecting evolving fake news patterns.

Some approaches also fuse content and propagation structures into heterogeneous graphs to enhance feature extraction. For instance, deep residual GCNs [36] and Graph Global Attention Networks with Memory (GANM) [12] integrate diverse graph features to improve the detection process.

Despite these advancements, GNN-based models still face significant limitations. Most of them primarily focus on the propagation structure of fake news, often neglecting key semantic features, such as emotionally charged keywords that can indicate misinformation. These gaps highlight the need for more comprehensive models that integrate both semantic analysis and temporal dynamics to effectively capture the full complexity of fake news spread. Table 1 provides a comparative summary of the contributions and limitations of the existing fake news detection models discussed in this section.

## 2.4 Current limitations and motivations for this study

Despite the progress made in recent years for fake news detection, existing methods struggle to capture temporal dynamics and semantic content simultaneously. Popular models like EBGCN and GCAN focus on static graph structures, overlooking the evolving fake news dynamics and their semantic features, thus leading to missed

**Table 1** Comparative Analysis of Machine Learning Models in Fake News Detection: Highlighting Datasets, Key Contributions, and Limitations

Model	Datasets	Strengths	Limitations
GCNs-MT [12]	FakenewsNet, Weibo	Structural formation of rumor propagation	Oversimplification of news propagation dynamics
GACN [34]	Twitter15, Twitter16	Captures deep-level node attributes	Poor user interaction and emotionally content
CHGA [35]	Twitter15-16, PHEME	Undirected interaction graphs news	Limited adaption to dynamic online environments
HGNN [15]	Weibo	Fuses content and propagation into a graph	Neglect time dependencies and high-dimensional data
EBGCN [16]	Twitter15, PHEME	Models uncertain relationships in edges	Overemphasises propagation; limited explainability
GCAN [2]	Twitter15, Twitter16	Capturing correlation in tweets and retweet	Less attention to temporal features and news semantics
Word2Vec-LSTM [24]	ISOT, T1-CNN	Extract salient features to decide news veracity	Static Word Embeddings lead to less Vocabulary Coverage
FakeBERT [22]	DFE	Preserve integrity and promotes decision-making	Tested on limited dataset; low contextual features
CNNs-RNNs [26]	ISOT, FakeNewsNet	Leverages both spatial and temporal features	Overlooks semantics and contextual features
GBCA [14]	Twitter15-16, PHEME	Enhances GNN to handle unstructured data	Less attention to temporal dependency and relationships

dynamic patterns and emotionally charged keywords crucial for detection. While GBCA combines static GCN for propagation modelling with fixed BERT embeddings via a single coattention layer, it still lacks temporal sensitivity and adaptive refinement. Similarly, TEH-GCN relies on predefined topic-event hierarchies and rigid temporal segmentation, aggregating features without real-time semantic alignment or dynamic node updates.

In this light, we propose the TGCN-HCA model to address the limitations of these previous models by combining temporal graph convolution networks (TGCNs) with hierarchical co-attention (HCA) mechanisms. The TGCN models temporal dependencies and structural patterns in fake news propagation, while the HCA mechanism jointly learns dependencies between temporal graph features and semantic content from news, allowing the model to focus on the most relevant graph regions and news content. By integrating these two components, the TGCN-HCA model dynamically updates node representations based on real-time shifts in propagation patterns, such as engagement bursts. The model iteratively refines GPT-2 embeddings with evolving graph interactions, enabling real-time semantic-temporal co-evolution, unlike the static fusion in GBCA and pre-structured hierarchies in TEH-GCN. The result is better detection accuracy, greater generalizability, and earlier misinformation identification across platforms.

### 3 Materials and Methods

This section describes the proposed model for fake news detection, as shown in Fig 1. The Proposed approach integrates state-of-the-art techniques in graph neural networks, large language models, and data processing to effectively capture the complexities of fake news dissemination and content.

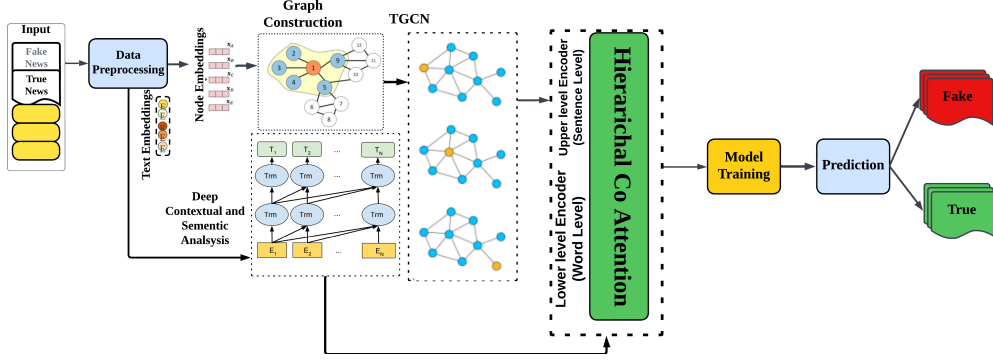


Fig. 1 Research Methodology Workflow Diagram

#### 3.1 Preliminaries Formulation

The dataset is represented as  $G = \{G_1, G_2, \dots, G_N\}$ , where  $G_i$  corresponds to the  $i$ -th propagation network of news, with  $N$  denoting the total number of news items in the dataset. Specifically, each  $G_i = \{X_i, E_i\}$  consists of the node feature set  $X_i = \{r_i, x_{i1}, x_{i2}, \dots, x_{in_i}\}$  and the edge set  $E_i = \{e_{s \rightarrow d} \mid s, d = 1, 2, \dots, n_i\}$ . In this context,  $r_i$  signifies the embedding of the root node, while  $x_{ij}$  represents the embedding of the  $j$ -th leaf node. The interaction between the  $s$ -th and  $d$ -th nodes is captured by  $e_{s \rightarrow d}$ , indicating that the  $d$ -th node responds to the  $s$ -th node. The parameter  $n_i$  denotes the total number of posts in  $G_i$ . The set of edges  $E_i$  is represented using an adjacency matrix  $A_i \in \{0, 1\}^{n_i \times n_i}$ , where  $a_{sd} = 1$  if  $e_{s \rightarrow d} \in E_i$ , and 0 otherwise. Each distinct graph  $G_i$  is associated with a ground-truth label  $y_i \in \{0, 1\}$ , where  $y_i = 1$  indicates true news, and  $y_i = 0$  indicates false news. To address the challenges as mentioned earlier, the formulation for constructing the propagation network is provided as:

$$G_i = (N_i, E_i) \quad \text{with} \quad N_i = \{c_{i0}, c_{i1}, \dots, c_{i(n_i-1)}\}, \quad E_i = \{e_{s \rightarrow d} \mid s, d = 0, \dots, n_i - 1\} \quad (1)$$

Defining the adjacency matrix:

$$A_i(p, q) = \begin{cases} 1 & \text{if } e_{p \rightarrow q} \in E_i \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

Text feature extraction:

$$X_i = [x_{i0}, x_{i1}, \dots, x_{i(n_i-1)}]^T \quad (3)$$

These preliminary steps set the stage for proposing a model to address the identified challenges in fake news detection, forming the basis for constructing time-evolving propagation networks, to be detailed in the subsequent section.

### 3.2 Graph Construction

Each news item is represented as a propagation network  $G_i = \{X_i, E_i\}$ , where  $X_i$  represents node characteristics, including node content and relevant temporal context such as time stamps of posts and interactions. The set  $E_i$  represents the edges that encode the relationships between nodes, including interactions such as retweets, replies, shares, and other user interactions.

The adjacency matrix  $A_i \in \mathbb{R}^{n_i \times n_i}$  encodes these relationships, where an edge  $e_{s \rightarrow d}$  exists if the  $d$ -th node responds to the  $s$ -th node (such as a retweet or reply), and the entry  $a_{sd}$  is assigned a weight reflecting the recency of the interaction. Specifically,  $a_{sd} = \exp(-\alpha(t_{\text{current}} - t_{sd}))$  if  $e_{s \rightarrow d} \in E_i$ , and  $a_{sd} = 0$  otherwise, with  $t_{sd}$  denoting the interaction time and  $\alpha$  as a decay parameter. This weighted adjacency matrix captures the temporal relationships between nodes, emphasising the importance of more recent interactions in the propagation chain. The ground truth labels  $y_i$  are encoded in binary format, with  $y_i = 1$  indicating true news and  $y_i = 0$  indicating false news. These labels serve as the target outputs during model training, allowing the model to learn to classify news based on its propagation dynamics.

Data is pre-processed for proper formatting in model training and evaluation, forming the basis for temporal graph convolution updates in the proposed model.

The Algorithm 1 outlines the graph construction process:

### 3.3 Graph Convolutional Networks

Graph Convolutional Networks (GCNs) are neural networks designed to process data structured as graphs, where nodes represent entities and edges capture their relationships [13]. GCNs excel in graph-structured data tasks by propagating and aggregating information across nodes to learn rich, contextual representations. A notable application of GCNs in fake news detection is the GBCA model; GCNs are combined with BERT and Co-Attention to enhance fake news detection by integrating graph structures and semantic features [14]. Similarly, GCNs are widely applied in other domains, such as recommendation systems, where GSIRec leverages graph side information to improve recommendation quality [37] and multimodal sequence analysis, where graph convolution and pooling techniques are used to capture complex dependencies [38]. GCNs also excel in tasks like traffic flow forecasting, where attention mechanisms enhance predictive performance [39]. In the context of our work, GCNs capture the propagation structure of news items by modelling interactions between users and posts. As fake news spreads, the relationships between the source and users interacting with it form a graph. GCNs can leverage these graph structures to enhance detection accuracy.

---

**Algorithm 1** Graph Construction Algorithm

---

```
1: Input: Raw dataset  $D$ 
2: Initialize: Load GPT-2 model and tokenizer
   from transformers import GPT2Tokenizer, GPT2Model
   tokenizer = GPT2Tokenizer.from_pretrained('gpt2')
   model = GPT2Model.from_pretrained('gpt2')
3: for each news item  $d_i \in D$  do
4:   preprocessing  $\leftarrow$  remove_special_chars(lowercase(remove_stopwords( $d_i$ )))
5:   cleaned_text  $\leftarrow$  GCN_tokenizer(preprocessing)
6:   tokens  $\leftarrow$  tokenizer(cleaned_text, return_tensors='pt')
7:   embeddings  $\leftarrow$  model(tokens).last_hidden_state
8:    $X_i \leftarrow \{(r_i, t_{i0}), (x_{i1}, t_{i1}), \dots, (x_{in_i}, t_{in_i})\}$   $\triangleright$  Node features with timestamps
9:   Initialize edge set  $E_i \leftarrow \emptyset$ 
10:  for each pair of nodes  $(s, d)$  in  $d_i$  do
11:    if  $d$  interacts with  $s$  (e.g., retweet, reply) then
12:      Add edge  $e_{s \rightarrow d}$  to  $E_i$ 
13:    end if
14:  end for
15:  Form adjacency matrix  $A_i$ 
16:  Construct propagation network  $G_i \leftarrow \{X_i, E_i, y_i\}$   $\triangleright y_i =$  ground truth for  $d_i$ 
17:  Add  $G_i$  to the set  $\mathcal{G}$ 
18: end for
19: return Propagation networks  $\mathcal{G}$   $\triangleright \mathcal{G} = \{G_1, G_2, \dots, G_N\}$ 
```

---

### 3.4 The Proposed Model

The proposed Temporal Graph Convolution Networks with Hierarchical Co-Attention Mechanism (TGCN-HCA) features the strengths of Graph Convolution Networks (GCNs) and GPT-2, which is integrated through a hierarchical co-attention mechanism to enhance feature representation and detection accuracy (Fig. 2). The input to the model consists of text content and its temporal propagation structure, which captures the dynamic flow of information over time.

In the TGCN component, a temporal propagation graph  $G_t = (V_t, E_t)$  is constructed at each time step  $t$ , where  $V_t$  and  $E_t$  represent the sets of nodes and edges, respectively, at that specific time. The node update module iteratively updates the node representations  $h_v(t)$  based on interactions with neighbouring nodes in previous time steps, using the update function:

$$h_v(t) = \sigma \left( \sum_{u \in N(v)} W_1 h_u(t-1) + W_2 h_v(t-1) + b \right) \quad (4)$$

where  $N(v)$  represents the neighbors of node  $v$  weighted by  $A_i$ ,  $W_1$  and  $W_2$  are weight matrices,  $b$  is a bias vector, and  $\sigma$  is an activation function, with  $h_v(0) = x_v$  initialized from node embeddings.

The edge weights  $e_{uv}(t)$  between nodes  $u$  and  $v$  are updated at each time step to reflect the temporal characteristics of the interactions. The edge weights are adjusted dynamically, with more recent interactions receiving higher weights to capture the recency and relevance of these interactions in the propagation process. This ensures that the temporal nature of news dissemination is effectively captured in the graph structure.

The proposed architecture adopts the pre-trained GPT-2 model (124M parameters, 12 transformer layers) as the foundation for semantic feature extraction, complemented by dynamically learned temporal and structural features from the TGCN.

### 3.4.1 Contextual Embedding Extraction

First, text preprocessing is performed on raw news content by removing URLs, mentions, hashtags, and special characters. The GPT-2 model is fine-tuned on our dataset to adapt its parameters, including the embeddings, for the specific task of fake news detection.

During fine-tuning, Layers 1-8 are frozen to preserve general language understanding and prevent overfitting, while Layers 9-12 are fine-tuned to learn task-specific semantic patterns such as sensationalism, emotional triggers, and deceptive phrasing, all while leveraging the foundational knowledge of the frozen layers. The model generates embeddings  $X_i = [x_{i0}, x_{i1}, \dots, x_{i(n_i-1)}]^T$ , which capture both deep semantic features and contextual information, representing the semantic essence of the news content. These embeddings are passed through the model’s layers to produce feature vectors  $V \in \mathbb{R}^{d \times N}$ , where  $d$  denotes the feature dimensionality and  $N$  denotes the number of news categories. The contextual embeddings  $z_{\text{GPT-2}}$  are extracted from the last hidden state (Layer 12). These 768-dimensional embeddings encapsulate semantic patterns such as sensationalism, emotional triggers, and deceptive phrasing, providing rich representations of local word semantics and global document context. This enables accurate fake news detection through subsequent graph processing. These embeddings provide a stable semantic foundation for further integration with the TGCN to learn the dynamics of temporal propagation and misinformation spread patterns. Once fine-tuned, the embeddings are frozen during training of the downstream TGCN and HCA components to maintain their task-specific semantic consistency, which reduces computational complexity and prevents overfitting.

During TGCN training, Layers 11-12 remain trainable, which enables the model to adapt high-level semantic features to temporal propagation patterns, allowing it to focus on learning task-specific dynamics while maintaining computational efficiency by reducing the number of trainable parameters. Meanwhile, the TGCN dynamically updates the graph’s node representations  $h_v(t)$  based on interactions with neighboring nodes over time, capturing the temporal propagation dynamics and structural relationships within the graph  $G_t = (V_t, E_t)$ . This dual approach ensures that the fine-tuned GPT-2 embeddings provide a solid semantic foundation, while the TGCN refines these representations to highlight task-specific patterns of misinformation spread, such as sudden engagement spikes.

### 3.4.2 Hierarchical co-attention Mechanism

The hierarchical co-attention mechanism is designed to jointly capture dependencies between the temporal graph features from TGCNs and the semantic features extracted from GPT-2. The mechanism targets specific graph regions and news content to better understand fake news spread. The outputs from TGCN, denoted as  $Q \in \mathbb{R}^{d \times N}$ , and from GPT-2, denoted as  $V \in \mathbb{R}^{d \times N}$ , are fed into the hierarchical co-attention mechanism. An affinity matrix  $C$ , defined in the space  $\mathbb{R}^{N \times N}$ , is calculated as shown in Eq. 5, where  $W_b \in \mathbb{R}^{d \times d}$  is a weight matrix and  $\sigma$  is the activation function of the sigmoid, which quantifies the similarity between temporal and semantic features, thereby initiating the fusion of dynamic propagation data with static content analysis.

$$C = \sigma(Q^\top W_b V), \quad (5)$$

Next, intermediate representations  $Z$  are calculated to refine this mapping, derived using the ReLU activation function applied to sigmoid-transformed weighted sums, as defined in Eq. 6, with weight matrices  $W_i^{(1)}$ ,  $W_j^{(1)}$  and bias vectors  $b_i^{(1)}$  facilitating the computation. This step enhances the model’s ability to focus on relevant patterns by combining temporal node updates  $h_v(t)$  with semantic insights.

$$Z_i = \text{ReLU}(\sigma(W_i^{(1)} V_i + W_j^{(1)} V_j C^{T_i})) + b_i^{(1)}, \quad (6)$$

where  $i, j \in \{v, q\}$  with  $i \neq j$ , and  $T_i$  is the transpose indicator ( $T_v = I$ ,  $T_q = \top$ ), allowing  $V_i, V_j \in \{V, Q\}$  to dynamically represent the appropriate input vectors.

The process continues with the transformation of  $Z$  into final hidden states  $H$  using an exponential function followed by ReLU activation, defined as Eq. 7, where  $i, j \in \{v, q\}$  with  $i \neq j$ , weight matrices  $W_i^{(2)}$ ,  $W_j^{(2)}$  and bias vector  $b_i^{(2)}$  drive the transformation, and  $T_i$  is the transpose indicator ( $T_v = I$ ,  $T_q = \top$ ).

$$H_i = \text{ReLU}(\exp(W_i^{(2)} Z_i + W_j^{(2)} Z_j C^{T_i})) + b_i^{(2)}, \quad (7)$$

These parameters, distinct from the initial  $W_b$  in the affinity matrix  $C$ , enable the model to refine feature integration by amplifying significant temporal and semantic shifts—such as sudden engagement spikes—through multiple layers of transformation. Specifically,  $W_i^{(2)}$  and  $W_j^{(2)}$  enhance the attention mechanism’s focus on dynamic propagation trends, prioritizing relevant graph regions and content for attention probabilities  $a_v$  and  $a_q$ . This progressive refinement optimizes the representation of complex relationships, preparing the data for subsequent processing.

To enhance feature extraction, max-pooling layers are applied to the transformed representations, reducing noise and emphasizing the most significant features, such as peak engagement events as Eq. 8-9. This max-pooling operation refines the hidden states by attenuating less relevant temporal and semantic variations, thereby focusing on key moments of interaction that indicate potential misinformation spread. This process prepares the data for subsequent attention weighting, ensuring that the model prioritizes the most impactful features in fake news detection.

$$H_v^{(4)C} = \text{MaxPooling}(\sigma(H_v)), \quad (8)$$

$$H_q^{(4)C} = \text{MaxPooling}(\sigma(H_q)), \quad (9)$$

Finally, attention probabilities  $a_v$  and  $a_q$  are computed using the softmax function to prioritize relevant graph regions and content:

$$a_v = \text{softmax}(w_h^v H_v^{(4)C}), \quad (10)$$

$$a_q = \text{softmax}(w_h^q H_q^{(4)C}), \quad (11)$$

where  $w_h^v, w_h^q \in \mathbb{R}^k$  are weight parameters. The TGCN output  $Q$  is then weighted and optimized to produce the final output vector:

$$\hat{q} = \sum_{t=1}^T a_q^t q_t, \quad (12)$$

where  $\hat{q}$  represents the refined feature set for classification. This step ensures the model focuses on the most relevant temporal and semantic cues, improving detection accuracy.

Although we have labelled data, there are complex, latent patterns within the news propagation structure that are not fully captured by the labels. These include the dynamics of user interactions, sudden spikes in engagement, and how false information spreads over time. Many of these patterns do not have explicit labels but provide crucial context for identifying fake news. By incorporating unsupervised loss, the model can learn from these hidden structures and relationships within the data. This improves its ability to generalize and make accurate predictions even in cases where the labelled data may be noisy, incomplete, or ambiguous. The unsupervised loss is a regularizer, encouraging the model to discover underlying trends in the propagation network that complement the labelled information. This is especially important in domains like fake news detection, where misinformation can exhibit subtle and complex propagation behaviours that are not always evident from labelled examples alone. In the training phase, unsupervised consistency training is applied through the Bayes by Backprop loss for latent relations without labels, denoted by:

$$L_{\text{unsup}} = \mathbb{E}_{q(z|x)} [\log p(x|z)] - \text{KL}(q(z|x) \parallel p(z)) \quad (13)$$

where  $L_{\text{unsup}}$  is the unsupervised loss,  $z$  are latent variables,  $x$  are observed variables, and KL denotes the Kullback-Leibler divergence. The model is then optimized by minimizing the weighted sum of the unsupervised and supervised loss, where  $\lambda$  is a balancing factor:

$$L = \lambda L_{\text{unsup}} + (1 - \lambda) L_{\text{sup}} \quad (14)$$

By integrating temporal structural characteristics from TGCNs and deep contextual features from GPT-2 through a hierarchical co-attention mechanism, the TGCN-HCA algorithm 2 aims to improve the accuracy of fake news detection by learning from both labeled data (supervised loss) and latent patterns (unsupervised loss).

---

**Algorithm 2** TGCN-HCA Algorithm

---

```
1: Input: Dataset  $\mathcal{D} = \{G_1, \dots, G_N\}$ , where  $G_i = (V_i, E_i)$  is the  $i$ -th propagation graph;  
    $V_i = \{v_{i1}, \dots, v_{in_i}\}$  are nodes (tweets),  $E_i$  are retweet edges.  
2: for  $i = 1$  to  $N$  do  
3:   if  $G_i$  then  
4:      $X_i \leftarrow \text{GPT2}(G_i)$   $\triangleright$  Process text content to capture semantic features  
5:      $G_i \leftarrow (X_i, \text{retweet\_relation}(X_i))$ ,  $A_i \in \{0, 1\}^{n_i \times n_i}$   
6:      $F_i \leftarrow \text{GPT2}(r_i)$   $\triangleright$  Generate output vectors from root node embeddings  
7:      $G'_i \leftarrow (\{node_1^i, \dots, node_j^i\}, \text{co\_occurrence\_relation}(\cdot))$ ,  $A'_i \in \{0, 1\}^{r \times r}$   
8:   end if  
9: end for  
10: for  $i = 1$  to  $N$  do  
11:    $X_i \leftarrow \text{TGCN}(X_i)$   $\triangleright$  Update node representations based on temporal neighbor  
   interactions  
12: end for  
13: for  $i = 1$  to  $N$  do  
14:   for  $j = 1$  to  $p$  do  
15:     if  $node_j^i \in r_i$  then  
16:        $z_j^i \leftarrow f_j + r_i$   $\triangleright$  Update feature vector based on node interactions  
17:     end if  
18:   end for  
19: end for  
20:  $F_i \leftarrow [z_1^i, z_2^i, \dots, z_p^i]^T$   
21: for  $i = 1$  to  $N$  do  
22:    $F_i \leftarrow \text{GCN}(F_i)$   $\triangleright$  Further refine feature vectors using a Graph Neural Network  
23: end for  
24: for  $i = 1$  to  $N$  do  
25:    $F_i \leftarrow \text{avg}(F_i) + \text{maxpooling}(F_i)$   $\triangleright$  Aggregate features using average and max pooling  
26: end for  
27: for  $i = 1$  to  $N$  do  
28:    $Q_i, V_i \leftarrow \text{TGCN}(X_i), \text{GPT2}(X_i)$   $\triangleright$  Obtain features from TGCN and GPT-2  
29:    $C_i \leftarrow \text{calculate\_affinity}(Q_i, V_i)$   $\triangleright$  Calculate affinity matrix  
30:    $Z_v, Z_q \leftarrow \text{compute\_intermediate\_representations}(Q_i, V_i, C_i)$   
31:    $H_v, H_q \leftarrow \text{transform\_representations}(Z_v, Z_q, C_i)$   $\triangleright$  Transform representations  
32:    $H_v\_pool, H_q\_pool \leftarrow \text{maxpooling}(H_v), \text{maxpooling}(H_q)$   $\triangleright$  Apply max-pooling  
33:    $a_v, a_q \leftarrow \text{softmax}(H_v\_pool), \text{softmax}(H_q\_pool)$   $\triangleright$  Calculate attention probabilities  
34:    $\hat{q} \leftarrow \text{compute\_final\_output}(a_q, Q_i)$   $\triangleright$  Compute final output vector  
35: end for  
36: for  $i = 1$  to  $N$  do  
37:    $\hat{y} \leftarrow \text{softmax}(F_i)$   $\triangleright$  Predict label  
38: end for return  $\hat{y}$   $\triangleright$  predicted label of an event
```

---

## 4 Experimental Setup

All experiments are performed using Python and PyTorch frameworks. The models are trained and evaluated on a server equipped with an NVIDIA RTX 6000 Ada GPU, offering 81.4 TFLOPS and 48 GB of memory, along with an AMD EPYC 75F3 32-core

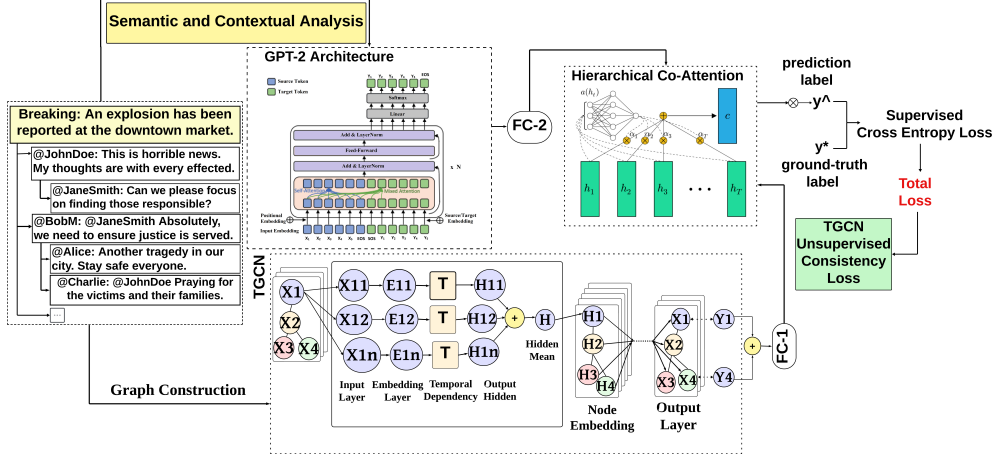


Fig. 2 The Proposed TGCN\_HCA Model Architecture

Processor and 128 GB of RAM. The server features PCIe 4.0 with 24.7 GB/s bandwidth, NVMe storage with 5223 MB/s read speed, and a maximum CUDA version of 12.2. This setup ensures efficient handling of large datasets and complex computations. The training environment is configured with CUDA version 11.1, optimizing performance and compatibility with the hardware.

#### 4.1 Datasets Description and Preprocessing

The efficacy of the proposed approach has been validated by conducting experiments on established benchmark datasets, including Twitter15, Twitter16, and Weibo [40], [41]. The statistics for these datasets are shown in Table 2 and Fig. 3. Each dataset consists of source tweets and their associated retweet user sequences. Specifically, Twitter15 includes 1,490 claims, Twitter16 has 818 claims, and Weibo contains 4,664 claims. Twitter15 and Twitter16 datasets are categorized with four ground truth veracity labels: true news (T), fake news (F), non-fake news (NF), and unverified news (U). Only "true" and "fake" labels are used for the Weibo dataset. Since the original data lacks user profiles, user information is collected using user IDs through a web crawler and the Twitter API. The datasets are divided into three parts: 70% of the data is randomly selected for training, 15% for testing, and 15% for validation.

First, text cleaning is performed, which involves removing URLs, mentions, hashtags, and special characters from the text. The cleaned text is then tokenized. GPT-2 is adopted to convert these tokens into their corresponding word embeddings. These embeddings capture the semantic meaning of the words and are used as input features for the model. The embeddings,  $X_i$ , are extracted from the last hidden state. Importantly, these embeddings are frozen throughout the training process and are used as static input features for the graph construction. This ensures that the embeddings capture the semantic meaning of the text consistently without being influenced by the randomly initialized parameters of the downstream model.

Table 2 Dataset Statistics

Feature	Twitter15	Twitter16	Weibo
<b>Total Claims</b>	1,490	818	4664
<b>True News (T)</b>	370	205	2351
<b>False News (F)</b>	374	204	2313
<b>Non-Fake News (NF)</b>	374	203	–
<b>Unverified News (U)</b>	372	206	–
<b>Total Postings</b>	331,612	204,820	3,805,656
<b>Users</b>	190,868	115,036	2,746,818
<b>Avg. Retweets per Story</b>	292.19	308.70	414
<b>Avg. Words per Source</b>	13.25	12.81	16.30
<b># Total Nodes</b>	912638	501032	2,856,741
<b># Total Edges</b>	697523	382936	2,183,388

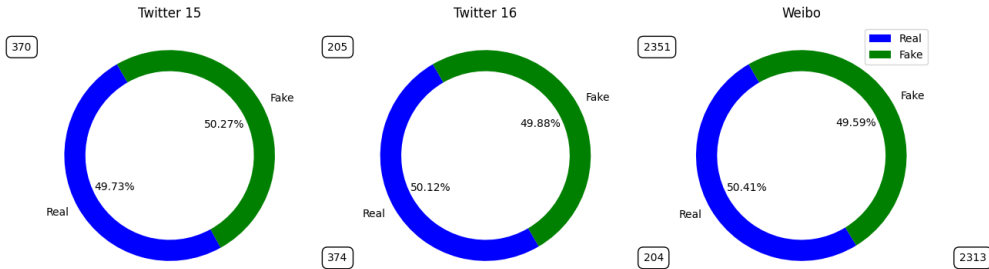


Fig. 3 Fake news and real news distribution in the datasets

## 4.2 Hyper parameters and Complexity Analysis

The settings for the TGCN component are shown in Table 3. TGCN operates with a hidden vector dimension of 64 and 3 layers, which are adjustable based on performance outcomes, effectively balancing model complexity with computational demands. Conversely, GPT-2 employs a hidden vector dimension of 768 across 12 layers, optimizing model effectiveness while managing complexity efficiently. Both models utilize 16-bit floating points to minimize size and enhance processing speed. A pruning rate of 10% in both models streamlines the architecture by removing non-critical weights, thus improving generalization and efficiency. The configurations of potential relations  $T$  range from 1 to 4, and coefficient weight  $\gamma$  from 0.1 to 0.9. Training is conducted using back-propagation with the Adam optimizer, setting learning rates at 0.001 for Twitter15, 0.0008 for Twitter16, and 0.01 for Weibo. Training extends over 25 epochs, with early stopping activated after 10 epochs of consistent loss decrease.

For the TGCN model, the attention mechanism uses weight matrices  $W_q, W_k, W_v$ , each with a dimension of  $64 \times 64$ , which results in 12,288 parameters per layer for a single attention head. With 3 layers, this contributes 36,864 parameters from the attention mechanisms alone, leading to a total of 49,920 parameters for TGCN when including additional layers and components. In GPT-2, each transformer block has a dimension of 768 and 12 attention heads. The parameters for each attention block (queries, keys,

and values) are calculated as  $3 \times 768^2 \times 12$ , which totals 4,226,048 parameters for the attention mechanisms. The total parameter count for GPT-2 reaches 4,614,912 when including feed-forward layers and additional network components.

**Table 3** Hyperparameter Tuning for TGCN-HCA Model

Parameter	TGCN	GPT-2
Hidden Vector Dimension	64	768
Number of Layers	3 (adjustable)	12
Quantization	16-bit floating points	16-bit floating points
Pruning Rate	10%	10%
Number of Potential Relations ( $T$ )	[1, 4]	-
Coefficient Weight ( $\gamma$ )	[0.1, 0.9]	-
Learning Rate (Twitter15)	0.001	$5 \times 10^{-5}$
Learning Rate (Twitter16)	0.0008	$5 \times 10^{-5}$
Learning Rate (Weibo)	0.01	$5 \times 10^{-5}$
Training Epochs	25	25
Batch Size	32	32
Early Stopping Patience	10	5
Number of Attention Heads	N/A	12
Dropout Value	0.5	0.3
Other Parameters	Consistent with GPT-2	Consistent with TGCN

### 4.3 Comparative Analysis with Existing Studies

The proposed model is compared with several existing models in both datasets. For the Twitter 15 dataset, the model is evaluated against the following state-of-the-art (SOTA) models:

- The DGCAN model utilizes a Dual-Channel Graph Convolutional Attention Network to improve fake news detection by capturing both semantic and dissemination patterns in a heterogeneous graph. It processes tweet, user, and word interactions while using attention mechanisms to prioritize essential features [42].
- The Graph-aware Co-attention Network (GCAN) predicts fake news based on the source tweet and its propagation, with a variant called GCAN-G that excludes the graph convolution part to test the effectiveness of graph-aware representation [2].
- The GNN-based REPORT model jointly learns user correlation and information propagation to detect rumours [43].
- EBGCN employs bidirectional GCNs to model trees in ascending and descending directions, considering edge relationships as uncertain [16].
- The deep residual network, GCRES, integrates four cascading graph convolutional networks to capture long-range dependencies and nonlinear features effectively [36].
- The hierarchical graph convolutional network TEH-GCN fuses topic and structural information of fake news [17].
- Lastly, the ClaHI-GAT model represents tweet contents and interaction graphs in a latent space to capture multi-level fake news features through claim-aware and inference-based attention [35].

The proposed model is evaluated against baseline models using the Weibo dataset, which is categorized into true news (T), fake news (F), and unverified news (U).

- The Graph Global Attention Network with Memory (GANM) is introduced for fake news detection and features methods like GAT-GANM, SAGE-GANM, and GCN-GANM [12].
- The GCN-MT approach utilizes NLP techniques to encode nodes with news context and user content [44]. It employs graph convolutional networks to extract key features from the news propagation network, integrating both user-generated and external information.
- The Hierarchically Aggregated Graph Neural Network (HAGNN) captures various granularities of high-level text content representations and merges rumour propagation structures [15].
- Test-time Adaptation for Rumor Detection (TARD) models news propagation as a graph and creates a test-time adaptation framework for handling distribution shifts [45].
- The GAME-ON framework is a GNN-based seq2seq trainable model that enables robust multimodal fake news detection through granular interactions within and across different modalities [46].

#### 4.4 Evaluation Metrics

Several key metrics are utilized to evaluate the performance of the proposed model for fake news detection. Accuracy measures the overall correctness of the model across all predictions. Precision measures the proportion of optimistic predictions among all positive predictions, while Recall emphasizes the model’s ability to capture all actual positives. F1 Score provides a balanced measure that combines Precision and Recall and offers a single metric for assessing a model’s accuracy in terms of both relevance and retrieval completeness. The Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC) curve is also critical. AUC measures the model’s ability to distinguish between classes, serving as a comprehensive performance indicator across various threshold levels.

## 5 Results and Discussion

The performance of the proposed TGCN-HCA model is evaluated for the four categories in Twitter15 and Twitter16 as mentioned in Section 4.1. The accuracy and F1 are calculated in two classes for the Weibo dataset. The results are presented in Table 4 and Table 5 for Twitter 15 - 16 datasets and Weibo datasets respectively. TGCN-HCA achieved an overall F1 of 95.81%, with F1 of 96.10% for non-fake news, 95.00% for fake news, 96.78% for true news, and 94.45% for unverified news On Twitter 15.

The TGCN-HCA model surpasses all comparative models in performance across three datasets. For Twitter 15, TGCN-HCA significantly surpasses state-of-the-art models, increasing accuracy by 5.37% and improving F1 scores for non-fake news, true news, fake news, and unverified news by 0.40%, 4.81%, 1.58%, and 4.85%, respectively. On Twitter16, TGCN-HCA enhances accuracy by 4.08%, with F1 score

**Table 4** Comparison of Proposed Model with Baseline and State of the art Models on Twitter 15 and Twitter 16 datasets (%).

Method	Year	Twitter15					Twitter16				
		Acc	NF	F	T	U	Acc	NF	F	T	U
			F1	F1	F1	F1		F1	F1	F1	F1
GCAN [2]	2020	82.50	82.95	82.57	87.67	88.21	90.84	75.93	76.32	75.94	78.76
GCAN-G [2]	2020	79.38	79.90	79.59	86.36	87.33	79.39	67.54	68.02	67.85	69.41
ClaHi-GAT [35]	2021	84.70	80.60	81.70	88.60	85.40	83.50	83.20	82.30	82.40	84.90
EBGCN [16]	2021	89.20	86.90	89.70	93.40	86.70	91.50	87.90	90.60	94.70	91.00
GCREs [36]	2023	85.30	85.50	85.80	90.30	74.60	88.80	80.10	87.70	91.20	91.90
TEH-GCN [17]	2023	88.80	95.70	87.40	89.40	81.40	88.70	95.10	90.10	80.30	86.00
GBCA [14]	2024	90.30	90.10	91.00	95.20	89.60	92.80	89.20	91.50	92.80	93.20
REPORT [43]	2024	86.00	93.80	82.80	87.70	79.70	85.30	84.10	83.10	90.30	83.50
DGCAN [42]	2024	92.30	96.40	92.30	92.40	87.80	94.00	94.60	90.90	95.80	94.50
<b>TGCN-HCA</b>		<b>95.67</b>	<b>96.10</b>	<b>95.81</b>	<b>96.78</b>	<b>94.45</b>	<b>96.58</b>	<b>96.11</b>	<b>97.60</b>	<b>96.90</b>	<b>96.00</b>

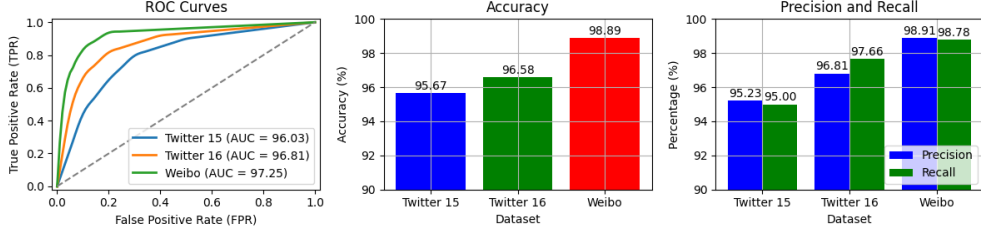
improvements of 1.01% for non-fake news, 6.10% for fake news, 2.20% for true news, and 2.80% for unverified news. For the Weibo dataset, TGCN-HCA shows substantial performance improvements over state-of-the-art models. Using Event+User+Network features, it increases accuracy by 1.03%, with an F1 improvement of 1.13% and a precision increase of 1.58% compared to the best-performing baseline model, GCN-MT. Additionally, with Event+User+Profile features, TGCN-HCA achieves an accuracy of 98.23% and an F1 of 98.87%, outperforming other models. The baseline results are referenced from the original papers.

**Table 5** Comparative Analysis of Proposed Model with Baseline and State of the art Models on the Weibo dataset (%).

Method	Year	Features	Accuracy	F1	Precision	Recall
HAGNN [15]	2023	Context	95.70	96.60	94.90	98.30
TARD [45]	2024	Context	73.70	72.70	71.50	72.00
GAT-GANM [12]	2024	Profile	87.13	87.19	87.10	87.00
SAGE-GANM [12]	2024	Profile	92.25	92.22	91.87	92.34
GCN-GANM [12]	2024	Profile	92.16	92.18	92.45	92.00
GCN-MT [44]	2024	Event+User+Network	97.86	97.87	97.33	98.41
GAME-ON [46]	2024	Content+Context	88.90	89.10	89.00	89.00
<b>TGCN-HCA</b>		<b>Event+User+Network</b>	<b>98.89</b>	<b>99.00</b>	<b>98.91</b>	<b>98.78</b>
<b>TGCN-HCA</b>		<b>Event+User+Profile</b>	<b>98.23</b>	<b>98.87</b>	<b>98.89</b>	<b>98.63</b>

The accuracy is 95.67%, precision 95.23%, and recall is 95.00%. For Twitter 16, the overall F1 is 97.60%, with non-fake news at 96.11%, fake news at 97.60%, true news at 96.90%, and unverified news at 96.00%. The accuracy is 96.58%, precision is 96.81%, and recall is 97.66%. On the Weibo dataset, TGCN-HCA achieved the highest performance with an overall F1 score of 99.00%, an accuracy of 98.89%, a

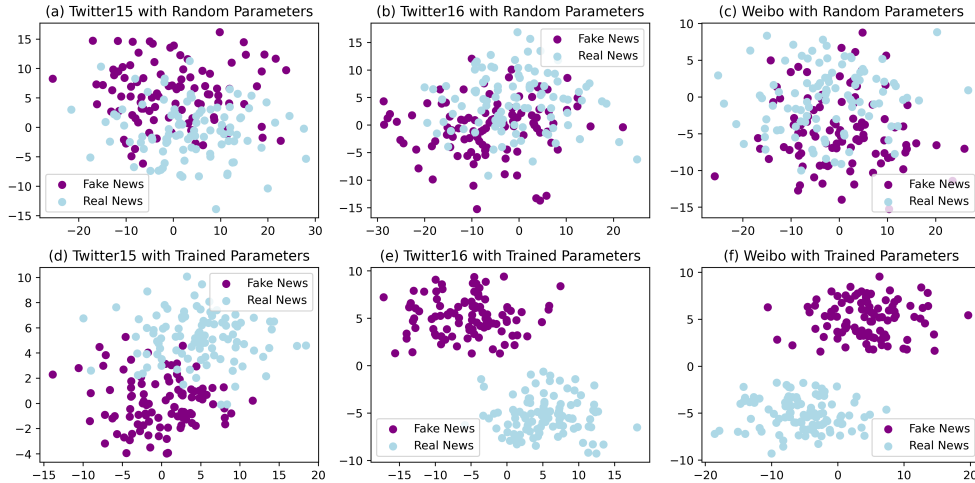
precision of 98.91%, and a recall of 98.70%. The ROC curves for all datasets in Fig. 4 illustrate the TGCN-HCA model’s high performance distinguishing between true and false news. The Area Under the Curve (AUC) values are 96.03 for Twitter 15, 96.81 for Twitter 16, and 97.25 for Weibo, indicating excellent classification capability across all datasets. These high AUC values demonstrate the model’s robust generalization and effective feature extraction. It ensures reliable fake news detection across diverse social media platforms. These results validate the model’s ability to effectively capture the temporal evolution of news dissemination and the importance of different graph regions and content features for incorporating semantic features.



**Fig. 4** Assessing Model Efficacy: ROC Curves, Accuracy, and Precision-Recall Metrics

The temporal dynamics of news propagation vary across platforms. Twitter has a faster and more dynamic spread of information, which creates challenges for the model to accurately capture the temporal dependencies, especially in unverified or quickly spreading fake news. Weibo’s more structured and possibly slower propagation patterns have allowed the model to capture temporal dependencies more effectively, contributing to its higher performance on the Weibo dataset. The observed differences in the model’s performance can be attributed to these variations across the datasets. For instance, the model achieved a higher F1 score and accuracy on the Weibo dataset due to the more balanced data distribution, simpler graph structure, and richer semantic features in longer posts. Meanwhile, the slightly lower performance on Twitter15 and Twitter16 reflects the greater complexity in user interaction patterns, shorter post lengths, and more nuanced news categories like unverified news, which challenge the model’s ability to distinguish between them. These factors highlight the importance of dataset-specific characteristics in determining the effectiveness of fake news detection models like TGCN-HCA.

The TGCN-HCA model, with randomly initialised weight parameters and no gradient descent optimisation, shows a significant improvement in separating fake and real news data points for the Twitter15, Twitter1 [40] and Weibo [41] datasets, as visualised in Fig. 5 using t-SNE. This improved separation underscores the model’s robustness and effectiveness in identifying fake news across different datasets.



**Fig. 5** Analysing the separation of fake news and real news on datasets through scatter plots with random and trained parameter settings

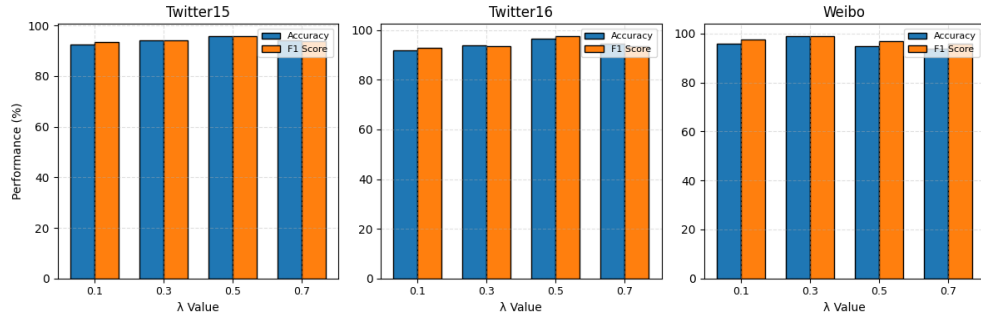
## 5.1 Sensitivity Analysis of $\lambda$ Parameter

The visualization in Fig. 6 presents the model’s performance across three datasets Twitter15, Twitter16, and Weibo at varying  $\lambda$  values (0.1, 0.3, 0.5, 0.7). Each dataset is analyzed through Accuracy and F1 Score which is represented by the blue and orange bars, respectively. For the Twitter15 dataset, the model achieves an accuracy of 95.67% and an F1 score of 95.81% at  $\lambda$  value of 0.5. Similarly, the Twitter16 dataset shows optimal performance with an accuracy of 96.58% and an F1 score of 97.60% at  $\lambda$  value of 0.5. However, performance declines when  $\lambda$  is increased to 0.7 or decreased to 0.1 and 0.3, indicating that a heavier emphasis on unsupervised loss may hinder the model’s ability to effectively utilize the supervised signals in these datasets.

In contrast, the Weibo dataset reaches its peak performance at  $\lambda$  value of 0.3, where it records an accuracy of 98.89% and an F1 score of 99.00%. At  $\lambda$  value of 0.5, the accuracy is still high as 98.00%, but there is a noticeable drop in performance at  $\lambda$  value of 0.7. It suggests that the unique characteristics of the Weibo dataset require a different balance of supervised and unsupervised learning compared to the Twitter datasets. These results highlight the model’s sensitivity to the  $\lambda$  parameter and underscore its critical role in optimizing performance across diverse datasets.

## 5.2 Model Behavior

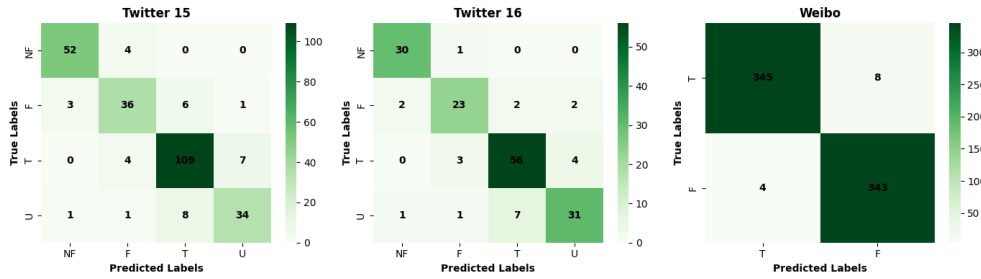
The confusion matrices in Fig. 7 for the Twitter 15, Twitter 16, and Weibo datasets demonstrate the model’s strong classification performance across various news categories. In the Twitter 15 dataset, the model effectively identifies True News, correctly classifying 109 instances, with only 4 misclassified as False News and 7 as Unverified News. This reflects the model’s high accuracy in distinguishing reliable information. For False News, 36 instances are correctly classified, while 6 are misclassified as True



**Fig. 6** Sensitivity Analysis of Model Performance with Varying  $\lambda$  Values

News and 1 as Unverified News. Similarly, in the Twitter 16 dataset, the model shows a high level of accuracy in classifying True News, with 56 instances correctly identified and only 3 misclassified as False News and 4 as Unverified News. The classification of False News is also effective, with 23 instances correctly classified and minor misclassifications (2 instances each) into True News and Unverified News. The model achieves excellent performance in the Weibo dataset, which consists solely of True and False News, correctly classifying 345 instances of True News and 343 instances of False News, with only 8 and 4 misclassifications, respectively. The binary nature of the Weibo dataset enables the model to achieve high accuracy with minimal errors.

The variation in results across the datasets can be attributed to the different classification challenges each presents. While the Weibo dataset involves a simpler binary classification between True and False News, the Twitter 15 and Twitter 16 datasets introduce more complexity with additional categories, such as Unverified News. Due to overlapping content and dissemination patterns, this complexity leads to slightly more misclassifications, particularly between False News and Unverified News. These results highlight the model’s ability to adapt to both binary and multi-class classification tasks, performing effectively across diverse datasets.



**Fig. 7** Confusion Matrices for Twitter 15, Twitter 16, and Weibo.

The training and validation accuracy and loss graphs in Fig 5.2 further support the model’s effectiveness, maintaining higher accuracy and lower loss throughout the training process without signs of overfitting.

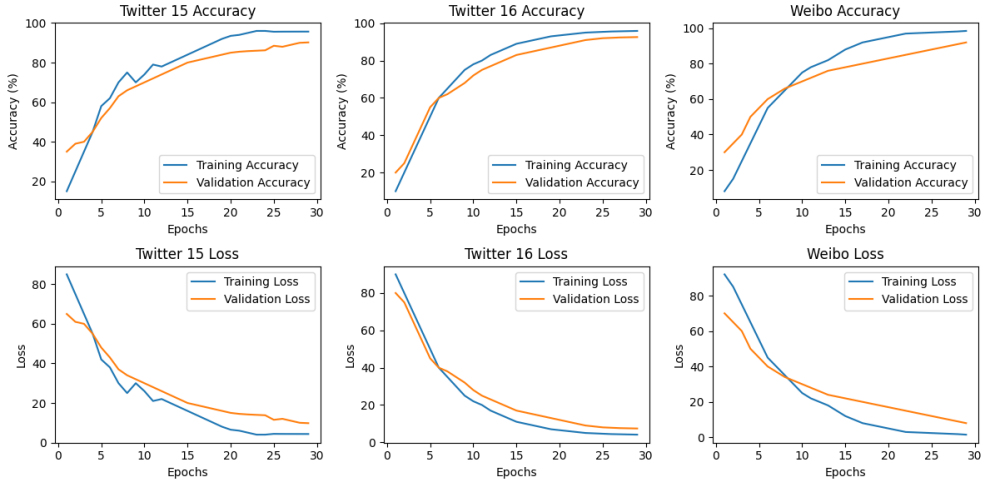


Fig. 8 TGCN-HCA Training Validation Accuracy and Loss

### 5.3 Cross Dataset Performance

The proposed TGCN-HCA model’s performance is evaluated through cross-dataset experiments using the same setup as Section 4. The results show robust accuracy and F1 scores across datasets: Trained on Twitter 15, it achieves 92.82% accuracy and 90.77% F1 score on Twitter 16, and 86.97% accuracy and 87.07% F1 score on Weibo. Trained on Twitter 16, it reaches 93.97% accuracy and 94.10% F1 score on Twitter 15, and 89.97% accuracy and 90.07% F1 score on Weibo. Trained on Weibo, it achieves 85.97% accuracy and 86.10% F1 score on Twitter 15, and 84.82% accuracy and 85.77% F1 score on Twitter 16. These results, presented in Table 6, demonstrate the model’s strong generalization capabilities in cross-dataset analysis. While a slight performance drop occurs when applied to datasets with different characteristics, the TGCN-HCA model adapts well across datasets, showcasing its robustness.

### 5.4 Ablation Study

To comprehensively evaluate the effectiveness of the proposed TGCN-HCA model, an ablation study is conducted by systematically removing or adding key components and observing the impact on performance across the datasets in Table 7. It focused on the Temporal Graph Convolution Networks (TGCN), Hierarchical Co-Attention Mechanism (HAN), GPT-2, and the combined TGCN-HCA model. The configurations

**Table 6** Detailed Cross-Dataset Model Performance with F1 Scores and Accuracy

Trained Dataset	Tested Dataset	Accuracy	F1	F1 (T)	F1 (F)	F1 (U)
Twitter 15	Twitter 16	92.82	90.77	93.77	89.79	88.82
	Weibo	86.97	87.07	88.27	91.08	89.09
Twitter 16	Twitter 15	93.97	94.10	94.19	95.15	95.29
	Weibo	89.97	90.07	91.70	90.08	90.39
Weibo	Twitter 15	85.97	86.10	86.22	88.15	87.19
	Twitter 16	84.82	85.77	86.11	87.79	86.82

tested included using TGCN with other models like Text CNN [47], TextRNN [48], and BERT [49], with and without GPT-2 and HAN.

Results revealed that TGCN alone is not fully effective and lacks semantic feature extraction, which GPT-2 enriched models compensated for by improving performance. The absence of GPT-2 embeddings, graph convolutions, or co-attention mechanisms reduced performance, emphasizing their critical roles. TextCNN efficiently processed parallel text, TextRNN adeptly captured temporal dynamics, and BERT excelled in extracting deep text semantics and contextual understanding. TextRNN tends to outperform TextCNN because TextCNN struggles with capturing information in longer text sequences.

In contrast, TextRNN effectively considers the order of text and captures temporal features, which is crucial in classification tasks. Consequently, while the fusion of BERT models delivers the best results due to their deep semantic and contextual learning capabilities, it is essential to note that we exclude the HCA mechanism in the configurations using TextCNN, TextRNN, and BERT. This exclusion allowed us to precisely assess the impact of the HCA when integrated with GPT-2 in the full model. Additionally, we replaced the unsupervised loss function with the supervised loss function to assess the impact of the unsupervised loss. The results from this configuration and a comparison to the version that included the unsupervised loss were incorporated into the study. The findings indicate that the unsupervised loss is impactful, as the TGCN-HCA model, which consists of the unsupervised loss, significantly outperforms the version without it, underscoring the importance of the unsupervised learning component in enhancing model performance.

The TGCN-HCA model captures complex relationships within longer texts, performs deep semantic analysis, and recognizes temporal dependencies. This comprehensive capability allows it to understand and analyze intricate data interactions effectively.

## 5.5 Computational Studies

The table 8 compares various fake news detection models by focusing on their total parameters and accuracy across different datasets. The GCN+BERT model, despite having a high parameter count of 77,885,952, shows moderate accuracy between 90%-92% for Twitter 15 and Twitter 16 datasets. In contrast, the GCN-MT model, with 5,505,024 parameters, achieves high accuracy of up to 97% for the Weibo dataset, indicating a more efficient performance relative to its parameter size. Models like

**Table 7** Ablation Study Results on Different Datasets

Configuration	Twitter15		Twitter16		Weibo	
	Accuracy (%)	F1 (%)	Accuracy (%)	F1 (%)	Accuracy (%)	F1 (%)
TGCN Only	87.50	87.60	86.70	86.90	92.80	93.00
TGCN w/ GPT-2	89.20	89.30	88.10	88.20	94.50	94.70
TGCN w/ supervised loss	90.37	90.11	89.43	89.37	94.79	95.23
TGCN + TextRNN	92.00	92.30	90.80	91.00	96.00	96.20
TGCN + TextCNN	91.30	91.50	89.90	90.10	95.80	96.00
TGCN + BERT	93.00	93.20	91.50	91.70	97.00	97.20
<b>TGCN-HCA</b>	<b>95.67</b>	<b>95.81</b>	<b>96.58</b>	<b>97.60</b>	<b>98.89</b>	<b>99.00</b>

EBGCN and GCRES, with lower parameter counts, still maintain commendable accuracy, demonstrating that smaller models can be quite effective. Notably, TGCN-HCA, with relatively few parameters (4,664,832), achieves impressive accuracy, particularly on Weibo (98%).

The TGCN-HCA model achieves high accuracy and efficiency by integrating efficient graph-based architectures with advanced semantic features like GPT-2, while maintaining a low parameter count. By streamlining the model architecture and leveraging weight pruning and layer freezing techniques, we reduce the total parameter count to 4,664,832 without sacrificing performance. This efficient design enables the model to effectively process complex data and detect nuanced patterns in fake news, making it highly effective while being computationally efficient compared to bulkier models with higher parameter counts. It effectively processes complex data and nuanced language, which are essential for detecting fake news while keeping parameter counts low compared to bulkier models like GCN-GANM. This makes our model highly effective and practical for real-world applications, blending cutting-edge text processing with robust architectural efficiency.

**Table 8** Comparison of Model Parameters, Estimated Execution Time, and FLOP with Existing Models

Model	Total Parameters	Est. Time (s)	Accuracy (%)		
			Twitter 15	Twitter 16	Weibo
GCN+BERT[14]	77,885,952	0.050	0.90	0.92	–
GCAN [50]	3,279,168	0.002	0.82	0.90	–
EBGCN[16]	3,042,506	0.002	0.89	0.91	0.93
GCN-MT [44]	5,505,024	0.004	–	–	0.97
GCRES [36]	82,692	0.0001	0.88	0.90	0.97
Report[43]	92,712	0.0001	0.86	0.85	0.88
GCN-GANM[12]	110,909,600	0.071	0.92	0.91	0.92
GAME-ON[46]	66,012,672	0.042	0.88	0.91	0.88
<b>TGCN-HCA</b>	<b>4,664,832</b>	<b>0.003</b>	<b>0.95</b>	<b>0.96</b>	<b>0.98</b>

## 5.6 Discussion

The TGCN-HCA model demonstrates impressive performance due to its effective integration of Temporal Graph Convolution Networks (TGCNs) with Hierarchical Co-Attention Mechanisms (HCA). This hybrid approach allows the model to leverage both temporal dependencies in news propagation and semantic features from text data, enabling it to capture complex dynamics that traditional models struggle with. By incorporating fine-tuned GPT-2 for contextual analysis and focusing on multi-level interactions between graph structures and textual content, the model effectively addresses the evolving nature of fake news spread. The hierarchical co-attention mechanism optimizes the model’s focus, helping it to prioritize significant features and improve classification accuracy, particularly in distinguishing between true and false news. This ability to dynamically adjust to changing patterns in news dissemination makes TGCN-HCA a robust solution for fake news detection across various social media platforms.

The model’s efficiency in handling the computational challenges posed by graph dynamics and semantic analysis is achieved through various optimizations. Graph convolution operations are handled efficiently by using the TGCN architecture, which aggregates features from neighboring nodes and updates node representations dynamically. Additionally, the integration of GPT-2 embeddings is managed by freezing the initial layers of GPT-2 during training, which prevents overfitting and reduces computational complexity. This approach ensures that the contextual embeddings generated by GPT-2 remain task-specific and consistent, while the model refines higher-level embeddings during training. This combination of layer freezing in GPT-2 and dynamic feature aggregation from TGCN helps the model to maintain a balance between accuracy and computational efficiency, even when processing large and complex datasets.

The model has limitations, facing challenges with data set variability, including content structure, user behavior, and interaction pattern differences between platforms like Twitter and Weibo. These variations slightly reduce performance across datasets as the model struggles to generalize to new content formats or user behaviors. For instance, shorter Twitter posts and longer Weibo posts present structural differences that challenge the model, reducing accuracy on unseen platforms.

The model’s lower performance on unverified news is due to the lack of clear indicators such as factual accuracy or sensationalist language, making it difficult to classify. Unverified news shares characteristics with both true and fake news, leading to frequent misclassifications. Additionally, conflicting user reactions and incomplete content contribute to the challenge. In Twitter15, the performance drop for unverified news is 4.5%, while in Twitter16, the drop is 5%. Improved handling of content ambiguity, authoritativeness, and timing could enhance classification in future models.

In cross-dataset performance (e.g., Twitter  $\rightarrow$  Weibo), the balanced true-false ratio in both datasets is complicated by the presence of unverified news. The structural differences—shorter Twitter posts versus Weibo’s longer content—make it harder for the model to generalize. The performance drops by 5.5% when trained on Twitter15 and tested on Weibo and 6% when trained on Twitter16 and tested on Weibo, due to these content discrepancies. Moreover, user interaction patterns vary between platforms,

leading to a 5.2% accuracy drop when trained on Weibo and tested on Twitter15, and a 6.3% drop when trained on Weibo and tested on Twitter16.

Although the true-false ratio is balanced, the unverified news category adds complexity, particularly in Twitter datasets, where distinguishing between false and unverified news is challenging. This contributes to a 4%–6% performance drop when tested on Weibo, where the distinction between true and false news is clearer. Addressing platform-specific differences and better handling of unverified content will improve model generalization.

These observations highlight the importance of addressing platform-specific differences, early-stage propagation signals, and ambiguous categories like unverified news. To improve generalization across datasets, further research into adaptive learning techniques, feature normalization, and enhanced handling of uncertain content will be essential.

## 6 Conclusion

This paper presents a Temporal Graph Convolution Network with Hierarchical Co-Attention (TGCN-HCA) model for fake news detection. The model captures semantic properties and intricate temporal propagation networks. Our model captures the dynamic nature of news patterns by modelling temporal interaction events of nodes. GPT-2 transforms each token into a high-dimensional vector. These embeddings capture the relationships between tokens and the semantic structure. The adjacency matrix and propagation network aid in comprehending the contextual relationships. A temporal propagation graph refines each node’s representation to reflect the dynamic context by integrating information from nearby nodes. Dependencies between temporal graph features from TGCNs and semantic features taken from GPT-2 are grasped by the hierarchical co-attention. It increases the model’s ability to analyze this complex phenomenon. The performance of TGCN-HCA is demonstrated by extensive testing and the model obtains an accuracy of 95.67%, 96.58%, and 98.89% on Twitter15, Twitter16, and Weibo respectively. It demonstrates notable gains in F1 scores in every category and outperforms the most recent baselines. Training and validation graphs verify that the model performs consistently and effectively without overfitting.

We envisage two primary ways for advancing this research in future studies to delve into temporal heterogeneous graphs that include information about social media users.

A first approach would be to propose the deployment of instruction-tuned LLMs to generate synthetic AI-composed misinformation, such as through prompts like “*Compose a sensational false headline*”, as well as authentic human-like news articles for the purpose of controlled experimental assessments. This methodology will facilitate a rigorous appraisal of detection models in the context of realistic and heterogeneous misinformation scenarios.

A subsequent work will explore the use of advanced LLMs, such as GPT-4 and Grok 4, either as standalone systems or in hybrid forms for detecting fake news. This will lead to their evaluation as effective detector of semantic deception and temporal patterns when combined with graph-based frameworks, which may lead to the discovery of new multimodal detection paradigms.

## 7 Competing Interests

The authors declare no competing interests.

## 8 Availability of data and materials

We have utilized publicly available datasets appropriately cited within the article.

## 9 Code Availability

This is ongoing research with future expansion plans. The code will be made available on a reputable public platform. In the meantime, the Code developed for this article is available upon reasonable request from the corresponding author.

## 10 Author contribution

Muhammad Abdullah<sup>1†</sup>, Zan Hongying<sup>1\*†</sup>, Arifa Javed<sup>1†</sup>, <sup>1†</sup>, 2

software: Muhammad Abdullah, Arifa Javed; Conceptualization: Zan Hongying, Junaid Abdul Wahid, Kanwal Ahmed, Fabio Caraffini, Hassan Eshkiki; Methodology: Muhammad Abdullah; Formal analysis and investigation: Muhammad Abdullah; Visualization: Muhammad Abdullah, Arifa Javed; Writing - original draft preparation: Muhammad Abdullah, Arifa Javed; Writing - review and editing: Fabio Caraffini, Hassan Eshkiki, Kanwal Ahmed; Supervision: Zan Hongying, Junaid Abdul Wahid.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

- [1] Ahmed, K., Khan, M.A., Haq, I., Al Mazroa, A., MS, S., Innab, N., Alajmi, M., Alkahtani, H.K.: Social media's dark secrets: A propagation, lexical and psycholinguistic oriented deep learning approach for fake news proliferation. *Expert Systems with Applications* **255**, 124650 (2024)
- [2] Lu, Y.-J., Li, C.-T.: GCAN: Graph-aware Co-Attention Networks for Explainable Fake News Detection on Social Media (2020)
- [3] Odum, F.: Covid conspiracy narratives: Dissecting the origins of misinformation in digital space (2021)
- [4] Bruns, A., Hurcombe, E., Harrington, S.: Covering conspiracy: Approaches to reporting the covid/5g conspiracy theory. *Digital Journalism* **10**(6), 930–951 (2022)

- [5] Chen, Y.-P., Chen, Y.-Y., Yang, K.-C., Lai, F., Huang, C.-H., Chen, Y.-N., Tu, Y.-C.: The prevalence and impact of fake news on covid-19 vaccination in taiwan: retrospective study of digital media. *Journal of Medical Internet Research* **24**(4), 36830 (2022)
- [6] Yang, Y., Zhou, Y., Ying, Q., Qian, Z., Zeng, D., Liu, L.: Fact-checking based fake news detection: a review. *arXiv preprint arXiv:2401.01717* (2024)
- [7] Gongane, V.U., Munot, M.V., Anuse, A.: Machine learning approaches for rumor detection on social media platforms: a comprehensive survey. *Advanced machine intelligence and signal processing*, 649–663 (2022)
- [8] Yadav, A., Gupta, A.: An emotion-driven, transformer-based network for multimodal fake news detection. *International Journal of Multimedia Information Retrieval* **13**(1), 1–16 (2024)
- [9] Merryton, A.R., Gethsiyal Augasta, M.: An attribute-wise attention model with bilstm for an efficient fake news detection. *Multimedia Tools and Applications* **83**(13), 38109–38126 (2024)
- [10] Javed, A., Zan, H., Mamyrbayev, O., Abdullah, M., Ahmed, K., Oralbekova, D., Dinara, K., Akhmediyarova, A.: Transformer-based re-ranking model for enhancing contextual and syntactic translation in low-resource neural machine translation. *Electronics* **14**(2), 243 (2025)
- [11] Farokhian, M., Rafe, V., Veisi, H.: Fake news detection using dual bert deep neural networks. *Multimedia Tools and Applications* **83**(15), 43831–43848 (2024)
- [12] Chang, Q., Li, X., Duan, Z.: Graph global attention network with memory: A deep learning approach for fake news detection. *Neural Networks* **172**, 106115 (2024)
- [13] Soga, K., Yoshida, S., Muneyasu, M.: Exploiting stance similarity and graph neural networks for fake news detection. *Pattern Recognition Letters* **177**, 26–32 (2024)
- [14] Zhang, Z., Lv, Q., Jia, X., Yun, W., Miao, G., Mao, Z., Wu, G.: Gbca: Graph convolution network and bert combined with co-attention for fake news detection. *Pattern Recognition Letters* (2024)
- [15] Xu, S., Liu, X., Ma, K., Dong, F., Riskhan, B., Xiang, S., Bing, C.: Rumor detection on social media using hierarchically aggregated feature via graph neural networks. *Applied Intelligence* **53**(3), 3136–3149 (2023)
- [16] Wei, L., Hu, D., Zhou, W., Yue, Z., Hu, S.: Towards Propagation Uncertainty: Edge-enhanced Bayesian Graph Convolutional Networks for Rumor Detection (2021)

- [17] Luo, Z., Li, P., Qian, Z., Zhu, X.: Teh-gcn: Topic-event based hierarchical graph convolutional networks for rumor detection. In: 2023 International Joint Conference on Neural Networks (IJCNN), pp. 1–8 (2023). IEEE
- [18] Abdullah, M., Hongying, Z., Javed, A., Mamyrbayev, O., Caraffini, F., Eshkiki, H.: A joint learning framework for fake news detection. *Displays*, 103154 (2025)
- [19] Rodríguez-Ferrándiz, R.: The plandemic and its apostles: Conspiracy theories in pandemic mode. In: *Digital Totalitarianism*, pp. 62–83. Routledge, ??? (2022)
- [20] Damisa, A.: *Fake news: Finding truth in strategic communication* (2024)
- [21] Dheyaa Radhi, A., Ali Hussein Al Naffakh, H., Fuqdan, A.-I., A Hakim, B., Al-Attar, B.: A comprehensive review of machine learning-based models for fake news detection. In: *BIO Web of Conferences*, vol. 97, p. 00123 (2024). EDP Sciences
- [22] Nanade, A., Kumar, A.: Combating fake news on twitter: A machine learning approach for detection and classification of fake tweets. *International Journal of Intelligent Systems and Applications in Engineering* **12**(1), 424–436 (2024)
- [23] Abd El-Mageed, A.A., Abohany, A.A., Ali, A.H., Hosny, K.M.: An adaptive hybrid african vultures-aquila optimizer with xgb-tree algorithm for fake news detection. *Journal of Big Data* **11**(1), 41 (2024)
- [24] Mallik, A., Kumar, S.: Word2vec and lstm based deep learning technique for context-free fake news detection. *Multimedia Tools and Applications* **83**(1), 919–940 (2024)
- [25] Chen, S.: Research on fine-grained classification of rumors in public crisis—take the covid-19 incident as an example. In: *E3S Web of Conferences*, vol. 179, p. 02027 (2020). E3S Web of Conferences
- [26] Sastrawan, I.K., Bayupati, I.P.A., Arsa, D.M.S.: Detection of fake news using deep learning cnn–rnn based methods. *ICT express* **8**(3), 396–408 (2022)
- [27] Deng, Z., Ma, W., Han, Q.-L., Zhou, W., Zhu, X., Wen, S., Xiang, Y.: Exploring deepseek: A survey on advances, applications, challenges and future directions. *IEEE/CAA Journal of Automatica Sinica* **12**(5), 872–893 (2025)
- [28] Deng, Z., Guo, Y., Han, C., Ma, W., Xiong, J., Wen, S., Xiang, Y.: Ai agents under threat: A survey of key security challenges and future pathways. *ACM Computing Surveys* **57**(7), 1–36 (2025)
- [29] Hu, X., Ma, W., Chen, C., Wen, S., Zhang, J., Xiang, Y., Fei, G.: Event detection in online social network: Methodologies, state-of-art, and evolution. *Computer Science Review* **46**, 100500 (2022)

- [30] Mehta, D., Dwivedi, A., Patra, A., Anand Kumar, M.: A transformer-based architecture for fake news classification. *Social network analysis and mining* **11**, 1–12 (2021)
- [31] Hongying, Z., Javed, A., Abdullah, M., Rashid, J., Faheem, M.: Large language models with contrastive decoding algorithm for hallucination mitigation in low-resource languages. *CAAI Transactions on Intelligence Technology* (2025)
- [32] Phan, H.T., Nguyen, N.T., Hwang, D.: Fake news detection: A survey of graph neural network methods. *Applied Soft Computing*, 110235 (2023)
- [33] Wei, P., Xu, N., Mao, W.: Modeling Conversation Structure and Temporal Dynamics for Jointly Predicting Rumor Stance and Veracity (2019)
- [34] Yang, P., Leng, J., Zhao, G., Li, W., Fang, H.: Rumor detection driven by graph attention capsule network on dynamic propagation structures. *The Journal of Supercomputing* **79**(5), 5201–5222 (2023)
- [35] Lin, H., Ma, J., Cheng, M., Yang, Z., Chen, L., Chen, G.: Rumor detection on twitter with claim-guided hierarchical graph attention networks. *arXiv preprint arXiv:2110.04522* (2021)
- [36] Ye, N., Yu, D., Zhou, Y., Shang, K.-k., Zhang, S.: Graph convolutional-based deep residual modeling for rumor detection on social media. *Mathematics* **11**(15), 3393 (2023)
- [37] Li, A., Yang, B.: Gsirec: Learning with graph side information for recommendation. *World Wide Web* **24**(5), 1411–1437 (2021)
- [38] Mai, S., Xing, S., He, J., Zeng, Y., Hu, H.: Multimodal graph for unaligned multimodal sequence analysis via graph convolution and graph pooling. *ACM Transactions on Multimedia Computing, Communications and Applications* **19**(2), 1–24 (2023)
- [39] Zhang, H., Chen, L., Cao, J., Zhang, X., Kan, S.: A combined traffic flow forecasting model based on graph convolutional network and attention mechanism. *International Journal of Modern Physics C* **32**(12), 2150158 (2021)
- [40] Ma, J., Gao, W., Mitra, P., Kwon, S., Jansen, B.J., Wong, K.-F., Cha, M.: Detecting rumors from microblogs with recurrent neural networks (2016)
- [41] Sheng, Q., Cao, J., Bernard, H.R., Shu, K., Li, J., Liu, H.: Characterizing multi-domain false news and underlying user effects on chinese weibo. *Information Processing & Management* **59**(4), 102959 (2022)
- [42] Zhao, M., Zhang, Y., Rao, G.: Fake news detection based on dual-channel graph convolutional attention network. *The Journal of Supercomputing*, 1–22 (2024)

- [43] Liu, T., Cai, Q., Xu, C., Zhou, Z., Ni, F., Qiao, Y., Yang, T.: Rumor detection with a novel graph neural network approach. arXiv preprint arXiv:2403.16206 (2024)
- [44] Chang, Q., Li, X., Duan, Z.: A novel approach for rumor detection in social platforms: Memory-augmented transformer with graph convolutional networks. *Knowledge-Based Systems* **292**, 111625 (2024)
- [45] Tao, X., Zhang, M., Liu, Q., Wu, S., Wang, L.: Out-of-distribution rumor detection via test-time adaptation. arXiv preprint arXiv:2403.17735 (2024)
- [46] Dhawan, M., Sharma, S., Kadam, A., Sharma, R., Kumaraguru, P.: Game-on: Graph attention network based multimodal fusion for fake news detection. *Social Network Analysis and Mining* **14**(1), 114 (2024)
- [47] Alshubaily, I.: Textcnn with attention for text classification. arXiv preprint arXiv:2108.01921 (2021)
- [48] Xia, F., *et al.*: Label oriented hierarchical attention neural network for short text classification. *Academic Journal of Engineering and Technology Science* **5**(1), 53–62 (2022)
- [49] Devlin, J., Chang, M.-W., Lee, K., Toutanova, K.: BERT: Pre-training of deep bidirectional transformers for language understanding, 4171–4186 (2019) <https://doi.org/10.18653/v1/N19-1423>
- [50] Bian, T., Xiao, X., Xu, T., Zhao, P., Huang, W., Rong, Y., Huang, J.: Rumor detection on social media with bi-directional graph convolutional networks. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, pp. 549–556 (2020)