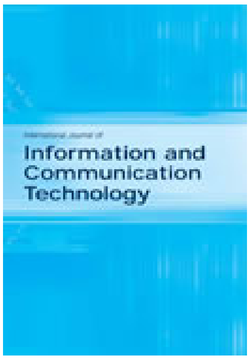


International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

Spatio-temporal convolutional networks empowering political ideology trend prediction on social media platforms

Wencheng Liu, Xiaojuan Deng

DOI: 10.1504/IJICT.2025.10074811

Article History:

Received:	08 September 2025
Last revised:	18 October 2025
Accepted:	20 October 2025
Published online:	12 December 2025

Spatio-temporal convolutional networks empowering political ideology trend prediction on social media platforms

Wencheng Liu

School of Marxism,
Hunan Open University,
Changsha, 410004, China
Email: liuwencheng@hnou.edu.cn

Xiaojuan Deng*

School of Economics and Management,
Shaoyang University,
Shaoyang, 422000, China
Email: dengxiaojuan218@163.com
*Corresponding author

Abstract: This paper proposes a prediction model based on a spatio-temporal graph convolutional network to capture the spatio-temporal dependencies of user interactions and improve the accuracy of predicting trends in the dissemination of ideological and political themes. Experimental results show that this method achieves an average improvement of 6.8% in the F1 score and a reduction of 7.2% in the prediction root mean square error for key metrics such as the probability of ideological and political hot topics appearing in the coming week and changes in regional sentiment distribution. This model effectively integrates the spatial structural information of social networks with dynamic temporal features, providing a reliable computational tool for quantitative analysis and forward-looking assessment of ideological and political dynamics in social media environments. These aid relevant departments in timely sensing and guiding the online ideological and political ecosystem.

Keywords: spatio-temporal convolutional networks; social media platforms; ideological and political communication.

Reference to this paper should be made as follows: Liu, W. and Deng, X. (2025) 'Spatio-temporal convolutional networks empowering political ideology trend prediction on social media platforms', *Int. J. Information and Communication Technology*, Vol. 26, No. 44, pp.39–57.

Biographical notes: Wencheng Liu obtained her Doctoral degree from Hunan University, China in 2024. Currently, she is a Lecturer at the School of Marxism, Hunan Open University. Her research interests include big data management, ideological and political education.

Xiaojuan Deng obtained her Master's degree from Xiamen University, China in 2010. Currently, she is an Associate Professor at the School of Economics and Management, Shaoyang University. Her research interests include big data management, ideological and political education.

1 Introduction

The rapid development of social media platforms has brought unprecedented opportunities and challenges to ideological and political education work. As an important arena for information dissemination and the exchange of viewpoints, the dissemination of ideological and political content on social media platforms exhibits characteristics of dynamism, complexity, and concealment intertwined with one another (Yu, 2022). The spatio-temporal network formed by the real-time interaction of a massive user base makes it difficult to accurately capture and predict the evolution of the popularity of ideological and political topics, the diffusion of emotional tendencies, and regional aggregation effects using traditional methods. Timely and accurate understanding of the evolution patterns of these ideological and political dynamics is of great practical significance for relevant departments to proactively perceive online public opinion trends, optimise the precise delivery of ideological and political education content, and effectively guide the dissemination of mainstream values. Therefore, developing computational models that integrate the spatio-temporal characteristics of social networks and achieve high-precision prediction of ideological and political dynamics has become a critical need in the field of intelligent governance of online ideological and political work (Hu and Li, 2018).

In recent years, research on using artificial intelligence technology to analyse social media data for predicting public opinion trends has made significant progress. In terms of time-series modelling, recurrent neural networks, long short-term memory networks, and their variants have been widely used for user behaviour time-series prediction and topic popularity trend analysis due to their strong sequence modelling capabilities. Mu et al. (2023) introduced an improved inertial weight and adaptive mutation operation to enhance the particle swarm optimisation algorithm; then, the improved PSO algorithm was used to optimise the parameters of the long short-term memory neural network; finally, an IPSO-LSTM hybrid prediction model was constructed to predict and analyse the propagation trends of sudden public opinion. However, such models often treat users as isolated individuals, ignoring the rich spatial structural information contained in social relationship networks. In terms of spatial dimension modelling, graph neural networks, particularly graph convolutional networks and graph attention networks, demonstrate superior performance in modelling social influence and opinion propagation among users by aggregating information from neighbouring nodes. Jain et al. (2023) proposed a graph neural network model specifically designed for opinion leader identification, which leverages the powerful capabilities of graph neural networks to classify opinion leaders and their influence on online social networks. However, most of these models handle static graph structures and struggle to effectively capture the dynamic interaction patterns that evolve over time in social networks, as well as the spatiotemporal coupling effects of ideological and political content dissemination that result from these patterns.

The development of spatio-temporal graph neural networks (STGNNs) has provided new solutions for integrating the spatio-temporal characteristics of social networks. These models have achieved significant success in fields such as traffic flow prediction and crowd movement prediction by combining graph convolutional networks with sequence models within a unified framework. However, applying them to the prediction of ideological and political dynamics on social platforms still presents unique challenges: first, the dissemination of ideological and political content is not only influenced by explicit social relationships but is also deeply rooted in users' implicit ideological associations and community structures; second, the evolution of ideological and political topics often involves complex social-emotional interactions, with their spatio-temporal diffusion patterns exhibiting strong nonlinearity and suddenness; third, the lack of large-scale, high-quality, publicly available annotated social network spatio-temporal datasets related to ideological and political content constrains model training and validation.

To address these challenges, this paper proposes an enhanced spatio-temporal graph convolutional network (enhanced ST-GCN) framework aimed at accurately predicting the dynamic evolution of ideological and political themes on social platforms. Our core contributions are:

1. designing a dual graph structure construction method that combines explicit social links with implicit semantic associations based on ideological and political content similarity, providing a more comprehensive representation of the spatial networks influencing ideological and political dissemination
2. introducing an adaptive temporal convolution module and gating mechanism to enhance the model's ability to capture nonlinear temporal patterns and sudden events in the dissemination of ideological and political topics
3. utilising a large-scale public Weibo dataset for model training and validation, filling the gap in high-quality benchmark datasets in this field.

2 Theoretical foundations of spatio-temporal convolutional networks

2.1 Graph structure data and graph convolution basics

Many complex systems in the real world, such as social networks, transportation networks, or biomolecular structures, can be naturally abstracted into graph-structured data (Dong et al., 2021). In graph representations, entities are abstracted as nodes, and the relationships between entities are represented as edges connecting the nodes. Graph data exhibits non-Euclidean properties, meaning there is no fixed grid-based arrangement order or uniform distance metric between nodes. This makes traditional convolutional neural networks, which rely on regular grid structures, inapplicable. Graph convolutional networks were developed to address this challenge. Their core idea is to extend the concept of convolution operations from regular grids to irregular graph structures, enabling effective feature learning on graph data (Meng et al., 2023).

The essence of graph convolution operations lies in aggregating the information of a node's neighbours to update its own representation. Based on the implementation path, they are primarily divided into spectral domain methods and spatial domain methods. Spectral domain methods are based on graph signal processing theory, utilising the mathematical representation of graphs (such as the Laplacian matrix) to define filter operations in the transform domain. This method has a solid mathematical foundation but typically involves complex computations (Amador et al., 2022). Spatial domain methods are more intuitive, directly defining convolution on the graph's topological structure: each node collects feature information from its first-order or higher-order neighbours via an aggregation function (e.g., averaging, maxing, or weighted summation), then updates its representation by combining this information with its own features. Spatial domain methods are more popular due to their flexibility, efficiency, and scalability to large-scale graphs. Variants such as graph attention networks (GAT) further introduce attention mechanisms, enabling the model to learn the importance weights of different neighbouring nodes relative to the central node, thereby achieving more refined information aggregation (Liu et al., 2021).

2.2 Challenges in spatio-temporal data modelling

Many dynamic systems not only contain complex spatial correlation structures, but their node attributes and edge relationships also evolve over time, forming spatiotemporal graph data. Modelling such data presents three intertwined core challenges: spatial dependency, temporal dependency, and spatiotemporal coupling (Sharafi et al., 2022).

Spatial dependency refers to the fact that the state or behaviour of a node is often influenced by the states of its neighbouring nodes and does not exist in isolation (Yu et al., 2022). Temporal dependency manifests as nodes exhibiting sequential characteristics, where the current state is closely related to both its own past states and those of its neighbours (Shi et al., 2023). Most critically, spatio-temporal coupling implies that spatial dependencies change over time, while temporal evolution patterns are strongly constrained by the spatial network structure. Traditional sequence models excel at capturing temporal sequence patterns but struggle to effectively incorporate and utilise spatial structural relationships between nodes. Static graph neural networks can effectively model spatial dependencies but cannot handle dynamic changes in graph structures or node features over time. Therefore, there is an urgent need for a new model framework that can uniformly capture complex spatio-temporal dependencies.

2.3 Core framework of spatio-temporal convolutional networks

The ST-GCN is a deep learning architecture designed to address the aforementioned challenges (Turhan et al., 2022). Its core concept is to synergistically integrate graph convolutional operations to capture spatial dependencies and sequence modelling operations to capture temporal dependencies within a unified model. A typical ST-GCN model consists of multiple stacked spatio-temporal convolutional modules, each responsible for extracting spatio-temporal features of different levels from the input data (Zhao et al., 2022).

A spatiotemporal graph typically includes the following key elements: a fixed set of nodes; a set of edges defining spatial connections between nodes; adjacency information describing the strength or existence of these connections; and most importantly, a node feature tensor that records the attribute or state vector of each node at every time point. This feature tensor serves as the foundational data for the model to learn spatiotemporal evolution patterns (Zhang et al., 2023a).

Spatial graph convolution is one of the core components of ST-GCN, whose role is to simulate how information propagates and aggregates within the spatial structure network of the graph at a specific time point or within a time segment. It draws inspiration from static graph convolution but applies it to each ‘time slice’ of the spatio-temporal graph (Lu et al., 2024). Specifically, for a given time point, the spatial convolution operation examines the direct neighbours of each node. It aggregates the feature information of neighbouring nodes through a predefined or learnable aggregation function, fuses it with the central node’s own features, and generates a new representation of the node at the current time point after spatial information enhancement. This process enables each node to perceive the state of its local network environment (Ye and Toyama, 2022).

After the spatial convolution provides node representations containing neighbour information for each time point, the temporal modelling component begins to function (Yao et al., 2024). Its task is to capture the evolutionary patterns and trends of each node's own feature sequence. Commonly used temporal modelling techniques primarily fall into three categories: first, one-dimensional convolutional neural networks, which use filters sliding along the time axis to efficiently capture local short-term patterns of node features; second, recurrent neural networks and their gated variants, which store historical information through internal states and are particularly adept at modelling long-range temporal dependencies and complex sequence dynamics; and third, temporal convolutional networks, which utilise specialised convolutional structures to effectively handle long-sequence dependencies while offering parallel computing advantages (Hedegaard et al., 2023). Regardless of the technique employed, the goal of temporal modelling is to understand how node states evolve and change over time.

Spatial convolution and temporal modelling are not performed in isolation but are carefully organised within spatio-temporal convolution blocks to achieve tight coupling (Ren et al., 2024). The workflow of a standard block typically follows two main strategies: the first is the 'space-first, time-second' strategy, where spatial graph convolutions are first applied to each independent time point within the input time interval to obtain a sequence of node representations rich in spatial relationships, followed by time modelling operations on this sequence to extract temporal evolution features (Zhang et al., 2021). The second is the 'time-first, space-second' strategy, which involves first performing smoothing or feature extraction on each node's individual time series, then applying spatial graph convolution to the resulting features to aggregate neighbour information. Regardless of the strategy, the output of each spatio-temporal convolution block is a comprehensive representation of each node at the current time, incorporating its historical state and the current/historical states of its neighbours. To improve model performance and stability, nonlinear activation functions and residual connection techniques are typically introduced within blocks (Zhang et al., 2023b).

Multiple spatio-temporal convolution blocks stacked together form the encoder of ST-GCN. The encoder processes the input historical spatio-temporal graph sequence data, extracting spatio-temporal features through layer-by-layer abstraction, and ultimately learns node or graph representations that highly summarise historical dynamics. For future prediction tasks, a decoder must be connected after the encoder (Andrade-Ambriz et al., 2022). The simplest decoder may be a linear mapping layer that directly maps the final encoding to the predicted value, suitable for simple short-term predictions. More commonly, a recursive prediction strategy is employed: the model uses the prediction output at the current time as part of the input for the next time step, iteratively generating multi-step future predictions. For more complex long-term predictions, the sequence-to-sequence architecture is often adopted, which uses an independent decoder network to gradually decode and generate the entire future prediction sequence based on the 'context vector' output by the encoder as the initial condition. The introduction of attention mechanisms can help the decoder focus more precisely on key segments in the historical sequence when generating future states (Ali et al., 2022).

ST-GCN provides a powerful theoretical foundation for solving the modelling and prediction challenges of spatiotemporal graph data by innovatively integrating the spatial information processing capabilities of graph convolutions with the temporal dynamic capture capabilities of sequence models within an end-to-end framework (Ma et al., 2023). Its core value lies in its ability to uniformly and collaboratively learn the complex spatiotemporal interaction patterns and evolutionary laws embedded in the data, overcoming the limitations of traditional methods that treat spatiotemporal dimensions in isolation. From the fundamental principles of graph convolutions to flexible temporal modelling technique selection, and from the coupling strategies of spatio-temporal modules to the design of prediction architectures, the ST-GCN framework demonstrates high adaptability and scalability, making it a mainstream and foundational method for numerous real-world tasks involving complex spatio-temporal dynamics, such as traffic flow prediction, human behaviour recognition, social network information propagation analysis, and weather forecasting (Luo et al., 2024).

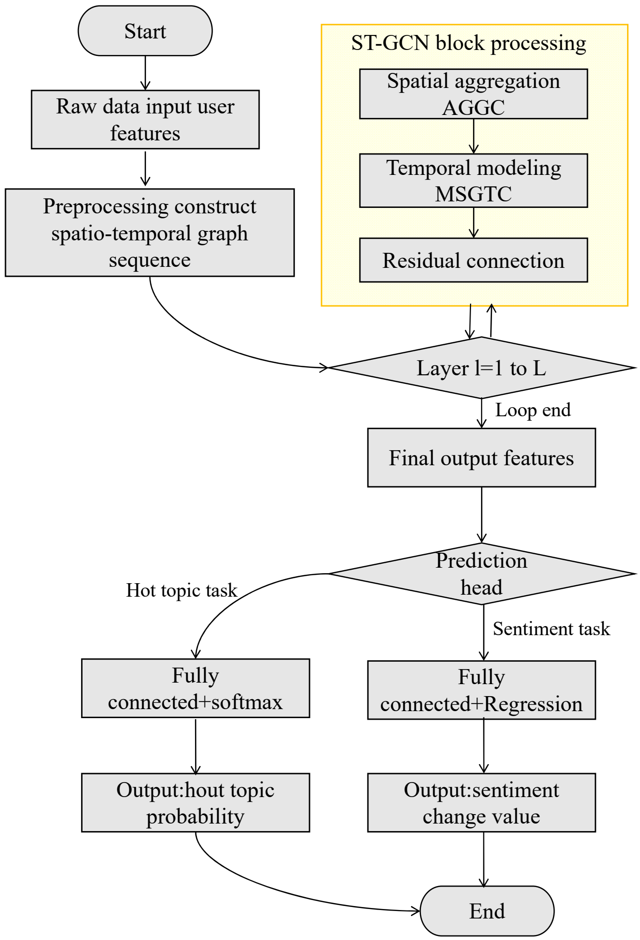
3 Enhanced ST-GCN model architecture

The core challenges faced in predicting ideological trends on social media platforms lie in effectively capturing the implicit ideological connections between users, addressing the highly nonlinear and sudden nature of topic evolution, and fully utilising limited high-quality labelled data (Weismueller et al., 2022). To address these issues, this chapter proposes an improved spatio-temporal graph convolutional network model – enhanced ST-GCN. As shown in Figure 1, the model takes user interaction data and ideological and political content data from social media platforms within a historical timeframe as input. Its core innovations are reflected in three aspects: first, by constructing a dual graph structure that integrates explicit social relationships and implicit semantic associations, it provides a more comprehensive representational foundation for the spatial network of ideological and political dissemination; second, it designs an adaptive gated graph convolution (AGGC) operator, significantly enhancing the flexibility of spatial information aggregation and the ability to perceive dynamic relationships; finally, it introduces a multi-scale gated temporal convolution (MSGTC) module specifically designed to capture multi-granularity temporal patterns and sudden event characteristics in the dissemination of ideological and political topics. The model's final output predicts the probability distribution of ideological and political hot topics and changes in user sentiment across regions within a specific future time window.

3.1 Construction of a space-time map

The dissemination mechanism of ideological and political content on social media platforms is complex, driven not only by explicit social connections between users but also by implicit semantic associations such as ideological similarity and overlapping topic interests (Chauhan et al., 2021). To comprehensively characterise this spatial network, enhanced ST-GCN constructs a dual graph structure, which is composed of an explicit social graph and an implicit semantic association graph.

Figure 1 Enhanced ST-GCN flowchart (see online version for colours)



The explicit social graph directly reflects the observable social interactions between users. The node set V represents all users, with the total number of users denoted as N, hence $|V| = N$. The edge set E^{sp} is defined based on the explicit social relationships between users. This connection relationship is quantified using an adjacency matrix $A^{sp} \in \mathbb{R}^{N \times N}$, where the value of matrix element A_{uv}^{sp} is defined as: if there is an explicit social link between users u and v, then $A_{uv}^{sp} = 1$; if not, then $A_{uv}^{sp} = 0$. In the initial stage, we use this binary adjacency matrix for simplified representation.

The implicit semantic correlation graph (ISC graph) aims to capture the underlying similarities between users based on ideological and political content (Cann et al., 2021). Its node set V is the same as that of the explicit social graph. The existence and strength of edge $e_{uv}^{se} \in E^{se}$ depend on the semantic similarity between users u and v in terms of ideological and political content. The process of calculating this similarity is as follows: first, extract all text content of posts labelled as ideological and political-related that user u has posted within the historical time window T_{hist}; second, use a pre-trained language model to encode these posts, obtaining the embedding vector for each post; next, by averaging all the ideological and political post embedding vectors of user u or applying importance-based attention aggregation, the final ideological and political semantic representation vector $s_u \in \mathbb{R}^d$ of the user is generated, where d is the dimension of the semantic vector. The semantic similarity sim(u, v) between users u and v is obtained by calculating the cosine similarity between their semantic representation vectors s_u and s_v:

$$sim(u,v) = \frac{s_u \cdot s_v}{\|s_u\|2\|s_v\|2} \tag{1}$$

In order to construct a sparse and more meaningful semantic association graph, we set a threshold τ and introduce an indicator function $\mathbb{I}$. Finally, the elements of the adjacency matrix $A^{se} \in \mathbb{R}^{N\times N}$ of the implicit semantic association graph are defined as:

$$A_{uv}^{se} = I(sim(u,v) > \tau) \cdot sim(u,v) \tag{2}$$

where $\mathbb{I}(\text{sim}(u,v) > \tau)$ is set to 1 at time sim(u, v) > τ, and 0 otherwise, meaning that only edges with similarity scores above the threshold τ are retained, and their weights are set to the actual similarity values.

The final fused spatiotemporal graph treats users as nodes, and the features of a node at time step t typically $x_u^{(t)} \in \mathbb{R}^D$ include key dynamic indicators such as the emotional orientation value and topic popularity value of the ideological and political content posted or interacted with by the user at that moment. The spatial structure is jointly defined by the explicit social adjacency matrix $x_u^{(t)} \in \mathbb{R}^D$ and the implicit semantic adjacency matrix A^{se}. In the model processing, these two graph structures will work together in the spatial information aggregation process.

3.2 Adaptive gate-controlled convolution operator

Standard graph convolution operations often assume that adjacency relationships are static and uniformly weighted when processing social networks, which makes it difficult to adapt to the dynamic changes in user interaction intensity and the complexity of implicit associations in the dissemination of ideological and political topics (Scheffauer et al., 2021). To address this issue, we designed an adaptive gated graph convolution operator (AGGC), which integrates attention mechanisms and gating mechanisms to dynamically learn the importance weights of neighbouring nodes and control information flow.

AGGC operates on the input spatial graph and node feature matrix $X^{(t)} \in \mathbb{R}^{N\times D_{in}}$ (where D_{in} is the input feature dimension). For the target node u and its neighbouring nodes $v \in N(u)$ (including u itself), AGGC first computes an adaptive attention weight α_{uv}:

$$e_{uv} = LeakyReLU\left(a^T\left[Wx_u^{(t)} \oplus Wx_v^{(t)}\right]\right) \tag{3}$$

$$\alpha_{uv} = \frac{\exp(e_{uv})}{\sum_{k\in N(u)}\exp(e_{uk})} \tag{4}$$

where $W \in \mathbb{R}^{D_{out} \times D_{in}}$ is a shared learnable weight matrix used for linear transformation of node features; $a \in \mathbb{R}^{2D_{out}}$ is a learnable attention vector; $\oplus$ represents vector concatenation; LeakyReLU is a linear rectification activation function with leakage. α_{uv} quantifies the importance of neighbour v's information for the ideological and political dynamics prediction of central node u at the current time t.

Next, AGGC dynamically adjusts the contribution of aggregated information through a gating mechanism. It calculates an information control gate g_u:

$$g_u = \sigma\left(U_g x_u^{(t)} + b_g\right) \tag{5}$$

where $\sigma(\cdot)$ is the sigmoid activation function, $U_g \in \mathbb{R}^{D_{gate} \times D_{in}}$ and $b_g \in \mathbb{R}^{D_{gate}}$ are learnable parameters, and D_{gate} is the dimension of the gate vector. The values of gate vector D_{gate} range from 0 to 1, and each dimension controls the information throughput of the corresponding feature channel. Finally, the spatial feature representation $hu^{(sp)} \in \mathbb{R}^{Dout}$ of node u after AGGC update is calculated as follows:

$$h_u^{(sp)} = g_u \odot \left(\sum\nolimits_{v \in N(u)} \alpha_{uv} W x_v^{(t)}\right) + \left(1 - g_u\right) \odot W x_u^{(t)} \tag{6}$$

where $\odot$ denotes element-wise multiplication. The equation means: the transformed features $Wx_v^{(t)}$ of neighbour v are weighted and aggregated using attention weights α_{uv}, then the aggregated information is selectively absorbed using a gate vector g_u, while retaining a portion of node u's own features $Wx_u^{(t)}$. This design endows the model with the ability to dynamically adjust the information flow during spatial aggregation, making it particularly uitable for capturing sudden shifts in user attention or dynamic changes in the influence of key opinion leaders in the dissemination of ideological and political topics.

In enhanced ST-GCN, AGGC is applied in parallel to the explicit social graph G^{sp} and the implicit semantic graph G^{se}:

$$h_u^{(sp,exp)} = AGGC\left(A^{sp}, X^{(t)}, u\right) \tag{7}$$

$$h_u^{(sp,imp)} = AGGC\left(A^{se}, X^{(t)}, u\right) \tag{8}$$

Then, the representations from the two spatial views are fused to generate the final spatial enhanced feature $h_u^{(sp)} \in \mathbb{R}^{Dout}$ of node u:

$$h_u^{(sp)} = ReLU\left(W_{fuse}\left[h_u^{(sp,exp)} \oplus h_u^{(sp,imp)}\right] + b_{fuse}\right) \tag{9}$$

where $W_{fuse} \in \mathbb{R}^{D_{out} \times 2D_{out}}$ and $b_{fuse} \in \mathbb{R}^{D_{out}}$ are learnable parameters of the fusion layer, and ReLU is the activation function.

3.3 Multi-scale gate-controlled time convolution module

The dissemination and evolution of ideological and political topics typically involve multiple temporal scales: long-term trends (such as the propagation of mainstream values guided by policy), medium-term cycles (such as the phased popularity of specific topics), and short-term fluctuations and sudden events (such as the instantaneous outbreak of public opinion triggered by hot news) (Tokita et al., 2021). To effectively capture this multi-scale temporal dependency and handle sudden events, we designed a MSGTC.

The input to MSGTC is a sequence of spatially enhanced features for a node u within a historical time window $[t-K+1, t]$: $H_u^{(sp)} = [h_u^{(sp,t-K+1)}, h_u^{(sp,t-K+2)}, \ldots, h_u^{(sp,t)}] \in \mathbb{R}^{K\times D_{out}}$. The core of this module is the parallel application of multiple gated temporal convolution (GTC) branches with different receptive fields, each focusing on a specific time scale.

The structure of a single GTC branch is as follows: first, a one-dimensional causal convolution (causal convolution) is performed on the input sequence $H_u^{(sp)}$, ensuring that the output at the current time step t depends only on the inputs up to and including t, in accordance with the causal requirements of time series prediction. The convolution kernel size is k_s (where s represents the scale index of the branch), and the convolution kernel weights are $W_{conv}^{(s)} \in \mathbb{R}^{k_s\times D_{out}\times D_{hidden}}$. The convolution operation yields an initial feature map $Z^{(s)} \in \mathbb{R}^{K\times D_{hidden}}$. Next, GTC introduces a gating mechanism implemented by two parallel convolutional layers:

$$A^{(s)} = \tanh\left(W_{gateA}^{(s)} * H_u^{(sp)}\right) \tag{10}$$

$$B^{(s)} = \sigma\left(W_{gateB}^{(s)} * H_u^{(sp)}\right) \tag{11}$$

where $*$ denotes a one-dimensional causal convolution operation, tanh is a hyperbolic tangent activation function, and σ is a sigmoid activation function. Ultimately, the output $o_u^{(s,t)} \in \mathbb{R}^{D_{hidden}}$ of this branch at time step t is given by the gated modulated features:

$$o_u^{(s,t)} = A^{(s)}[t] \odot B^{(s)}[t] \tag{12}$$

where $A^{(s)}[t]$ and $B^{(s)}[t]$ represent slices of $A^{(s)}$ and $B^{(s)}$ at time t, respectively. Gate $B^{(s)}$ controls which temporal pattern features are activated or suppressed, which is particularly important for filtering noise and enhancing key event signals (such as sudden hotspots).

MSGTC parallels the deployment of S GTC branches with different convolution kernel sizes $k_1, k_2, \ldots, k_S$. Each branch independently processes the input sequence $H_u^{(sp)}$ and produces its corresponding scale output $o_u^{(s,t)}$. These multi-scale outputs are then concatenated:

$$o_u^{(t)} = o_u^{(1,t)} \oplus o_u^{(2,t)} \oplus \cdots \oplus o_u^{(S,t)} \tag{13}$$

Finally, a linear projection layer generates the spatiotemporal fusion feature $z_u^{(t)} \in \mathbb{R}^{D_{st}}$ of node u at time t:

$$z_u^{(t)} = W_{proj} o_u^{(t)} + b_{proj} \tag{14}$$

where $W_{proj} \in \mathbb{R}^{D_{st}\times(S\cdot D_{hidden})}$ and $b_{proj} \in \mathbb{R}^{D_{st}}$ are learnable parameters. $z_u^{(t)}$ is the output of the enhanced ST-GCN spatio-temporal block at the current time step t, which integrates information refined through dual graph spatial aggregation and multi-scale temporal convolution.

3.4 Stacking spatio-temporal blocks and predicting outputs

The enhanced ST-GCN model is composed of L layers of identical spatio-temporal blocks stacked on top of each other. The input to the l^{th} layer block is the output feature sequence from the previous layer $Z^{(l-1)} = [z^{(t-K+1)}, \ldots, z^{(t)}] \in \mathbb{R}^{K \times N \times D_{st}^{(l-1)}}$. The processing flow for each layer block is consistent:

1 Spatial graph convolution: apply the AGGC operator described in Section 3.2 to the feature slice $Z^{(l-1,t')} \in \mathbb{R}^{N \times D_{st}^{(l-1)}}$ at each time step $t' \in [t - K + 1, t]$ in the input sequence, outputting the spatially enhanced features $H^{(sp,l,t')}$ at that time step.

$$H^{(sp,l,t')} = AGGC^{(l)}\left(A^{sp}, A^{se}, Z^{(l-1,t')}\right) \tag{15}$$

2 For each node u, input its spatial enhanced feature sequence $[h_u^{(sp,l,t-K+1)}, \ldots, h_u^{(sp,l,t)}]$ in time window $[t - K + 1, t]$ into the MSGTC module described in Section 3.3, and output the spatio-temporal fusion feature $z_u^{(l,t)}$ of that node at the current time t.

$$z_u^{(l,t)} = MSGTC^{(l)}\left(\left[h_u^{(sp,l,t-K+1)}, \ldots, h_u^{(sp,l,t)}\right]\right) \tag{16}$$

3 Residual connections and layer outputs: residual connections are introduced to alleviate the vanishing gradient problem in deep network training and preserve low-level features. If the input feature dimension $D_{st}^{(l-1)}$ and output dimension $D_{st}^{(l)}$ of layer l are the same, then input $z_u^{(l-1,t)}$ is directly added to output $z_u^{(l,t)}$, followed by an activation function:

$$z_u^{(l,t)} = ReLU\left(z_u^{(l,t)} + z_u^{(l-1,t)}\right) \tag{17}$$

If the dimensions are different, the input dimensions need to be adjusted through a linear projection $W_{res}^{(l)} \in \mathbb{R}^{D_{st}^{(l)} \times D_{st}^{(l-1)}}$:

$$z_u^{(l,t)} = ReLU\left(z_u^{(l,t)} + W_{res}^{(l)} z_u^{(l-1,t)}\right) \tag{18}$$

All nodes $z_u^{(l,t)}$ form the output feature matrix $z^{(l,t)} \in \mathbb{R}^{N \times D_{st}^{(l)}}$ of layer l at time t.

After layer-by-layer abstraction of L layers of spatiotemporal blocks, the model ultimately obtains high-level spatiotemporal feature representations $z^{(L,t)} \in \mathbb{R}^{N \times D_{st}^{(L)}}$. For prediction tasks:

Prediction of the probability of hot topics in political ideology appearing: this is a node-level classification task (predicting the probability of each node's region/group appearing in hot topics in political ideology in the future). Input $z_u^{(L,t)}$ into a fully connected layer + softmax:

$$\hat{y}_u^{topic} = Softmax\left(W_{topic} z_u^{(L,t)} + b_{topic}\right) \tag{19}$$

where $\hat{y}_u^{topic}$ is a probability vector.

Prediction of regional sentiment trend changes: this is a regional-level regression task. First, the features of all nodes $u \in R_r$ belonging to the same region r are averaged and pooled to obtain the regional representation $z_r^{(t)} = \frac{1}{|R_r|} \sum u \in R_r z_u^{(L,t)}$. Then, through a regression layer:

$$\Delta \hat{s}^r = w_{sent}^T z_r^{(t)} + b_{sent} \tag{20}$$

where $\Delta \hat{s}_r$ is the predicted emotional tendency change value of region *r*.

The total loss function *L* of the model is the weighted sum of the two prediction tasks:

$$L = \lambda L_{topic} + (1 - \lambda) L_{sent} \tag{21}$$

where L_{topic} is the cross-entropy loss for hot topic prediction, L_{sent} is the mean squared error loss for sentiment trend prediction, and λ is the balancing hyperparameter. By optimising *L*, the entire enhanced ST-GCN model can be trained end-to-end.

4 Experiments and analysis of results

4.1 Experimental setup

To comprehensively evaluate the performance of the enhanced ST-GCN model in the task of predicting ideological and political trends on social media platforms, we designed and implemented a series of rigorous experiments. This section will provide a detailed overview of the datasets used in the experiments, the evaluation metrics, the baseline models for comparison, and the specific experimental implementation details.

The dataset used in the experiments is sourced from Sina Weibo's public dataset. This dataset has undergone strict anonymisation processing and includes manually annotated ideological and political relevance labels. We selected approximately 1.2 million original posts labelled as highly relevant to ideological and political themes and crawled the follow relationship network among the users who posted these posts, forming a social graph containing 10,000 active user nodes. The daily features of each user node over a 92-day period include the average sentiment orientation of all ideology-related posts published by that user and the topic popularity value for that day. We divided users into different regional groups based on the geographical location information in their public profiles, for a total of 50 regions.

Two key metrics were used to evaluate model performance. For the prediction of the emergence of ideological and political hot topics, we used the macro F1 score, which balances the accuracy and recall rates of each topic category, avoiding evaluation biases caused by category imbalance. For the prediction of regional sentiment changes, we used the root mean square error.

The selected baseline models encompass the current mainstream methods for time series prediction, graph networks, and spatio-temporal prediction:

1 LSTM: the classic long short-term memory network, which models only time series and ignores spatial relationships

2 GCN-LSTM: first uses a graph convolutional network to aggregate spatial information, then inputs each node's sequence into an LSTM for temporal modelling, forming a classic sequential spatio-temporal model

3 ST-GCN: a classic spatio-temporal graph convolutional network that uses Chebyshev graph convolutions and temporal one-dimensional convolutions for joint modelling

4 GraphWaveNet: an advanced spatio-temporal forecasting model that combines adaptive adjacency matrices and expanded causal convolutions, offering strong expressive capabilities.

Our enhanced ST-GCN model employs the Adam optimiser with an initial learning rate of 0.001 and applies a learning rate decay strategy. Model hyperparameters are determined via grid search on the validation set. All experiments were conducted on a server equipped with an NVIDIA GeForce RTX 3090 GPU. Each model was run five times to obtain the average results, ensuring the stability and reliability of the experiments.

4.2 Overall performance comparison analysis

To validate the overall effectiveness of the enhanced ST-GCN model, we compared its performance with all of the above baseline models on the same test set. As shown in Table 1, the results clearly demonstrate the quantitative performance of each model in the two prediction tasks.

Table 1 Overall performance comparison of various models in the task of predicting ideological and political trends

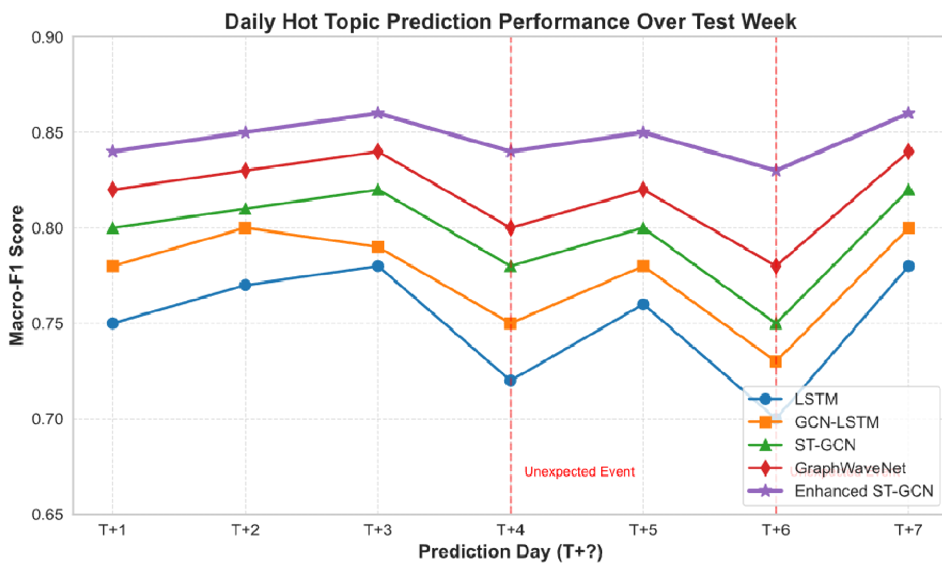
Model	*Hot topic predictions*	*Predicting emotional changes*
LSTM	0.781	0.152
GCN-LSTM	0.798	0.146
ST-GCN	0.812	0.141
GraphWaveNet	0.826	0.138
Enhanced ST-GCN	0.852	0.131

The following conclusions can be drawn from the data in Table 1. First, the performance of LSTM, which uses only a temporal model, is relatively the lowest, strongly demonstrating that incorporating spatial social relationships between users is crucial for predicting ideological and political dynamics. Second, both GCN-LSTM and ST-GCN outperform LSTM, indicating the effectiveness of jointly modelling spatio-temporal dependencies. Among these, ST-GCN outperforms GCN-LSTM with its more tightly coupled spatiotemporal design. GraphWaveNet achieves superior performance compared to traditional ST-GCN due to its adaptive adjacency matrix and powerful temporal convolution module, demonstrating the necessity of handling complex spatiotemporal relationships.

Ultimately, our proposed enhanced ST-GCN model achieved optimal performance on both tasks. In the hot topic prediction task, its macro-F1-score reached 0.852, an improvement of 3.1% over the strongest baseline model, GraphWaveNet. In the sentiment trend prediction task, its RMSE was reduced to 0.131, an improvement of 5.1% over GraphWaveNet. This significant performance improvement validates the effectiveness of the model's core innovations: the dual graph structure that integrates implicit semantic associations, the adaptive gated graph convolution operator, and the multi-scale gated temporal convolution module work together to more accurately capture the complex spatio-temporal patterns of ideological and political dynamics.

To more intuitively demonstrate the model's performance stability throughout the testing period, we plotted the daily F1-score changes of each model over a 7-day testing window, as shown in Figure 2. As can be observed from the figure, the performance curve of enhanced ST-GCN consistently outperforms other baseline models and exhibits the highest stability over time, with minimal fluctuations. Especially on days $T + 4$ and $T + 6$, when a sudden news event caused minor fluctuations in public opinion, enhanced ST-GCN demonstrated stronger robustness, with a much smaller performance decline than other models, thanks to its MSGTC module's effective capture of sudden time patterns.

Figure 2 Daily hotspot prediction F1-score change curves for each model during the test week (see online version for colours)



4.3 *Melting experiments and model analysis*

To further analyse the specific contributions of each core component in the enhanced ST-GCN model, we designed a comprehensive ablation experiment. We sequentially removed key modules from the model, constructing four variant models:

1 w/o semantic graph: removing the implicit semantic association graph and using only the explicit social graph

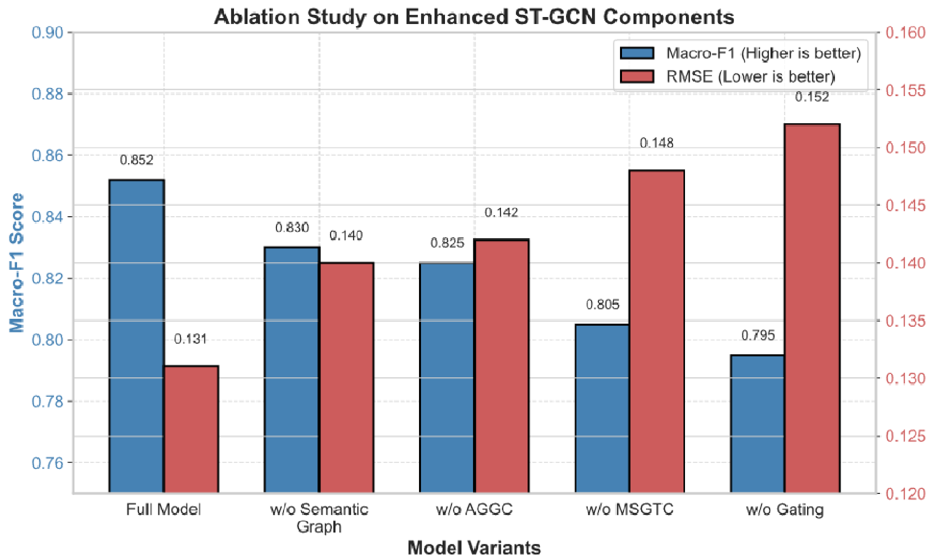
2 w/o AGGC: replacing the adaptive gated graph convolutional layer with a standard static graph convolutional layer

3 w/o MSGTC: replacing the multi-scale gated temporal convolutional module with a standard one-dimensional temporal convolutional layer

4 w/o gating: removing all gating mechanisms from AGGC and MSGTC.

The results of the ablation experiments are presented in a bar chart in Figure 3.

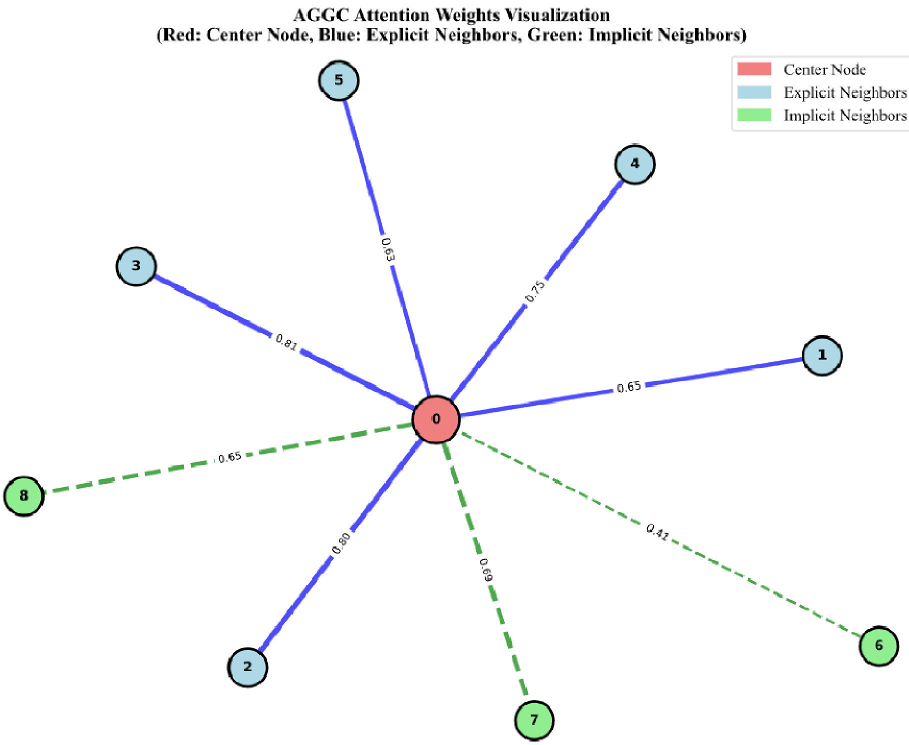
Figure 3 Bar chart of ablation experiment results for the enhanced ST-GCN model (see online version for colours)



The experimental results clearly demonstrate that each component is indispensable. First, removing the implicit semantic association graph leads to a significant decline in performance, which fully demonstrates that explicit social relationships and implicit semantic associations are complementary in ideological and political communication, and relying on only one type of relationship cannot fully characterise the spatial path of information dissemination. Second, removing the adaptive attention and gating mechanism from AGGC also resulted in a significant decline in model performance, indicating that the ability to dynamically adjust neighbour weights and control information flow is crucial for handling the complexity of user interactions. Third, replacing MSGTC with a simple temporal convolution resulted in the greatest performance loss, particularly in the RMSE metric, strongly validating our core hypothesis: the evolution of ideological and political topics involves multi-scale temporal patterns and sudden changes, necessitating the use of specially designed temporal modules with different receptive fields and gating mechanisms to effectively capture these phenomena. Finally, the variant that removes all gating mechanisms performs the worst, even below some baseline models, highlighting the central role of gating mechanisms in the entire model. They are not only critical to AGGC but also essential for MSGTC to effectively filter out noise and focus on key events.

To gain a deeper understanding of the working mechanism of the AGGC module, we randomly selected a hot political education event from the test set regarding the announcement of a major policy and visualised the attention weights obtained by a core user node and its neighbours after AGGC computation, as shown in Figure 4. As shown in the figure, the model not only assigns high attention weights to users with explicit strong connections (such as close friends) but, more importantly, also assigns relatively high weights to users with weak or no explicit connections but high semantic relevance (such as strangers who follow the same topic).

Figure 4 AGGC attention weight visualisation example (see online version for colours)



Based on the above experimental results, we can draw the following conclusions: the enhanced ST-GCN model demonstrates outstanding and stable performance in the task of predicting ideological and political dynamics on social media platforms through its innovative architectural design. Its success is primarily attributed to a deep understanding of the unique nature of ideological and political communication and targeted modelling: the dual graph structure addresses the complexity and implicit relationships of the communication space; the adaptive gated graph convolution enhances the model's ability to perceive and adapt to spatial dependencies in dynamic changes; while multi-scale gated temporal convolutions are specifically designed to capture the coexisting long-term trends, medium-term cycles, and short-term sudden events in topic evolution.

However, this study still has certain limitations. First, the model's performance depends to some extent on the quality of the implicit semantic association graph construction, which in turn depends on the choice of pre-trained language models and similarity thresholds. Future work could explore dynamic end-to-end learning of semantic associations. Second, the model currently focuses primarily on text content; future work could incorporate multimodal information for more comprehensive perception of ideological and political dynamics. Finally, experiments were conducted solely on the Weibo dataset, and the model's generalisability requires further validation on more diverse social media platform datasets.

Despite these limitations, the experimental results demonstrate that enhanced ST-GCN provides a powerful tool for precise prediction and in-depth analysis of ideological and political dynamics on social media platforms, aiding relevant authorities in better understanding public opinion trends and achieving precise guidance and governance.

5 Conclusions

This paper focuses on the cutting-edge research topic of predicting ideological and political trends on social media platforms, which has significant practical implications. Addressing the unique challenges posed by the complex implicit relationships, nonlinear evolution, and high volatility in the dissemination of ideological and political topics, we propose a novel enhanced spatio-temporal graph convolutional network model (Enhanced ST-GCN).

The main contributions of this study are threefold: first, in terms of spatial modelling, we innovatively constructed a dual graph structure that integrates explicit social relationships and implicit semantic associations, overcoming the limitations of traditional methods that rely solely on single explicit relationships. This provides a more comprehensive data foundation for depicting the complex spatial paths of ideological and political education dissemination. Second, in terms of graph convolution operator design, an adaptive gated graph convolution is proposed, which integrates attention mechanisms and gated units to achieve precise perception and flexible control of dynamically changing spatial dependencies and information flows. Finally, in terms of temporal dynamic modelling, a multi-scale gated temporal convolution module was designed, utilising a parallel multi-branch structure and gating mechanism to effectively capture multi-scale temporal patterns such as long-term trends, medium-term cycles, and short-term sudden fluctuations in the dissemination of ideological and political topics.

In summary, the enhanced ST-GCN model proposed in this paper successfully achieves more accurate and stable predictions of ideological and political dynamics on social media platforms through its systematic innovative design. This not only provides an advanced spatio-temporal data mining tool for online public opinion analysis but also provides reliable decision support and technical basis for relevant departments to dynamically perceive ideological and political trends, anticipate potential risks, and achieve precise guidance and intelligent governance.

Looking ahead, this study still has room for expansion, such as exploring end-to-end implicit relationship graph learning mechanisms, integrating multi-modal information for ideological and political content understanding, and validating the model's universality in more diversified social media platform scenarios. These will be the key research directions for the next phase.

Declarations

This work is supported by the Research Project on Ideological and Political Education in Hunan Provincial Universities in 2025 (No. 25A97), the Key Project of Teaching Reform Research in Hunan Provincial Ordinary Undergraduate Colleges and Universities in 2024 (No. 202401000474), the Teaching Reform Research Project in Hunan Provincial Vocational Colleges and Universities in 2023 (No. ZJGB2023239), and the Comprehensive Development Studio for College Counselors in Hunan Provincial Universities (No. 202302).

All authors declare that they have no conflicts of interest.

References

Ali, A., Zhu, Y. and Zakarya, M. (2022) 'Exploiting dynamic spatio-temporal graph convolutional neural networks for citywide traffic flows prediction', *Neural networks*, Vol. 145, No. 1, pp.233–247.

Amador, K., Wilms, M., Winder, A., Fiehler, J. and Forkert, N.D. (2022) 'Predicting treatment-specific lesion outcomes in acute ischemic stroke from 4D CT perfusion imaging using spatio-temporal convolutional neural networks', *Medical Image Analysis*, Vol. 82, No. 1, p.102610.

Andrade-Ambriz, Y.A., Ledesma, S., Ibarra-Manzano, M-A., Oros-Flores, M.I. and Almanza-Ojeda, D-L. (2022) 'Human activity recognition using temporal convolutional neural network architecture', *Expert Systems with Applications*, Vol. 191, No. 1, p.116287.

Cann, T.J., Weaver, I.S. and Williams, H.T. (2021) 'Ideological biases in social sharing of online information about climate change', *PloS One*, Vol. 16, No. 4, p.e0250656.

Chauhan, P., Sharma, N. and Sikka, G. (2021) 'The emergence of social media data and sentiment analysis in election prediction', *Journal of Ambient Intelligence and Humanized Computing*, Vol. 12, No. 2, pp.2601–2627.

Dong, X., Sun, Y., Li, Y., Wang, X. and Pu, T. (2021) 'Spatio-temporal convolutional network based power forecasting of multiple wind farms', *Journal of Modern Power Systems and Clean Energy*, Vol. 10, No. 2, pp.388–398.

Hedegaard, L., Heidari, N. and Iosifidis, A. (2023) 'Continual spatio-temporal graph convolutional networks', *Pattern Recognition*, Vol. 140, No. 1, p.109528.

Hu, Z. and Li, J. (2018) 'Innovative methods for ideological and political education of college students', *Educational Sciences: Theory & Practice*, Vol. 18, No. 5, p.14612.

Jain, L., Katarya, R. and Sachdeva, S. (2023) 'Opinion leaders for information diffusion using graph neural network in online social networks', *ACM Transactions on the Web*, Vol. 17, No. 2, pp.1–37.

Liu, S., Wang, X., Zhao, L., Li, B., Hu, W., Yu, J. and Zhang, Y-D. (2021) '3DCANN: a spatio-temporal convolution attention neural network for EEG emotion recognition', *IEEE Journal of Biomedical and Health Informatics*, Vol. 26, No. 11, pp.5321–5331.

Lu, Z., Luo, Y., Xu, L., Hu, Y., Zheng, X. and Sun, L. (2024) 'Channel spatio-temporal convolutional network for pedestrian trajectory prediction', *International Journal of Machine Learning and Cybernetics*, Vol. 15, No. 11, pp.5395–5413.

Luo, Y., Zheng, J., Wang, X., Tao, Y. and Jiang, X. (2024) 'GT-LSTM: a spatio-temporal ensemble network for traffic flow prediction', *Neural Networks*, Vol. 171, No. 1, pp.251–262.

Ma, Q., Sun, W., Gao, J., Ma, P. and Shi, M. (2023) 'Spatio-temporal adaptive graph convolutional networks for traffic flow forecasting', *IET Intelligent Transport Systems*, Vol. 17, No. 4, pp.691–703.

Meng, F., Li, J., Zhang, Y., Qi, S. and Tang, Y. (2023) 'Transforming unmanned pineapple picking with spatio-temporal convolutional neural networks', *Computers and Electronics in Agriculture*, Vol. 214, No. 1, p.108298.

Mu, G., Liao, Z., Li, J., Qin, N. and Yang, Z. (2023) 'IPSO-LSTM hybrid model for predicting online public opinion trends in emergencies', *PloS One*, Vol. 18, No. 10, p.e0292677.

Ren, X., Zhang, F., Sun, Y. and Liu, Y. (2024) 'A novel dual-channel temporal convolutional network for photovoltaic power forecasting', *Energies*, Vol. 17, No. 3, p.698.

Scheffauer, R., Goyanes, M. and Gil de Zuniga, H. (2021) 'Beyond social media news use algorithms: how political discussion and network heterogeneity clarify incidental news exposure', *Online Information Review*, Vol. 45, No. 3, pp.633–650.

Sharafi, M., Yazdchi, M., Rasti, R. and Nasimi, F. (2022) 'A novel spatio-temporal convolutional neural framework for multimodal emotion recognition', *Biomedical Signal Processing and Control*, Vol. 78, No. 1, p.103970.

Shi, X., Li, B., Wang, W., Qin, Y., Wang, H. and Wang, X. (2023) 'Classification algorithm for electroencephalogram-based motor imagery using hybrid neural network with spatio-temporal convolution and multi-head attention mechanism', *Neuroscience*, Vol. 527, No. 1, pp.64–73.

Tokita, C.K., Guess, A.M. and Tarnita, C.E. (2021) 'Polarized information ecosystems can reorganize social networks via information cascades', *Proceedings of the National Academy of Sciences*, Vol. 118, No. 50, p.e2102147118.

Turhan, G., Küçük, H. and Isik, E.O. (2022) 'Spatio-temporal convolution for classification of alzheimer disease and mild cognitive impairment', *Computer Methods and Programs in Biomedicine*, Vol. 221, p.106825.

Weismueller, J., Harrigan, P., Coussement, K. and Tessitore, T. (2022) 'What makes people share political content on social media? The role of emotion, authority and ideology', *Computers in Human Behavior*, Vol. 129, No. 1, p.107150.

Yao, Y., Chen, L., Wang, X. and Wu, X. (2024) 'Research on traffic flow forecasting of spatio-temporal convolutional networks with auto-correlation', *International Journal of Automotive Technology*, Vol. 1, No. 1, pp.1–10.

Ye, J. and Toyama, N. (2022) 'Automatic defect detection for ultrasonic wave propagation imaging method using spatio-temporal convolution neural networks', *Structural Health Monitoring*, Vol. 21, No. 6, pp.2750–2767.

Yu, C., Gu, Z., Li, D., Wang, G., Wang, A. and Li, E. (2022) 'STSC-SNN: spatio-temporal synaptic connection with temporal convolution and attention for spiking neural networks', *Frontiers in Neuroscience*, Vol. 16, No. 1, p.1079357.

Yu, Y. (2022) 'On the ideological and political education of college students in the new media era', *Open Journal of Social Sciences*, Vol. 10, No. 1, pp.1–14.

Zhang, K., He, F., Zhang, Z., Lin, X. and Li, M. (2021) 'Graph attention temporal convolutional network for traffic speed forecasting on road networks', *Transportmetrica B: Transport Dynamics*, Vol. 9, No. 1, pp.153–171.

Zhang, X., Liu, J., Gong, P., Chen, C., Han, B. and Wu, Z. (2023a) 'Trajectory prediction of seagoing ships in dynamic traffic scenes via a gated spatio-temporal graph aggregation network', *Ocean Engineering*, Vol. 287, No. 1, p.115886.

Zhang, Z., Gong, S., Liu, Z. and Chen, D. (2023b) 'A novel hybrid framework based on temporal convolution network and transformer for network traffic prediction', *PloS One*, Vol. 18, No. 9, p.e0288935.

Zhao, P., Fan, Z., Cao, Z. and Li, X. (2022) 'Intrusion detection model using temporal convolutional network blend into attention mechanism', *International Journal of Information Security and Privacy (IJISP)*, Vol. 16, No. 1, pp.1–20.