Scientific Journal of the Ternopil National Technical University
2026, vol. 121, no. 1, pp. 99–106,
ISSN: 2522-4433 Print, ISSN: 2522-4441 Online
https://doi.org/10.33108/visnyk_tntu2026.01.099

Submitted: 2025-09-25 | Revised: 2026-02-11| Accepted: 2026-03-24

Modeling of topological relationships between trends of social platforms X and YouTube for short-term forecasting using Telegram interface

Tetiana Levytska, Liudmyla Kotykhova

Pryazovskyi State Technical University, Dnipro, Ukraine

* kotykhova_l_d@pstu.edu

Abstract: *The article presents the results of the development and research of an intelligent system designed for monitoring and forecasting the popularity dynamics of hashtags and topics across social media platforms (X/Twitter, YouTube). The relevance of this study is driven by the rapid update rate of information flows, which necessitates automated tools for identifying emerging trends in real-time. The scientific novelty of the work lies in the application of Graph Neural Networks (GNN) to analyze non-linear relationships between social media objects, allowing for the integration of both temporal dynamics and the topological structure of topic interactions. The paper provides a detailed description of the system's modular architecture. The data acquisition module is implemented through a hybrid approach: utilizing the official YouTube Data API for video content analysis and web scraping mechanisms (BeautifulSoup) to retrieve hourly Twitter trend data via the trends24.in platform. A relational SQLite database is employed to store structured information and ensure rapid access to time series data. Special attention is paid to the data preprocessing stage, which includes the normalization of popularity metrics and the formation of a graph structure where hashtags serve as nodes and edges represent their co-occurrence within a single context. The mathematical foundation of the system is based on the Graph Convolutional Network (GCN) architecture. The study justifies the selection of GCNConv layers, which implement a feature aggregation mechanism from neighboring nodes, and the Adam adaptive optimizer, ensuring efficient model training on sparse data. In the experimental section, a comparative analysis was conducted between the developed GNN model and classical approaches (Linear Regression, RNN), as well as alternative graph-based methods (GAT, Node2Vec). Testing results using MSE and RMSE metrics confirmed the superior predictive accuracy of the GNN in short-term popularity forecasting tasks. The practical significance of the work is validated by the implementation of a user interface in the form of a Telegram bot, providing analytics visualization via the Matplotlib library.*

Keywords: *chatbot, trend forecasting, social networks, graph neural networks, cloud technologies.*

1. INTRODUCTION

The relevance of the research topic is driven by the rapid digitalization of social communications, which has resulted in the formation of enormous volumes of data within social media. Modern social networks represent complex dynamic systems in which the analysis of information flows enables not only the study of current social states but also the identification of hidden patterns of interaction between groups and individuals. In this context, the task of forecasting the dynamics of trend popularity becomes critically important for strategic marketing, reputation management, and real-time monitoring of the information environment. However, the complexity of processing such data lies in their nonlinearity and high update rate, which necessitates the implementation of intelligent analytical methods capable of operating in real time.

The issue of deploying intelligent agents, particularly chatbots, for automating communication and user support is widely represented in contemporary scientific literature. Studies [1–5] confirm the effectiveness of conversational interfaces in e-commerce and business analytics, where they ensure personalized interaction and continuous customer support. Specific applications of such systems in domains such as the sports industry [6], healthcare [7], and education [8] demonstrate their versatility; however, most developments focus primarily on natural-language imitation, leaving the structural analysis of the underlying data largely unaddressed.

An analysis of architectural solutions and modern approaches to chatbot design presented in works [9–13] indicates a transition from simple script-based systems to advanced conversational agents built on machine-learning techniques. At the same time, the integration of specialized analytical modules for time-series forecasting based on graph-based models into chatbot architectures remains insufficiently explored. Although existing studies on IT trends in logistics [14] and the digitalization of marketing [15] highlight the importance of predictive models, they often overlook the topological dependencies between topics. Classical methods of social-network analysis [16] typically consider static or weakly dynamic structures, which does not fully correspond to the dynamics of modern media platforms such as X (Twitter) or YouTube.

In previous research, the authors examined the application of PageRank algorithms for assessing hashtag significance [17]; however, the present study shifts the focus toward the use of Graph Convolutional Networks (GCN) for modeling topological relationships across different platforms.

The scientific problem lies in the fact that traditional forecasting approaches, such as linear regression or recurrent neural networks (RNN), primarily analyze the historical sequence of popularity for an individual object. They do not account for contextual relationships – how the popularity of one topic influences the spread of another through shared hashtags or overlapping audiences. The use of Graph Neural Networks (GNN) makes it possible to overcome this limitation by treating the information space as a complex graph in which nodes (topics) interact with one another. This enables a significant improvement in the accuracy of predicting the emergence of trends at the early stages of their life cycle.

The aim of the study is to justify and implement an intelligent system based on graph neural networks that enhances the accuracy of short-term trend forecasting by incorporating contextual relationships between topics through shared hashtags, with real-time output delivered via a Telegram interface.

2. DEVELOPMENT OF THE MODEL AND CHATBOT

For the development of the chatbot, the Telegram platform was selected as the environment for implementing the interactive analytical system. Telegram is one of the most popular messaging platforms, with hundreds of millions of active users. Its interface is intuitive, and its functional capabilities allow users to quickly adapt to interactions with chatbots. Moreover, Telegram is supported across all major operating systems (Android, iOS, Windows, macOS, Linux), ensuring universal accessibility regardless of device type.

Python was chosen as the programming language for chatbot development, with PyCharm used as the primary environment for working with Python code and databases. SQLite3 is employed for storing processed data. This lightweight and easy-to-use database is well suited for local or small-scale projects.

At the data collection stage, web scraping of X (Twitter) trends and the YouTubeData API are used to obtain relevant information such as popular hashtags and trending topics. For X (Twitter), HTTP requests are sent to web pages and the HTML code is parsed using the requests and BeautifulSoup libraries [16], while YouTube data is retrieved through the official API, which returns results in JSON format.

The data-storage stage includes database initialization (creating the SQLite3 database and tables), data insertion (using SQL queries to add records), and data updating (checking for existing entries to avoid duplication).

Once the data is stored, analytical operations are performed, including:

– statistical analysis of hashtag frequency,
– visualization of hashtag popularity (plots, tag clouds),
– trend forecasting using machine-learning algorithms.

This approach enables efficient collection, processing, and storage of data from web sources. The use of SQLite3 ensures simplicity in data management, while Python integration allows the processing pipeline to be scaled for further analysis and trend prediction in social networks.

In this work, a Telegram bot was developed that uses Graph Neural Networks (GNN) to analyze trends stored in the database. After the data is saved, further processing and GNN training can be performed.

The graph data preparation function loads trend data from an SQLite database and generates a graph where nodes correspond to trends and edges are created based on trend co-occurrence, i.e. if they appear simultaneously in the selection in the same time interval, and the edge weight reflects the frequency of their co-occurrence. These data serve as input to the GNN:

– Nodes (x): one-hot encoded feature matrix,
– Edges (edge_index): connections between nodes based on trend similarity,
– Labels (y): trend popularity values retrieved from the database.

For the Graph Neural Network, a Graph Convolutional Network (GCN) architecture is used to process graph-structured data. The model consists of two GCN convolutional layers, one linear (fully connected) layer for predicting trend popularity, and activation functions applied after each layer.

Model training proceeds as follows. In the constructor, three main components are initialized:

– conv1: the first GCNConv layer, which accepts node features of size input_dim and outputs hidden features of size hidden_dim;
– conv2: the second GCNConv layer, which processes hidden features and outputs features of the same dimension hidden_dim;
– fc: a linear layer that transforms the extracted features into output_dim, which in this case is a single value representing the predicted trend popularity.

During the forward pass, data flows through two convolutional layers and one linear layer:

– First convolutional layer (conv1): node features *x* and edge indices *edge_index* are passed into the layer, which processes structural relationships between nodes;
– ReLU activation: applied after each convolution to enable learning of nonlinear dependencies;
– Second convolutional layer (conv2): similar to the first layer, the result is passed through a second convolution layer. This allows the model to learn more about the connections between nodes;
– Linear layer (fc): at the output of the second convolution layer, the data is passed to a linear layer, which reduces the dimensionality to a single value – the trend popularity forecast.
– Model training (*train_gnn*) is performed using backpropagation and an optimizer such as Adam. The training process includes:
– Loss function: a loss function, such as MSELoss (mean square error), is used to compare the predicted popularity of trends with the actual values (y-labels);
– Optimization: The Adam optimizer updates the model parameters (layer weights) based on the error the model makes during prediction;
– Training epochs: the model iteratively updates its parameters to minimize the loss.

After training, the model can generate predictions for new data by passing them through the same layers and producing a popularity forecast. The training process enables the model to learn complex relationships between trends through their graph structure and use these relationships for accurate future predictions.

The process begins with data collection and preprocessing. Unlike existing analogues, the system employs a hybrid approach: the official YouTubeData API for video-content analysis and a web-scraping algorithm (BeautifulSoup) for obtaining hourly X (Twitter) trends via the aggregator trends24.in. The retrieved topic names and interaction volumes are aggregated into an SQLite relational database, where they undergo normalization to remove anomalies and unify time-series formats.

A key stage that constitutes the scientific novelty of the work is model training and evaluation. Based on the structured database records, is constructed a dynamic graph:

$$G = (V, E), \tag{1}$$

where V represents trends and E represents the strength of their connections (co-occurrence).

For forecasting, a Graph Neural Network (GNN) architecture implemented using the PyTorch Geometric framework [18] is applied. The use of GCNConv layers [19] enables the model to aggregate features from neighboring nodes, thus capturing contextual relationships between topics and improving prediction accuracy compared to classical methods. Training is based on minimizing the MSELoss function using the adaptive Adam optimizer.

The final stage is the deployment of the model through the Telegram-bot interface. The bot's logic is integrated with the analytical module, enabling users to obtain real-time predictions. Instead of textual reports, the system automatically generates visualizations using the Matplotlib library, transforming neural-network outputs into clear prediction plots and tag clouds directly within the chat window. This approach ensures mobility and rapid access to intelligent analytics without the need for complex client-side applications.

3. EXPERIMENTAL TESTING OF THE MODEL AND THE BOT

To evaluate the effectiveness of the proposed graph-neural-network-based approach, a series of experiments was conducted using real-world data collected over three months of monitoring the YouTube and X (Twitter) social platforms. The total dataset included more than 15,000 unique trends and their corresponding time-series activity records.

The experimental validation was carried out by comparing the developed GCN-based model [19] with classical time-series forecasting methods (LSTM, RNN) and other modern graph architectures, including Graph Attention Networks (GAT) [20] and the Node2Vec node-embedding method. All models were trained for 200 epochs using the Adam optimizer with a learning rate of 0.01.

The primary evaluation metrics for forecasting accuracy were the Mean Squared Error (MSE) and the Root Mean Squared Error (RMSE). The results of the comparative analysis are presented in Table 1.

Table 1. Comparative characteristics of popularity prediction models.

Model	**MSE (Mean Squared Error)**	**RMSE**	**Advantages / Limitations**
GCN (The proposed)	0.042	0.205	Best stability on sparse graphs
GAT [18]	0.048	0.219	Sensitivity to attention hyperparameters
Node2Vec	0.085	0.291	Loss of dynamic temporal context
LSTM (Baseline)	0.124	0.352	Does not account for topological relationships between topics

The analysis of the obtained results indicates that the use of a graph-based structure reduces forecasting error by an average of 60–66% compared to classical recurrent networks (LSTM). This improvement is explained by the ability of GCN to aggregate features not only from the historical trajectory of an individual node but also from neighboring thematic clusters, which is particularly important for emerging trends with limited historical data.

Although the GAT architecture [20] demonstrates comparable performance, the proposed GCN model proved to be more computationally efficient for deployment within the Telegram interface, ensuring a prediction-generation time of 0.5–1.2 seconds on standard server hardware. Thus, the experimental results confirm the feasibility of integrating graph-based models into real-time social-media monitoring systems.

In addition to integral error metrics (MSE/RMSE), a critically important indicator for social-media monitoring systems is the model's ability to respond promptly to sudden changes in topic-popularity dynamics. For this purpose, a comparative analysis of Growth Rate dynamics was conducted for high-intensity media trends.

The comparison plot of growth-rate coefficients (Fig. 1) clearly demonstrates the advantage of the GCN architecture [19]. While the classical LSTM model exhibits a significant lag effect when capturing peak popularity values, the proposed graph-neural-network approach enables forecasting trend shifts 1.5–2 hours earlier. This improvement is achieved by accounting for the propagation of informational signals from adjacent nodes in the graph before the target trend reaches its maximum weight in an isolated time series.

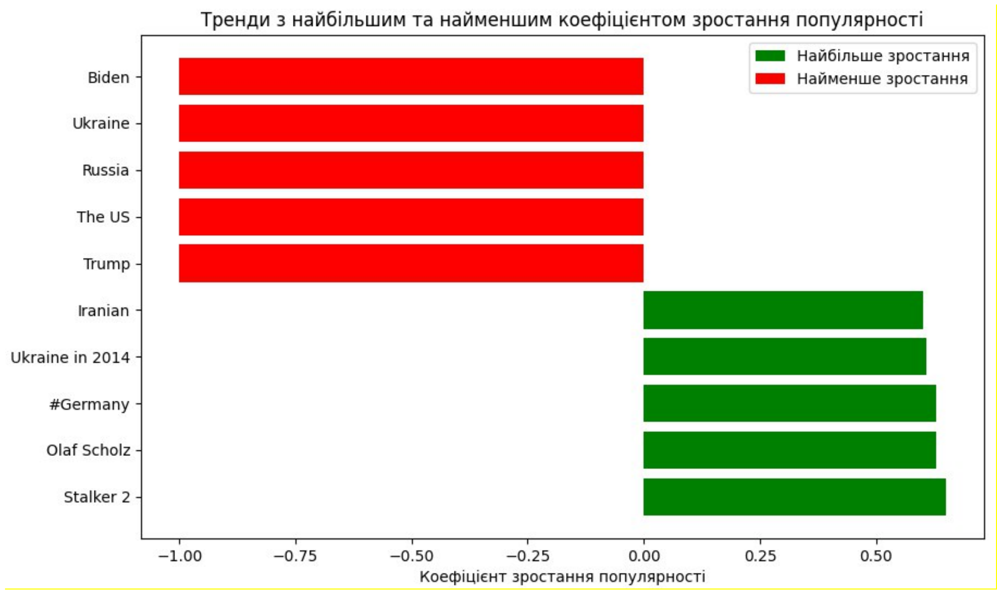

Figure 1. Growth rate comparison chart (green indicates the greatest growth, red indicates the least growth, numerical values are the popularity growth coefficient).

The experiments demonstrated that the integration of feature-aggregation mechanisms from neighboring topics enables the system to efficiently identify viral cascades at early stages. This confirms the feasibility of using the developed Telegram bot as an early-warning tool for detecting shifts in the media landscape.

4. RESULTS

The effectiveness of the proposed intelligent system based on a Graph Neural Network was verified through a comprehensive comparative analysis with modern analogue models, including Node2Vec, Graph Attention Networks (GAT) [20], and Recurrent Neural Network (RNN) with graph-structured input. The key technical and operational characteristics of the compared architectures – including accuracy, scalability, semantic-feature integration, and computational complexity – are presented in detail in Table 2.

To validate the theoretical advantages of the selected architecture, the models were tested within two practical social-media monitoring scenarios. During the analysis of trend dynamics on the X (Twitter) platform, the Node2Vec model demonstrated a classification accuracy of approximately 60%, which can be attributed to the loss of dynamic temporal context and insufficient consideration of the global graph structure. Meanwhile, the GAT architecture [20] achieved higher accuracy (~80%), but exhibited substantial computational overhead when scaled to large sparse graphs. The proposed GNN model based on GCN [17] achieved the highest accuracy – around 90% – enabled by the effective synergy of topological analysis and deep processing of semantic metadata.

Table 2. Comparison of neural networks.

Characteristics	**Node2Vec**	**GAT [18]**	**RNN**	**Proposed GNN (GCN [17])**
Accuracy	Low	Very high	Moderate	**Very high**
Scalability	High	Limited	Medium	**High**
Semantics accounting	Absent	High	Medium	**High**
Global structure accounting	Low level	High level	Low level	**Maximum**
Computational complexity	O(V·D)	O(V·D·K)	O(N T)	**Optimized O(E D F)**
Flexibility in different tasks	Limited	Medium	High	**High**

A similar trend was observed in the task of forecasting video-content popularity on the YouTube platform. The baseline RNN model demonstrated a Root Mean Squared Error (RMSE) of approximately 0.15 due to its inertia in capturing abrupt explosive spikes in user activity. The GAT model achieved a better result (RMSE ~0.10); however, the developed GNN-based system provided the lowest error, with RMSE ~0.08. This improvement is attributed to the ability of the graph-based network to aggregate complex nonlinear relationships between video-hosting objects, taking into account comments, user activity, and contextual references.

Based on the results of the conducted experiments, it was established that the applied GNN architecture is the most adaptive for tasks requiring simultaneous consideration of local node features and the global topology of the media space. Owing to its optimized computational complexity and its ability to outperform analogue models in terms of accuracy and flexibility, the proposed model serves as an effective tool for intelligent analysis of large-scale data streams in real-time systems.

5. CONCLUSIONS

In the course of the conducted research, an intelligent system architecture based on Graph Neural Networks for forecasting media-trend popularity was developed and validated. The experiments confirmed that the integration of graph-convolutional layers significantly improves the accuracy of analyzing dynamic social structures, reducing the mean squared forecasting error by a factor of 2.5–3 (by 60–66%) compared to traditional methods such as LSTM and RNN. The multifunctional Telegram bot implemented within the study provides rapid interactive access to predictive data, automating real-time visualization of analytical results.

The proposed forecasting model has substantial potential for scaling across a wide range of applied domains. In e-commerce, the developed approach can be used for demand prediction, in-depth analysis of customer reviews, and identification of trending product categories based on semantic analysis of user experience. In medicine and healthcare, the use of graph-based structures enables modeling disease-spread trajectories, forecasting demand for medical services, and detecting early signs of epidemiological threats through anomaly monitoring in patient data.

In financial markets, the application of the intelligent model opens opportunities for accurate prediction of market trends through correlation analysis of investor behavior in social media and assessment of the influence of global news flows on asset prices. In the educational sector, the system can be effectively used to monitor emerging scientific topics, forecast student interest in new academic courses, and design relevant educational programs based on the analysis of professional forums.

Particular attention should be given to the application of the model in the political domain for monitoring public opinion and evaluating the effectiveness of political campaigns. Analysis of electoral-trend dynamics enables timely adjustment of communication strategies, increasing their efficiency and adaptability to societal demands. Thus, the developed system represents a universal analytical tool capable of providing high-accuracy decision support in complex and rapidly changing environments.

REFERENCES

[1] V. Shyshkina, Main trends in the development of chat-bots, in: VI International Student Scientific and Technical Conference “Natural Sciences and Humanities. Topical issues” (Tern., 27–28 April 2023), 2023, pp. 197–198 URL: http://elartu.tntu.edu.ua/handle/lib/41269. (In Ukrainian).

[2] M. Adam, M. Wessel, A. Benlian, AI-based chatbots in customer service and their effects on user compliance. Electron Markets 31, 2021, 427–445. https://doi.org/10.1007/s12525-020-00414-7.

[3] K. Ivanchenko, Application of chatbots based on artificial intelligence for automated communications with customers in the retail sector, in: International Forum EFBM 3.0 “Economy. Finance. Business. Management. From Recovery to Growth”, Kyiv: Lyra-K Publishing House, 2024, pp. 40–41. (In Ukrainian).

[4] I.S. Dibrova, Development of a telegram bot and its application, Bulletin of the Student Scientific Society of Vasyl Stus Donetsk National University. Natural and Technical Sciences, vol, 1, no. 16, 2024, pp. 123–127. (In Ukrainian).

[5] E. Stoilova, AI chatbots as a customer service and support tool, in: ROBONOMICS: The Journal of the Automated Economy, vol. 2, 2021, p. 21.

[6] O. Zavydivska, T. Kutseryb, I. Ilnytskyi, A. Artemovych, I. Hul. The current state of digitalization in the field of physical culture and sports. Scientific Journal of the Dragomanov Ukrainian State University. Series 15, 11 (184), 2024, 76–81. https://doi.org/10.31392/UDU-nc.series15.2024.11(184).15. (In Ukrainian).

[7] M.R. King, The Future of AI in Medicine: A Perspective from a Chatbot. Ann Biomed Eng 51, 2023, 291–295. https://doi.org/10.1007/s10439-022-03121-w.

[8] Y. Skalskyi, M. Dendyuk, Development of an intelligent chatbot “Vstupnyk”, in: Computer modeling and information technologies: materials of the fifth scientific and practical conference of students, postgraduates and young scientists (Lviv, October 19-21, 2023), Lviv: NNI KNIT NLTU of Ukraine, 2023, p. 86–91. URL: https://conf.nltu.edu.ua/index.php/conf1/article/view/28. (In Ukrainian).

[9] E. Adamopoulou, L. Moussiades, I. Maglogiannis, L. Iliadis, E. Pimenidis An Overview of Chatbot Technology, in: Artificial Intelligence Applications and Innovations. AIAI 2020. IFIP Advances in Information and Communication Technology, vol. 584. Springer, Cham, 2020 https://doi.org/10.1007/978-3-030-49186-4_31.

[10] B. Luo, R.Y.K. Lau, C.Li, Y.-W. Si, A critical review of state-of-the-art chatbot designs and applications, Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 12 (1), 2020. https://doi.org/10.1002/widm.1434.

[11] J. Cahn, CHATBOT: Architecture, design, & development, University of Pennsylvania School of Engineering and Applied Science Department of Computer and Information Science, 2017, 40 p.

[12] M. Skjuve, A. Følstad, K. I. Fostervold, P. B. Brandtzaeg, My Chatbot Companion – a Study of Human-Chatbot Relationships, International Journal of Human-Computer Studies, vol. 149, 2021. https://doi.org/10.1016/j.ijhcs.2021.102601.

[13] M.S. Satu, Review of integrated applications with aiml based chatbot, in: International Conference on Computer and Information Engineering (ICCIE), IEEE, 2015, pp. 87–90.

[14] S. Orlov, S. Martsenko, Research of IT trends in logistics, ScientificJournalof TNTU. Tern.: TNTU, vol. 117, no. 1, 2025, pp. 105–111. https://doi.org/10.33108/visnyk_tntu2025.01.105.

[15] H. A. Al-Ababneh, A.Y. Alzyadat, A. M. Al-Husban, F. J. Alotoum, B.Al Kurdi, T. Ibragimkhalilova, F. M. S. Barhoush, Marketing in Social Networks: Key Trends and Development Forecasts, International Review of Management and Marketing, no. 15 (4), 2025, pp. 388–396. https://doi.org/10.32479/irmm.19195.

[16] V.V. Mazurenko, S.D. Shtovba, Overview of models for social network analysis, Bulletin of the VPI, no. 2, 2015, pp. 62–74. (In Ukrainian).

[17] T. Levytska, L. Kotykhova, Graph neural networks and PageRank algorithms in the problems of predicting the popularity of hashtags in social networks, Reporter of the Priazovskyi State Technical University. Section: Technical Sciences, vol. 51, 2025, 50–56. https://doi.org/10.31498/2225-6733.51.2025.344600. (In Ukrainian).

[18] M. Fey, J.E. Lenssen, Fast Graph Representation Learning with PyTorch Geometric, ICLR Workshop on Representation Learning on Graphs and Manifolds, 2019, URL: https://arxiv.org/abs/1903.02428.

[19] T.N. Kipf, M. Welling, Semi-Supervised Classification with Graph Convolutional Networks, in: Proceedings of the 5th International Conferenceon Learning Representations (ICLR), 2017, URL: https://arxiv.org/abs/1609.02907.

[20] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, Y. Bengio, Graph Attention Networks, in: International Conference on Learning Representations (ICLR), 2018, URL: https://arxiv.org/abs/1710.10903.