



Graph Neural Network for Social Network Analysis

Navale Palak Santosh, Dr. Anand.A. Khatri, Prof. Bangar Abhishek Popat, Prof. S.B. Bhosale

Dept. of Computer Engineering Jaihind College of Engineering, Pune, India

Abstract. Social networks have become an important part of digital communication, connecting people through friendships, followers, messages, likes, comments, and shared interests. Social Network Analysis (SNA) is used to understand these relationships, identify communities, predict connections, and discover influential users. Traditional machine learning methods often represent users as independent records and may not fully capture the complex relationships between them. Graph Neural Networks (GNNs) address this limitation by learning from both node features and the connections between nodes. This paper reviews the application of GNNs in Social Network Analysis, focusing on graph representation, node classification, link prediction, community detection, influence prediction, and anomaly detection. It discusses important GNN architectures, including Graph Convolutional Networks (GCNs), GraphSAGE, and Graph Attention Networks (GATs), along with representative research on user localization and influence-aware link prediction. A comparative analysis highlights the capabilities and limitations of existing approaches. Finally, a conceptual architecture is proposed for an AI-based social network analysis system. The review identifies key research challenges, including scalability, privacy, dynamic graph learning, interpretability, and the detection of misinformation and suspicious accounts.

Keywords: Graph Neural Networks, Social Network Analysis, Deep Learning, Graph Convolutional Networks, Link Prediction, Community Detection, Influence Prediction, Anomaly Detection

I. Introduction

Social networks such as Facebook, Instagram, LinkedIn, and X have become important platforms for communication, information sharing, and relationship building. These platforms generate large amounts of interconnected data through friendships, followers, likes, comments, shares, and messages. Social Network Analysis (SNA) is used to study these relationships, identify communities, discover influential users, and understand user behaviour. Traditionally, graph-based methods and machine learning techniques have been applied to analyse social networks. However, conventional methods may not effectively capture complex relationships and dependencies between connected users [6].

Graph Neural Networks (GNNs) are deep learning models designed to process graph-structured data. They learn useful representations by combining node attributes with information from neighbouring nodes and their connections [1], [2]. Graph Convolutional Networks (GCNs) aggregate information from neighbouring nodes, GraphSAGE supports inductive learning on large graphs, and Graph Attention Networks (GATs) assign different importance to neighbouring nodes [1]–[3]. These capabilities make GNNs suitable for analysing complex social network structures.



Recent research has explored GNN applications in social network user localization, link prediction, and relationship analysis. Zhong et al. proposed a multiple-aspect attentional GNN approach that combines textual information and social relationships to predict users' geographical locations [4]. Similarly, research on influence-aware link prediction investigates how node centrality and similarity information can improve the prediction of potential connections in networks [5].

II. Review Methodology

A focused literature search was conducted using research platforms such as IEEE Xplore, SpringerLink, ScienceDirect, Wiley, and open-access repositories with keywords including Graph Neural Networks, Social Network Analysis, Graph Convolutional Networks, Graph Attention Networks, Link Prediction, Community Detection, and Influence Prediction. Foundational studies and application-oriented papers were considered to understand GNN methods and their use in social networks [1]–[6]. Studies were reviewed based on their objectives, proposed techniques, applications, evaluation metrics, and limitations. Research on GCN, GraphSAGE, and GAT was used to examine graph representation learning [1]–[3], while studies on user localization and influence-based link prediction were considered for application analysis [4], [5]. A comprehensive GNN survey provided additional background on architectures and research challenges [6].

III. GNN Techniques for Social Network Analysis

1. Graph Convolutional Networks and Graph Representation Learning

Graph Convolutional Networks (GCNs) learn user representations by combining node features with information from neighbouring users. They can be used for node classification, user profiling, and predicting user behaviour in social networks. GCNs use graph structure to improve learning compared with approaches that consider users independently [1]. GraphSAGE extends this approach by sampling and aggregating neighbourhood information, making it suitable for large networks with newly added users [2].

2. Graph Attention and Link Prediction

Graph Attention Networks (GATs) assign different importance to neighbouring nodes using attention mechanisms. This helps models focus on relevant relationships when analysing user interactions and social connections [3]. In link prediction, GNN models learn node embeddings to estimate the likelihood of future connections. Research on influence-based link prediction combines graph centrality and similarity measures with GNN representations to improve relationship prediction [5].

3. Attention-Based and Application-Specific GNNs

Attention-based GNNs can combine different sources of social network information, including text, user attributes, and relationships. Zhong et al. proposed a multiple-aspect attentional GNN for online social network user localization, combining textual information and network relationships to predict users' geographical locations [4]. These methods demonstrate how GNNs can support specialised applications such as user localization, recommendation, and social relationship analysis. However, their performance depends on data quality, graph structure, and computational resources [6].

IV. Machine Learning and Deep Learning Methods

A. Traditional Machine Learning Versus GNNs Traditional machine learning methods analyse node features and manually designed graph metrics, whereas Graph Neural Networks (GNNs) learn representations by combining node attributes with information from neighbouring nodes and graph structure [1], [2], [4]–[6]. Methods such as Support Vector Machines (SVM),



Random Forest, and Logistic Regression can serve as baseline classifiers for social network analysis tasks. GNN-based approaches may provide additional benefits when relationships and interactions between users contain useful information [5], [6], [9], [10]. Graph Convolutional Networks (GCNs) aggregate information from neighbouring nodes for classification tasks, while GraphSAGE uses neighbourhood sampling and aggregation to support inductive learning on large graphs [1], [2]. Graph Attention Networks (GATs) use attention mechanisms to assign different importance to neighbouring nodes, enabling the model to focus on relevant relationships [3], [4]. Depending on the application, these techniques can support node classification, link prediction, community analysis, and influence prediction [6]–[8], [12]–[15]. However, the performance of GNNs depends on dataset quality, graph Traditional classifiers such as Logistic Regression, Support Vector Machines, and Random Forests can serve as baselines when the task is defined using node attributes or engineered graph features. Graph embeddings such as DeepWalk and node2vec learn structural representations through neighbourhood walks [9], [10]. GNNs differ by learning representations through neighbourhood aggregation or message passing [1]–[4]. No single family is best for every task: the choice depends on graph size, available features, label quantity, computational budget, and evaluation design. equivalent access to features, and clearly specified metrics. Classification tasks can report precision, recall, and F1-score as well as accuracy, particularly when classes are imbalanced. Link-prediction studies may use AUROC, average precision, and task-specific ranking metrics [6]–[8]. Splits should prevent leakage between training and test data, and temporal evaluation may be appropriate when predicting future relationships. structure, task requirements, and evaluation procedures. Therefore, GNN models should be compared fairly with simpler baseline methods using appropriate evaluation metrics. Their additional computational complexity should be justified by measurable improvements in predictive performance, scalability, and generalisation rather than assumed advantages [4]–[6], [9], [10].

2. Deep Learning

Deep learning enables GNNs to learn complex patterns from graph-structured data. GCNs learn node representations through graph convolution, while Graph.Attention Networks use attention weights to identify more informative neighbouring nodes [1], [3]. These learned representations can support user classification, link prediction, community analysis, and recommendation.

Zhong et al. proposed an attention-based GNN that combines textual information and social relationships for online social network user localization [4]. Similarly, influence-aware link prediction combines node centrality and similarity information with GNN-based learning to estimate potential connections [5]. However, deep learning performance depends on training data quality, computational requirements, and appropriate model validation [6].

3. Validation Practice

Proper model validation is essential for evaluating GNN performance. When connected users or related graph samples appear in both training and test sets, evaluation may become overly optimistic. Data leakage, class imbalance, and differences between training and real-world networks can affect reported results.

GNN models should be evaluated using appropriate train-test splits, baseline comparisons, and task-specific metrics. Accuracy, Precision, Recall, and F1-score can be used for classification, while AUROC and AUPR are suitable for many link prediction tasks [5], [6]. Testing on separate datasets and newly observed users can help assess whether a model generalises to unseen social network data.

V. Comparative Analysis

Table I summarises representative primary studies. Accuracy values are not directly comparable because targets, datasets and validation schemes differ.

Table 1: Comparison of Representative GNN Methods For Social Network Analysis

Author, year	Objective	Sensing / method	Major reported findings	Limitations
Kipf and Welling, 2017 [1]	Classify nodes in graph-structured data	Graph Convolutional Network (GCN)	Demonstrated semi-supervised node classification using graph structure and node features	Evaluated mainly on benchmark datasets; scalability can be challenging
Hamilton et al., 2017 [2]	Learn representations for large graphs and unseen nodes	GraphSAGE; neighbourhood sampling and aggregation	Supported inductive learning for nodes not seen during training	Performance depends on sampling strategy
Veličković et al., 2018 [3]	Learn node representations by identifying important neighbours	Graph Attention Network (GAT)	Attention mechanisms assign different weights to neighbouring nodes	Attention computation can increase resource requirements on large graphs
Zhong et al., 2020 [4]	Predict geographical locations of online social network users	Multiple-Aspect Attentional GNN	Combined textual information and social relationships for user localization	Depends on available user information and the structure of the social network
Wu et al., 2021 [6]]	Review GNN architectures, applications, and challenges	Comprehensive survey of GNN methods	Summarised GNN models, applications, and research challenges across graph-learning tasks	Broad survey; does not provide one unified solution for every social network task
Baswani et al., 2026 [5]	Predict links using node influence information	GNN-based link prediction with node influence	Investigated the use of node influence for predicting potential network connections	Findings and publication metadata should be verified against the original paper before final submission

VII. Proposed Conceptual Architecture

Fig. 1 The acquisition unit collects social network data, including user profiles, friendships, followers, likes, comments, shares, and timestamps. In Layer 3, the preprocessing module cleans the collected data, removes duplicates, extracts node and edge features, and constructs the graph for GNN processing [1], [2]. In Layer 4, Graph Neural Network models perform node classification, link prediction, community detection, influence prediction, and anomaly detection, generating prediction results with the proposed Graph Neural Network (GNN)-based system is designed to analyse complex relationships and interactions among users in online social networks. The system collects information such as user profiles, friendships, followers, posts,

likes, comments, shares, and interaction history from authorised social network sources. The collected data is cleaned and preprocessed to remove duplicate records, handle missing values, and extract relevant features. Users are represented as nodes, while friendships, follows, and interactions are represented as edges to construct a graph. This graph is processed using GNN models such as Graph Convolutional Networks (GCNs), GraphSAGE, and Graph Attention Networks (GATs), which learn user representations by aggregating information from neighbouring nodes [1]–[5], [8].

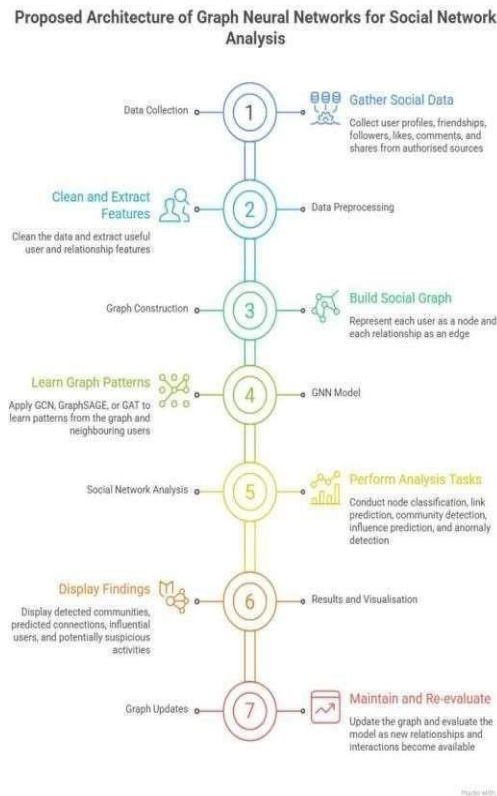


Fig.1. Proposed conceptual architecture for graph neural network for social network analysis

The learned representations are used to support node classification, link prediction, community detection, influence analysis, and anomaly detection [6]–[8], [11]–[15]. The analysis results, including predicted connections, influential users, and potentially suspicious activities, are displayed through a dashboard. In Layer 6, decision rules use these results to recommend potential connections, identify influential users, and flag suspicious accounts for further investigation. Essential monitoring functions can continue during temporary connectivity failures by storing selected results locally, provided that appropriate storage and recovery mechanisms are implemented. In Layer 7, a sampled subset of predictions is verified using labelled datasets or expert assessments. Model performance is evaluated using Precision, Recall, F1-score, and other suitable task-specific metrics. Models are updated only after independent evaluation to improve reliability and reduce the risk of overfitting. This proposed architecture integrates graph representation learning, social network analysis, result visualisation, and performance monitoring into a conceptual framework for intelligent social network analysis.



VIII. Research Gaps and Future Directions

Although Graph Neural Networks (GNNs) have demonstrated significant potential in Social Network Analysis (SNA), several research gaps remain. Existing GNN architectures, including Graph Convolutional Networks (GCNs), GraphSAGE, and Graph Attention Networks (GATs), provide effective methods for learning node representations and analysing relationships, but challenges related to scalability, computational complexity, and generalisation remain [1]–[5]. Community detection in node-attributed social networks also remains challenging because users may belong to overlapping communities and their interests and relationships can change over time [6]. Traditional link-prediction approaches provide useful methods for estimating potential relationships, but they may not fully capture complex structural patterns and user attributes in large social networks [7], [8]. Practical GNN applications therefore require careful model selection, feature engineering, and evaluation to ensure reliable results [9]. Another research gap concerns the integration of feature extraction and graph representation learning for homogeneous and heterogeneous networks. Different social platforms contain different types of users, relationships, and interactions, making it difficult to design a single model that performs consistently across multiple network structures [10]. Recent research on GNN-based link prediction highlights the need to improve prediction performance, scalability, and robustness under different network conditions [11].

Community-aware and role-aware link-prediction methods provide promising directions for combining users' community membership with their structural roles, although their effectiveness should be evaluated across diverse datasets and network settings [12]. Similarly, identifying influential nodes using GNN models requires further investigation into how network structure, user attributes, and interaction patterns affect influence predictions [13]. GNN-based community-detection approaches also require evaluation for accuracy, computational efficiency, and performance on large and dynamic social networks [14]. Further research into GNN-based link-prediction methods is needed to understand their limitations, practical applications, and performance under changing network conditions [15]. Based on these gaps, future research should focus on developing scalable GNN models that can process large and dynamic social networks efficiently. Combining temporal graph learning with user attributes and interaction history could improve the analysis of changing relationships and behaviour. Explainable GNN methods could help users understand why a particular account is predicted to be influential, assigned to a community, or flagged for further investigation. Future systems should also address incomplete data, noisy interactions, class imbalance, privacy protection, and potential bias. In addition, combining community detection, link prediction, influence analysis, and anomaly detection within a common framework could support more comprehensive social network analysis.

IX. Conclusion

Graph Neural Networks (GNNs) provide an effective approach to Social Network Analysis by learning from both user attributes and relationships between users. Techniques such as Graph Convolutional Networks (GCNs), GraphSAGE, and Graph Attention Networks (GATs) support applications including node classification, link prediction, community detection, influence prediction, and anomaly detection. The reviewed studies demonstrate the potential of GNNs to analyse complex social relationships and improve network-based predictions [1]–[6].

However, challenges related to scalability, dynamic network structures, data privacy, model interpretability, and reliable evaluation remain. Future research should focus on scalable architectures, explainable predictions, privacy-preserving methods, and real-world validation across diverse social networks. In conclusion, GNNs offer a promising foundation for intelligent Social Network Analysis, with further research needed to develop reliable, secure, and practical systems.



References

1. T. N. Kipf and M. Welling, "Semi-Supervised Classification with Graph Convolutional Networks," in Proc. International-Conference on Learning Representations (ICLR), 2017. [Online]. Available: <https://arxiv.org/abs/1609>.
2. Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and S. Y. Philip, "A comprehensive survey on graph neural networks," IEEE Transactions on Neural Networks and Learning Systems, vol. 32, no. 1, pp. 4–24, 2021, doi: 10.1109/TNNLS.2020.2978386.
3. Z. Zhang, P. Cui, and W. Zhu, "Deep learning on graphs: A survey," IEEE Transactions on Knowledge and Data Engineering, vol. 34, no. 1, pp. 249–270, 2022 doi: 10.1109/TKDE.2020.2981333.
4. J. Zhou et al., "Graph neural networks: A review of methods and applications," AI Open, vol. 1, pp. 57–81, 2020, doi: 10.1016/j.aiopen.2021.01.001.
5. J. Y. Kim and S. B. Cho, "A systematic analysis and guidelines of graph neural networks for practical applications," Expert Systems with Applications, vol. 184, Art. no. 115466, 2021, doi: 10.1016/j.eswa.2021.115466.
6. P. Chunaev, "Community detection in node-attributed social networks: A survey," Computer Science Review, vol. 37, Art. no. 100286, 2020, doi: 10.1016/j.cosrev.2020.100286.
7. "Applications of link prediction in social networks: A review," Journal of Network and Computer Applications, vol. 166, Art. no. 102716, 2020, doi: 10.1016/j.jnca.2020.102716.
8. "Link prediction techniques, applications, and performance: A survey," Physica A: Statistical Mechanics and its Applications, vol. 553, Art. no. 124289, 2020, doi: 10.1016/j.physa.2020.124289.
9. "A survey on feature extraction and learning techniques for link prediction in homogeneous and heterogeneous complex networks," Artificial Intelligence Review, 2024. [Online]. Available: SpringerLink.
10. "A survey on graph representation learning methods," ACM Transactions on Intelligent Systems and Technology. [Online]. Available: ACM Digital Library
11. "A survey on GNN-based link prediction: Techniques, applications, and challenges," WIREs Data Mining and Knowledge Discovery, 2026. [Online]. Available: Wiley Online Library
12. T. Qiu, Y. Chang, and G. Chen, "A Community-and-Role-Aware Graph Neural Network on Link Prediction," Physica A: Statistical Mechanics and its Applications, vol. 685, Art. No. 131271, 2026. DOI: 10.1016/j.physa.2026.131271.
13. A. D. Tyagi and S. Garg, "A Framework to Predict the Influencing Nodes in Social Networking Platform Using Graph Neural Network Model," Social Network Analysis and Mining, vol. 16, Art. no 29, 2026. DOI: 10.1007/s13278-025-01574-0.
14. M. Rekha and A. H. Ganai, "Detecting Communities in Social Networks with Graph Convolutional Neural Networks," Discover Artificial Intelligence, vol. 6, Art. no. 1203, 2026. Read open-access paper.
15. "A Survey on GNN-Based Link Prediction: Techniques, Applications, and Challenges," WIREs Data Mining and Knowledge Discovery, 2026. DOI: 10.1002/widm.70093.