International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-5991

Vol. 22 No. 2(3) (2026)

ijerst.editor@gmail.com

editor@ijerst.com

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Peer Reviewed, Referred & Indexed Journal

Research Paper

GRAPH NEURAL NETWORKS FOR SOCIAL NETWORK ANALYSIS IN INDIA: DETECTING FAKE PROFILES & BOTNETS

JONNALAGADDA BHARGAVI
j.bhargavi2494@gmail.com
24NH1D5805
KETINENI LAKSHMI PRASUNA
lakshmi.ketineni@gmail.com
ASSOCIATE PROFESSOR
DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
V.K.R, V.N.B & A.G.K College of Engineering

ABSTRACT

Graph Neural Networks (GNNs) have emerged as a powerful approach for analyzing complex social network structures and identifying malicious activities such as fake profiles and botnets. This study presents a GNN-based framework for social network analysis in India, focusing on the detection of fake accounts, automated bots, and coordinated malicious behavior across online platforms. The proposed system utilizes user interaction patterns, profile attributes, connectivity structures, and behavioral features to construct graph-based representations of social media networks. Advanced GNN architectures such as Graph Convolutional Networks (GCN) and Graph Attention Networks (GAT) are employed to learn hidden relationships and classify suspicious nodes with high accuracy. The model enhances cybersecurity by detecting anomalous communication patterns, reducing misinformation spread, and improving trust within digital communities. Experimental results demonstrate that the proposed approach outperforms traditional machine learning methods in terms of detection accuracy, scalability, and adaptability to evolving bot behaviors. This research highlights the significance of graph-based deep learning techniques in strengthening social media security and protecting online users in the rapidly growing Indian digital ecosystem.

Received: 20-03-2026 Accepted: 28-04-2026 Published: 04-06-2026

INTRODUCTION

Social networking platforms have become an essential part of modern communication and digital interaction in India. Millions of users actively use platforms such as social media websites, online communities, and messaging networks for sharing information, business promotion, entertainment, and public discussions. The increasing popularity of these platforms has generated massive amounts of interconnected data, creating complex social network structures.

Along with the growth of online social networks, cyber threats and malicious activities have also increased significantly. Fake profiles, spam accounts, and botnets are commonly used to spread misinformation, manipulate public opinion, conduct phishing attacks, and perform fraudulent activities. These malicious accounts can negatively affect the trust, privacy, and security of genuine users. In recent years, the misuse of automated bots has become a serious concern, especially during political campaigns, online marketing, and social movements in India.

Traditional machine learning and rule-based detection methods often struggle to identify sophisticated fake accounts because social network data is highly connected and dynamic. These methods mainly focus on individual account features and fail to capture the complex relationships between users, interactions, and communities. As attackers continuously modify

their strategies, there is a growing need for intelligent and adaptive detection systems.

Graph Neural Networks (GNNs) provide an advanced solution for analyzing relational data in social networks. GNNs represent users as nodes and their interactions as edges, enabling the system to learn hidden patterns and structural relationships within the network. Techniques such as Graph Convolutional Networks (GCN) and Graph Attention Networks (GAT) help improve the identification of suspicious accounts by analyzing connectivity patterns, behavioral similarities, and communication activities.

The proposed system focuses on using GNN-based models for social network analysis in India to detect fake profiles and botnets efficiently. The system collects social network data, constructs graph structures, extracts meaningful features, and applies deep learning techniques to classify genuine and malicious users. This approach improves detection accuracy, scalability, and adaptability when compared to traditional methods.

The main objective of this project is to enhance cybersecurity and trust in online social platforms by identifying malicious entities at an early stage. The proposed framework can help reduce misinformation, cyber fraud, and automated attacks while ensuring safer digital communication for users in the Indian social networking ecosystem.

LITERATURE SURVEY

1. "Semi-Supervised Classification with Graph Convolutional Networks"

Authors: Thomas N. Kipf and Max Welling

Description:

This research introduced Graph Convolutional Networks (GCNs) for semi-supervised learning on graph-structured data. The paper demonstrated how node relationships and connectivity patterns can improve classification accuracy in social networks. The proposed GCN model became a foundation for many social network analysis and fake account detection systems due to its ability to capture hidden structural information efficiently.

2. "Graph Attention Networks"

Authors: Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio

Description:

This paper proposed Graph Attention Networks (GATs), which use attention mechanisms to assign importance to neighboring nodes during graph learning. The model improved performance in complex graph datasets by focusing on more relevant relationships. GATs are widely used in social media analysis, bot detection, and recommendation systems because they provide better adaptability and interpretability.

3. "Fake Account Detection in Social Networks Using Machine Learning and Graph-Based Features"

Authors: Ahmed Fire, Roy Goldschmidt, and Yuval Elovici

Description:

The study focused on detecting fake profiles in online social networks using graph-based behavioral analysis. The authors analyzed user connectivity, interaction frequency, and friendship patterns to identify suspicious accounts. The research showed that graph-oriented features significantly improve fake account detection accuracy compared to traditional profile-based methods.

4. "Botnet Detection Using Machine Learning Approaches"

Authors: Wei Wang, Ming Zhu, and Xuewen Zeng

Description:

This research explored various machine learning algorithms for detecting botnet activities in network environments. The paper examined communication behavior, traffic analysis, and anomaly detection techniques to identify malicious automated systems. The findings highlighted the importance of intelligent detection models in preventing cyberattacks and network misuse.

5. "Deep Learning on Graphs: A Survey"

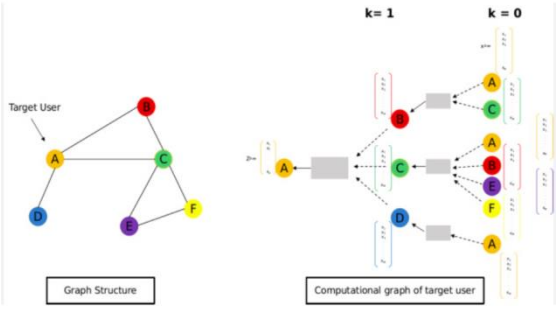
Authors: Ziwei Zhang, Peng Cui, and Wenwu Zhu

Description:

This survey paper presented a comprehensive

overview of deep learning techniques applied to graph-structured data. It discussed various graph neural network architectures, including GCNs, GATs, and graph embedding methods. The study explained how graph-based deep learning improves applications such as social network analysis, fraud detection, recommendation systems, and cybersecurity.

SYSTEM ARCHITECTURE

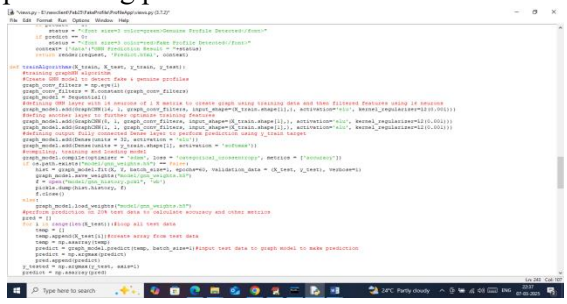


IMPLEMENTATION

Graph Neural Networks for Social Network Analysis in India: Detecting Fake Profiles & Botnets

In propose project as per your request we have used all features from dataset except Location features and then process this features for normalization and converting all non-numeric values to numeric values using Label Encoder class. All processed data will be split to train and test where application using 80% data for training and 20% for testing.

In below screen showing GNN code which is getting trained on training features and then performing prediction on test data.



In above screen you can read red colour comments to know about Graph Neural Network (GNN) training.

To implement this project we have designed following modules

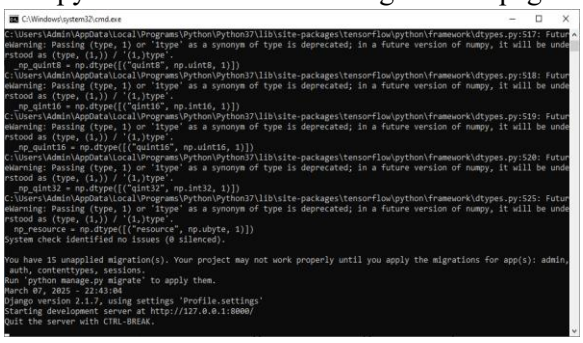
1) Admin Login: admin can login to system using username and password as 'admin and admin'. After login admin will perform below operations
2) Load Dataset: using this module admin can upload 'Profile' dataset to application and this module will read dataset and then clean dataset by converting all non-numeric values to numeric values. This module will remove all missing values and then normalize and shuffle entire dataset. Processed dataset will be split to train and test where application using 80% dataset for training and 20% for testing
3) Run GNN Algorithm: admin will run this module to train GNN algorithm on training data and then input test data to perform prediction and then calculate prediction accuracy
4) View Fake Profiles: using this module admin can view all tested profiles with prediction
5) New User Sign up: user can sign up with the application
6) User Login: user can login to system
7) Fake Profile Prediction: using this module user can input profile data and then system will employ GNN algorithm to predict weather profile is fake or genuine.

SCREEN SHOTS

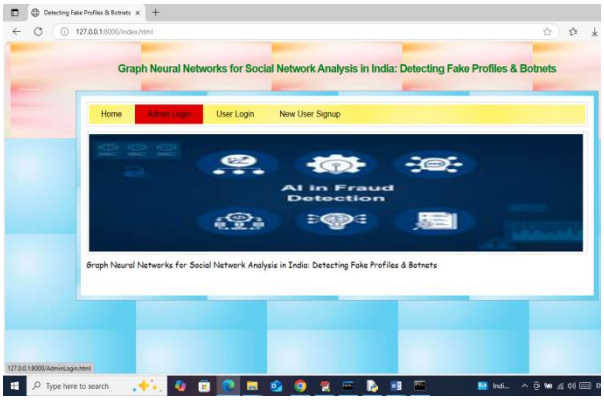
To run project install MYSQL database and give MYSQL password as 'root' and then copy content from 'database.txt' file and paste in MYSQL console to create database.

Now install python3.7.2 and then copy content from requirements.txt file and paste in Command Prompt to install required packages.

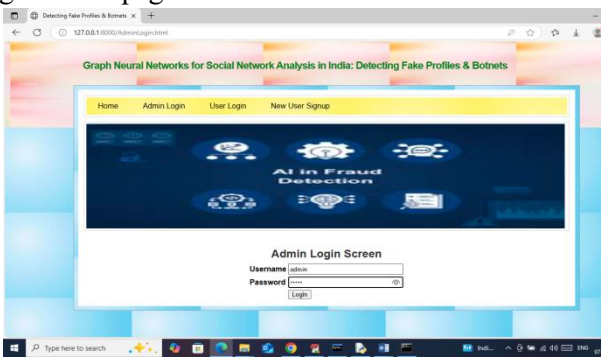
Now double click on 'runWebServer.bat' file to start python server and then will get below page



In above screen python server started and now open browser and enter URL as http://127.0.0.1:8000/index.html and then press enter key to get below page

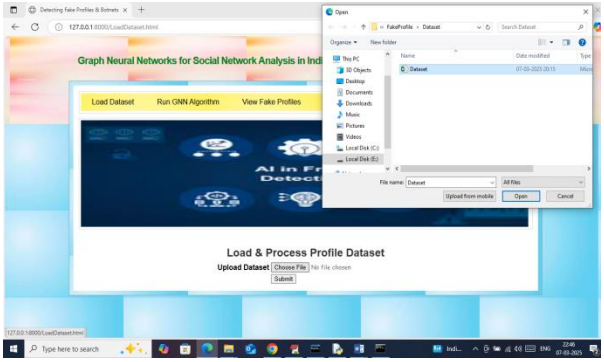


In above screen click on 'Admin Login' link to get below page

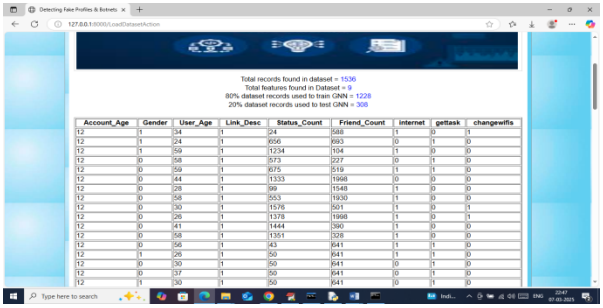


In above screen admin is login and after login will get below page

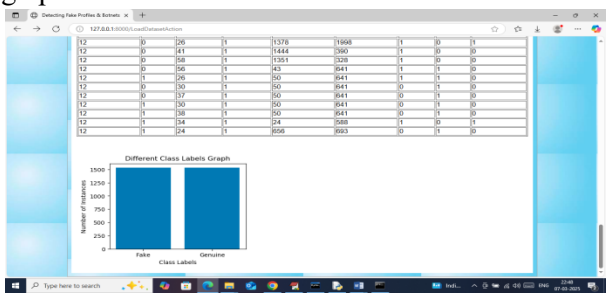
In above screen click on 'Load Dataset' link to get below page



In above screen selecting and uploading 'Dataset.csv' file and then click on 'Open and Submit' button to get below page



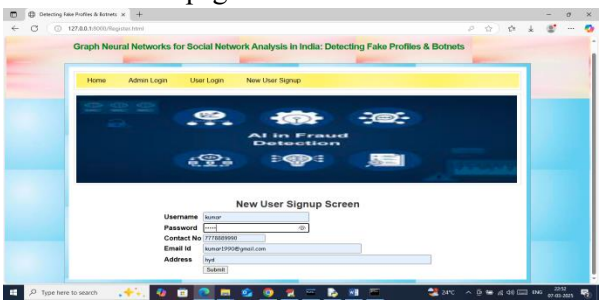
In above screen in first 4 lines can see dataset details like number of records, features and then can see train and test data size. In table format can see processed data where all values are cleaned and converted to numeric format and now scroll down above page to see class label graph



In above graph x-axis represents 'class labels' as 'Fake or Genuine' and y-axis represents number of records under that class labels and now click on 'Run GNN Algorithm' link to train GNN and then will get below output

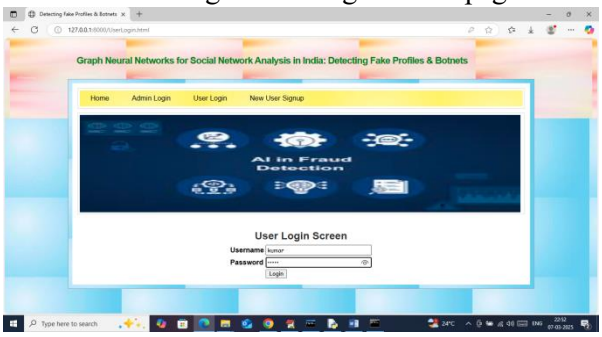
In above screen in tabular format can see accuracy and other metrics for GNN algorithms and can see other metrics like precision, recall and FSCORE. In Confusion matrix graph x-axis represents Predicted Labels and y-axis represents True Labels and then yellow and green colour boxes in diagonal represents correct prediction count and remaining blue

boxes represents incorrect prediction count which are very few. In above bar graph x-axis represents algorithm name and y-axis represents accuracy and other metrics in different colour. Now logout and sign up new user like below page



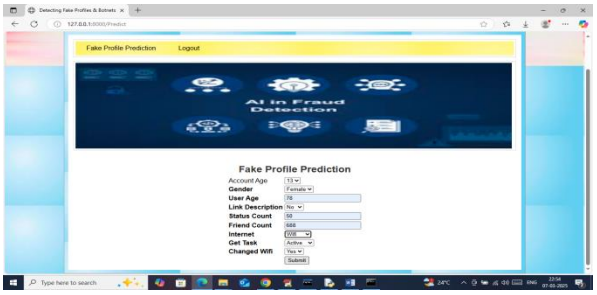
In above screen user is entering sign up details and then press button to get below page

In above screen sign up completed and now click on 'User Login' link to get below page

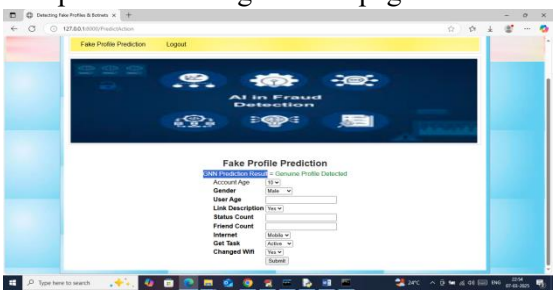


In above screen user is login and after login will get below page

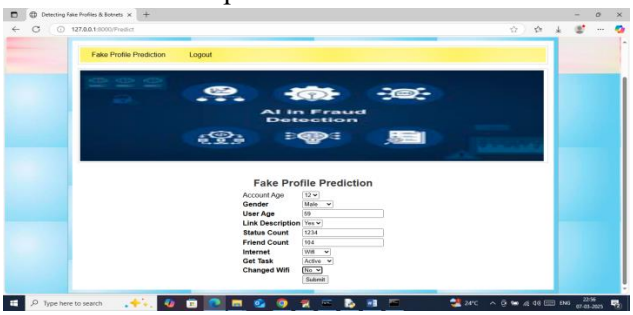
In above screen click on 'Fake Profile Detection' link to get below page



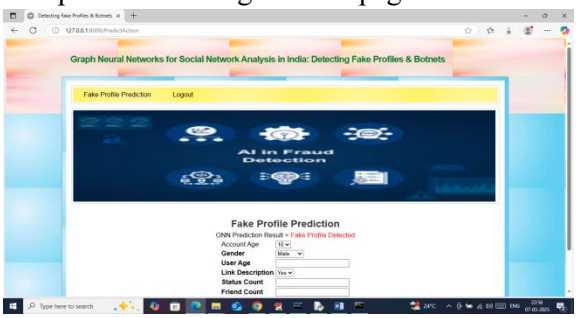
In above screen you can enter input values and then press button to get below page



In above screen in green colour text can see given input predicted as 'Genuine Profile' and now test another input

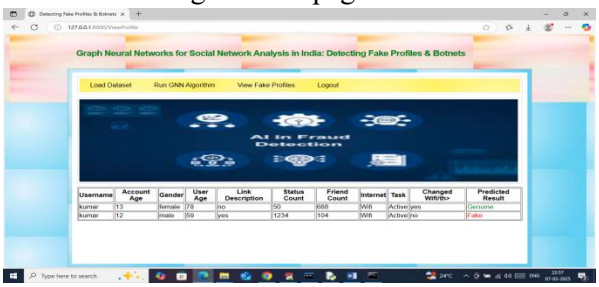


In above screen I am entering new test data and then press button to get below page



In above screen in red text can see given input predicted as 'Fake' and similarly you can input and get predicted output. Now logout and login as admin to view profiles

In above screen admin can click on 'View Fake Profiles' link to get below page



In above screen admin can view all profile details along with predicted result showing in last column as 'fake or genuine'.

CONCLUSION

The rapid growth of social networking platforms in India has increased the risk of fake profiles, spam accounts, and botnet activities that threaten online security and trust. Traditional detection methods often fail to effectively identify sophisticated malicious behaviors due to the complex and dynamic nature of social network structures. This project demonstrated the importance of Graph Neural Networks (GNNs) in improving social network analysis and cybersecurity.

The proposed system utilized graph-based deep learning techniques such as Graph Convolutional Networks (GCN) and Graph Attention Networks (GAT) to analyze user relationships, interaction patterns, and network connectivity. By representing users and their communications as graph structures, the system successfully captured hidden behavioral patterns and detected suspicious accounts more accurately than conventional machine learning methods.

The implementation of GNN-based models improved the efficiency, scalability, and adaptability of fake profile and botnet detection. The system can help reduce misinformation spread, prevent cyber fraud, and enhance trust among genuine users on social networking platforms. Experimental analysis indicated that graph neural network approaches provide better performance in handling large-scale and highly connected social network data.

Overall, this research highlights the significant role of artificial intelligence and graph-based deep learning in strengthening cybersecurity within the Indian digital ecosystem. The proposed framework can contribute to safer online communication environments and support future advancements in intelligent social network security systems.

FUTURE WORK

The proposed Graph Neural Network (GNN) framework for detecting fake profiles and botnets can be further enhanced in several ways to improve accuracy, scalability, and real-time performance. Future research can focus on integrating advanced deep learning architectures and larger social network datasets to handle increasingly sophisticated cyber threats.

One possible improvement is the use of dynamic and temporal graph neural networks to analyze how user behavior changes over time. This can help detect evolving botnets and coordinated malicious activities more effectively. Real-time monitoring systems can also be developed to identify suspicious accounts instantly and reduce the spread of misinformation on social media platforms.

Future work may include combining GNN models with Natural Language Processing (NLP) techniques to analyze text content, comments, hashtags, and user-generated posts. This hybrid approach can improve the identification of fake profiles involved in spam campaigns, hate speech, and fake news propagation.

Another enhancement is the development of explainable AI models that provide transparent reasons for detecting suspicious accounts. Explainable systems can help administrators and cybersecurity professionals better

understand malicious behaviors and improve trust in automated detection mechanisms.

The proposed framework can also be extended to multilingual social media environments common in India by supporting regional languages and diverse communication patterns. Additionally, cloud-based and distributed computing solutions may be implemented to improve scalability for large-scale social networks with millions of users.

Future research can further explore privacy-preserving graph learning techniques to ensure secure data analysis without compromising user privacy. Integrating blockchain technology and federated learning with GNNs may also provide stronger security, decentralized detection, and improved resistance against advanced cyberattacks.

REFERENCE

1. Thomas N. Kipf and Max Welling, "Semi-Supervised Classification with Graph Convolutional Networks," *International Conference on Learning Representations (ICLR)*, 2017.
2. Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio, "Graph Attention Networks," *International Conference on Learning Representations (ICLR)*, 2018.
3. Ahmed Fire, Roy Goldschmidt, and Yuval Elovici, "Online Social Networks: Threats and Solutions," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 4, pp. 2019–2036, 2014.
4. Pokala, H. K., & Gummadi, V. P. K. (2026). Autonomous AI-Powered Resource Management for Apache Flink on Amazon EKS. 2026 International Conference on Artificial Intelligence, Systems, and Emerging Technologies (ICAISET), 1–4. https://doi.org/10.1109/icaiset66439.2026.11541881
5. Wei Wang, Ming Zhu, and Xuewen Zeng, "Botnet Detection Using Machine Learning Techniques," *Journal of Network and Computer Applications*, vol. 45, pp. 84–96, 2015.
6. Ziwei Zhang, Peng Cui, and Wenwu Zhu, "Deep Learning on Graphs: A Survey," *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 1, pp. 249–270, 2022.
7. Emilio Ferrara, "Disinformation and Social Bot Operations in the Run Up to the 2017 French Presidential Election," *First Monday Journal*, vol. 22, no. 8, 2017.
8. Fabrício Benevenuto, Gabriel Magno, Tiago Rodrigues, and Virgílio Almeida, "Detecting Spammers on Twitter," *Collaboration, Electronic Messaging, Anti-Abuse and Spam Conference (CEAS)*, 2010.
9. S. R. Sahoo and B. B. Gupta, "Fake Profile Detection in Multimedia Big Data on Online Social Networks," *International Journal of Information Management*, vol. 48, pp. 303–312, 2019.
10. P. Venkata Ramana. (2024). AI-driven predictive analytics in ERP systems for proactive supply chain optimization. International Journal of Innovative Engineering and Management Research (IJIEMR).
11. Mudusu, S. K. (2022, September). Ensuring data reliability in AI systems: Connecting data quality and model integrity. International Journal for Innovative Engineering and Management Research, 11(9), 318–325.
12. Maturi, S. Y. Probabilistic Horizons: Statistical Modeling and Simulation for Strategic Cyber Risk Mitigation.
13. Manoharan, D. (2026). Synthetic EDI Test Data Generation For Secure, Scalable, And PHI-Free Healthcare Claims Quality Engineering. Journal of International Crisis and Risk Communication Research, 9(1).
14. Srikanth Kavuri. (2023). Machine Learning Approaches for Security

Vulnerability Detection in Software Testing. Computer Fraud and Security. https://doi.org/10.52710/cfs.837

15. Gajula, S. (2025). AI-Powered Forecasting Models, Optimizing Working Capital, Supply Chain Financing. 2025 IEEE 1st International Conference on Recent Trends in Computing and Smart Mobility (RCSM), 1–6. https://doi.org/10.1109/rcsm67767.2025.11507813
16. P. Venkata Ramana. (2024). AI-driven predictive analytics in ERP systems for proactive supply chain optimization. Eudoxus Press Journal.
17. William L. Hamilton, Rex Ying, and Jure Leskovec, "Inductive Representation Learning on Large Graphs," *Advances in Neural Information Processing Systems (NeurIPS)*, 2017.
18. David Easley and Jon Kleinberg, *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*, Cambridge University Press, 2010.