Modeling Movie Success from Collaboration Networks Using Graph Neural Networks

David Corcoran
Gentry Lamb
Adam Stein

https://www.shutterstock.com/image-vector/movie-cinema-elements-collection-tickets-popcorn-2442422335

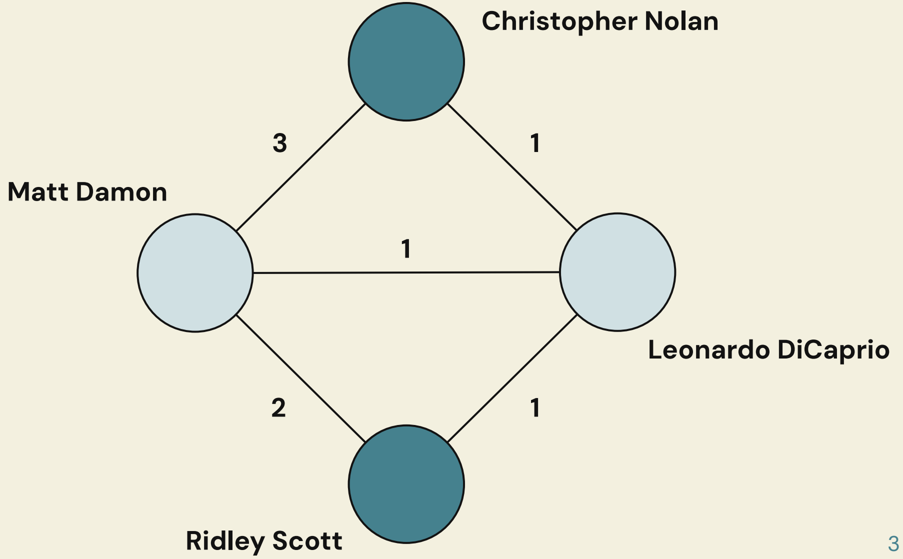
7 AUGUST 2025

Agenda

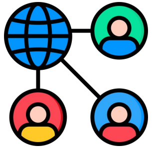
- Research Topic
- Research Objectives
- Dataset Overview
- Methods
- Results
- Conclusions

Research Topic

- ❖ Explore how the structure of professional collaboration networks (actors, directors, producers) affects the success of movies at the box office.
- ❖ **Hypothesis** – The network structure of a movie's team is a strong predictor of its commercial and critical performance.
- ❖ **Approach**
  - ➢ **Nodes:** Actors, Directors, Producers
  - ➢ **Edges:** Collaborations
  - ➢ Use Graph Neural Networks to predict Box Office performance
- ❖ Network analysis allows us to model this relational structure and uncover hidden patterns.



Research Objectives

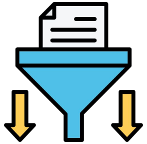


Build a Movie Collaboration Graph

- **Nodes**: Actors, Directors, Producers
- **Edges**: Collaboration edges between cast/crew who worked on the same film
- Include metadata such as genre, year, studio, and success outcome (revenue/ratings)

Model Success Using GNNs and Other ML Models

- Define prediction targets:
  - Box office success (e.g., high/mid/low revenue tier)
  - Critical acclaim (e.g., ratings score buckets or award nominations)
- Train GNNs on collaboration subgraphs to classify movie outcomes



Extract and Evaluate Network-Based Features

- Hand-crafted metrics:
  - Node-level: Degree, betweenness, PageRank
  - Graph-level: Clustering coefficient, modularity
- Learnable features:
  - GNN node embeddings
  - Edge strength (e.g., frequency of past collaboration)

Analyze Patterns Behind Predictions

- Use explainability tools to identify important nodes/edges
- Cluster movies by graph structure to discover common success motifs
- Compare successful and unsuccessful teams on key metrics

Dataset Overview

Data Sources

TMDB API:

- Queried through 18 genres to collect 19 base features

OMDB API:

- Iterated through movie IDs already collected from TMDB to obtain 8 additional features

Data Cleaning

Raw Data:

- 7333 movies
- 28 movie attributes

Processed Data:

- Converted data types, dealt with duplicate titles, dropped unnecessary rows
- 6940 movies
- 34 movies attributes

Target Variables

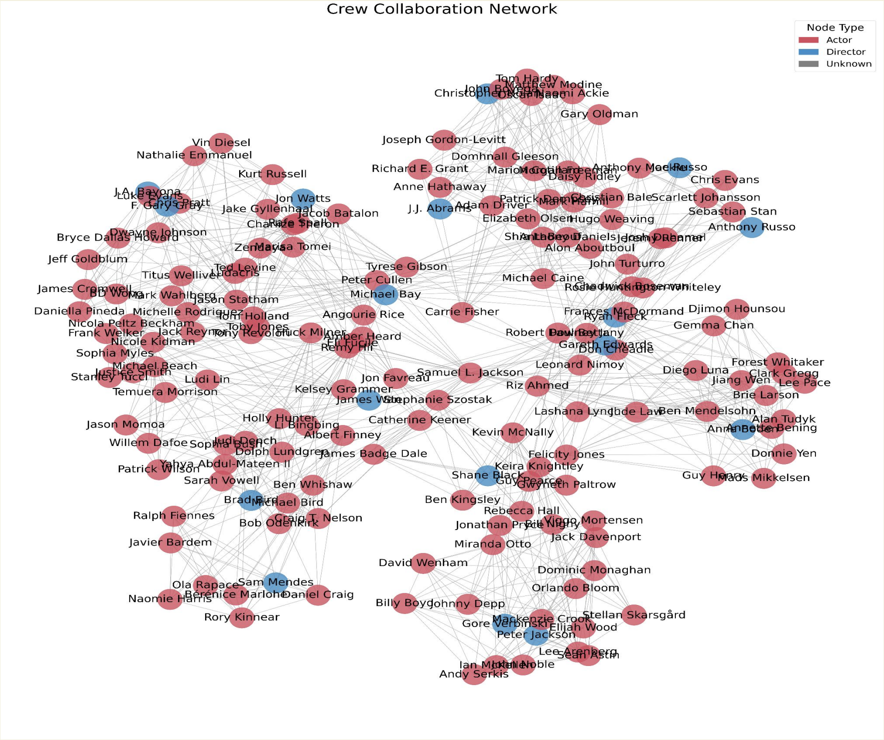
Financial Success

- Box Office Revenue

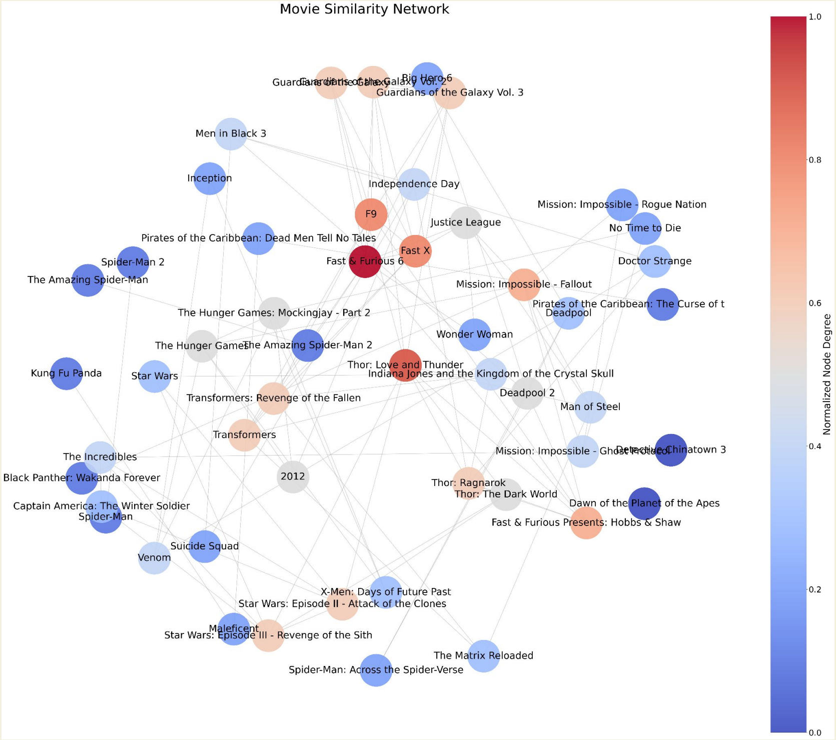
Critical Success

- Ratings

Methods - Network Construction



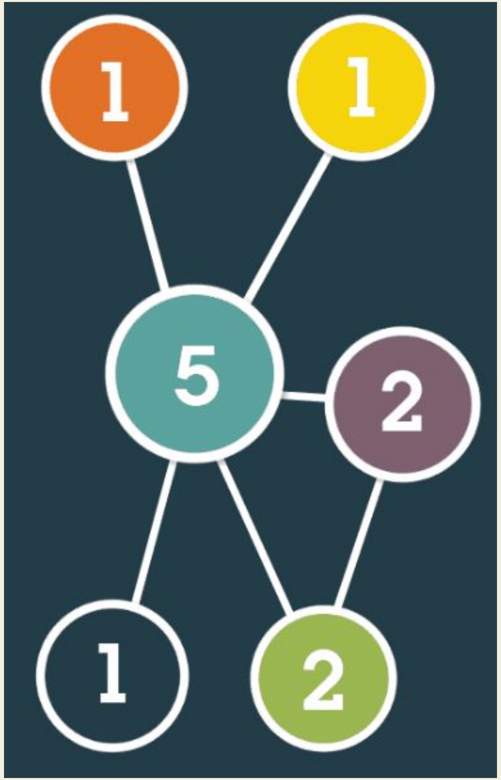
Crew Collaboration Network (15 movie sample):
nodes = people; edges = worked on same movie



Movie Similarity Network (50 movie sample):
nodes = movies; edges = shared collaborators

Methods - Centrality Measures

- Using **crew collaboration network** calculated node centrality measures to capture influence of collaborators
- Centrality Metrics Added:
  - **Degree Centrality:** extent of a contributor's collaborations
  - **Closeness Centrality:** how near a contributor is to others in the network
  - **Betweenness Centrality:** a measure of how often a contributor connects otherwise separate groups
- Integration into Film Dataset:
  - For each movie, aggregated collaborators' scores
  - *Mean* and *Maximum* values calculated



Methods - Regression

- **Two target variables:**
  - Revenue_Normalized (standardized box office performance)
  - Average_Rating (mean of critic and audience ratings)
- **Modeling setup:**
  - 80/20 train-test split
  - Ordinary Least Squares (OLS) via LinearRegression in scikit-learn
  - Evaluated with MSE, RMSE, MAE, and R^2
- **Incorporated network features:**
  - Added Degree, Closeness, and Betweenness Centrality (mean & max per film)
  - Centrality measures derived from cast & crew collaboration networks

Methods - Classification

- **Two binary targets:**
  - Success_Financial (e.g., ROI > 1 = success)
  - Success_Critical (e.g., top rating quartile = success)
- **Feature sets:**
  - Baseline model: traditional film features (e.g., budget, genre, runtime)
  - Extended model: added network centrality metrics (Degree, Closeness, Betweenness)
- **Modeling setup:**
  - Logistic Regression for interpretability and binary classification suitability
  - Stratified 80/20 train-test split to maintain class balance
  - Accuracy, Precision, Recall, F1-score for evaluation

Methods - GNN

Graph Convolutional Network (GCN) vs Graph Attention Network (GAT)

Target Variables

Revenue (Normalized): Standardized reported box office performance

- Financial success indicator

Average Rating: Mean of critic and audience ratings

- Critical success indicator

Modeling Setup

Created labels by dividing the dataset into quantiles based on target values

- **Revenue:** 4 labels
- **Rating:** 3 labels

80/10/10 train/validation/test dataset split

Model Selection

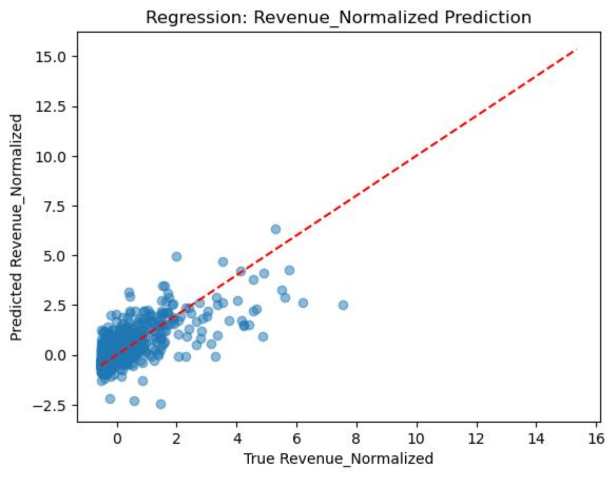
Hyperparameter Tuning: Explored combinations of

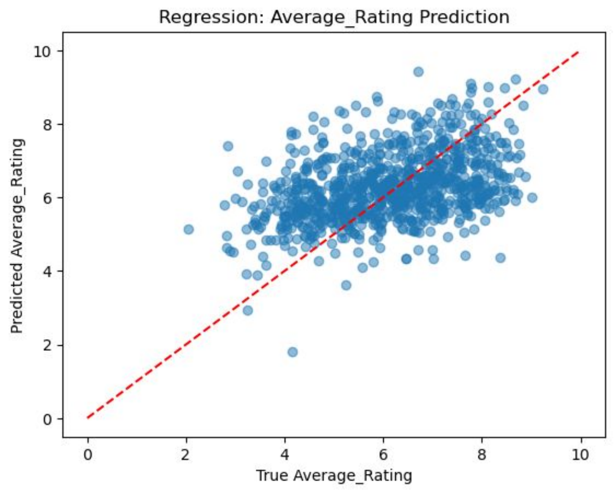
- Learning rate
- Hidden dimensions
- Weight decay
- Optimizer
- Dropout

Selection Criteria: Model configuration with lowest validation loss

Results – Regression

Target	Model Specification	MSE	RMSE	MAE	R^2
Revenue_Normalized	Baseline	0.5421	0.7363	0.4761	**0.4513**
Revenue_Normalized	With Centrality (All)	0.4853	0.6967	0.4639	**0.5088**
Revenue_Normalized	With Centrality (Max Only)	0.4811	0.6936	0.4573	**0.5130**
Average_Rating	Baseline	1.8077	1.3445	1.0871	**0.0476**
Average_Rating	With Centrality (All)	1.7052	1.3058	1.0599	**0.1016**
Average_Rating	With Centrality (Max Only)	1.6857	1.2983	1.0522	**0.1119**





Key Takeaways:

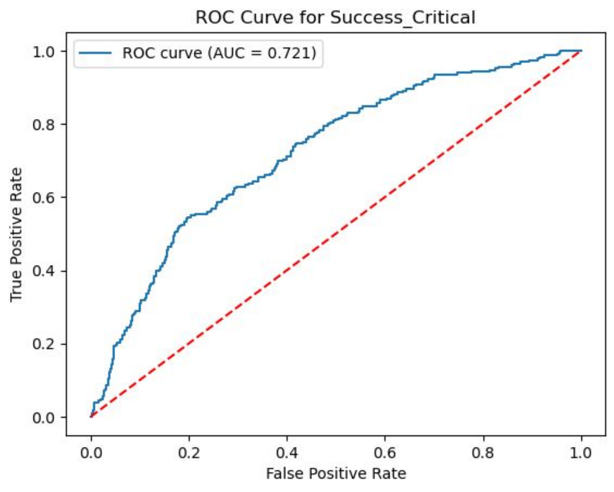
- ❖ Network features help explain financial performance
- ❖ Limited utility for predicting critical reception

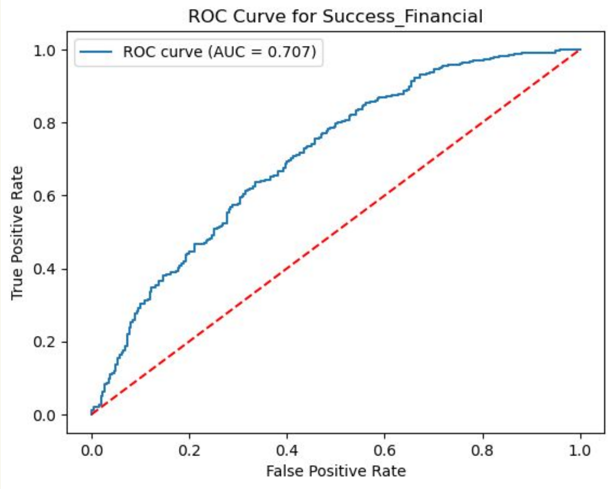
Results – Classification

Target	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Success_Critical	0.76	0.55	0.21	0.30	0.71
Success_Financial	0.76	0.77	0.95	0.85	0.71
Success_Critical	0.76	0.56	0.21	0.31	0.74
Success_Financial	0.75	0.78	0.90	0.83	0.71

Key Takeaways:

- ❖ Adding centrality measures yielded modest improvements
- ❖ Persistent challenge of detecting critically successful films
- ❖ Difficulty predicting critical acclaim compared to financial outcomes, even with network features





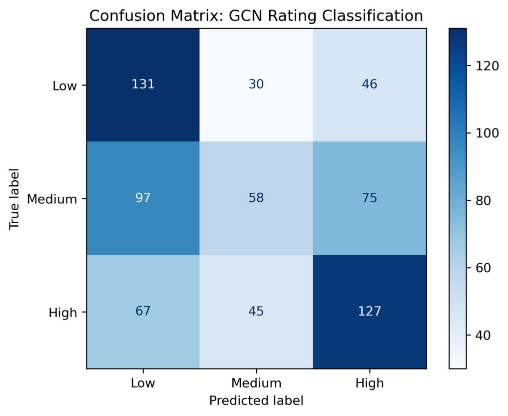
Results - GCN (Rating)

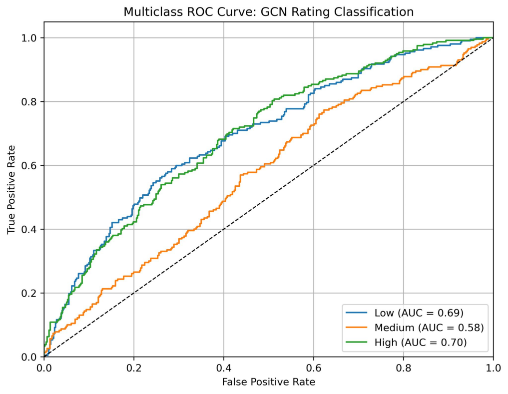
Final Model Configuration

Hidden Layer Size	[128]
Learning Rate	0.01
Weight Decay	0.001
Dropout Rate	0.2
Optimizer	Adam

Best Validation Loss: 0.9928

Test Accuracy: 46.75%



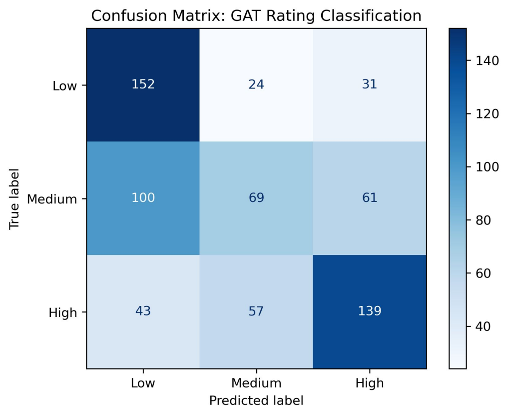


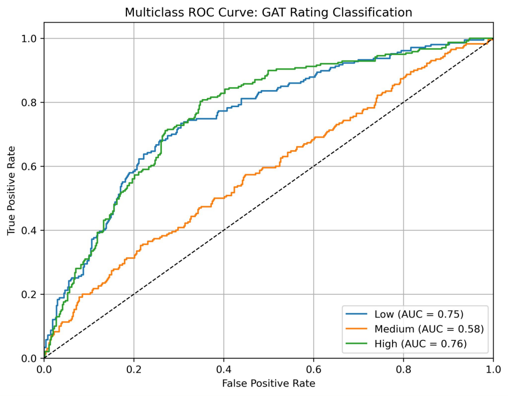
Results - GAT (Rating)

Final Model Configuration	
Hidden Layer Size	[128]
Learning Rate	0.01
Weight Decay	0.001
Dropout Rate	0.2
Optimizer	Adam

Best Validation Loss: 0.9445

Test Accuracy: 53.25%





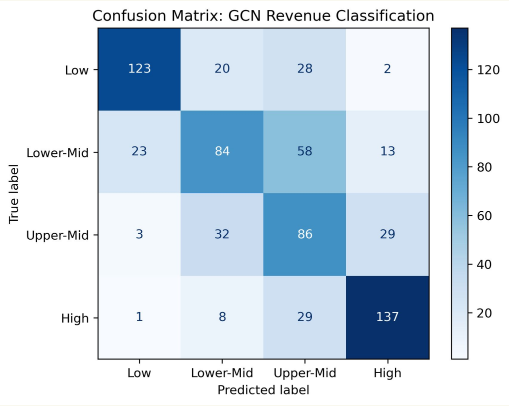
Results – GCN (Revenue)

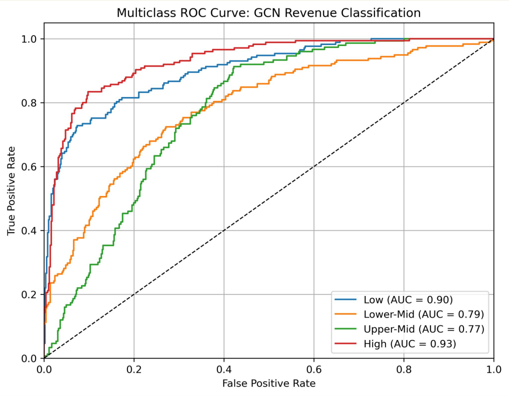
Final Model Configuration

Hidden Layer Size	[128]
Learning Rate	0.01
Weight Decay	0.001
Dropout Rate	0.1
Optimizer	Adam

Best Validation Loss: 0.8937

Test Accuracy: 63.61%



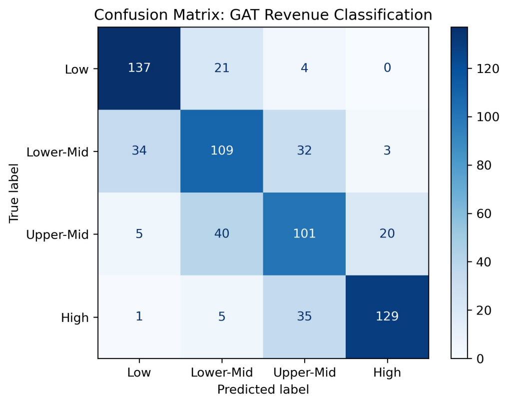


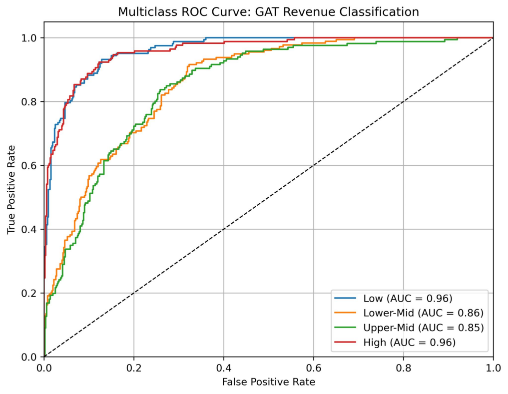
Results - GAT (Revenue)

Final Model Configuration	
Hidden Layer Size	[64]
Learning Rate	0.01
Weight Decay	0.001
Dropout Rate	0.1
Optimizer	Adam

Best Validation Loss: 0.7249

Test Accuracy: 70.41%





Results – GNN Overview

Classification Target	Model Type	Validation Loss	Accuracy
Rating	GCN	0.9928	46.75%
Rating	**GAT**	**0.9445**	**53.25%**
Revenue	GCN	0.8937	63.61%
Revenue	**GAT**	**0.7249**	**70.41%**

GAT outperforms GCN for both tasks

Limitations

1. **Lack of Temporal Dynamics:** A static graph cannot model trends, career trajectories, or shifts in genre popularity, which may be key predictors of success
2. **Arbitrary Buckets:** Classes might not reflect meaningful distinctions (6.9 vs 7.0 rating)
3. **Edge Weight Quality:** Edge weights may oversimplify complex relationships (hard to encode the importance or influence of each collaborator properly)

Future Directions

1. **Include Sentiment or Review Aggregates:** Incorporate NLP-based sentiment analysis of critic or audience reviews as features to better model audience perception
2. **Regression Instead of Classification Using GNNs:** Move from bucketing revenue into classes to predicting continuous revenue or ROI
3. **Apply to Other Forms of Media:** Expand the dataset to include other formats, such as TV shows, video games, etc

References

- Borsboom, D., Deserno, M. K., Rhemtulla, M., et al. (2021). Network analysis of multivariate data in psychological science. Nature Reviews Methods Primers, 1, 58.
- Chaudhary, S., Giri, S., & Gautam, B. (2025). Analyzing social networks of actors in movies and TV shows. In International Conference on Applied Algorithms. Cham: Springer Nature Switzerland.
- Gao, C., Zheng, Y., Li, N., Li, Y., Qin, Y., Piao, J., Quan, Y., Chang, J., Jin, D., He, X., & Li, Y. (2023). A survey of graph neural networks for recommender systems: Challenges, methods, and directions. ACM Transactions on Recommender Systems, 1(1), Article 1.
- Mondal, P., Das, A., Mishra, S., & Debnath, N. C. (2023). Graph convolutional neural network for multimodal movie recommendation. In Proceedings of the 38th ACM/SIGAPP Symposium on Applied Computing (SAC '23). ACM.
- Nesmaoui, R., Louhichi, M., & Lazaar, M. (2023). A collaborative filtering movies recommendation system based on graph neural network. Procedia Computer Science, 220, 456–461.
- Vitelli, M. (2015). Predicting box office revenue for movies. Stanford Network Analysis Project.
- Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. Nature, 393(6684), 440–442.
- Yang, H., Li, J., & Jahng, S. (2024). The application of knowledge graph convolutional network-based film and television interaction under artificial intelligence. IEEE Access.
- Zhang, Z., Meng, Y., & Xiao, D. (2024). Prediction techniques of movie box office using neural networks and emotional mining. Scientific Reports, 14, 21209.
- The Movie Database (TMDB), www.themoviedb.org/?language=en-US. Accessed 4 Aug. 2025.
- "OMDb API." OMDb API - The Open Movie Database, www.omdbapi.com/. Accessed 4 Aug. 2025.

Questions