

# From Text to Graph: Combining LLMs, TV Tropes, and Node2Vec for Movie Recommendation

Ngan Thao Huynh  
Department of Computer Science  
Georgia Southern University  
Savannah, United States  
th25692@georgiasouthern.edu

**Abstract**—This project explores the use of TV Tropes data to enhance movie recommendation systems by focusing on shared storytelling patterns. By leveraging language models, dimensionality reduction techniques, and graph-based embeddings, we aim to capture relationships between tropes and media. Our approach moves beyond traditional genre-based recommendations by analyzing structural connections between storytelling elements. This work lays the foundation for a more interpretable and narrative-driven recommendation system, offering insights into how tropes influence audience engagement and movie similarity.

## GitHub Repository:

[https://github.com/nganhhuynhgs/DS\\_ML\\_Project\\_colab-integration](https://github.com/nganhhuynhgs/DS_ML_Project_colab-integration)

**Index Terms**—graph neural network, movie recommendation system

## I. INTRODUCTION

TV Tropes is a collaborative database that catalogs recurring storytelling devices used across media such as movies, TV shows, and books. Unlike traditional categorization methods, tropes offer nuanced insights into why audiences are drawn to particular narratives.

Recommending movies based on shared tropes presents challenges due to the high-dimensional nature of textual data and the complex relationships between tropes and media. Existing recommendation systems often rely on predefined genres, which fail to capture the narrative intricacies that tropes provide.

This project aims to:

- Generate embeddings of TV Tropes using pre-trained large language models (LLMs).
- Apply UMAP for dimensionality reduction to enhance computational efficiency and visualization.
- Construct a graph-based recommendation system using Node2Vec to capture structural relationships.
- Recommend movies based on user-selected favorite movies by leveraging shared trope embeddings.

## II. BACKGROUND AND RELATED WORK

This section reviews current technologies used in modern movie recommendation systems and their limitations. It also examines the techniques we will apply and how they can enhance the system.

## III. ALGORITHMS AND EXPERIMENTS

This section outlines our architecture and the experiments conducted to enhance our algorithms.

## IV. CONCLUSION

The project demonstrates how leveraging TV Tropes data, combined with LLM embeddings, dimensionality reduction, and graph-based techniques like Node2Vec, can provide a novel and interpretable approach to movie recommendation, paving the way for future improvements in narrative-driven recommendation systems.

## REFERENCES

- [1] A. Grover and J. Leskovec, “Node2vec: Scalable Feature Learning for Networks,” in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2016.
- [2] L. McInnes, J. Healy, and J. Melville, “UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction,” 2018.
- [3] “Sentence Transformers Documentation,” Available at: <https://www.sbert.net>.