Movie Casting Problems

Alireza Torabian, Hossein Naderibeni

EECS 6414 Data Analytics and Visualization | 2021

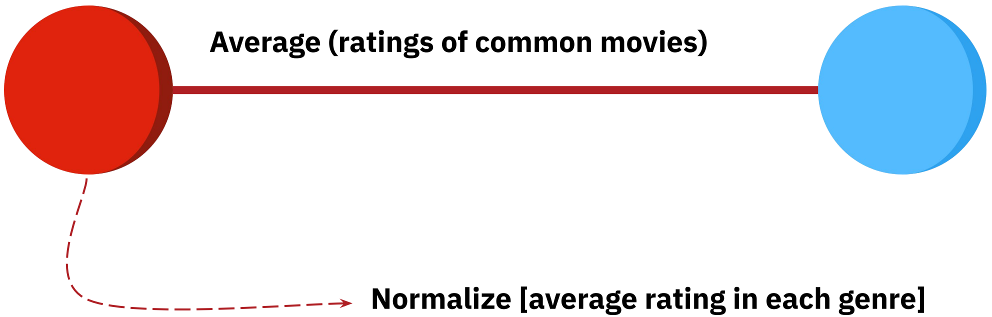
YORK U

Motivation

> Who is the best alternative actor?
> Which directors and actors fit together?
> Which actors perfectly match each other to play the main roles of a movie?
> How to compare two actors group to choose them for a movie?

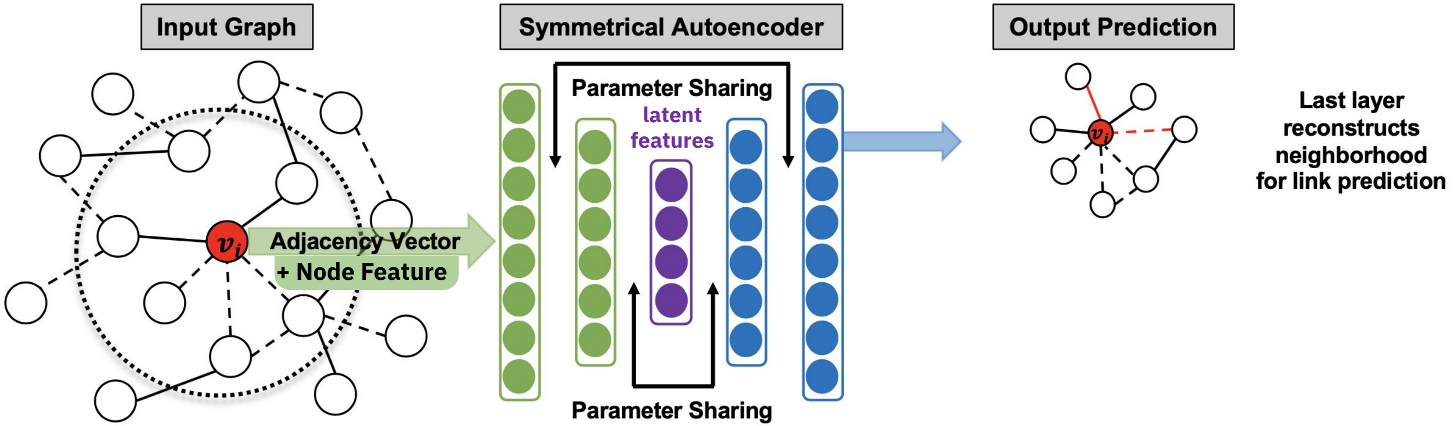
Actors Network

- Using the main actors (first five actors) of each movie
- A weighted graph with node features

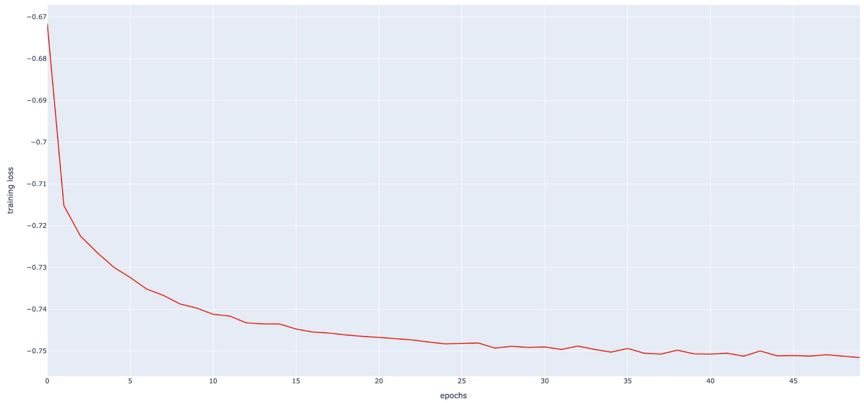


Graph Autoencoder on Actors Network

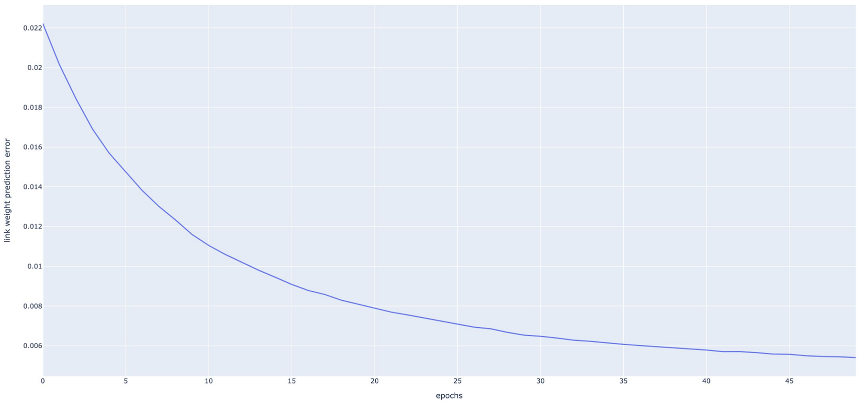
- Learning a joint representation of both local graph structure and node features
- MSE on link weight prediction task 0.005406 (for edge weights from 0. to 1.)



Graph Autoencoder on Actors Network

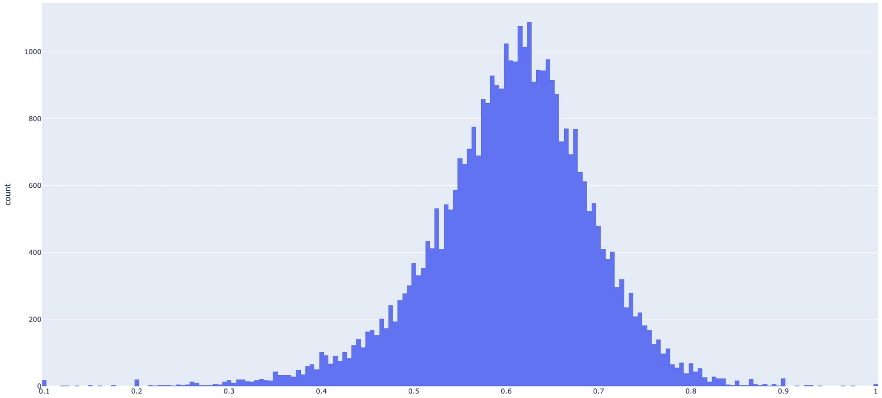


Training Loss

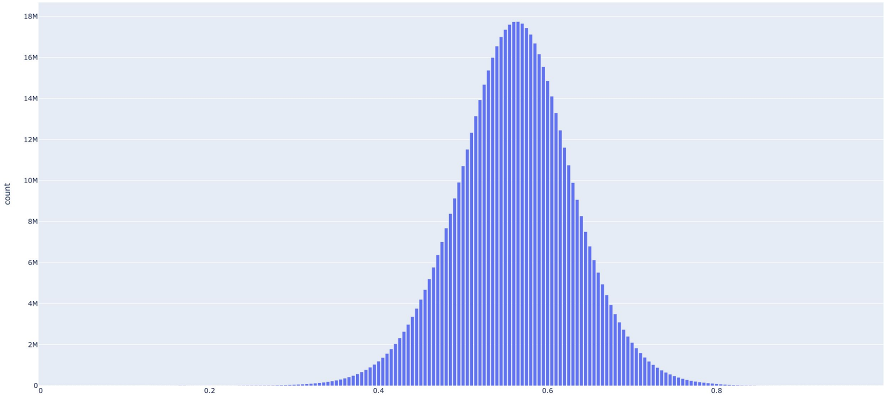


Link weight prediction error

Graph Autoencoder on Actors Network



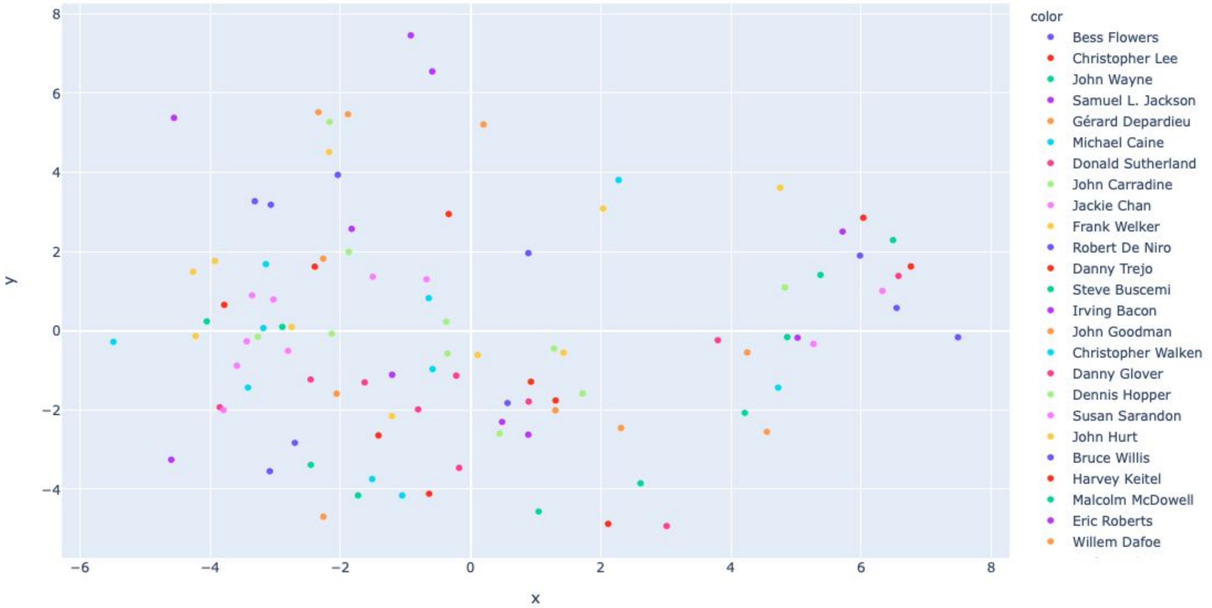
Actors Network Weights

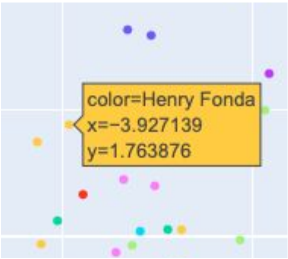


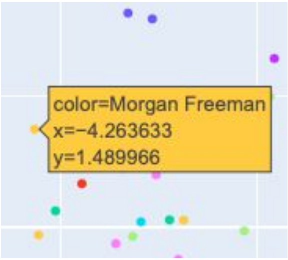
Predicted Actors Network Weights

Latent Vector Space Visualisation

- High Dimensionality Problem
- Principal Component Analysis
- Good Enough?



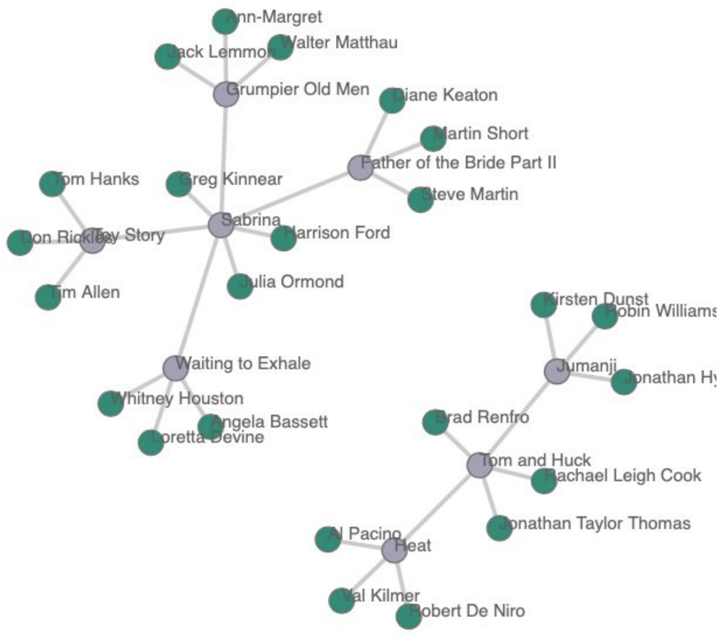




Alternative Actor Problem

First Approach Using Movie Similarity - Midterm Report

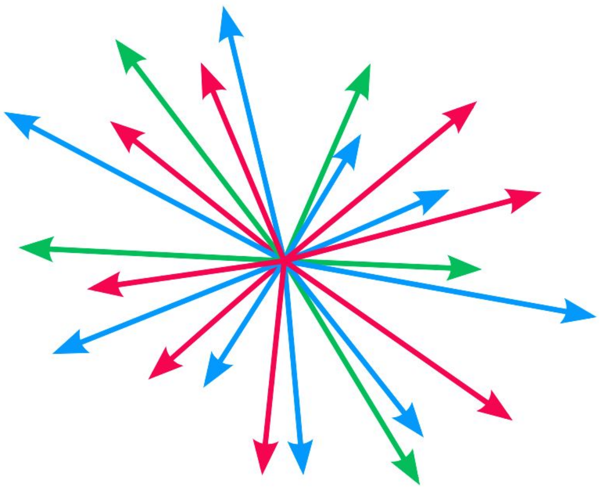
- If a selected actor is not available for casting, who is a good replacement?
- Recommend alternative actor based on the most common similar movies
- Two movies are similar if they have close content (genres & keywords)
- Measuring Distance using Cosine Similarity



YORK U

Alternative Actor Problem
Second Approach using Latent Vector Space

- In our graph model we assign a vector to each actor in a 128-d space
- Close vectors show similarity among encoded properties
- Distance measure? Algorithm?

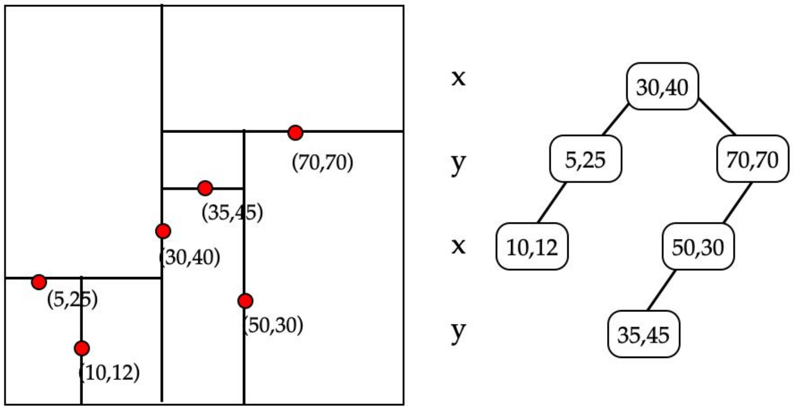


Nearest Neighbour Problem

- In our approach we decide Euclidean Distance
- Queries: Add some point in a k-d space

  Find nearest neighbors to a given point
- K-d tree with Amortized O(log n) Steps

insert: (30,40), (5,25), (10,12), (70,70), (50,30), (35,45)



YORK U

Movie Cast Rating

- To evaluate the quality of a group of actors
- Using the predicted weights and node features in the output of the GAE
- For a movie we find the following values for each pair of actors according to its genres:
  - *s(a,b): how good is co-acting of actors a & b in our graph model*
  - *rating_g(a): how good does actor a star in movies with genre g*

$$w(a,b) = \sum_{g \in genres} rating_g(a) + rating_g(b)$$

- Cast rating is the weighted average of *s* with the weights *w*:

$$cast\ rating = \frac{\sum_{a,b \in cast} s(a,b)\, w(a,b)}{\sum_{a,b \in cast} w(a,b)}$$

Movie Cast Rating

> Evaluating based on several well-known movies with perfect cast [1]

The Big Lebowski: 0.647, percentile 93%

The Godfather: 0.658, percentile 96%

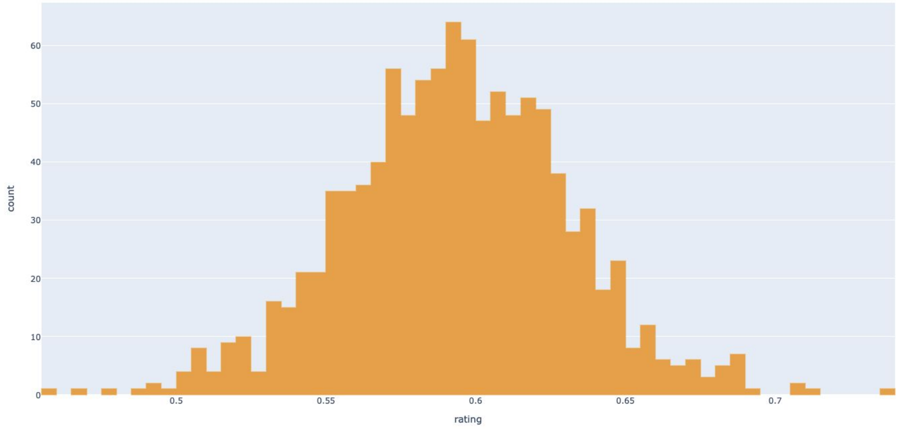
12 Angry Men: 0.597, percentile 55%

The Departed: 0.661, percentile 96%

Inception: 0.666, percentile 97%

Pulp Fiction: 0.599, percentile 56%

American Hustle: 0.696, percentile 99%



Cast rating on random movies

1 We have used the lists in the webpages thetoptens.com/greatest-movie-casts-of-all-time, and tasteofcinema.com/2018/10-great-movies-with-the-most-perfect-cast

Comparison between two Approaches for Alternative Actors

> Problem for Evaluation: No training data on how good two actors are as alternatives

> Our Ideas to solve this

- Compute movie cast score replacing alternative actor and the original cast and compare results of the approaches
- Using intuitive data based on lists on the Internet

1.Using Movie Similarity among common movies played

Pros: We have a sense of why these two actors are recommended and can check if it is accurate or not

Cons: Very slow

Some results:

- Similar Movies
  Balto -> Toy Story
  Nixon -> Carrington
  Cutthroat Island -> Sudden Death
  Casino -> Dangerous Minds

- Alternative Actors
  Mel Brooks ? Steve Martin, Martin Short, Diane Keaton
  Al Pacino ? Julianne Moore, Sylvester Stallone
  Tom Hanks ? Kevin Bacon, Bob Hoskins, Bridget Fonda

2.Using Distance in latent vector space among actors

Pros: Better results with cast scoring and faster queries

Some results:

- Alternative Actors
  Mel Brooks ? Jack Lemmon, Anthony Perkins
  Al Pacino ? Trevor Howard, Shirley MacLaine
  Tom Hanks ? Tony Curtis, Peter Lawford

Co-Star Suggestion

- Using the predicted weights between the actors.
- Evaluating rating predict among some successful duos [1]

Chris Evans,Scarlett Johansson: 0.6687

Anne Hathaway,Jake Gyllenhaal: 0.7095

Ryan Gosling,Emma Stone: 0.6432

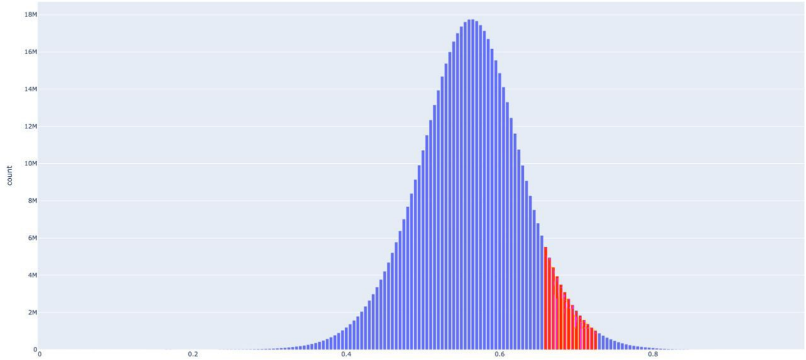
Steve Carell,Christian Bale: 0.6629

Will Ferrell,John C. Reilly: 0.5880

Bradley Cooper,Jennifer Lawrence: 0.6485

Johnny Depp,Helena Bonham Carter: 0.6566

Al Pacino,Robert De Niro: 0.6995



Rating Predictor Distribution

1 We have used the lists in the webpages screenrant.com/iconic-actor-duos-ranking/, and brightside.me/wonder-films/15-acting-duos

Conclusion

- Based on movie-related information provided on the web, we have evaluated our models.
- The mean-squared-error on the link weight prediction task was close to zero, which leads us to high quality results in the co-star suggestion task.
- Alternative Actor Problem approaches have their own pros and cons. Reducing the problem to a well studied problem like nearest neighbor problem leads us to a practical solution.
- We can improve our model and results using more attribute of movies and actors.

References

[1] Rounak Banik. 2018. The Movies Dataset. https://www.kaggle.com/rounakbanik/the-movies-dataset

[2] Bernard Kurka. 2019. Building a movie recommender system. https://medium.com/@bkexcel2014

[3] Stefano Leone. 2020. IMDB Movies Extensive Dataset. https://www.kaggle.com/stefanoleone992/imdb-extensive- dataset

[4] Xin Li and Hsinchun Chen. 2009. Recommendation as Link Prediction: A Graph Kernel-Based Machine Learning Approach. In Proceedings of the 9th ACM/IEEE-CS Joint Conference on Digital Libraries (Austin, TX, USA) (JCDL '09). Association for Computing Machinery, New York, NY, USA, 213–216.

[5] Xin Li and Hsinchun Chen. 2013. Recommendation as link prediction in bipartite graphs: A graph kernel-based machine learning approach. Decision Support Systems 54, 2 (2013), 880–890.

[6] Phi Vu Tran. 2018a. Learning to make predictions on graphs with autoencoders. In 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA), pages 237–245. IEEE.