Barcelona School of Economics

Data Science Methodology Program

21D009 Networks: Concepts and Algorithms

Movie recommendations using networks

Authors:
Codd, Jonny
Chen, Joshua
Gallegos, Rafael
Pérez, Carlos

Professors:
Milán, Pau
Komander, Björn

December 20^{th}, 2023

Contents

List of Figures

List of Tables

1 Introduction

The travelling salesman problem (TSP) is an archetypal problem in linear programming. The setup tells the reader to imagine a salesman who must visit n cities, then asks them to find the shortest path the salesman can take, visiting each city exactly once before returning to the starting city given the distance between all pairs of cities. Its deceptively simple formulation lies in stark contrast to the vast computational expense required to solve it; it is NP-hard, and despite nearly one-hundred years of literature, we rely on heuristic algorithms to find good solutions.

2 Literature Review

3 Dataset

MovieLens is a recommender system that was developed by GroupLens, a computer science research lab at the University of Minnesota. The goal of this challenge is to recommend movies to its users based on their movie ratings. Group Lens offers datasets of different sizes and their datasets are widely used in research and teaching contexts.

The selected dataset consists mainly on two files: movies.csv and ratings.csv. Movies dataset has 9742 unique films and a column indicating the genres of the film. All possible genres are: 'Romance', 'Musical', 'Children', 'Documentary', 'Sci-Fi', 'Film-Noir', '(no genres listed)', 'Crime', 'Mystery', 'Drama', 'Western', 'Fantasy', 'Animation', 'Thriller', 'War', 'Action', 'Adventure', 'IMAX', 'Comedy', 'Horror'. The number of movies per genre is represented in Figure 3.2.

Ratings dataset consists of 100836 ratings with 610 unique users that rated 9724 movies. As it can be observed in Figure **??**, the ratings from users are high, which suggests that users tend to enter or register their rating on movies that they probably have liked.

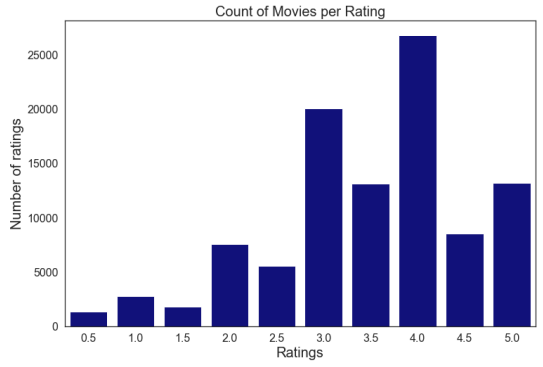

Figure 3.1: Count of Movies per Rating

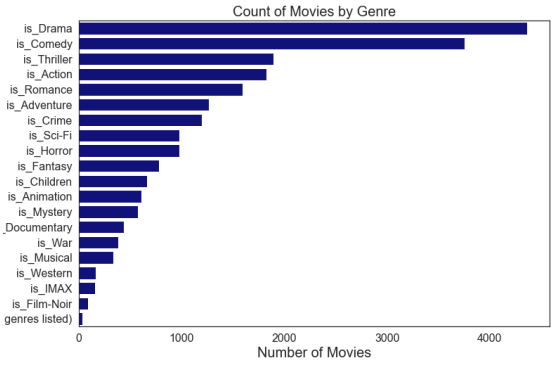

Figure 3.2: Count of Movies per Genre

7 Conclusions