

Ghanim Alghanim
Ali Altamimi

IMDb Rating Predictions

IMDb movies dataset will be used for this project. The goal is to apply data preparation knowledge to prepare the data for the model. The model that will be used is linear regression to predict the rating based on age, gender and genre. The dataset is not found online, therefore, scraping the data from IMDb website directly will be mandatory for the analysis. The goal is to scrap the top 250 rated movie's important features. Moreover, achieving that goal multiple pages will be visited which theoretically might add up to 500 pages to obtain all the features necessary for this project.