Deep learners: Topic proposal

IMDB Feature Film Analysis

Dataset: <https://www.kaggle.com/stefanoleone992/imdb-extensive-dataset>

Our objective is to predict the movie ratings by analyzing the popularity of movies and the aspects that make them successful by looking at relationships between user movie ratings and various movie features.

Our dataset comprises 85,855 movie instances and 297,705 cast instances corresponding to those movies. The attributes available include the movie title, year, date published, genre, duration, country, language, director, actor, gender, age, ratings, etc.

Our research on the dataset will answer the following questions:

* What model would be built to predict movie ratings, with achieving higher accuracy?
* What features from the dataset would be selected to build the model?
* Data reduction? unbalanced data check?
* Are the movie ratings normalized?
* Train and test data?