Movie Score Prediction

Karl Soosalu, Mark Alexander, Marten Mark

Introduction

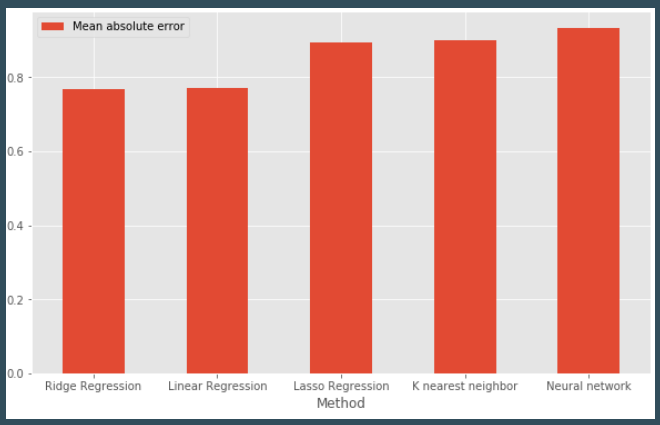
New movies are always rated on the internet. Some have higher ratings than others so we thought what if we could predict the movie's score by its details before it is released to the public. In this project, we sought to find if we really can do that.

Objectives

Our main objective was to predict the film's score through its attributes. Secondly we wanted to find out the best method for the prediction and lastly we wanted to find out the biggest factor impacting the movie's score.

Results

What we found out was that the best method for predicting the score was using the ridge regression and its MAE was 0.768. There were no noticeable individual factors impacting the movie's score more than others.



Conclusion

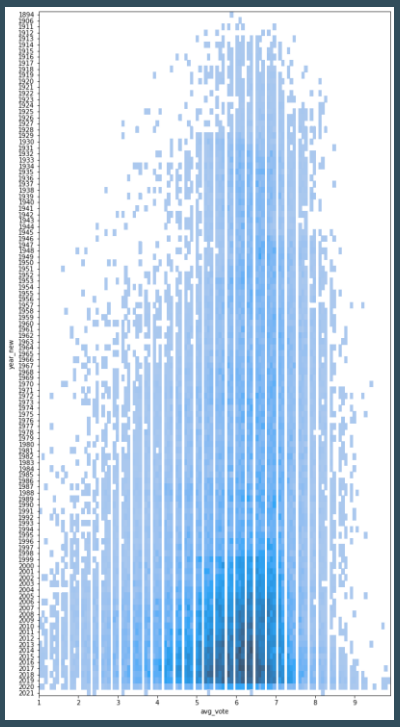
Yes, we can predict a movie's score by its attributes to a certain degree.

Data

In this project, we used the data from kaggle which was scraped from IMDb-s website, to predict the score of the movie it would get on IMDb. The data contains details about over 85 500 movies such as their release year, genre, ratings etc.

Methods

The methods used to predict the movie's score were linear regression, ridge regression, lasso regression, K nearest neighbor regression and neural network.



Sources:

https://www.kaggle.com/stefanoleone992/imdb-extensive-dataset?select=IMDb+movies.csv

github.com/MarkMarten/MSP