Individual Report

Jinshun Su

I. Introduction

This study is aimed to draw a prediction of the rating of a movie based on its properties. With this model applied to commercial movies shall one be able to give an estimation on the rating those movies may generate. Considering the study's predictive nature, a regression analysis is to be expected. Consequently, the tasks lie in finding out the potential factors that contribute to the rating of a movie and building a regression model that predicts the outcome as satisfactory as possible when compared to the targets of test inputs.

II. Individual Work

1) Based on the general code made by my group member Maeshal Hijazi, I utilized **sklearn.standardizer** to standardize the values of input features. The code is shown below:

```
from sklearn.preprocessing import StandardScaler
sc = StandardScaler().fit(X_train)
X_train2 = sc.transform(X_train)
X_test2 = sc.transform(X_test)
```

2) When using MLPREGRESSOR, I modified the algorithm with 'activation=relu'.

3) I tested the model in different cases and obtain the value of mean squared error at each case. Case I includes two input features of runtime and votes, case II involves two input features runtime and metascores, and Case III consists of two input features votes and metascores. In case IV, three input features (runtime, votes, and metascores) are taken into account. The values of mean squared error are shown below:

	Case I	Case II	Case III	Case IV
Values	0.586	0.589	0.518	0.469

III. Conclusion

Based on the study cases, the model was relatively satisfactory when all properties (runtime, metascore and votes) are considered as factors to contribute to the rating of a movie.