ANALISIS PREDIKSI *THE MOVIE DATABASE RATING FILM* DENGAN *ENSEMBLE LEARNING* MENGGUNAKAN METODE *RANDOM FOREST REGRESSION*

Nuravifah Novembriana Marpid

ABSTRAK

Industri film sudah menjadi industri yang sangat menguntungkan. Namun masa pandemi covid-19, industri film mengalami dampak tidak menguntungkan dengan tertahannya jadwal penayangan film-film baru, banyak bioskop-bioskop dilarang untuk beroperasi sehingga tutup total dan sulitnya mendapat izin untuk melakukan proses pembuatan film. Untuk bertahan dalam industri ini dari dampak pandemi covid-19, diperlukan pertimbangan beberapa faktor seperti penjadwalan film yang sesuai, cara promosi yang tepat sasaran bahkan dengan menggunakan prediksi untuk mengetahui perkembangan pada pasar. Prediksi keberhasilan film sangat membantu untuk mengetahui peringkat keberhasilan dan kualitas film yang akan dirilis. Maka dibuatlah analisis prediksi rating film untuk mendapatkan metode prediksi yang akurat terhadap film yang akan dirilis. Pada penelitian ini diterapkan dengan Ensemble Learning menggunakan metode Random Forest Regression. Data yang digunakan pada penelitian ini adalah *The Movie Dataset* (TMDB). Penelitian ini menggunakan 6134 *rows* dan 4 *attributes* setelah melakukan proses *data preprocessing*. Dari hasil pengajian yang dilakukan dengan metode Random Forest Regression setelah pembagian data 80:20 didapatkan nilai R^2 *Score* 86%, nilai MSE 12%, nilai RMSE 35%, nilai MAE 22% dan nilai *10-Fold Cross Validation* sebesar 85%.

Kata Kunci : *10-Fold Cross Validation,* Prediksi, *Random Forest Regression, Rating Film*, TMDB

ANALYSIS OF THE MOVIE DATABASE FILM RATING PREDICTION WITH ENSEMBLE LEARNING USING RANDOM FOREST REGRESSION METHOD

Nuravifah Novembriana Marpid

ABSTRACT

The film industry has become a very profitable industry. However, during the COVID-19 pandemic, the film industry experienced an unfavorable impact with new film screening schedules being delayed, many cinemas being prohibited from operating so they were completely closed and it was difficult to obtain permission to carry out the filmmaking process. To survive in this industry from the impact of the COVID-19 pandemic, it is necessary to consider factors such as appropriate film scheduling, targeted promotional methods and even using predictions to determine developments in the market. Predicting the success of films is very helpful in knowing the success ranking and quality of films that will be released. So a film rating prediction analysis was made to get an accurate prediction method for films that will be released. In this research, Ensemble Learning was implemented using the Random Forest Regression method. The data used in this research is The Movie Dataset (TMDB). This research uses 6134 rows and 4 attributes after carrying out data preprocessing. From the results of the study carried out using the Random Forest Regression method after dividing the data 80:20, the R2 Score value was 86%, the MSE value was 12%, the RMSE value was 35%, the MAE value was 22% and the 10-Fold Cross Validation value was 85%.

***Keywords :** 10-Fold Cross Validation, Prediction, Random Forest Regression, Rating Film, TMDB*