Model Development Phase Template

Date	23 September 2024
Team ID	LTVIP2024TMID24986
Project Title	Movie Box Office Gross Prediction using Machine Learning
Maximum Marks	5 Marks

Feature Selection Report Template

In the forthcoming update, each feature will be accompanied by a brief description. Users will indicate whether it's selected or not, providing reasoning for their decision. This process will streamline decision-making and enhance transparency in feature selection.

Feature	Description	Selected (Yes/No)	Reasoning
Movie_ID	Unique identifier for each movie	No	Not required for predicting box office revenue, as it serves only as an identifier.
Budget	The production budget of the movie	Yes	Critical factor influencing revenue; higher budgets often correlate with higher potential earnings.
Genres	Movie genres (e.g., action, drama, comedy)	Yes	Different genres can attract different audiences, impacting box office performance.
Cast	List of actors in the movie	Yes	Well-known actors often drive higher audience attendance, influencing box office revenue.

Director	The director of the movie	Yes	Renowned directors can attract larger audiences and potentially increase revenue.
Release Date	The date the movie was released	Yes	Timing of release (e.g., holiday seasons) can significantly affect box office success.
Production Company	The company that produced the movie	Yes	Large, well-known production companies may have better distribution networks, impacting revenue.
Runtime	The duration of the movie	Yes	The length of the movie can influence audience retention and showtimes, which may affect revenue.
Language	The primary language in which the movie was made	Yes	Language can determine the size of the target audience and accessibility in different regions.
Budget to Revenue Ratio	The ratio of budget to revenue achieved by the movie	No	Although relevant for analysis, it's not a predictor for the model, as it's derived from the target variable (revenue).
Revenue	The total box office revenue generated by the movie	Yes	The target variable for predictive modeling – essential for the project's goal.