Title: Movie Success Prediction and Sentiment Study

Abstract:

This project predicts movie success using IMDB/Kaggle datasets and analyzes viewer sentiment from user reviews. A regression-based predictive model combined with VADER sentiment analysis helps understand the correlation between reviews and box office performance.

Introduction:

Movies generate large volumes of audience feedback and financial data. Predicting success using data-driven approaches enables better decision-making in production and marketing. This study investigates box office prediction and sentiment trends using real-world movie datasets.

Tools Used:

• Anaconda – Environment & package management

• Jupyter Notebook – Interactive development

• Python – NLTK, VADER, Scikit-learn

• Excel – Data preprocessing

• Pandas, Matplotlib – Data handling & visualization

Steps Involved:

1. Imported and cleaned IMDB/Kaggle movie data (title, genre, budget, box office, reviews).

2. Performed exploratory data analysis in Excel/Python using Jupyter Notebook.

3. Applied VADER sentiment analysis on user reviews to compute sentiment scores.

4. Built a regression model using features like budget, genre, rating, sentiment score.

5. Predicted box office success and analyzed genre-wise sentiment trends.

Conclusion:

The model showed meaningful correlations between review sentiment and movie performance. Sentiment trends revealed audience preference patterns across genres. This approach can support movie producers in predicting success and improving strategic decisions.