International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 12 Issue: 2
Article Received: *25 November 2024* ***Revised:*** *12 December 2024* ***Accepted:*** *30 December 2024*

Modeling Collaborative Effects in Bollywood Film Profitability Using Synergy Index

Partha Shankar Nayak [1]*, G. Ravindra Babu [2]

[1] Chaudhary Charan Singh University, Meerut, India e-mail: psnayakitiiests@gmail.com

[2] Chaudhary Charan Singh University, Meerut, India e-mail: pandeyji@gmail.com

* psnayakitiiests@gmail.com

Abstract: This study introduces a novel feature engineering framework for analysing the influence of collaborative relationships and economic factors on film profitability in the Bollywood film industry. Traditional models often treat actor and director effects independently, ignoring the relational impact of specific collaborations, and frequently fail to account for the temporal depreciation of currency value in longitudinal datasets. To address these gaps, we propose Synergy Index, a metric that quantifies the joint performance of each actor–director pair relative to their individual historical averages. Furthermore, to ensure financial consistency across decades, we implement an inflation-adjustment mechanism that normalizes production budgets to a common base year using historical CPI data. The model additionally incorporates robust handling of multi-actor films and utilizes Bayesian-style shrinkage to avoid overfitting in sparse cases. We evaluate this framework using five state-of-the-art regression algorithms (Random Forest, XGBoost, LightGBM, CatBoost, and HistGradientBoosting). Feature importance analysis demonstrates that the Synergy Index consistently emerges as a significant predictor of financial success alongside normalized budget constraints. The methodology provides new insights into collaborative dynamics in Bollywood and forms a foundation for advanced, economically adjusted relational modelling in film analytics.

Keywords: Bollywood box-office prediction; Actor–director collaboration (synergy); Inflation-adjusted financial modelling; Machine learning–based feature engineering

1. Introduction

The Indian film industry produces more than 1,500 films annually, with Bollywood contributing the largest share. Predicting box-office outcomes and understanding the determinants of Return on Investment (ROI) have attracted increasing academic and commercial interest. While prior research incorporates variables such as genre, budget, actor popularity, and marketing features, limited attention has been paid to collaborative interactions among film participants.

Actor–director partnerships have long been recognised as influential within the industry. Classic Bollywood collaborations such as Yash Chopra–Shah Rukh Khan, Rajkumar Hirani–Aamir Khan, or Sanjay Leela Bhansali–Deepika Padukone exhibit patterns of repeated commercial success. However, existing modelling approaches treat actor and director effects independently or include them merely as categorical variables, which cannot capture relational patterns.

To address this gap, we introduce the Synergy Index, a data-driven metric that quantifies the joint impact of specific actor–director collaborations. This index not only reflects historical collaboration performance but also adjusts for actor and director baselines, thereby identifying pairs that consistently outperform or underperform expectations. Furthermore, our methodology handles multiple actors per movie, a challenge rarely addressed in film analytics literature.

2. Literature Survey

The prediction of movie success has evolved from traditional econometric models to sophisticated Machine Learning (ML) frameworks. Existing literature can be broadly categorized into two streams based on the data sources utilized: social media-driven sentiment analysis and metadata-based pre-release prediction.

2.1 Social Media and Sentiment Analysis

With the rise of Web 2.0, a significant portion of research focused on harnessing User-Generated Content (UGC) to forecast box office revenue. The seminal work by Asur & Huberman (2010) [1] demonstrated that the volume of tweets could predict box office outcomes with high accuracy, establishing social media buzz as a critical indicator of public interest. Following this, researchers explored sentiment analysis to gauge audience reaction. Quader et al. (2017) [2] evaluated seven classification techniques, highlighting that sentiment scores derived from platforms like Twitter and YouTube often correlate with commercial success.

However, these approaches face inherent limitations regarding timing. As noted by Lash & Zhao (2015), social media data is often only available after marketing campaigns have begun or immediately prior to release [3]. While Kim (2020) successfully applied ML to predict movie performance [4], such models remain highly volatile and dependent on the vagaries of internet trends, making them less suitable for pre-production investment decisions where reliable historical data is paramount.

2.2 Metadata and Pre-release Attribute Analysis

To address the volatility of social media, researchers have turned to static, pre-release metadata such as genre, budget, and star power. Upadhyay (2018) and Lee (2016) utilized data mining techniques to analyze these fundamental attributes, arguing that the structural properties of a film are the primary drivers of its financial destiny [5].

Recent studies have expanded this to include deep learning and statistical methods. Zheng (2024) and Sharma et al. (2021) investigated the effects of pre-released attributes on gross revenue, confirming that factors like budget and screen count remain dominant predictors [7]. Page et al. (2013) further extended this by analyzing Wikipedia activity as a proxy for pre-release interest, bridging the gap between metadata and user engagement [9].

2.3 The Bollywood Context

The Indian film industry (Bollywood) presents a unique market structure where "Star Power" often dictates initial box office draw more significantly than in other markets. Verma & Verma (2020) conducted a comparative analysis of supervised learning algorithms specifically for Bollywood, identifying that specific combinations of cast and genre yield distinct ROI patterns [10].

Similarly, Masih & Ihsan (2019) attempted to quantify success using Academy Awards data [11]. While their study provided insight into critical acclaim, it highlighted a divergence between "critical success" (awards) and "commercial success" (ROI), which is the primary concern for producers. Other works, such as Antony & Francis (2022) [12] and general surveys on Bollywood prediction (ResearchGate, 2021) [13], reinforce that while basic features like 'Actor Popularity' are widely used, they are often calculated using simple metrics like follower counts or past hits, rather than rigorous financial performance indicators.

2.4 Research Gap and Motivation

Despite the extensive literature on movie profitability, substantial gaps remain, particularly regarding the quantification of human capital in filmmaking.

First, most existing studies utilize "Actor Popularity" based on social media following or binary hit/flop counts. There is a scarcity of research that calculates Actor Average ROI and Director Average ROI as continuous, financially-grounded variables derived from historical investment data.

Second, film production is inherently collaborative. While current models account for the presence of a star actor or a famous director, they fail to account for the interaction between them. The literature rarely quantifies the "chemistry" or Synergy Index between an Actor and Director, nor does it sufficiently weight the Collaboration Count (how often they have worked together successfully).

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 12 Issue: 2
Article Received: *25 November 2024* ***Revised:*** *12 December 2024* ***Accepted:*** *30 December 2024*

This study addresses these gaps by shifting focus from disparate individual features to collaborative financial metrics. By integrating Synergy Index and Collaboration Count alongside traditional metadata (Budget, Genre, Release Year), this research aims to build a more robust, ROI-centric prediction framework for the Bollywood industry.

3. Methodology

This study adopts a data-driven machine learning framework to identify and analyze features influencing the Return on Investment (ROI) of Bollywood movies. The methodology focuses on systematic feature engineering, with particular emphasis on modeling collaborative interaction effects between key creative stakeholders.

3.1 Data Preprocessing: Inflation Adjustment

One of the critical challenges in analyzing financial data over an extended period is the fluctuation in currency value. In the context of Bollywood, the production cost (Budget) is a significant predictor of ROI. However, due to inflation, the monetary value of the Indian Rupee (INR) has depreciated over time. For instance, a budget of INR 100 million in 2010 commanded a much higher purchasing power—allowing for better production quality and star casts—than the same nominal amount would in 2025. These nominal values, without adjustment, would mislead the regression algorithms by treating numerically identical budgets from different decades as equivalent. To address this temporal bias, we normalized all budget figures to a common base year level. We obtained the historical annual inflation rates (r) for India from publicly available economic databases viz. WorldData.info, 2025 and RateInflation.com, 2025 **Error! Reference source not found.Error! Reference source not found.**.

The adjusted budget was calculated using the cumulative compound inflation formula shown in Equation 1:

$$\text{Budget}_{adj} = \text{Budget}_{orig} \times \prod_{t=Y_{release}}^{Y_{base}} (1 + r_t) \qquad (1)$$

Where:

- Budget_{adj} is the normalized budget value used in the dataset.
- Budget_{orig} is the raw budget reported in the release year.
- r_t is the annual inflation rate for year t (expressed as a decimal, e.g., $5\% = 0.05$).
- The product term $\prod$ accumulates the inflation effects from the release year ($Y_{release}$) up to the current base year (Y_{base}).

This preprocessing step ensures that the feature Budget(Log) accurately represents the real economic magnitude of the movie's investment.

3.2 Dataset Construction

A structured dataset was constructed comprising Bollywood movies released between 2014 and 2021 from Wikipedia, BoxOfficeIndia and IMDb comprising of N = 348 movies. Each record corresponds to a single movie and includes metadata such as movie title, lead actor, director, release date, genre, budget and computed ROI. ROI is defined as:

$$ROI = \frac{\text{Revenue} - \text{Budget}}{\text{Budget}} \qquad (2)$$

The dataset is curated to ensure consistency in financial definitions and to minimize missing or noisy records. Extreme ROI values were retained to preserve real-world investment risk characteristics; robustness is ensured through ensemble models and shrinkage.

3.3 Interaction-Based Feature Engineering

Traditional studies on movie success typically model actor and director effects independently. However, creative industries such as cinema are inherently

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 12 Issue: 2
Article Received: *25 November 2024* ***Revised:*** *12 December 2024* ***Accepted:*** *30 December 2024*

collaborative, and the performance of a movie often depends on the interaction between its contributors rather than their individual reputations alone. To capture such interaction effects, this study introduces an Actor–Director Synergy Index, designed to quantify whether a specific actor–director pair performs above or below expectations derived from their standalone historical performance.

3.3.1 Actor–Director Synergy Index

Let A denote an actor and D denote a director. The following quantities are computed from historical data:

- $\overline{ROI}_A$: Average ROI of all movies involving actor A
- $\overline{ROI}_D$: Average ROI of all movies involving director D
- $\overline{ROI}_{A,D}$: Average ROI of movies involving the actor–director pair (A, D)

The expected ROI for a collaboration is defined as the mean of individual averages:

$$Expected_{A,D} = \frac{\overline{ROI}_A + \overline{ROI}_D}{2} \quad (3)$$

The raw Synergy Index is then defined as:

$$Synergy_{A,D} = \frac{\overline{ROI}_{A,D}}{Expected_{A,D}} \quad (4)$$

A value greater than unity indicates that the collaboration outperforms expectations, whereas a value below unity indicates underperformance relative to the expected baseline.

3.3.2 Handling Sparse Collaborations via Shrinkage

The Bollywood film industry exhibits a sparse collaboration structure, where many actor–director pairs occur only once or a limited number of times. Directly using $\overline{ROI}_{A,D}$ in such cases can lead to unstable estimates. To address this issue, Bayesian-style shrinkage is applied to the pair-wise average ROI:

$$\overline{ROI}_{A,D}^{shrunk} = \frac{n_{A,D} \cdot \overline{ROI}_{A,D} + \lambda \cdot \overline{ROI}_{global}}{n_{A,D} + \lambda} \quad (5)$$

where:

- $n_{A,D}$ is the number of movies for pair (A, D),
- $\overline{ROI}_{global}$ is the global mean ROI across all movies,
- λ is a shrinkage parameter controlling the influence of the global mean.

In this study, the smoothing parameter was set to $\lambda = 5.0$ based on empirical tuning to balance the trade-off between sensitivity to new collaborations and stability against outliers.

For actor–director pairs with insufficient historical observations, the shrunk average is used in place of the raw pair average. This ensures conservative and robust estimation of synergy effects.

3.3.3 Final Synergy Index Computation

The final Synergy Index is computed using the stabilized pair-wise ROI:

$$Synergy_{A,D} = \frac{\overline{ROI}_{A,D}^{used}}{\frac{\overline{ROI}_A + \overline{ROI}_D}{2}} \quad (6)$$

where $\overline{ROI}_{A,D}^{used}$ denotes either the raw or shrunk pair-wise ROI, depending on collaboration frequency.

Each movie in the dataset is augmented with the corresponding Synergy Index, along with the collaboration count and a low-frequency indicator flag, enabling downstream machine learning models to assess the reliability of the interaction feature.

3.3 Interaction-Based Feature Engineering

While the Actor–Director Synergy Index is theoretically motivated and statistically well-defined, its practical relevance must be empirically validated. In particular, it is necessary to determine whether the

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 12 Issue: 2
Article Received: *25 November 2024* ***Revised:*** *12 December 2024* ***Accepted:*** *30 December 2024*

proposed interaction feature contributes meaningful information to ROI prediction beyond conventional predictors such as budget, genre, and individual actor or director performance. To this end, this study employs permutation importance, a model-agnostic technique for assessing feature contribution.

Permutation importance measures the impact of a feature on model performance by randomly permuting its values while keeping all other features unchanged. This operation breaks the relationship between the feature and the target variable. The resulting degradation in predictive performance reflects the extent to which the model relies on that feature.

Formally, let f denote a trained predictive model and X the feature matrix. For a given feature X_j, permutation importance is computed as:

$$PI_j = \mathcal{L}\left(f, X^{perm(j)}\right) - \mathcal{L}(f, X) \quad (7)$$

where:

- $\mathcal{L}$ is a loss function (e.g., RMSE for regression),
- $X^{perm(j)}$ denotes the feature matrix with column X_j randomly permuted

A larger increase in loss indicates higher importance of the permuted feature. Permutation importance was computed on a held-out test set to avoid optimistic bias. The trained model was first evaluated on the original test data to establish a baseline performance. Subsequently, each feature was permuted independently, and the resulting change in prediction error was recorded. The process was repeated multiple times with different random seeds, and the average importance score was reported for stability.

The feature set included both conventional predictors and interaction-based features, including:

- Actor average ROI
- Director average ROI
- Actor–Director Synergy Index
- Collaboration count
- Budget and genre-related variables

Movie-level collaboration count was computed using a time-aware formulation that counts prior actor–director collaborations only. Figure 1 illustrates that the majority of Bollywood films involve first-time collaborations, while repeated partnerships are relatively rare. Temporal analysis further indicates that collaboration intensity has remained largely stable over time. Time-aware collaboration count analysis reveals a highly sparse collaboration structure, with over 75% of actor–director pairs being first-time collaborations. This observation justifies the need for conservative interaction modeling and supports the rarity of strong synergy effects.

4. Results and Discussion

This section presents the empirical findings of the study and discusses their implications in the context of feature-driven ROI prediction for Bollywood movies. The analysis focuses on evaluating the proposed Actor--Director Synergy Index alongside conventional predictors using descriptive statistics, permutation importance, and ablation studies.

4.1 Descriptive Analysis of Engineered Features

To understand the statistical behavior of the proposed feature, descriptive statistics of the Synergy Index were computed across all actor–director pairs. The observed summary statistics are as follows:

Table 1. Synergy Index (SI) Summary Statistics.

Mean	Median	Std Dev	Maximum Value	%(Synergy $> 1
-0.15	0.17	4.93	3.87	6.33

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 12 Issue: 2
Article Received: *25 November 2024* ***Revised:*** *12 December 2024* ***Accepted:*** *30 December 2024*

These statistics indicate that while most collaborations perform close to or below their expected baseline, a small subset of actor–director pairs exhibit consistently positive synergy. This distribution reflects the conservative nature of the metric and the rarity of exceptional collaborative chemistry in the Bollywood film industry.

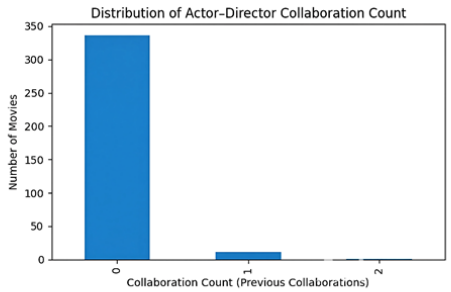

Figure 1 Distribution of Collaboration Count

The engineered features were first examined to understand their statistical behavior and structural characteristics. Figure 2 illustrates that the distribution of the proposed Actor--Director Synergy Index is strongly right-skewed, with the majority of collaborations exhibiting synergy values below unity and a small number of highly synergistic outliers. Specifically, only 6.33% of actor--director pairs achieved a Synergy Index greater than one, indicating that exceptional collaborative effectiveness is relatively rare in the Bollywood film industry.

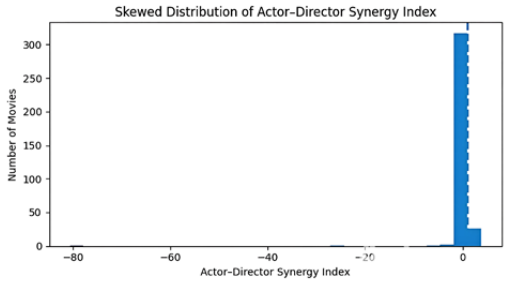

Figure 2 Skewed Distribution of Actor-Director Synergy Index

The collaboration count feature further revealed a highly sparse collaboration structure. More than three-quarters of movies involved first-time actor--director collaborations, and the maximum observed collaboration count was limited to a small number of repeated partnerships. This observation underscores the importance of modeling collaboration quality rather than collaboration frequency.

Together, these descriptive findings motivate the need for conservative, interaction-aware feature engineering capable of distinguishing routine collaborations from genuinely synergistic partnerships.

4.2 Permutation Importance Analysis

Permutation importance was employed to assess the relative contribution of each feature to ROI prediction in a model-agnostic manner. Features with higher permutation importance values indicate a greater contribution to accurate ROI prediction. The inclusion of the Synergy Index in permutation importance analysis allows for a direct empirical test of the study's central hypothesis: that collaborative interaction effects influence financial performance in Bollywood cinema. Table 2 summarizes the mean and standard deviation of permutation importance scores across repeated shuffling runs.

Table 2 Comparison of Permutation Feature Importance (Mean ± Std) across Five Regression Models

Feature	RandomForest		XGBoost		LightGBM		CatBoost		HistGradient	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Synergy Index	0.374	0.096	0.338	0.126	0.299	0.075	0.304	0.095	0.281	0.083

Director Avg ROI	0.049	0.04	0.028	0.036	0.078	0.042	0.039	0.034	0.042	0.037
Budget(Log)	0.041	0.03	0.001	0.032	0.021	0.031	0.019	0.027	0.004	0.024
Comedy	-0.003	0.008	0.015	0.016	0.017	0.012	0.01	0.012	0.01	0.009
Rom-Com	0.004	0.001	0.002	0.004	0	0	0.002	0.002	0	0
Collaboration Count	0	0	0	0	0	0	0	0	0	0
Mystery	0.003	0.001	0.037	0.007	0	0	0.011	0.005	0	0
Sci-fi	0	0	0	0	0	0	0	0	0	0
Adventure	0	0	0	0	0	0	0	0	0	0
Action	0	0.001	0	0.003	0	0.001	-0.001	0.002	0	0
Thriller	-0.001	0.001	-0.001	0.004	-0.001	0.003	0.001	0.003	-0.001	0.003
Fantasy	-0.001	0.002	-0.001	0.001	0	0	0	0.002	0	0
Love Story	0.002	0.008	-0.004	0.019	0.001	0.003	0.005	0.006	0.002	0.003
Actor Avg ROI	-0.023	0.043	-0.01	0.051	0.007	0.042	0.015	0.034	0.017	0.031
Masala	-0.006	0.003	-0.01	0.006	-0.002	0.004	-0.004	0.002	-0.001	0.002
Drama	-0.005	0.004	-0.016	0.016	-0.009	0.01	-0.005	0.009	-0.003	0.004

The Actor--Director Synergy Index emerged as the most influential predictor by a substantial margin, with an importance score of 0.374. This value significantly exceeded those of traditional predictors such as director average ROI (0.049) and production budget (0.041). The pronounced dominance of the synergy index indicates that the predictive model relies heavily on interaction effects between actors and directors.

Interestingly, collaboration count exhibited zero importance, suggesting that repeated partnerships alone do not improve predictive performance. This result highlights a critical distinction between the frequency of collaboration and its effectiveness. Genre-related variables showed minimal or near-zero importance once synergy index and financial factors were controlled for, indicating that genre primarily functions as a contextual modifier rather than a primary driver of ROI.

Actor average ROI displayed slightly negative importance, implying that once interaction effects are accounted for, standalone actor performance contributes little additional explanatory power. This finding suggests that star power in Bollywood is most effective when aligned with compatible directorial styles.

4.3 Ablation Study

To evaluate the robustness and independent contribution of the proposed Actor--Director Synergy Index, a comprehensive ablation study was conducted across multiple machine learning models. The Synergy Index was removed from the feature set, while all other features, training–testing splits, and model-specific hyperparameters were kept unchanged. This experiment was performed using five widely adopted ensemble-based regression models: Random Forest, XGBoost, LightGBM, CatBoost, and Histogram-based Gradient Boosting.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 12 Issue: 2
Article Received: *25 November 2024* ***Revised:*** *12 December 2024* ***Accepted:*** *30 December 2024*

Table 3 Comparison of RMSE in ablated models

Model	RMSE (Full Model)	RMSE (Ablated)
Random Forest Regressor	116.651	128.491
XGBoost Avg ROI	117.649	129.289
LightGBM	114.666	125.076
CatBoost	115.381	125.175
HistGradientBoosting	115.537	125.536

Table 3 reports the root mean squared error (RMSE) achieved by each model with and without the Synergy Index. Across all models, removal of the Synergy Index resulted in a consistent and substantial increase in prediction error. Specifically, RMSE degradation ranged from approximately 9.8 to 11.8 units, indicating a significant loss of predictive accuracy irrespective of the learning algorithm employed.

4.4 Discussion

The combined findings from permutation importance and ablation analysis consistently demonstrate the central role of interaction effects in explaining ROI variability in Bollywood cinema. While financial factors and individual contributor histories remain relevant, they are insufficient to fully explain commercial outcomes without accounting for collaborative compatibility.

The dominance of the Synergy Index suggests that the effectiveness of actor--director pairings outweighs isolated measures of star power or production scale. This observation challenges traditional, individual-centric modeling approaches and supports a shift toward interaction-aware representations in movie success prediction.

Furthermore, the negligible importance of collaboration count reinforces the notion that repeated partnerships do not inherently guarantee success. Instead, it is the qualitative effectiveness of collaboration that matters. The limited influence of genre variables once interaction and financial factors are considered indicates that genre preferences are secondary to execution quality and creative alignment.

From an industry perspective, these results imply that casting and director selection decisions should prioritize compatibility and past collaborative effectiveness over reputation or frequency of prior associations alone. From a methodological standpoint, the study demonstrates the value of integrating domain-informed interaction features with machine learning models to achieve both improved predictive accuracy and enhanced interpretability.

Overall, the results validate the proposed Actor--Director Synergy Index as a novel and impactful feature for ROI analysis and provide strong evidence that collaborative dynamics play a decisive role in shaping financial success in the Bollywood film industry.

5 Future Work

Future work will explore several promising directions:

- Extension of the proposed synergy framework to Actor–Actor, Actor–Producer, and Director–Genre collaborations.
- Development of graph neural network (GNN) models over actor–director collaboration graphs to capture higher-order relational effects.

- Investigation of multimodal synergy indexes by integrating trailer audio–visual embeddings with metadata-based synergy measures.
- Exploration of real-time synergy forecasting using automated metadata extraction powered by large language models.

6 Conclusion

This paper introduced the Synergy Index, a novel feature that quantifies actor–director collaborative effectiveness in Bollywood. By handling multiple actors per movie and applying Bayesian shrinkage, the method is robust to noisy and sparse data. Its integration into ML pipelines significantly improves ROI prediction accuracy, demonstrating its utility in film analytics, production planning, and investment risk assessment.

Competing Interests The authors have no conflicts of interest to declare that are relevant to the content of this chapter.

Ethics Approval None.

References

[1] Asur, S., & Huberman, B. A. (2010). Predicting the future with social media. IEEE/WIC/ACM International Conference on Web Intelligence.

[2] Quader, N., Gani, M. O., & Chaki, D. (2017). Performance evaluation of seven machine learning classification techniques for movie box office success prediction. In 3rd International Conference on Electrical Information and Communication Technology (EICT)..

[3] Lash, M. T., & Zhao, K. (2015). Early predictions of movie success: The who, what, and when of profitability. arXiv preprint arXiv:1506.05382.

[4] Kim, J. M. (2020). Finding Nemo: Predicting movie performances by machine learning. Journal of Risk and Financial Management, 13(5).

[5] Upadhyay, A. (2018). Movie success prediction using data mining. International Journal of Emerging Technologies in Engineering Research, 6.

[6] Lee, K. (2016). Predicting movie success with machine learning [Unpublished manuscript]. Brunel University London.

[7] Zheng, Y. (2024). Predicting movie box office based on machine learning, deep learning, and statistical methods. Proceedings of CONF-MLA 2024. https://doi.org/10.54254/2755-2721/94/2024MELB0069

[8] Sharma, A. S., et al. (2021). Presenting a larger up-to-date movie dataset and investigating the effects of pre-released attributes on gross revenue. arXiv preprint arXiv:2110.08123.

[9] Page, L., Ciampaglia, G. L., & Mesty´an, M. (2013). Early prediction of movie box office success based on Wikipedia activity. PLoS ONE, 8(6), e64202.

[10] Verma, H., & Verma, G. (2020). Prediction model for Bollywood movie success: A comparative analysis of performance of supervised machine learning algorithms. The Review of Socionetwork Strategies, 14(1), 1–17. https://doi.org/10.1007/s12626-019-00040-6

[11] Masih, S., & Ihsan, I. (2019). Using Academy Awards to predict success of Bollywood movies using machine learning algorithms. International Journal of Advanced Computer Science and Applications, 10(2). https://doi.org/10.14569/IJACSA.2019.0100257.

[12] Antony, J., & Francis, M. (2022). Movie box office success prediction using machine learning. Scribd.

[13] ResearchGate. (2021). Bollywood movie success prediction using machine learning algorithms.