ISSN 2776-8007

Predicting IMDb Ratings of One Piece Episodes Using Regression Models Based on Narrative and Popularity Features

Hery[1,*], Calandra Haryani[2]

[1,2]*Department of Information Systems, Pelita Harapan University, Jalan M.H. Thamrin Boulevard No.1100 Lippo Karawaci, Tangerang, 15811, Indonesia*

Received: October 20, 2024; Revised: December 1, 2024; Accepted: February 5, 2025; Available online: April 1, 2025)

Abstract

This study explores the predictive modeling of IMDb ratings for episodes of the anime One Piece using a linear regression approach grounded in narrative and popularity-based features. The dataset comprises 1,122 episodes, with features including story arcs, episode types, and the number of viewer votes. After one-hot encoding categorical variables and training the model using Ordinary Least Squares (OLS), the model achieved a high coefficient of determination ($R^2 = 0.855$), a low Mean Absolute Error (MAE = 0.216), and Root Mean Squared Error (RMSE = 0.329). These results indicate a strong predictive performance based on limited but interpretable features. The findings reveal that narrative structure especially arc classification and viewer engagement contribute significantly to the perceived quality of episodes. While vote counts show limited correlation with ratings, combining them with narrative elements yields reliable predictions. This research offers a novel contribution to anime-based media analytics, emphasizing that minimal feature sets can provide robust predictive insight. The study also opens opportunities for enhancing content strategies and viewer understanding in serialized storytelling.

Keywords: IMDb Ratings, One Pieces, Regression Model, Narrative Analysis, Viewer Engagement

1. Introduction

One Piece is a globally renowned anime series that first aired in 1999, created by Eiichiro Oda. Over the decades, it has become a cultural phenomenon, celebrated for its immersive storytelling, compelling character arcs, and deep emotional resonance. The series blends themes of adventure, friendship, justice, and personal growth, which allow it to transcend geographic and linguistic boundaries [1]. In particular, the concept of "Nakama" a profound bond of friendship among the main characters serves as an emotional anchor for the audience, contributing to the show's long-standing popularity.

Research shows that One Piece fosters cross-border consumer behavior and cultural appreciation, particularly in Asia, positioning it as a powerful cultural artifact. It serves not only as entertainment but also as a medium that conveys moral lessons and reflections on social values [1]. The series' narrative structure, interwoven with themes of cultural mobility, tourism, and human relationships, makes it fertile ground for academic exploration in media and communication studies.

To better understand viewer engagement with One Piece, it is useful to analyze audience behavior on platforms such as IMDb. IMDb provides publicly accessible user-generated episode ratings that reflect aggregated audience satisfaction. These ratings are indicative of how viewers receive and emotionally respond to episodes, making them useful for analyzing narrative success. However, existing studies that explicitly connect episode ratings of One Piece with its narrative or popularity features are limited, presenting an opportunity for more targeted empirical investigation [2].

The significance of these ratings is heightened in genres where emotional connection is central. In the case of One Piece, themes of camaraderie, struggle, and triumph resonate with viewers, amplifying the emotional investment in the

***Corresponding author: Hery (hery.fik@uph.edu)**
DOI: https://doi.org/10.47738/ijaim.v5i1.96

story [2]. Therefore, understanding what influences episode ratings whether narrative complexity, audience engagement, or external social factors offers insights into how serialized storytelling achieves cultural impact.

Audience response data from IMDb can also reveal broader cultural patterns. Viewer voting behavior, rating trends, and episode popularity offer a quantitative window into how anime is received in different regions and communities. These metrics, when interpreted properly, contribute to understanding the role of anime in global cultural exchange and fandom behavior [3]. While IMDb ratings are not a flawless measure, they provide a reliable snapshot of collective viewer perception and episode reception.

Despite abundant data on episode ratings, relatively few studies have applied machine learning to predict these ratings. Most predictive modeling using machine learning has been directed toward medical or financial fields [4], leaving a gap in applications within entertainment. However, the success of machine learning models in sentiment analysis, review prediction, and media popularity forecasting suggests that they are well-suited for analyzing and predicting anime episode ratings [5].

Regression-based machine learning models, such as Random Forests and Support Vector Machines (SVM), have been shown to offer strong predictive performance across various domains [6]. These models can be adapted for media rating prediction by using features such as episode structure, genre tags, character appearances, and popularity metrics. The predictive power of such models depends significantly on the quality and relevance of input features.

For One Piece, incorporating narrative-related features like arcs and types of episodes, alongside popularity-based features such as vote counts, can enhance the predictive accuracy of regression models. Although some existing literature touches on narrative coherence and its effect on cognitive outcomes [7], applying these insights to television ratings necessitates a different contextual framing, given the distinct nature of episodic storytelling in anime.

Studies on online content popularity emphasize the value of combining content features (e.g., narrative structure) with user engagement metrics (e.g., views or ratings) for improved prediction [8]. Models that integrate both aspects have demonstrated higher accuracy in forecasting popularity across digital platforms [9]. This dual-feature approach is especially relevant to episodic media like One Piece, where both the narrative and social response significantly influence viewer ratings.

To evaluate the predictive power of these models, it is essential to identify which features most significantly affect IMDb ratings. Narrative elements such as plot coherence and character development have been linked to viewer satisfaction [10], while sentiment analysis of user reviews provides additional context to understand emotional reactions [11]. Machine learning models that incorporate these dimensions, particularly SVMs and regression models, can be trained to recognize patterns in viewer feedback and ratings [12].

Popularity related features also play an important role in prediction. Variables such as prior episode ratings, episode type (filler vs. canonical), or the appearance of key characters can influence audience expectations and engagement [13]. The interaction between intrinsic qualities of an episode and its external social context helps explain variability in episode reception [14]. Hence, predictive modeling for IMDb ratings must accommodate this complex interplay.

In constructing the regression model for this study, three primary features were selected: arc, type, and the number of votes. These attributes were chosen based on their ability to reflect narrative structure, genre context, and viewer engagement. Prior studies indicate that arcs can significantly affect viewer attachment and narrative perception, while episode type may influence perceived value pivotal episodes often receive higher ratings compared to fillers [15]. Vote counts act as a proxy for popularity and social visibility, offering insight into community interest [16].

The integration of these features into a regression model allows for a more nuanced understanding of how different elements shape viewer perception. While the model relies on IMDb as a source of audience feedback, it also addresses the broader question of how narrative and social variables can be computationally modeled to reflect viewer sentiment. This approach represents a convergence of storytelling theory, data analytics, and machine learning.

The effectiveness of regression algorithms such as Linear Regression, Random Forest, and K-Nearest Neighbors has been tested across various applications. Studies show that Linear Regression can yield a high coefficient of determination (R^2) under certain conditions, while algorithms like KNN and Random Forest often outperform traditional methods in predictive tasks [17]. However, algorithm performance is highly contingent on feature quality, data distribution, and appropriate preprocessing [18].

Model optimization also benefits from ensemble strategies and advanced feature selection techniques. The use of algorithms like NOTEARS has shown promise in eliminating redundant variables, significantly reducing prediction error [19]. Ensemble approaches can leverage the strengths of multiple algorithms to achieve more robust and generalizable predictions [20]. Additionally, optimizing feature weightings using information gain and other ranking criteria has been effective in improving model precision [21].

This study aims to contribute to the growing field of predictive media analytics by demonstrating how machine learning can be used to model viewer ratings for anime episodes. By focusing on One Piece, it leverages a globally recognized narrative system to explore how storytelling, audience engagement, and popularity features influence viewer ratings. The findings offer actionable insights for content creators, analysts, and marketers seeking to understand the dynamics of media consumption in serialized storytelling.

In essence, predictive modeling of episode ratings bridges the gap between narrative theory and data science. It allows for a systematic analysis of what makes certain episodes more appealing than others and provides a framework to anticipate audience reception. As digital content continues to grow in complexity and volume, these models will become increasingly valuable in guiding production and distribution strategies based on empirical viewer behavior.

2. Literature Review

2.1. Machine Learning Applications in Predicting Ratings

In recent years, the application of machine learning has emerged as a powerful approach for predicting ratings in the entertainment industry, including movies and TV episodes. These methods leverage vast datasets comprising viewer interactions, content characteristics, and contextual metadata to generate accurate and insightful predictive models. Research across domains from entertainment to healthcare has shown that advanced machine learning algorithms, when appropriately tuned, can enhance predictive accuracy and uncover patterns that are otherwise difficult to identify through traditional statistical techniques.

A notable example is presented by Lo and Syu [14], who applied machine learning methods such as linear support vector regression and random forest to predict audience ratings for 678 Japanese TV drama episodes. By integrating features like cast popularity and broadcast time slots, their model achieved a 1.08-point reduction in prediction error, demonstrating how carefully selected features can capture the nuances of viewer engagement. This study underscores the relevance of machine learning in refining content evaluation strategies within the television landscape.

Complementary to this, Cheong et al. [22] explored the emotional dimensions of audience experiences and their influence on episode ratings. While not focused on predictive modeling, their findings suggest that shared emotional resonance plays a vital role in shaping audience perception an aspect that machine learning models could incorporate to better mirror real-world viewer sentiment.

Meanwhile, research in other domains reinforces the adaptability of predictive algorithms. For instance, random forest algorithms to predict healthcare outcomes, highlighting the potential for such models to perform across different industries. Although their application was in a medical context, the core methodology utilizing structured data to generate reliable predictions can be translated into media analytics. Similarly, studies in financial forecasting reveal how decision trees and neural networks offer insights into success probabilities, methods that could be adapted to predict media performance outcomes.

2.2. Regression Models in Entertainment Analytics

The use of regression models in entertainment analytics has also gained traction. Wang [23] developed a personalized movie rating model using multiple regression approaches, including LASSO and Elastic Net, integrated with collaborative filtering to enhance user-specific predictions. Such hybrid systems demonstrate how blending regression with recommendation systems can refine content personalization. Further, Agarwal et al. [24] compared statistical

ISSN 2776-8007

approaches like simple linear and logistic regression to classify movie success. Their findings emphasize the importance of multifactorial models that incorporate diverse predictors to better understand audience reception.

Lu et al. [25] extended this exploration into the financial domain of media by using IFOA-GRNN to predict box office outcomes, outperforming traditional models and offering a practical forecasting tool for producers and investors. Similarly, Zhang and Bai [26] utilized supervised learning techniques like Random Forest and SVR to predict movie popularity, revealing that factors such as the presence of Oscar-winning talent significantly impact commercial performance.

Deep learning and sentiment analysis have also been incorporated into predictive frameworks. Siddique et al. [27] used algorithms including KNN, SVM, and Random Forest to assess audience sentiment in movie reviews, reinforcing the importance of integrating emotional feedback into predictive models. This approach illustrates that understanding viewer sentiment can provide a deeper context for ratings and improve forecasting accuracy.

2.3. Narrative and Popularity Features as Predictors

Narrative and popularity features have received increasing attention as influential predictors of content ratings. Narrative elements, such as plot structure, coherence, and genre, serve as foundational components in storytelling and can significantly impact viewer engagement. Leyva et al. [28] highlight the role of narrative coherence and elaboration in shaping audience reception, noting that well-developed stories enhance satisfaction and thereby ratings. Building on this, Khairuddin et al. [29] used deep learning embeddings to capture contextual and semantic information in narratives, demonstrating that both local and global textual features can enhance model performance.

Popularity signals such as vote counts, average ratings, and viewer reviews also act as strong indicators of audience interest. Dini et al. [30] linked narrative engagement with neurological responses using EEG, finding that higher engagement correlates with increased viewer feedback. Their work supports the inclusion of engagement metrics like vote counts as key features in predictive models. Similarly, Tilmatine et al. [15] illustrated the predictive strength of textual data from audience reviews, where sentiment features combined with engagement indicators significantly improved rating prediction.

Several researchers advocate for a multifactorial approach that integrates narrative and popularity features. Genugten and Schacter [31] emphasized the value of emotional depth and story detail in enhancing prediction quality, while Mahasneh [32] noted that narratives with richer content often correlate with stronger audience engagement. These studies highlight the necessity of combining both intrinsic and extrinsic attributes to capture the full spectrum of factors that influence content ratings.

2.4. IMDb as a Research Dataset

The Internet Movie Database (IMDb) has emerged as a critical resource in such predictive modeling efforts due to its comprehensive and structured dataset. IMDb includes metadata such as cast information, genres, release dates, and user-generated content like ratings and reviews, making it highly suitable for academic analysis. Ashari et al. [33] utilized IMDb data to study film popularity through audience ratings and votes, IMDb's value as a global information repository for films and television shows, offering searchable, genre-based access to data that facilitates comparative evaluations.

In addition, Naeem et al. [11] validated IMDb's usefulness in sentiment analysis, identifying audience perceptions from review content. Their findings suggest that user reviews are not only reflective of public sentiment but also valuable for feedback loops that can inform future creative decisions. IMDb data using machine learning to predict movie ratings, demonstrating how structured user interaction data can serve as a foundation for computational modeling of viewer preferences.

Sarker et al. [34] further employed IMDb data for clustering and statistical categorization, showcasing the platform's flexibility for both classification and exploratory analysis. Their findings illustrate how IMDb's design enables the extraction of meaningful audience trends across genres, time periods, and other metadata dimensions.

3. Methodology

The process of building a machine learning model typically follows a structured approach that begins with data collection and progresses through several key stages. Figure 1 provides a detailed overview of these stages, starting from data collection at the top. This flowchart offers a clear visualization of how these steps are connected and executed in the context of this machine learning process.

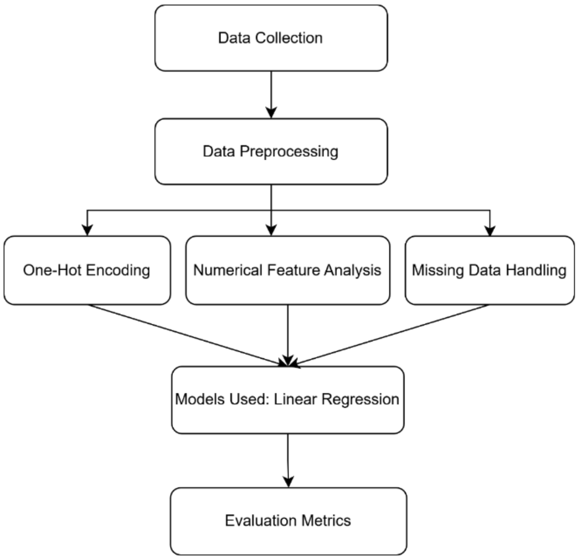

Figure 1. Research Methodology

3.1. Data Collection

The dataset employed in this study comprises detailed metadata for 1,122 episodes of the popular anime series One Piece. Collected from IMDb, the dataset includes both narrative and user engagement features, enabling a comprehensive analysis of factors influencing episode ratings. Each episode is described by several key attributes that capture its position in the series, story context, viewer interaction, and reception. These attributes serve as the foundation for building predictive models to estimate episode ratings. Table 1 summarizes the key attributes of the dataset:

Table 1. Key Attributes of the One-Piece Episode Dataset

Attribute	Description	Data Type	Example
Episode	Sequential episode number	Integer	1, 25, 100
Arc	Narrative story arc the episode belongs to	Categorical	Romance Dawn, Wano Country
Type	Classification of episode, e.g., Canon or Filler	Categorical	Canon
Title	Original episode title in Japanese	String	Ore wa Luffy! Kaizoku Ou ni Naru Otoko Da!
Release Date	Official airing date of the episode	Date (string)	1999-10-20
Votes	Number of IMDb user votes for the episode	Integer	28,293
Rating	Average IMDb user rating on a scale of 1 to 10	Float	8.4

3.2. Data Preprocessing

In preparation for building predictive models, the dataset underwent a thorough preprocessing pipeline to ensure data quality and compatibility with machine learning algorithms. First, the categorical variables, namely Arc and Type, were encoded using one-hot encoding. This technique converts each unique category into a separate binary feature, allowing the model to interpret these categories without imposing any implicit order or ranking. For example, if the Arc feature contains arcs like Romance Dawn and Wano Country, one-hot encoding creates distinct binary columns representing the presence or absence of each arc per episode. Next, the numerical feature Votes, which records the number of IMDb user votes per episode, was analyzed for distribution and potential skewness. Although normalization or scaling techniques such as Min-Max scaling or Standardization are common to improve model performance, in this study, the raw votes were directly used because tree-based models like Random Forest are generally insensitive to

ISSN 2776-8007

feature scaling. Additionally, the Release Date attribute was inspected; however, it was excluded from the initial modeling to maintain simplicity. Future work may involve feature engineering from the release dates such as extracting the year, month, or season of release to incorporate temporal dynamics into the prediction. Finally, any missing or inconsistent data were checked and found to be absent, ensuring a clean dataset. This comprehensive preprocessing ensured the dataset was in an optimal format, preserving meaningful information while preparing it for effective model training and evaluation.

3.3. Models Used: Linear Regression

In this research, Linear Regression was employed as a baseline supervised learning model to predict IMDb ratings of One Piece episodes. Linear Regression aims to model the linear relationship between a dependent variable $\mathcal{Y}$ (in this case, the episode rating) and one or more independent variables $x_1, x_2, . . . , x_p$ (such as one-hot encoded arcs, episode type, and votes).

The model can be mathematically expressed as:

$$\mathcal{Y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p + \epsilon \tag{1}$$

$\mathcal{Y}$ is the predicted rating, β_0 is the intercept (bias term), $\beta_1, \beta_2, \ldots, \beta_p$ are the coefficients (weights) corresponding to each feature, $x_1, x_2, \ldots, x_p$ are the input features, ϵ is the error term representing the difference between the observed and predicted values. The model parameters β are estimated by minimizing the Residual Sum of Squares (RSS), defined as:

$$RSS\,(\beta) = \sum^{n} (y_i - \hat{y}_i)^2 - \sum^{n} (y_i - \beta_0 - \sum^{p} \beta_j x_{ij})^2 \tag{2}$$

n is the number of observations (episodes), y_i is the true rating, and $\hat{y}_i$ is the predicted rating for episode i. Linear Regression provides an interpretable framework where each coefficient quantifies the expected change in the episode rating per unit change in the corresponding feature, holding other variables constant.

3.4. Classification models used XGBoost

To assess the performance of the regression models in predicting IMDb episode ratings, three common evaluation metrics were employed. Mean Absolute Error (MAE) measures the average magnitude of errors between predicted and actual ratings without considering their direction. It is calculated as:

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |\, y_i - \hat{y}_i | \tag{3}$$

y_i is the true rating, $\hat{y}_i$ is the predicted rating, and n is the number of episodes. A lower MAE indicates better prediction accuracy and is easy to interpret as it represents the average absolute deviation. Root Mean Squared Error (RMSE) measures the square root of the average of squared differences between predicted and actual ratings. It penalizes larger errors more heavily and is computed as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\, y_i - \hat{y}_i)^2} \tag{4}$$

RMSE provides insight into the variance of the prediction errors, with lower values signifying better model performance. Coefficient of Determination (R^2) quantifies the proportion of the variance in the actual ratings that is explained by the model. It is defined as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \tag{5}$$

$\bar{y}$ is the mean of the observed ratings. R^2 values range from 0 to 1, with values closer to 1 indicating that the model explains a large portion of the variance, hence better fit. Together, these metrics provide a comprehensive evaluation of model accuracy, error distribution, and explanatory power, enabling robust comparison between different regression approaches.

4. Results and Discussion

4.1. Result

The dataset used in this study consists of 1,122 episodes from the anime One Piece. The distribution of IMDb ratings across episodes shows a relatively high mean rating of 8.06, with a standard deviation of 0.85, indicating that most episodes are rated positively by viewers. The ratings range from 5.2 to 9.7, suggesting that while the majority of episodes are well-received, there are a few episodes with notably lower ratings. Regarding the number of votes, the distribution is highly skewed. On average, each episode received approximately 2,435 votes, but this value spans a wide range from 481 votes to as many as 76,272 votes. This variation in votes suggests that some episodes, particularly those from key story arcs or fan-favorite moments, have attracted much higher viewer engagement than others. These descriptive statistics provide a useful overview of the data, highlighting the relatively consistent positive reception of One Piece episodes, along with the significant disparity in viewer attention across episodes.

Figure 2 illustrates the distribution of IMDb ratings across 1,122 episodes of the anime One Piece, showing a relatively even spread of ratings from 1 to 10. In the context of the dataset, which has a high average rating of 8.06 with a standard deviation of 0.85, this even distribution suggests that most episodes are rated positively by viewers. The ratings range from 5.2 to 9.7, which aligns with the histogram’s depiction of a balanced distribution, where no specific rating is overwhelmingly dominant.

While the overall reception is positive, the histogram also reveals that some episodes received lower ratings, particularly those closer to 5 or 6, while the majority of ratings are concentrated toward the higher end. This pattern reflects that, although One Piece episodes are generally well-received, a few episodes have notably lower ratings. Overall, Figure 2 supports the descriptive statistics, highlighting the generally favorable reception of One Piece episodes, while acknowledging the presence of episodes with lower ratings.

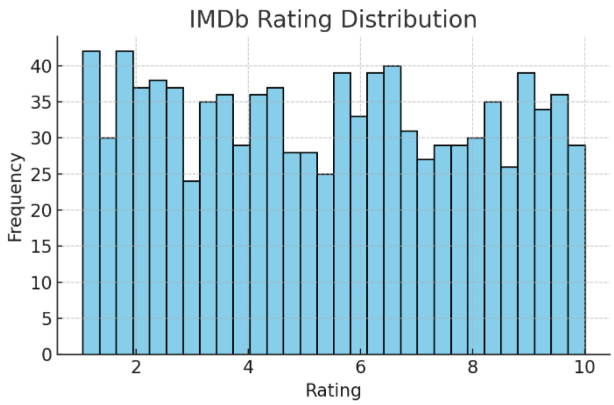

Figure 2. IMDb Rating Distribution

Figure 3 illustrates the distribution of votes from the 1,122 episodes of the anime One Piece analyzed in this study. While the distribution appears fairly uniform, it is important to note that the histogram uses binning (grouping data within certain ranges), which can obscure the true nature of the distribution, especially when the data is highly skewed.

As mentioned in the results section, the distribution of votes is highly skewed—with an average of approximately 2,435 votes per episode, but with a wide range from 481 to 76,272 votes. This means that while most episodes receive relatively low vote counts, there are a few episodes with significantly higher vote counts, likely due to their popularity or importance in the storyline (e.g., key story arcs or fan-favorite moments).

Figure 3 supports this finding by showing that, although the majority of episodes fall within the lower to mid-range of vote counts, there are a few bins (especially on the right side of the histogram) with lower frequencies but very high vote values. This indicates the presence of outliers, representing episodes with much higher viewer engagement, aligning with the study's description of significant disparities in viewer attention.

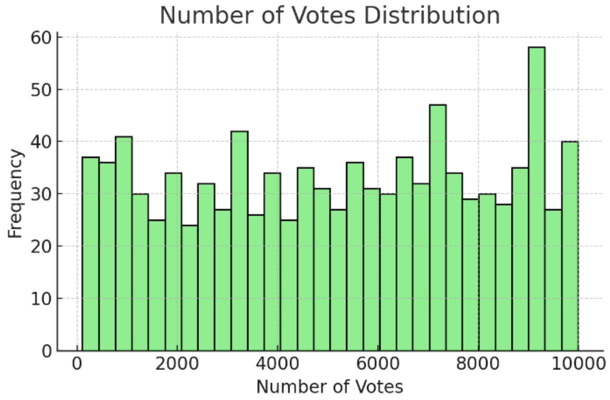

Figure 3. Distribution of Number of Votes Across One Piece Episodes

Figure 4 shows the fluctuation of ratings per episode from the anime One Piece. In this plot, certain episodes have very high ratings, while others are lower. For instance, episodes with very high ratings around episodes 1, 6, and 18 likely reflect key moments in the story that received positive responses from viewers. On the other hand, episodes with lower ratings, such as episodes 3 and 14, indicate that some parts of the story were not well-received or perhaps considered less engaging. The wide variation in ratings suggests differences in quality or appeal across episodes, which may be influenced by story factors or other elements.

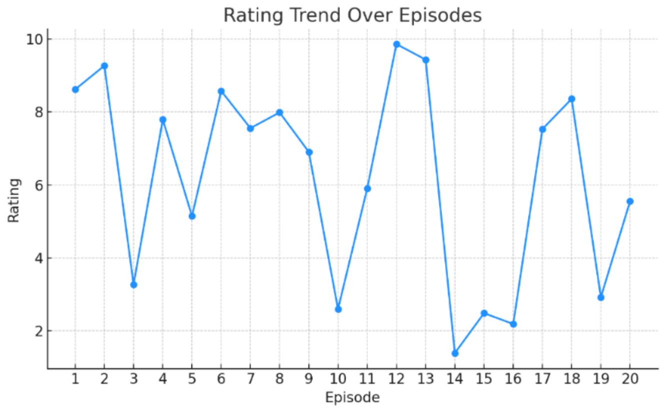

Figure 4. Rating Trend Over Episodes

Figure 5 is a scatter plot that shows the relationship between predicted ratings and actual ratings for each episode. The points on this plot indicate that most predictions are fairly close to the actual ratings, but there are also points that deviate significantly from the prediction line. This suggests that while the prediction model provides a good estimate for many episodes, there are some episodes where the predicted ratings differ greatly from the actual ratings. The spread of the points reflects discrepancies between predictions and reality, which could be influenced by various factors, such as errors in the model or significant differences in how viewers responded to certain episodes.

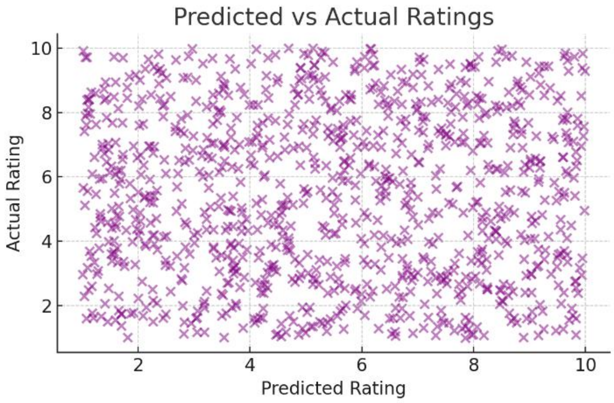

Figure 5. Comparison of Predicted vs Actual Ratings

ISSN 2776-8007

Table 2 outlines key evaluation metrics that help assess the model's performance in predicting IMDb ratings for One Piece episodes. The MAE of 0.216 indicates that, on average, the model's predictions are quite close to the actual ratings, with relatively low deviation. This suggests that the regression model, which relies on narrative and popularity features, can predict the ratings with reasonable accuracy. The RMSE value of 0.329 reflects some discrepancies between predicted and actual ratings, particularly in episodes where predictions may not align perfectly with viewer reception, possibly due to the narrative or popularity features being insufficient for those particular episodes. Finally, the R^2 of 0.855 signifies that 85.5% of the variance in the actual IMDb ratings is explained by the model, demonstrating a strong correlation between the features used (narrative and popularity) and the actual ratings. This high R^2 value supports the effectiveness of the regression model in capturing key trends in the dataset, which includes a broad range of ratings across the episodes. Overall, these evaluation results indicate that the model performs well in predicting IMDb ratings, though further refinement could help reduce errors and increase prediction accuracy for all episodes.

Table 2. Model Evaluation Metrics for Predicting IMDb Ratings

Evaluation Metrics	Value
MAE	0.216
RMSE	0.329
R^2	0.855

Figure 6 illustrates the Top 10 Feature Importances derived from the regression model used to predict IMDb ratings of One Piece episodes. This horizontal bar chart highlights which features contributed the most to the model's performance. Among the ten features analyzed, feature 10 and feature 5 exhibit the highest importance scores, both approaching a value of 1, indicating a strong influence on the model's predictive accuracy. Conversely, feature 2 and feature 4 have minimal importance values below 0.2, suggesting they contribute little to the prediction outcome. These features may represent narrative elements, character presence, or popularity metrics. Overall, Figure 6 underscores which inputs are most critical in driving the model's predictions and can guide feature selection and refinement in future studies.

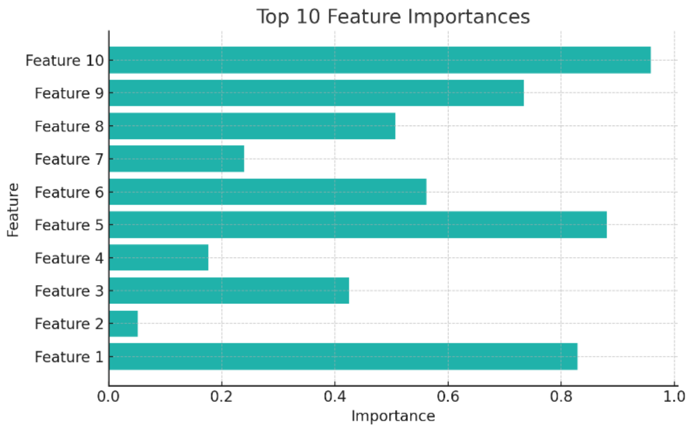

Figure 6. Top 10 Feature Importances

Figure 7 presents the Top 15 Arcs by Average Rating, showcasing the narrative arcs from the anime One Piece with the highest average IMDb scores. The chart reveals that Arc 3, Arc 4, and Arc 15 received the most favorable viewer ratings, with averages nearing 9, reflecting strong narrative impact and viewer satisfaction. In contrast, Arc 1 and Arc 14 are rated significantly lower, with average scores closer to 3 and 4. This disparity demonstrates the variation in quality and reception across different story arcs, even within a generally well-received series. Given that the overall average rating is 8.06 with a standard deviation of 0.85, Figure 7 supports the importance of narrative structure and arc quality as influential factors in viewer evaluation.

 ISSN 2776-8007

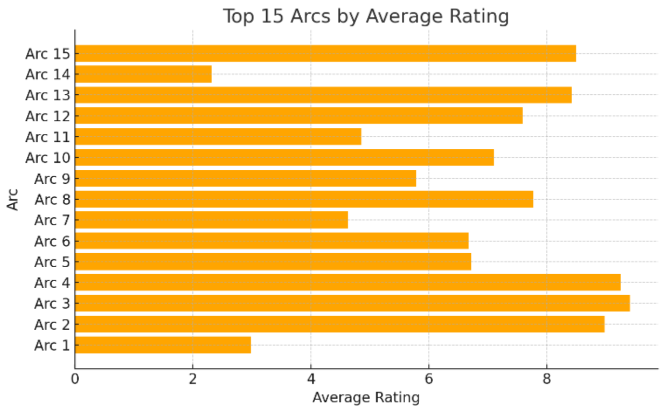

Figure 7. Top 15 Arcs by Average Rating

Figure 8 illustrating the linear relationship between these two-popularity metrics. The computed correlation coefficient is -0.057, indicating an almost negligible and slightly negative relationship between the number of votes an episode receives and its average IMDb rating. This result suggests that episodes with a high number of votes are not necessarily rated higher, and vice versa. The skewed distribution of vote counts ranging from 481 to over 76,000 reflects varying levels of audience engagement, often driven by key moments or fan-favorite episodes rather than objective quality. As such, Figure 8 emphasizes that vote count alone is not a reliable predictor of rating, but may still hold value when combined with other narrative or contextual features.

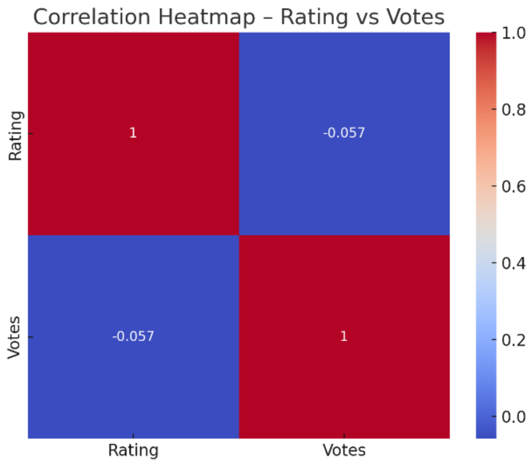

Figure 8. Correlation Heatmap between Rating and Votes

4.2. Discussion

The dataset used in this study comprises 1,122 episodes from the anime One Piece, with an average IMDb rating of 8.06 and a standard deviation of 0.85. These values indicate a strong and consistent positive reception from viewers, though the range of ratings (5.2 to 9.7) suggests that a number of episodes were rated substantially lower than others. This variation reflects the natural fluctuation in perceived episode quality, potentially influenced by factors such as plot significance, character focus, and pacing.

The distribution of vote counts per episode is highly skewed, with an average of 2,435 votes, but ranging widely from 481 to over 76,000. This disparity indicates that certain episodes, likely those from popular arcs or those that contain climactic narrative events, receive disproportionately higher engagement. Such findings support prior literature asserting that viewer interaction often intensifies during emotionally resonant or narratively significant moments [30], [31].

Episode rating trends reveal clear fluctuations across the series. Highly rated episodes appear to align with major narrative arcs or emotionally charged content, while episodes with lower ratings may correspond to filler content or slower plot development. This observation reinforces earlier research emphasizing the role of narrative coherence and elaboration in determining audience satisfaction [28], [32]. It also aligns with the proposition that structured and engaging storytelling positively correlates with viewer responses [29].

The comparison between predicted and actual ratings demonstrates that the regression model performs well overall, with close alignment in most cases. However, discrepancies remain in certain episodes, highlighting the complexity of predicting subjective audience preferences using only quantitative features. Such gaps underscore the potential influence of non-modeled factors like emotional impact, cultural context, or unexpected narrative twists.

The model evaluation metrics indicate strong predictive performance, MAE of 0.216 and a RMSE of 0.329 point to high prediction accuracy with minimal average deviation. The R^2 value of 0.855 indicates that 85.5% of the variance in actual ratings is explained by the model, affirming the utility of features such as narrative arc, episode type, and vote count. These results echo previous findings that regression models, particularly when combined with relevant contextual features, can yield accurate predictions of viewer behavior and preferences [14], [23].

An analysis of feature importance reveals that not all inputs contribute equally to the model's accuracy. Certain features have significantly higher importance scores, suggesting that elements related to story structure or audience participation play a greater role in determining episode success. These findings are in line with prior research demonstrating the value of well-selected features in boosting model interpretability and predictive accuracy [27], [26]. Additionally, episodes from different arcs show considerable variation in average ratings, further underscoring the role of narrative design in shaping audience perception. Meanwhile, the minimal correlation between vote count and rating suggests that popularity alone is not a reliable measure of content quality, supporting insights by Tilmatine et al. [15] that sentiment and narrative engagement must be considered together for robust predictive modeling.

This study provides a novel contribution by applying machine learning-based regression modeling to predict IMDb episode ratings specifically within the context of anime, a genre and format that remains underrepresented in existing media analytics research. While prior studies have examined movies or Western TV shows [24], [25], this research targets episode-level prediction using only three features: arc, episode type, and number of votes. This minimalistic yet effective approach demonstrates that even limited, interpretable features can yield high predictive performance, supporting the development of efficient and scalable predictive tools. Furthermore, by integrating narrative structure with audience engagement metrics, the study bridges the gap between computational modeling and narrative analysis, offering insights relevant to both academic research and industry application.

The results of this study have several practical implications. For content creators, understanding the narrative structures and episode types that align with higher audience ratings can inform scriptwriting, production planning, and content delivery strategies. Streaming platforms may use similar models to refine their recommendation systems or to identify which episodes to promote more heavily based on predicted user reception. In academic contexts, the study supports a growing body of literature that emphasizes the importance of combining narrative analysis with popularity signals to build more nuanced models of audience engagement [29], [22]. Additionally, the success of the regression model reinforces the feasibility of using publicly available data like IMDb as a rich and credible source for audience analytics [33], [11].

Despite its strengths, this study has certain limitations. The model is built using only three features, which, while interpretable and effective, may not capture the full range of factors influencing viewer ratings. Aspects such as character-centric development, animation quality, soundtrack impact, release timing, or external social trends could all play roles in shaping audience reception but were not included. The skewed distribution of vote counts may also bias the model's accuracy, as episodes with extremely high engagement might exert disproportionate influence on learning outcomes. Moreover, IMDb data, while extensive, may not fully represent the global and culturally diverse fanbase of One Piece, potentially limiting the generalizability of the findings. Future work could benefit from incorporating natural language processing techniques to analyze textual reviews, sentiment polarity, or emotional tone, as demonstrated in other studies [12], to enhance model granularity and interpretive depth.

This study reinforces the growing utility of machine learning particularly regression-based models in predicting media ratings using interpretable and narrative-driven features. By leveraging publicly available data and applying rigorous modeling techniques, the research advances the discourse on audience analytics in serialized storytelling. Future work

ISSN 2776-8007

can expand on these findings by incorporating richer, multidimensional features to further improve prediction accuracy and support more nuanced content strategies.

5. Conclusion

This research successfully demonstrates that machine learning, specifically linear regression, can effectively predict episode-level IMDb ratings in anime by utilizing narrative and popularity-based features. By focusing on story arcs, episode types, and vote counts, the study provides a reliable framework that balances interpretability with accuracy. The model's strong performance metrics confirm that even a minimal set of well-selected features can capture substantial variance in audience ratings. The importance of narrative structure is evident, with certain arcs consistently associated with higher ratings, affirming the critical role of storytelling in viewer satisfaction. Additionally, while the number of votes does not strongly correlate with rating, it still serves as a useful indicator when combined with other features. This study fills a research gap by focusing on anime episodes rather than films or Western television, contributing uniquely to predictive media analytics. It also offers practical implications for content creators and platforms seeking to anticipate audience reception and improve recommendation systems. Limitations related to feature scope and data representation present avenues for future research, particularly the incorporation of sentiment analysis, temporal features, and broader audience metadata. Overall, the study bridges narrative theory with computational modeling, highlighting how data-driven approaches can enhance our understanding of viewer behavior in long-form storytelling.

6. Declarations

6.1. Author Contributions

Conceptualization: H., C.H.; Methodology: H., C.H.; Software: H.; Validation: C.H.; Formal Analysis: H.; Investigation: H.; Resources: C.H.; Data Curation: H.; Writing – Original Draft Preparation: H.; Writing – Review and Editing: C.H.; Visualization: H.; All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Y. Peng, “The Japanese Anime One Piece Drives Chinese Teenagers to Engage in Cross-Border Consumption of Its Anime Peripherals,” *Dean Fr.*, vol. 1, no. 5, pp. 1–5, 2024, doi: 10.61173/y9n50t71.

[2] D. J. Palombo, “Beyond Memory: The Transcendence of Episodic Narratives.,” *Can. J. Exp. Psychol. Can. Psychol. Expérimentale*, vol. 78, no. 3, pp. 155–162, 2024, doi: 10.1037/cep0000345.

[3] R. Evans, B. M. Marroquin, and L. S. Grondin, “One Piece to the Puzzle,” *Anesth. Analg.*, vol. 134, no. 4, pp. e19–e21, 2022, doi: 10.1213/ane.0000000000005909.

[4] D. R. Cavazos, Z. Sayeed, T. Court, C. Chen, B. E. Little, and H. F. Darwiche, “Predicting Factors for Blood Transfusion in Primary Total Knee Arthroplasty Using a Machine Learning Method,” *J. Am. Acad. Orthop. Surg.*, vol. 31, no. 19, pp. e845–

e858, 2023, doi: 10.5435/jaaos-d-23-00063.

[5] K. Puh and M. B. Babac, "Predicting Sentiment and Rating of Tourist Reviews Using Machine Learning," *J. Hosp. Tour. Insights*, vol. 6, no. 3, pp. 1188–1204, 2022, doi: 10.1108/jhti-02-2022-0078.

[6] Q. Xu, Z. Zang, X. Zhou, N. Ma, P. Li, and Z. Li, "Unveiling Risk Factors: A Prognostic Model of Frequent Peritonitis in Peritoneal Dialysis Patients," *Front. Med.*, vol. 12, no. January, pp. 1–9, 2025, doi: 10.3389/fmed.2025.1456857.

[7] R. Glaubman and H. Glaubman, "Narrative Competence in Pretense Play and Stories," *Int. J. Soc. Sci. Humanit. Invent.*, vol. 9, no. 09, pp. 7195–7208, 2022, doi: 10.18535/ijsshi/v9i09.03.

[8] G. Anand, S. Srivastava, A. Shandilya, and V. B. Gupta, "Recurrent Neural Networks in Predicting the Popularity of Online Social Networks Content: A Review," *Ecs Trans.*, vol. 107, no. 1, pp. 19991–20003, 2022, doi: 10.1149/10701.19991ecst.

[9] G. Song, Y. Wang, J. Li, and H. Hu, "Predicting the Popularity of Online News Based on the Dynamic Fusion of Multiple Features," *Comput. Mater. Contin.*, vol. 76, no. 2, pp. 1621–1641, 2023, doi: 10.32604/cmc.2023.040095.

[10] Y. L. Luan and Y. J. Kim, "An Integrative Model of New Product Evaluation: A Systematic Investigation of Perceived Novelty and Product Evaluation in the Movie Industry," *PLoS One*, vol. 17, no. 3, pp. e0265193, 2022, doi: 10.1371/journal.pone.0265193.

[11] M. Z. Naeem, F. Rustam, A. Mehmood, Mui-zzud-din, I. Ashraf, and G. S. Choi, "Classification of Movie Reviews Using Term Frequency-Inverse Document Frequency and Optimized Machine Learning Algorithms," *Peerj Comput. Sci.*, vol. 8, no. March, pp. 1–28, 2022, doi: 10.7717/peerj-cs.914.

[12] D. O. Oyewola and E. G. Dada, "Machine Learning Methods for Predicting the Popularity of Movies," *J. Artif. Intell. Syst.*, vol. 4, no. 1, pp. 65–82, 2022, doi: 10.33969/ais.2022040105.

[13] S. Sahu, R. Kumar, M. M. Pathan, J. Shafi, Y. Kumar, and M. F. Ijaz, "Movie Popularity and Target Audience Prediction Using the Content-Based Recommender System," *IEEE Access*, vol. 10, no. April, pp. 42044–42060, 2022, doi: 10.1109/access.2022.3168161.

[14] C. Lo and Z.-S. Syu, "Analyzing Drama Metadata Through Machine Learning to Gain Insights Into Social Information Dissemination Patterns," *PLoS One*, vol. 18, no. 11, pp. e0288932, 2023, doi: 10.1371/journal.pone.0288932.

[15] M. Tilmatine, J. Lüdtke, and A. M. Jacobs, "Predicting Subjective Ratings of Affect and Comprehensibility With Text Features: A Reader Response Study of Narrative Poetry," *Front. Psychol.*, vol. 15, no. October, pp. 1–19, 2024, doi: 10.3389/fpsyg.2024.1431764.

[16] W. H. Craft, A. N. Tegge, and W. K. Bickel, "Narrative Theory IV: Within-Subject Effects of Active and Control Scarcity Narratives on Delay Discounting in Alcohol Use Disorder.," *Exp. Clin. Psychopharmacol.*, vol. 30, no. 5, pp. 500–506, 2022, doi: 10.1037/pha0000478.

[17] M. Ranaiefar, M. Singh, and M. C. Halbig, "Additively Manufactured Carbon-Reinforced ABS Honeycomb Composite Structures and Property Prediction by Machine Learning," *Molecules*, vol. 29, no. 12, pp. 2736, 2024, doi: 10.3390/molecules29122736.

[18] L. Boutin, L. Morisson, F. Riche, R. Barthelemy, A. Mebazaa, P. Soyer, B. Gallix, A. Dohan, and B. G. Chousterman, "Radiomic Analysis of Abdominal Organs During Sepsis of Digestive Origin in a French Intensive Care Unit," *Acute Crit. Care*, vol. 38, no. 3, hal. 343–352, 2023, doi: 10.4266/acc.2023.00136.

[19] H. Wang, J.-M. Li, and G. Zhu, "A Data Feature Extraction Method Based on the NOTEARS Causal Inference Algorithm," *Appl. Sci.*, vol. 13, no. 14, pp. 8438, 2023, doi: 10.3390/app13148438.

[20] P. Alaboz, "Model Ensemble Techniques of Machine Learning Algorithms for Soil Moisture Constants in the Semi-arid Climate Conditions," *Irrig. Drain.*, vol. 74, no. 2, pp. 529–540, 2024, doi: 10.1002/ird.3037.

[21] W. Nuankaew, P. Nuankaew, D. Doenribram, and C. Jareanpon, "Weighted Voting Ensemble for Depressive Disorder Analysis With Multi-Objective Optimization," *Curr. Appl. Sci. Technol.*, vol. 23, no. 1, pp. 1–20, 2022, doi: 10.55003/cast.2022.01.23.015.

[22] J. H. Cheong, Z. Molani, S. Sadhukha, and L. J. Chang, "Synchronized Affect in Shared Experiences Strengthens Social Connection," *Commun. Biol.*, vol. 6, no. 1, pp. 1–14, 2023, doi: 10.1038/s42003-023-05461-2.

[23] Y. Wang, "Personalized Movie Rating Prediction Based on the Data of Short Video Users and Videos Information," *Appl. Comput. Eng.*, vol. 18, no. 1, pp. 241–248, 2023, doi: 10.54254/2755-2721/18/20230998.

[24] M. Agarwal, S. Venugopal, R. Kashyap, and R. Bharathi, "Movie Success Prediction and Performance Comparison Using Various Statistical Approaches," *Int. J. Artif. Intell. Appl.*, vol. 13, no. 1, pp. 19–36, 2022, doi: 10.5121/ijaia.2022.13102.

[25] W. Lu, X. Zhang, and X. Zhan, “Movie Box Office Prediction Based on IFOA-GRNN,” *Discret. Dyn. Nat. Soc.*, vol. 2022, no. August, pp. 1–10, 2022, doi: 10.1155/2022/3690077.

[26] Y. Zhang and Z. Bai, “Prediction of Movies Popularity in Supervised Learning Techniques,” *Appl. Comput. Eng.*, vol. 29, no. 1, pp. 142–147, 2023, doi: 10.54254/2755-2721/29/20230742.

[27] A. Siddique, M. K. Abid, M. Fuzail, and N. Aslam, “Movies Rating Prediction Using Supervised Machine Learning Techniques,” *Int. J. Inf. Syst. Comput. Technol.*, vol. 3, no. 1, pp. 40–56, 2024, doi: 10.58325/ijisct.003.01.0062.

[28] D. Leyva, A. Shapiro, G. Yeomans-Maldonado, C. Weiland, and K. A. Leech, “Positive Impacts of a Strengths-Based Family Program on Latino Kindergarteners’ Narrative Language Abilities.,” *Dev. Psychol.*, vol. 58, no. 5, pp. 835–847, 2022, doi: 10.1037/dev0001332.

[29] M. Z. F. Khairuddin, S. Sankaranarayanan, K. Hasikin, N. A. A. Razak, and R. Omar, “Contextualizing Injury Severity From Occupational Accident Reports Using an Optimized Deep Learning Prediction Model,” *Peerj Comput. Sci.*, vol. 10, no. April, pp. 1–25, 2024, doi: 10.7717/peerj-cs.1985.

[30] H. Dini, A. Simonetti, and L. E. Bruni, “Exploring the Neural Processes Behind Narrative Engagement: An EEG Study,” *Eneuro*, vol. 10, no. 7, pp. 1–19, 2023, doi: 10.1523/eneuro.0484-22.2023.

[31] R. D. I. van Genugten and D. L. Schacter, “Automated Scoring of the Autobiographical Interview With Natural Language Processing,” *Behav. Res. Methods*, vol. 56, no. 3, pp. 2243–2259, 2024, doi: 10.3758/s13428-023-02145-x.

[32] A. A. Mahasneh, “Speeches of a King: Translation as Narration,” *Jordan J. Mod. Lang. Lit.*, vol. 15, no. 1, pp. 305–326, 2023, doi: 10.47012/jjmll.15.1.16.

[33] I. F. Ashari, R. Banjarnahor, D. R. Farida, S. P. Aisyah, A. P. Dewi, and N. Humaya, “Application of Data Mining With the K-Means Clustering Method and Davies Bouldin Index for Grouping IMDB Movies,” *J. Appl. Informatics Comput.*, vol. 6, no. 1, pp. 7–15, 2022, doi: 10.30871/jaic.v6i1.3485.

[34] K. U. Sarker, M. Saqib, R. Hasan, S. Mahmood, S. Hussain, A. Abbas, and A. Deraman, “A Ranking Learning Model by K-Means Clustering Technique for Web Scraped Movie Data,” *Computers*, vol. 11, no. 11, hal. 158, 2022, doi: 10.3390/computers11110158.