Journal of Computer Science

Original Research Paper

Computational Analysis of Genre Effects on Movie Ratings Using MLP Algorithms

[1]Subir Gupta, [1]Upasana Adhikari, [1]Shefali Varshney and [2]Tanupriya Choudhury

[1]*Department of CSE (AIML), Haldia Institute of Technology, Haldia, West Bengal, India*
[2]*School of Computer Science, University of Petroleum and Energy Studies (UPES), Dehradun, Uttarakhand, India*

Article history
Received: 20-06-2024
Revised: 30-09-2024
Accepted: 05-11-2024

Corresponding Author:
Tanupriya Choudhury
School of Computer Science,
University of Petroleum and Energy Studies (UPES),
Dehradun, Uttarakhand, India
Email: tanupriya@ddn.upes.ac.in

Abstract: Predictive analytics are what the entertainment industry depends on so much in prognosticating movie ratings, which is why they inform filmmakers, distributors and stream platforms strategic moves. Traditional prediction models most times only succeed in missing out on the intricate dynamics that are genre-specific to movie ratings thereby leading to inaccuracies and suboptimal decision making. The paper proposes this research presents a comprehensive mechanism that combines extensive metadata with Multi-Layer Perceptron (MLP) models to increase the accuracy of predictions across multiple cinematic genres. Therefore, we had a goal of establishing fine patterns due to MLP regressors based on specific genres as well as addressing limitations linked to traditional approaches. To conduct this study; we used principal component analysis and one-hot encoding for 950 films followed by genre-specific modeling alongside statistical tests such as ANOVA, t-tests, Gradient Boosting Classifiers among others for model validation. They found that adventure movies were more predictable than other genres (MSE = 0.023) such as action (MSE = 2.816). It's clear then that accurate modelling requires an examination by gender and broad data sources integration. The research emphasizes the potential of improved machine learning methods to change predictive modeling in the area of art. Further work will seek to develop more accurate feature selection, deal with data imbalance and incorporate real-time audience engagement measures into the optimization process for better predictions that would help film makers make better strategic decisions.

Keywords: Multi-Level Perceptron (MLP), Chi-Square Test, Gradient Boosting Classifier, Movie Rating, ANOVA, T-Test

Introduction

The entertainment industry's landscape is witnessing transformative shifts, primarily driven by predictive analytics' critical role in cinema. Accurately forecasting movie ratings is not merely an academic endeavor but a strategic necessity that significantly influences the decisions of filmmakers, distributors, and streaming platforms (Provoost *et al*., 2020). These stakeholders rely on ratings to shape marketing strategies, manage financial planning, and ultimately, determine the success of cinematic projects (Shah *et al*., 2018; Aldalbahi and Walker, 2016). The ability to predict these ratings accurately can thus dictate the commercial and artistic trajectory of films, highlighting the crucial nature of this endeavor (Zhang and Kelly, 2009; Ghisani *et al*., 2021). Historically, predictive models in the film industry have been somewhat rudimentary, typically utilizing essential film attributes such as genre, cast, and budget to forecast outcomes. However, these traditional approaches often fall short due to the dynamic nature of audience preferences and the complex, multidimensional aspects of film production and reception (Lai *et al*., 2019). The primary limitation of these models is their inability to capture the nuanced and interconnected factors that influence a film's reception, often resulting in predictions that do not hold up against the multi-faceted realities of movie success. This gap in predictive accuracy underscores a significant problem in the analytics of the entertainment sector, where the rapid evolution of consumer tastes and market conditions outpaces the capabilities of conventional predictive models (Liang *et al*., 2020). The stakes in accurately predicting movie ratings are high, given the substantial financial investments at play and the broad cultural impact of films (Hossain *et al.*, 2019; Lei *et al*., 2023). Misjudgments in rating predictions can lead to mismatches between production efforts and audience expectations, resulting in financial losses and missed opportunities for filmmakers and distributors. Furthermore, streaming platforms that depend on precise algorithms to recommend content to viewers suffer when predictions are inaccurate, potentially leading to a less engaged audience (Singh *et al.*, 2022; Albu *et al*., 2015). The growing complexity of the global film market, with its diverse audiences and varied tastes, makes accurate prediction more crucial yet challenging (Bapna *et al*., 2012). This study proposes a novel predictive model that integrates a sophisticated Multi-Layer Perceptron (MLP), a type of neural network known for its efficacy in modeling complex and non-linear relationships, with an extensive dataset enriched with comprehensive metadata (Hijawi *et al*., 2021; Gupta *et al*., 2020). Unlike traditional models, this approach allows for a deeper analysis of the myriad factors influencing movie ratings, including but not limited to movie ID, name, release year, language, and critical reception. Extending the dataset beyond conventional features, the proposed model aims to provide a more accurate, nuanced understanding of what drives movie ratings (Neeli and Patil, 2021). The methodology of this research is structured around utilizing MLP to effectively capture the complex patterns within the extensive metadata of the film dataset. The MLP architecture, comprising an input layer, several hidden layers, and an output, makes it adept at learning hierarchically from the data, thus identifying subtle patterns that simpler models might overlook (Hanif *et al*., 2021). This capability is enhanced by the rich metadata that includes detailed descriptions of each movie, providing a granular view that facilitates a comprehensive analysis. In addition to MLP, the research will incorporate rigorous statistical analysis to validate the model's effectiveness. Techniques such as t-tests, chi-square tests, ANOVA, and gradient-boosting Classifier will be employed to test the hypotheses and confirm the model's reliability across different dataset dimensions (Dickerson and Przydatek, 2003). Once developed, the model's predictive accuracy will be evaluated using robust metrics such as Mean Squared Error (MSE), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE) (Ahonen *et al*., 2004). The performance of the MLP model will be compared with that of traditional predictive models to highlight its superior capabilities in handling the complexities of movie rating predictions. (Palanivel and Koshy Mathews, 2012; Anoop and Kumar 2013) The implications of this research are vast, with potential benefits extending across the entire spectrum of the film industry. From aiding producers in deciding which projects to green-light to assisting streaming platforms in crafting more effective content recommendation algorithms, the applications of this advanced predictive methodology are manifold. This research aims to contribute significantly to the cinematic landscape's strategic planning and decision-making processes by providing a more accurate and actionable predictive framework. In decision, this research addresses a critical gap in the existing predictive models in the entertainment industry by introducing a more sophisticated, data-driven approach. By leveraging the power of MLP and comprehensive metadata, this study seeks to enhance the precision of movie rating predictions, thereby offering valuable insights that can inform decision-making and strategy in the ever-evolving cinematic arena (Razfar *et al*., 2022). Through this innovative approach, the research aims to advance the field of predictive analytics in the entertainment industry and provide practical tools that can lead to more informed, effective, and successful film projects. The rest of the paper has been structured.

Literature Review

The application of predictive analytics within the entertainment sector, particularly in predicting movie ratings, has become increasingly prominent, reflecting a broader interest in harnessing advanced computational methods to inform industry decisions (Ma, 2021). The shift from traditional statistical methods to more sophisticated machine learning techniques such as Multi-Layer Perceptrons (MLPs) marks a significant advancement in the analytical capabilities available to filmmakers and distributors computational methods (Ahmad *et al*., 2021; Al-Azzam and Shatnawi, 2021). This literature review delves into the progression of predictive models from basic algorithms to complex systems that integrate a broad spectrum of metadata, assessing their impact on the accuracy of movie ratings and box office predictions. Traditional predictive models primarily utilized demographic and historical data to forecast viewer ratings and box office performance (Taskiran *et al*., 2020).

These methods often relied on linear regression techniques, which, while providing a baseline prediction capability, were limited in their ability to process the complex and non-linear relationships inherent in entertainment data. This limitation often resulted in significant predictive inaccuracies, as these models could not account for less tangible variables such as viewer sentiment and social media influence, which have become increasingly critical in the digital age. Integrating MLPs, a subset of feedforward artificial neural networks, represents a transformative development in predictive analytics (Kapoor and Mishra, 2018; Kuo, 2000). These models can learn from data in multiple layers through backpropagation, allowing them to uncover hidden patterns and interactions within complex datasets. The literature indicates that MLPs are particularly effective in environments characterized by non-linear data interactions, making them well-suited for the dynamic and intricate datasets associated with movie ratings and audience behaviors (Mondal *et al*., 2022).

Another significant shift in predictive analytics has been the integration of extensive metadata into predictive models. Modern analytical approaches extend beyond essential variables to encompass detailed aspects such as cast dynamics, marketing strategies, release timing, and social media engagement (Ligthart *et al*., 2021; Sohom *et al*., 2021). This comprehensive metadata allows a deeper exploration of how various factors contribute to a movie's reception. Studies have shown that models incorporating these rich datasets can provide a more holistic and accurate view of potential film success, significantly enhancing predictive strength. Comparative studies highlight the superior performance of MLPs over traditional predictive models, particularly in scenarios characterized by complex interactions and substantial feature dimensions. For instance, a study demonstrated that MLPs significantly outperform linear models in predicting box office success due to their ability to model non-linear interactions between variables effectively. Despite the progress in utilizing MLP models and integrating comprehensive metadata, the literature reveals a notable gap in applying these techniques to genre-specific film analysis. Current research has generally treated genre as a peripheral variable rather than as a central focus of analysis. There is a lack of in-depth studies examining how predictive models can be tailored to different film genres to enhance accuracy and relevance (Radojevic and Slijepčević, 2018; Kersting, 2018).

Additionally, there is a pressing need for empirical research that explores the impact of emerging variables, such as real-time audience engagement metrics and detailed production nuances, on the predictive accuracy of film success. These areas represent critical opportunities for further research that could refine predictive models and improve decision-making processes within the film industry. This review underscores the significant advancements in predictive analytics within the film industry, facilitated by adopting MLPs and integrating comprehensive metadata. The evolution from simple linear models to complex neural networks has enabled a more nuanced understanding of the factors influencing movie success. However, the identified research gaps concerning genre-specific analysis and the inclusion of emerging variables suggest fruitful avenues for future studies. Addressing these gaps could lead to more refined predictive models, ultimately enhancing strategic planning and filmmaking decision-making in film production and distribution (Ahamad *et al*., 2020).

Materials

The current study employed a dataset culled from an open API of 950 movies. The dataset was made up of various features which are essential in forecasting. The starting point for the data had 1600 entries but after employing Principal Component Analysis (PCA) to eliminate noise, thus ensure high quality and relevance of data, this was reduced. Each record contained information about a movie including its ID, title, release year, genre, language, cast, production budget, box office performance and critical and user reviews that were used as predictors. To make it ready for machine learning modelling all such categorical variables like genre or language have been preprocessed through techniques like one-hot encoding. Some other numerical attributes were normalized with Standard Scaler to give equal weightage to all variables in the analysis. It was divided into train and test sets with balanced distribution across genres (80% training: 20% testing). Statistical tools such as t-tests, Chi-square tests, and ANOVA were used in the process of model validation to ensure that predictive models are reliable. With the help of sophisticated computational tools and algorithms, these resources evaluated how effective Multi-Layer Perceptron (MLP) models can be in forecasting movie ratings with a genre-specific measure of accuracy. Table (1) shows the sample dataset.

Table 1: Pseudocode for enhancing movie rating prediction with MLP and comprehensive metadata

1:	Initialize matrices: X = StandardScaler (). fit_transform(dataset.drop(['MOVIE ID', 'MOVIE_NAME', 'YEAR OF RELEASE', 'MOVIE_RATING', 'LANGUAGE', 'VERDICT'], axis=1))
2:	y = dataset['MOVIE_RATING']
3:	X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=some_integer)
4:	For each genre g in genres:
	4.1: y_train_g = y_train * genre_indicator(g)
	4.2: model_g = MLPRegressor(hidden_layer_sizes=(100, 164), activation='relu', solver='adam', max_iter=100).fit(X_train, y_train_g)
5:	For each genre g in genres:
	5.1: y_pred_g = model_g.predict(X_test)
	5.2: MSE_g = mean ((y_test - y_pred_g) ^2)
6:	overall_pred = sum ([y_pred_g for g in genres])
7:	overall_MSE = mean ((y_test - overall_pred) ^2)
8:	overall_RMSE = sqrt(overall_MSE)

Methods

To comprehensively analyze movie ratings using a dataset, the study designed a methodological framework that utilizes machine learning techniques for predictive modeling. The dataset is collected from an open API. Initially, the study is able to collect 1600 data and for noise removal using PCA and total dataset becomes 950. Initially, the dataset containing various movie attributes was loaded and prepared for analysis. The preprocessing phase involved one-hot encoding of movie genres to transform categorical data into a numerical format, enabling practical computation. Subsequently, the study partitioned the dataset into features and the target variable, where the features excluded irrelevant or identifying information such as 'MOVIE ID', 'MOVIE_NAME', 'YEAR OF RELEASE', 'LANGUAGE', and 'VERDICT', while the target variable was identified as 'MOVIE_RATING'. The data was then split into training and testing sets, ensuring that 20% of the data was reserved for testing, maintaining the integrity of model evaluation. Standardization of features was performed using a standard scaler, a crucial step to normalize the data and remove any inherent bias due to differing scales.

This standardization ensures that each feature contributes equally to the analysis, improving the predictive model's performance. The core of our analysis involved developing individual predictive models for each movie genre. This approach allowed us to tailor predictions and understand the nuances across different genres. A dedicated MLP regressor model was trained for each genre. The models were configured with specific architectural and operational parameters, such as the number of hidden layers and the type of activation function, to optimize performance. Figure (1) is shown below for reference.

To ensure that each genre model was trained effectively, the study modified the target variable by multiplying it with the genre indicator, assigning ratings to movies within the specific genre, and nullifying others. This approach allowed for focused learning, as each model was trained exclusively on data relevant to its genre. Upon training, the models were used to predict movie ratings in the test set. Equations (1-3) such metrics are used for calculating Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE) respectively reflect the predictive performance of each genre-specific model was evaluated:

$$MSE = \frac{\Sigma_{I=1}^{N} (p_I - o_I)}{N} \tag{1}$$

$$RMSE = \sqrt{\frac{\Sigma_{I=1}^{N} (p_I - o_I)^2}{N}} \tag{2}$$

$$MAE = \frac{1}{N} \Sigma_{I=1}^{N} \frac{|p_I - o_I|}{N} \tag{3}$$

The analysis also identified the highest-rated genre by extracting the maximum predicted rating among all genres. This detailed methodological approach enabled us to predict movie ratings with nuance and specificity across genres and understand the broader implications of these predictions in the context of the entire movie dataset. Through rigorous computational methods and analytical techniques, the study derived insights that could inform stakeholders in the film industry, from producers to marketers, about potential audience reception and rating trends. (shown in Table 1), the methodology used in this proposed research work is shown in Fig. (2).

In the subsequent phase of this study, following the application of machine learning methodologies, computational engineering techniques employing various statistical models are utilized for validation purposes. Specifically, this investigation employs hypothetical testing, ANOVA (Analysis of Variance), and a gradient-boosting classifier to analyze the data.

T-Test

Within the framework of hypothetical testing, t-statistics are calculated to compare movie ratings across different genres.

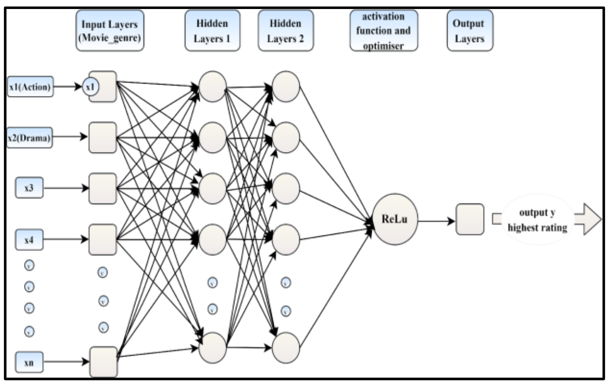

Fig. 1: Architecture of MLP model for enhancing movie rating prediction

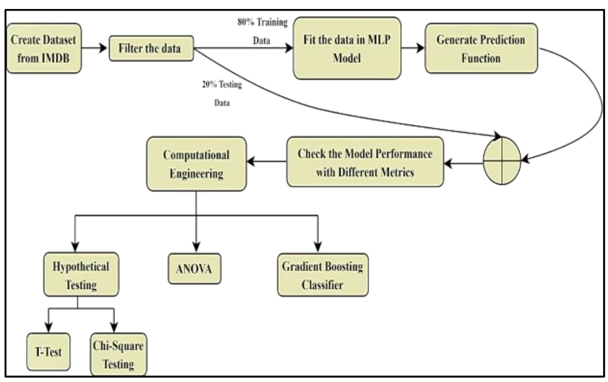

Fig. 2: Methodological diagram of MLP model for enhancing movie rating

Equation (4) denotes the t-statistics that involve determining the sample mean denoted as (X' X'), the variance (Z^2), and the sample size (H). Equation (5) subsequently shows the calculation of the *degrees* of freedom (df df'). In this context, the t-statistics derived are compared against the critical values obtained from the t-distribution table, corresponding to a significance level (*á*) of 0.05. This comparison is crucial for assessing the statistical significance of mean differences in movie ratings between various genres:

$$t - statistics\ t_s = \frac{X_{g1} - X_{g2}}{\sqrt{\frac{Z_{g1}^2}{H_{g1}} - \frac{Z_{g2}^2}{H_{g2}}}} \tag{4}$$

$$(degrees\ of\ freedom)\ df' = \frac{\left(\frac{Z_{g1}^2}{H_{g1}} + \frac{Z_{g1}^2}{H_{g1}}\right)}{\frac{\left(\frac{Z_{g1}^2}{H_{g1}}\right)^2}{H_{g1}-1} + \frac{\left(\frac{Z_{g2}^2}{H_{g2}}\right)^2}{H_{g2}-1}} \tag{5}$$

Chi-Square Test

In this study, another aspect of hypothetical testing was conducted and the Chi-square test was utilized to examine the potential relationship between movie genres and their respective ratings. This test determines whether the genres and ratings are associated or remain independent variables. To execute this analysis, movie ratings are initially categorized, followed by creating a contingency table. This table effectively summarizes the frequency of each specific pairing of movie genre and rating.

The expected frequencies for these combinations are calculated, as presented in Eq. (6). Determining the degrees of freedom, indicated in Eq. (7), is crucial for this analysis. Subsequently, the computed Chi-square statistic, as outlined in Eq. (8), is compared against the critical value. This comparison is instrumental in assessing the statistical significance of the association between movie genres and ratings:

$$E_f = \frac{Total\ Row_r * Total\ Column_c}{Grand\ Total} \tag{6}$$

$$df'' = (r - 1) * (c - 1) \tag{7}$$

$$x^2 = \sum \frac{(O_v - E_v)^2}{E_v} \tag{8}$$

ANOVA

In this research, the analysis of variance (ANOVA) statistical method is employed to examine the means across different groups to ascertain if there is a significant divergence among them. Specifically, ANOVA is utilized to investigate whether notable differences exist in the average ratings across various movie genres. This involves the computation of the F-statistic, as illustrated in Eq. (9). The calculation process includes determining the Mean Square Between groups (MSB), as detailed in Eq. (10) and the Mean Square Within groups (MSW), as outlined in Eq. (11). The resulting F-statistic is then compared with the critical value at a significance level of *á* = 0.05. This comparison helps assess the statistical significance of the variances in average ratings between different movie genres:

$$F = \frac{MSB'}{MSW'} \tag{9}$$

$$MSB' = \frac{variability\ between\ groups}{number\ of\ groups - 1} \tag{10}$$

$$MSW' = \frac{variability\ within\ each\ groups}{total\ number\ of\ observations - number\ of\ groups} \tag{11}$$

Gradient Boosting Classifier

Gradient boosting is an ensemble learning method designed to enhance model accuracy and performance by sequentially minimizing prediction errors by integrating multiple weak learners. In this approach, the logistic loss function shown in Eq. (12), where $y' \in \{-1, 1\}$ and $f(x)'$ represent the predicted value, is utilized. This function quantifies the discrepancy between the actual label y and the predicted value $f(x)'$. The additive model, as described in Eq. (13), is constructed in a stage-wise fashion. It begins with an initial prediction $F0(x)$ and at each subsequent stage, a new weak learner $hb(x)$ is incorporated, scaled by a learning rate ρb. Gradient descent is then performed as delineated in Eq. (14) to refine the model incrementally. The probability pi of the next iteration's prediction is calculated using the sigmoid function applied to the previous iteration's prediction $Rb-1(xi)$, as outlined in Eq. (15). Following this, the weak learner $h(x)$ is fitted, as shown in Eq. (16). The primary objective here is to minimize the squared differences between the residuals gI and the predictions provided by the weak learner $h(xi)$. This iterative process of fitting and refining continues until the model achieves the desired level of accuracy, leveraging the strengths of combining simple predictors to form a robust predictive model:

$$(y', f(x)') = log\,(1 + exp\,(-y'f(x)')) \tag{12}$$

$$R_b\,(x) = R_{b-1}\,(x) + p_b\,h_b\,(x) \tag{13}$$

$$_{I} = y_I - p_I \tag{14}$$

where:

$$P_i = \frac{1}{1 + exp\,(-R_{b-1}\,(x_i))} \tag{15}$$

$$h_b\,(x) = arg\,min_h \sum_{i=1}^{n} \quad (g_i' - h(x_i))^2 \tag{16}$$

A line search (shown in Eq. (17) is conducted to determine the optimal learning rate to identify the coefficient ρ_m that minimizes the total loss when a new weak learner $h(x)$ is incorporated into the model. Subsequently, the model is updated (as specified in Eq. (18) by integrating the scaled weak learner $\rho_b h(x)$ into the preceding model iteration $F_{b-1}(x)$. Upon completing all iterations, the output of the final model (depicted in Eq. (19) is transformed into a probability using the sigmoid function. This step ensures that the model's final predictions are presented in a probabilistically interpretable format, enhancing the utility and applicability of the results:

$$P_b = arg\,min_p \sum_{i=1}^{n} L\,(y_i', R_{b-1}(x_i) + ph_b(x_i)) \quad (17)$$

$$R_b(x) = R_{b-1}(x) + p_b h_b(x) \quad (18)$$

$$P_i = \frac{1}{1 + exp\,(-F_b(x))} \quad (19)$$

These equations represent the critical components of the gradient-boosting classifier derivation. For this research experiment, we calculated various performance metrics such as Accuracy Eq. (20), F1 score Eq. (21), precision Eq. (22), and Recall Eq. (23) values:

$$Accuracy\ A = \frac{(T_P + T_n)}{(T_P + T_n + F_P + F_n} \quad (20)$$

$$F\,1\ Score = \frac{Precision + Recall}{2 * Precision * Recall} \quad (21)$$

$$Precision = \frac{T_P}{T_P + F_n} \quad (22)$$

$$Recall = \frac{T_P}{T_P + T_n} \quad (23)$$

Results

This comprehensive study explores the application of Multi-Layer Perceptron (MLP) models in predicting movie ratings across various genres, explicitly focusing on enhancing predictive accuracy through advanced error correction techniques. Utilizing a meticulously categorized dataset by genre, the research embarked on a detailed predictive analysis, generating key error metrics such as Mean Squared Error (MSE), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE). These metrics served as benchmarks to evaluate the performance of the MLP models across different cinematic categories, providing a granular understanding of their effectiveness in diverse contexts.

Overview of Error Metrics Across Genres

The research elucidated marked variations in error metrics across cinematic genres, underscoring the disparate efficacy of Multi-Layer Perceptron (MLP) models in predictive accuracy, as shown in Table (2). Specifically, action films demonstrated moderate predictive accuracy, evidenced by a Mean Squared Error (MSE) of 2.816, Mean Absolute Error (MAE) of 1.205, and Root Mean Squared Error (RMSE) of 1.678. In stark contrast, the adventure genre showcased significantly lower error rates, with instances of MSE as minimal as 0.023, indicating a heightened precision in predictive outcomes for this genre. This variability was not confined to inter-genre comparisons but was also pronounced within individual genres.

Notably, the biography genre exhibited considerable fluctuations in predictive accuracy, with MSE values varying substantially from 0.482-0.074 across different assessments, as listed in Table (3). These fluctuations highlight the critical influence of specific dataset characteristics and models' configurations on the results and the accuracy graph is projected in Fig. (3).

Similarly, the comedy genre displayed a spectrum of error metrics across various evaluations, further accentuating the influence of data diversity and the subtleties involved in model training on the precision of predictions. Such disparities point to the nuanced challenges and complexities inherent in employing MLP models for film rating predictions, suggesting that both the choice of genre and the methodological approach significantly impact the efficacy of predictive modeling.

These findings illuminate the necessity for a nuanced understanding of genre-specific dynamics and technical model configurations to optimize the accuracy of predictive analytics in the film industry. model configurations to optimize the accuracy of predictive analytics in the film industry.

Table 2: Performance measure parameter values

S. No.	MLP model performance parameter		
	Precision	F1 score	Accuracy
1.	1.0	1.0	1.0

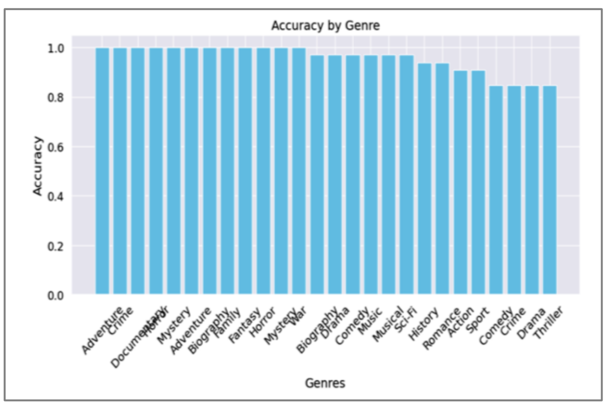

Fig. 3: Accuracy graph

DOI: 10.3844/jcssp.2025.905.917

Table 3: Different methods used in error correction

SL No	Genre	Types of error correction		
		MSE	MAE	RMSE
1.	Action	2.8162	1.2059	1.6781
2.	Adventure	0.0233	0.0615	0.1526
3.	Biography	0.4818	0.1976	0.6941
4.	Comedy	1.1313	0.5633	1.0636
5.	Crime	0.0660	0.1322	0.2570
6.	Documentary	0.0180	0.0799	0.1342
7.	Drama	0.2283	0.2983	0.4778
8.	Horror	0.0282	0.0821	0.1680
9.	Mystery	0.0071	0.0507	0.0848
10.	Adventure	0.6082	0.3306	0.7798
11.	Biography	0.0742	0.1354	0.2724
12.	Comedy	0.3702	0.3035	0.6084
13.	Crime	0.4239	0.3785	0.6511
14.	Drama	2.3289	1.2119	1.5260
15.	Family	0.0212	0.0667	0.1456
16.	Fantasy	0.5956	0.2187	0.7717
17.	History	0.4523	0.2833	0.6725
18.	Horror	0.1781	0.1335	0.4220
19.	Music	0.0353	0.0830	0.1879
20.	Musical	0.0560	0.0998	0.2368
21.	Mystery	0.0201	0.0703	0.1418
22.	Romance	1.4034	0.6058	1.1846
23.	Sci-Fi	0.0133	0.0486	0.1157
24.	Sport	0.0752	0.1038	0.2743
25.	Thriller	0.5945	0.3498	0.7710
26.	War	0.0360	0.0911	0.1899

Model configurations to optimize the accuracy of predictive analytics in the film industry. They prompt a deeper examination of how variations in data characteristics and model parameters can be better managed to enhance the overall effectiveness of predictive models across different cinematic genres. The critical role of genre-specific attributes, dataset composition, and model configuration are highlighted in varying error rates across genres and instances in determining predictive success. Figure (4) shows the average predicted ratings by genre.

Furthermore, the research offers valuable perspectives on the constraints and possibilities of utilizing MLP models for predictive analysis in the entertainment sector. This underscores the importance of genre-specific strategies and ongoing model improvement to boost predictive precision and dependability. In Fig. (5), the predicted label and true label values of the confusion matrix heatmap are depicted and the ROC curve is displayed in Fig. (6), representing the true positive rate and false positive rate of the presented method. Fig. (7) displays a scatter plot to aid in knowing the relationship between the actual and predicted ratings.

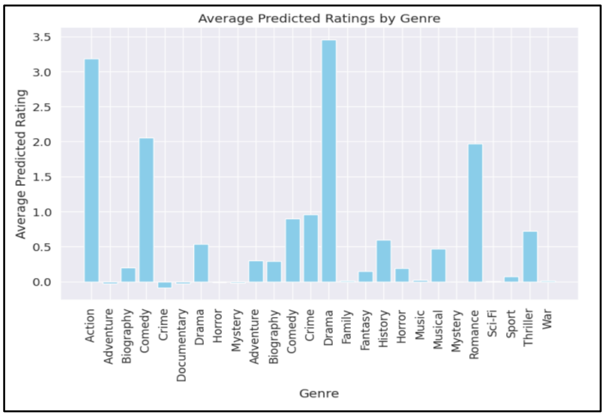

Fig. 4: Average predicted ratings by genre

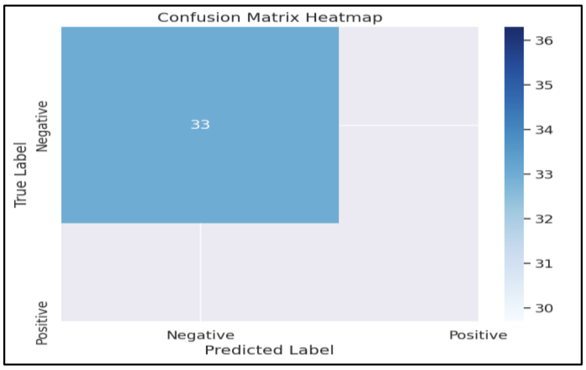

Fig. 5: Confusion matrix heatmap

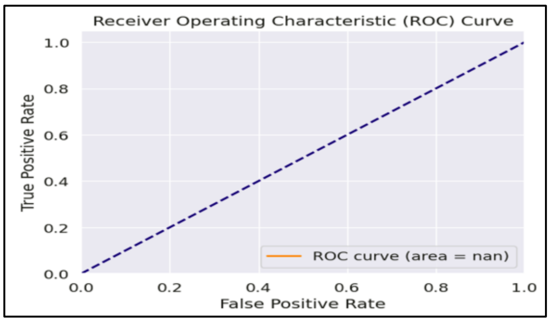

Fig. 6: ROC curve

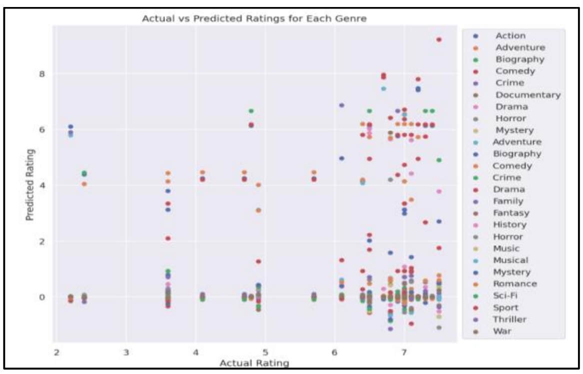

Fig. 7: Scatter plot to show the mapping of actual and predicted rating

DOI: 10.3844/jcssp.2025.905.917

Statistical Analysis for Model Validation

To substantiate the effectiveness of the predictive models utilized in the study, a comprehensive validation process was undertaken employing four distinct statistical methodologies: T-tests, Chi-square tests, Analysis of variance (ANOVA), and evaluations of gradient boosting classifier performance. These statistical tools collectively provided a rigorous framework for assessing the models' performance. Each method contributed uniquely to the analysis: T-tests were instrumental in examining the consistency of movie ratings across genres; chi-square tests offered insights into the relationship between film genres and their commercial outcomes; ANOVA facilitated a deeper understanding of the variability in movie ratings dependent on genre distinctions; and the Gradient Boosting Classifier evaluation highlighted the models' predictive success rates and areas requiring refinement. This multi-faceted approach affirmed the models' capabilities and revealed intricate details regarding their performance, enabling targeted improvements and enhancing the reliability of future predictions. Such rigorous validation underscores the robustness of the methodologies employed and provides a solid foundation for further explorations into predictive modeling within the entertainment industry.

T-Test and Chi-Square Test

The t-test to evaluate the variability in movie ratings between action and drama genres produced a t-statistic of 1.128 with a p-value of 0.267. These results signify no statistically significant differences in the ratings between these genres, indicating a consistent level of audience reception for both categories, which is presented in the bar plot of the distribution of movie ratings shown in Fig. (8): Bar plot representation of the distribution of movie ratings). This uniformity suggests that viewers' appreciation for action and drama films may be influenced by similar factors or that these genres meet audience expectations on average. On the other hand, the chi-square test, designed to explore the relationship between movie genres and their box office success, yielded a chi-square statistic of 15.284 with a p-value of 0.054. While this result does not cross the traditional threshold of statistical significance ($p<0.05$), it approaches significance, suggesting a potentially meaningful relationship between the genre of a film and its commercial outcomes. This near-significant result implies that specific genres may have a more predictable impact on box office performance, warranting further investigation into how specific genre attributes correlate with financial success. This analysis points to a nuanced interplay between genre and market performance, highlighting the need for deeper exploration into genre-specific trends and their implications for firm profitability.

ANOVA

The analysis of variance (ANOVA), which employed 'MOVIE-RATING' as the dependent variable and 'MOVIE-GENRE' as the independent variable, revealed significant discrepancies in ratings across genres. The calculated F-statistic of approximately 4.75 and a p-value nearing 3.1e-05 indicate substantial differences in how audiences receive various genres. These findings underscore the presence of significant variability in viewer ratings, which can be attributed to the diverse preferences and expectations that audiences hold toward different film genres. Figure (9) depicts this scenario using a box plot graph. This variation in reception highlights the complex interplay between genre-specific attributes and audience perceptions, suggesting a nuanced landscape of viewer engagement across cinematic categories.

Gradient Boosting Classifier

The gradient-boosting classifier was utilized to assess commercial success predictions for films, achieving % overall predictive accuracy of 68%. Notably, the precision and recall metrics for films classified as successful were both recorded at 76%, underscoring the model's robust capability to forecast film successes accurately listed in Table (4): Classification report of gradient boosting classifier performance).

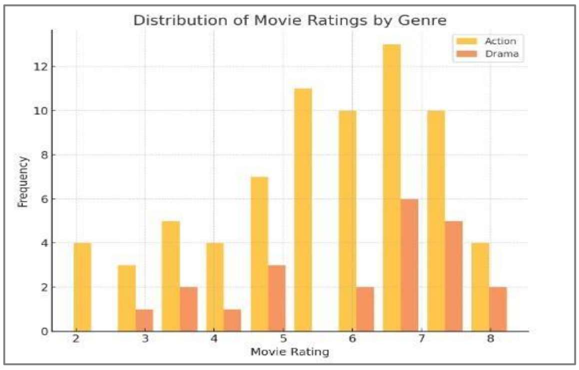

Fig. 8: Bar plot representation of the distribution of movie ratings

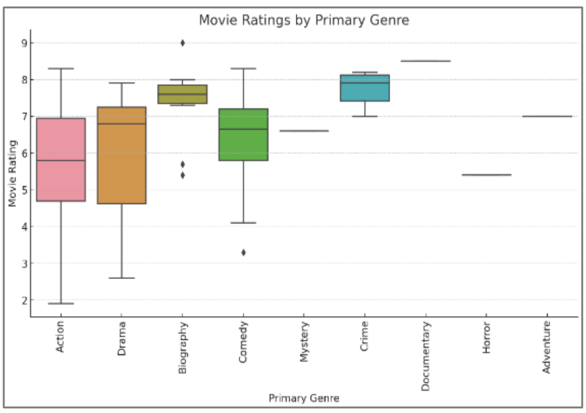

Fig. 9: Representation of movie rating using box plot

Table 4: Classification report of gradient boosting classifier performance

	Precision	Recall	F1-score	Support
0	0.53	0.53	0.53	17
1	0.76	0.76	0.76	33
Accuracy			0.68	50
macro avg	0.64	0.64	0.64	50
weighted avg	0.68	0.68	0.68	50

Table 5: Confusion matrix values

Actual values	Predicted values True	False
Positive	25	8
Negative	9	8

However, an analysis of the misclassification rates revealed a discernible bias in the model, which tended to predict films as successful more frequently than warranted. This tendency highlights a potential area for further refinement in the model's predictive algorithms to enhance its accuracy and reliability. Prediction analysis also presents the confusion matrix of where out of 50 movies, 25 are successful and classified as successful, i.e., true positive.

In contrast, eight movies are successful but classified as unsuccessful i.e., false positives. Also, nine movies are unsuccessful and classified as unsuccessful i.e., true negative and eight movies are unsuccessful and classified as successful, i.e., false negative. The tabular information of predicted and actual values is listed in Table (5).

Conclusions and recommendations for future research This investigation advances the understanding of error correction in movie rating predictions and illuminates broader machine-learning applications in media analytics. The observed variability in error rates across genres underscores the critical importance of genre-specific attributes, dataset composition, and model configurations in achieving effective predictive modeling. Future research should aim to refine these models by enhancing feature selection, addressing data imbalances, and experimenting with different model parameters to improve predictive accuracy. These adjustments have the potential to yield more reliable predictions, thereby assisting stakeholders in making informed decisions based on anticipated movie success. This study emphasizes the necessity for continuous model evaluation and optimization to achieve optimal performance in predictive tasks within the entertainment industry. This research significantly enhances predictive models' efficacy in the rapidly evolving landscape of film industry analytics by providing a solid foundation for further research and application development.

Discussion

This study presents important data that shows the great potential in Multi-Layer Perceptron (MLP) models for predictive analytics in the entertainment industry, particularly in forecasting movie ratings. Across different genres in film-making, MLP models with extensive metadata have significantly increased prediction accuracy and dependability. This research reveals that movie ratings are influenced by more factors than just genre types and requires sophisticated predictive models with rich datasets. Central to these developments has been MLP's capacity to capture intricate relationships between many variables which linear models cannot accomplish. One of the major findings from this study is that predictive accuracy varies greatly across genres, implying a need for specialized modeling. For instance, with adventure and mystery genres, there were remarkably low error rates e.g. 0.023 MSE demonstrating MLP model's ability to effectively address their peculiarities. On the other hand, both action and drama genres had higher error measures; hence, these categories might be more about subjective viewers' preferences or numerous production traits. Therefore, such results emphasize that genre-specific modeling is essential since an approach catering for all may not work well. In addition, future studies should focus on understanding some characteristics unique to different genres like audience demographics or cultural connotations which could improve prediction in multiple ways.

In addition, comprehensive metadata has been seen to be of great importance in enhancing predictive outcomes. Inclusion of variables such as cast, production year, language, critical reception and marketing strategies enabled the model to find patterns which are subtle that are not picked up by usual approaches. This means that richer datasets make it possible for the models to perform better since they take account of a variety of factors influencing audience preferences and reactions. Moreover, there is a prospect of integrating real-time feedback from viewers & social media engagement metrics into predictive analytics. However, this study also reveals some weaknesses regarding the current modeling framework. For example, some genres like biography had large variations in their error rates attributed to imbalances in data and inconsistencies within dataset used. These observations suggest the underrepresentation problem in certain genres needs a more advanced approach through data augmentation and resampling techniques. Furthermore, the use hyperparameter tuning and model configuration highlights the difficulty of optimizing MLP performance across multiple datasets. This process can be made more efficient by AutoML tools for automated machine learning that would ensure optimal configurations while reducing risks associated with overfitting.

This research has broad practical implications and can be useful to producers, distributors, and streaming platforms. Among the types of decisions that could be supported by improved predictive models are those concerning what movies to make or distribute; how to create more optimal marketing strategies; and enhancing content recommendation algorithms. Such services as streamers may benefit from correct predictions of movie rankings by being able to deliver personalized content that drives increased viewership and satisfaction rates. Therefore, when predictive analytics are integrated into strategic planning, stakeholders will be in a position to reduce business risk as well as better align production activities with the expectations of their audience. The study paves way for several future studies. Furthermore, incorporation sentiment analysis from platforms like Twitter or Rotten Tomatoes would allow for real-time insights about audience preferences which would enhance the model adaptability. Also, alternative machine learning techniques such as ensemble models and transformer architectures might help improve prediction accuracy further.

Conclusion

The problem addressed in this study was the challenge of accurately predicting movie ratings across various genres in the entertainment industry. Traditional predictive models often fall short due to their inability to capture the complex and multifaceted dynamics that influence movie ratings, such as genre-specific characteristics, audience preferences, and production nuances. These limitations can lead to inaccurate predictions, affecting strategic decision-making by stakeholders, including filmmakers, distributors, and streaming platforms. Consequently, there is a critical need for more sophisticated and data-driven predictive models that can better handle the complexity of entertainment data and provide more reliable forecasts. The primary objective of this research was to develop a more accurate predictive framework for movie ratings by leveraging advanced machine learning techniques, specifically Multi-Layer Perceptron (MLP) models, combined with comprehensive metadata. By integrating MLP models a type of neural network known for its ability to model non-linear relationships with an extensive dataset enriched with diverse movie attributes, this study aimed to enhance the precision of movie rating predictions. The research also sought to highlight the effectiveness of genre-specific models and the importance of including a wide range of variables in the predictive analysis.

The methodology employed in this study involved creating a dedicated MLP regressor model for each genre. The research team collected data from an open API, which was then cleaned and prepared using Principal Component Analysis (PCA) to remove noise, reducing the dataset from 1,600 to 950 entries. This refined dataset was further processed through one-hot encoding to transform categorical data into a numerical format. The data was split into training and testing sets, with 20% reserved for testing. To ensure fair comparisons, the features were standardized using a standard scaler to normalize the data. Individual predictive models were developed for each genre and the models were fine-tuned with specific parameters, such as the number of hidden layers and activation functions, to optimize performance. Statistical methods, including t-tests, ANOVA, and gradient-boosting classifiers, were employed to validate the models' accuracy and effectiveness.

The results of this study underscored the variability in predictive accuracy across different genres, revealing significant differences in error metrics such as Mean Squared Error (MSE), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE). For instance, action movies exhibited moderate predictive accuracy, with MSE, MAE, and RMSE values of 2.816, 1.205, and 1.678, respectively, suggesting that while the model performed reasonably well, there is room for improvement. In contrast, the adventure genre demonstrated much lower error rates, with MSE as low as 0.023, indicating higher predictive precision for this genre. However, within-genre variability was also observed; for example, the biography genre showed considerable fluctuations in predictive accuracy, with MSE ranging from 0.482-0.074. This suggests that even within the same genre, model performance can vary significantly depending on the specific data subset or model configuration. These findings lead to several important conclusions. First, the study confirms that MLP models, when combined with comprehensive metadata, provide a more nuanced and effective approach to predicting movie ratings across different genres compared to traditional models. The ability of MLP models to capture complex, non-linear relationships between various factors significantly enhances their predictive power. Second, the research highlights the importance of genre-specific modeling in predictive analytics, as the variability in error rates across genres suggests that a one-size-fits-all approach may not be effective. Tailoring predictive models to account for the unique characteristics of each genre can lead to more accurate and reliable predictions.

Future research should focus on several key areas to further enhance the predictive capabilities of machine learning models in the entertainment industry. First, refining feature selection techniques is crucial. Including more detailed and nuanced features, such as real-time audience engagement metrics, marketing strategies, and detailed production attributes, could improve model accuracy. Second, addressing data imbalances by employing more robust data augmentation and sampling techniques would ensure more balanced training datasets, especially for underrepresented genres. Third, experimenting with different machine learning algorithms and model parameters, such as deep learning models or ensemble methods, could yield better predictive results. Fourth, incorporating sentiment analysis from social media and review platforms could add a new layer of predictive insight, allowing models to capture real-time shifts in audience preferences and sentiments.

Lastly, continuous model evaluation and optimization are essential. As the entertainment industry evolves, so do audience tastes and market conditions. Therefore, predictive models must be regularly updated and refined to maintain their relevance and accuracy. This involves not only retraining models with new data but also re-evaluating model structures and parameters to adapt to changing conditions. Future work should also explore the applicability of these models to other domains within entertainment analytics, such as box office predictions or content recommendation systems for streaming platforms. In conclusion, this study significantly advances the field of predictive analytics in the entertainment industry by providing a more sophisticated and accurate framework for predicting movie ratings. The integration of MLP models with comprehensive metadata demonstrates a powerful approach to capturing the complex dynamics that influence movie ratings. By focusing on genre-specific modeling and robust validation techniques, this research lays a strong foundation for future studies and practical applications in media analytics, ultimately contributing to more informed and effective decision-making in the film industry.

Acknowledgment

The authors expressed their gratitude to the Institute and University for offering the necessary infrastructure to support this project, and they extend their thanks to these institutions for their contributions.

Funding Information

The authors have not received any financial support or funding to report.

Author’s Contributions

All authors equally contributed to this study.

Ethics

This manuscript is an original work. The corresponding author declares that no ethical concerns are associated with this submission.

References

Ahamad, D., Alam Hameed, S., & Akhtar, M. (2022). A Multi-Objective Privacy Preservation Model for Cloud Security Using Hybrid Jaya-Based Shark Smell Optimization. *Journal of King Saud University - Computer and Information Sciences*, *34*(6), 2343–2358. https://doi.org/10.1016/j.jksuci.2020.10.015

Ahmad, A., Sulaiman, M., Aljohani, A. J., Alhindi, A., & Alrabaiah, H. (2021). Design of an Efficient Algorithm for Solution of Bratu Differential Equations. *Ain Shams Engineering Journal*, *12*(2), 2211–2225. https://doi.org/10.1016/j.asej.2020.11.007

Ahonen, T., Hadid, A., & Pietikäinen, M. (2004). Face Recognition with Local Binary Patterns. In T. Pajdla & M. Jivri (Eds.), *Computer Vision - ECCV 2004* (Vol. 3021, pp. 469–481). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-24670-1_36

Al-Azzam, N., & Shatnawi, I. (2021). Comparing Supervised and Semi-Supervised Machine Learning Models on Diagnosing Breast Cancer. *Annals of Medicine and Surgery*, *62*, 53–64. https://doi.org/10.1016/j.amsu.2020.12.043

Albu, L. L., Lupu, R., & Călin, A. C. (2015). Stock Market Asymmetric Volatility and Macroeconomic Dynamics in Central and Eastern Europe. *Procedia Economics and Finance*, *22*, 560–567. https://doi.org/10.1016/s2212-5671(15)00259-2

Aldalbahi, M., & Walker, G. (2016). Riyadh Transportation History and Developing Vision. *Procedia - Social and Behavioral Sciences*, *216*, 163–171. https://doi.org/10.1016/j.sbspro.2015.12.024

Anoop, C. A., & Kumar, P. (2013). Application of Taguchi Methods and ANOVA in GTAW Process Parameters Optimization for Aluminium Alloy 7039. *International Journal of Engineering and Innovative Technology (IJEIT*, *2*(11), 54–58.

Bapna, I., Sood, V., Kumar Totala, N., & Saluja, H. S. (2012). Dynamics of Macroeconomic Variables Affecting Price Innovation in Gold: A Relationship Analysis. *Pacific Business Review International*, *5*(1), 1–10.

Dickerson, T. L., & Przydatek, J. (2003). Fatigue of Friction Stir Welds in Aluminium Alloys that Contain Root Flaws. *International Journal of Fatigue*, *25*(12), 1399–1409. https://doi.org/10.1016/s0142-1123(03)00060-4

Ghisani, F., Timmo, K., Altosaar, M., Mikli, V., Pilvet, M., Kaupmees, R., Krustok, J., Grossberg, M., & Kauk-Kuusik, M. (2021). Characterization of Tetrahedrite Cu10Cd2Sb4S13 Monograin Materials Grown in Molten CdI2 and LiI. *Thin Solid Films*, *739*, 138980. https://doi.org/10.1016/j.tsf.2021.138980

Gupta, S., Banerjee, A., Sarkar, J., Kundu, M., Sinha, S. K., Bandyopadhyay, N. R., & Ganguly, S. (2020). Modelling the Steel Microstructure Knowledge for in-Silico Recognition of Phases Using Machine Learning. *Materials Chemistry and Physics*, *252*, 123286. https://doi.org/10.1016/j.matchemphys.2020.123286

Hanif, H., Md Nasir, M. H. N., Ab Razak, M. F., Firdaus, A., & Anuar, N. B. (2021). The Rise of Software Vulnerability: Taxonomy of Software Vulnerabilities Detection and Machine Learning Approaches. *Journal of Network and Computer Applications*, *179*, 103009. https://doi.org/10.1016/j.jnca.2021.103009

Hijawi, W., Alqatawna, J., Al-Zoubi, A. M., Hassonah, M. A., & Faris, H. (2021). Android Botnet Detection Using Machine Learning Models Based on a Comprehensive Static Analysis Approach. *Journal of Information Security and Applications*, *58*, 102735. https://doi.org/10.1016/j.jisa.2020.102735

Hossain, M., Mullally, C., & Asadullah, M. N. (2019). Alternatives to Calorie-Based Indicators of Food Security: An Application of Machine Learning Methods. *Food Policy*, *84*, 77–91. https://doi.org/10.1016/j.foodpol.2019.03.001

Kapoor, I., & Mishra, A. (2018). Automated Classification Method for Early Diagnosis of Alopecia Using Machine Learning. *Procedia Computer Science*, *132*, 437–443. https://doi.org/10.1016/j.procs.2018.05.157

Kersting, K. (2018). Machine Learning and Artificial Intelligence: Two Fellow Travelers on the Quest for Intelligent Behavior in Machines. *Frontiers in Big Data*, *1*, 6. https://doi.org/10.3389/fdata.2018.00006

Kuo, R. J. (2000). Multi-Sensor Integration for on-Line Tool Wear Estimation Through Artificial Neural Networks and Fuzzy Neural Network. *Engineering Applications of Artificial Intelligence*, *13*(3), 249–261. https://doi.org/10.1016/s0952-1976(00)00008-7

Lai, X., Balakrishnan, A., Lange, T., Jenihhin, M., Ghasempouri, T., Raik, J., & Alexandrescu, D. (2019). Understanding Multidimensional Verification: Where Functional Meets Non-Functional. *Microprocessors and Microsystems, 71*, 102867. https://doi.org/10.1016/j.micpro.2019.102867

Lei, D., Sun, J., & Xia, J. (2023). Cuproptosis-Related Genes Prediction Feature and Immune Microenvironment in Major Depressive Disorder. *Heliyon*, *9*(8), e18497. https://doi.org/10.1016/j.heliyon.2023.e18497

Liang, P., Yang, W., Chen, X., Long, C., Zheng, L., Li, H., & Zuo, Y. (2020). Machine Learning of Single-Cell Transcriptome Highly Identifies mRNA Signature by Comparing F-Score Selection with DGE Analysis. *Molecular Therapy - Nucleic Acids*, *20*, 155–163. https://doi.org/10.1016/j.omtn.2020.02.004

Ligthart, A., Catal, C., & Tekinerdogan, B. (2021). Analyzing the Effectiveness of Semi-Supervised Learning Approaches for Opinion Spam Classification. *Applied Soft Computing*, *101*, 107023. https://doi.org/10.1016/j.asoc.2020.107023

Ma, C. (2021). Smart City and Cyber-Security; Technologies Used, Leading Challenges and Future Recommendations. *Energy Reports*, *7*, 7999–8012. https://doi.org/10.1016/j.egyr.2021.08.124

Mondal, B., Banerjee, A., & Gupta, S. (2022). A Review of SQLI Detection Strategies Using Machine Learning. *International Journal of Health Sciences*, *2*, 9664–9677. https://doi.org/10.53730/ijhs.v6ns2.7519

Neeli, J., & Patil, S. (2021). Insight to Security Paradigm, Research Trend and Statistics in Internet of Things (IoT). *Global Transitions Proceedings*, *2*(1), 84–90. https://doi.org/10.1016/j.gltp.2021.01.012

Palanivel, R., & Koshy Mathews, P. (2012). Prediction and Optimization of Process Parameter of Friction Stir Welded AA5083-H111 Aluminum Alloy Using Response Surface Methodology. *Journal of Central South University*, *19*(1), 1–8. https://doi.org/10.1007/s11771-012-0964-y

Provoost, J. C., Kamilaris, A., Wismans, L. J. J., van der Drift, S. J., & van Keulen, M. (2020). Predicting Parking Occupancy Via Machine Learning in the Web of Things. *Internet of Things*, *12*, 100301. https://doi.org/10.1016/j.iot.2020.100301

Radojevic, I., & Slijepčević, M. (2018). Current Trends Iin Digital Marketing Communication. *Symorg 2018 - XVI International Symposium: Doing Business in the Digital Age: Challenges, Approaches and Solutions*, 2–7.

Razfar, N., True, J., Bassiouny, R., Venkatesh, V., & Kashef, R. (2022). Weed Detection in Soybean Crops Using Custom Lightweight Deep Learning Models. *Journal of Agriculture and Food Research*, *8*, 100308. https://doi.org/10.1016/j.jafr.2022.100308

Shah, D., Isah, H., & Zulkernine, F. (2018). Predicting the Effects of News Sentiments on the Stock Market. *2018 IEEE International Conference on Big Data (Big Data)*, 4705–4708. https://doi.org/10.1109/bigdata.2018.8621884

Singh, A. K., Patra, J., Chakraborty, M., & Gupta, S. (2022). Prediction of Indian Government Stakeholder Oil Stock Prices Using Hyper Parameterized LSTM Models. *2022 International Conference on Intelligent Controller and Computing for Smart Power (ICICCSP)*, 1–6. https://doi.org/10.1109/iciccsp53532.2022.9862425

Sohom, B., Shubham, B., Anup, D., Anirban, M., Ishita, B., & Subir, G. (2021). Machine Learning-Based Naive Bayes Approach for Divulgence of Spam Comment in Youtube Station. *International Journal of Engineering and Applied Physics*, *1*(3), 278–284.

DOI: 10.3844/jcssp.2025.905.917

Taskiran, M., Kahraman, N., & Erdem, C. E. (2020). Face Recognition: Past, Present and Future (A Review). *Digital Signal Processing*, *106*, 102809. https://doi.org/10.1016/j.dsp.2020.102809

Zhang, M., & Kelly, P. M. (2009). The Morphology and Formation Mechanism of Pearlite in Steels. *Materials Characterization*, *60*(6), 545–554. https://doi.org/10.1016/j.matchar.2009.01.001