Predicting Oscar Nominees: A Multi-Labeling Approach with Logistic Regression, Multi-Layer Perceptron, and Random Forest Techniques

Victor Wu
wu.victo@northeastern.edu

Macarious Kin Fung Hui
hui.mac@northeastern.edu

Tianyi Zhang
zhang.tianyi9@northeastern.edu

Abstract

The allure of an Oscar nomination is a pinnacle achievement in the film industry, representing artistic and technical excellence. In this study, we present a exploration of predicting Oscar nominations across 13 major categories using machine learning techniques. Our multi-label classification approach considers diverse features such as awards, movie attributes, and critical acclaim. We deploy Logistic Regression, Neural Networks, and Random Forest models to predict the likelihood of nominations. Each algorithm brought unique strengths to the table, and the choice of model could be tailored to the specific goals and priorities of the analysis. Logistic Regression excelled in interpretability, providing actionable insights into specific features' impacts. Neural Networks, with their capacity to capture complex relationships. Random Forest, with its ensemble approach, showcased robustness across diverse datasets. The evaluation, instead of traditional accuracy, encompasses balanced accuracy, precision, recall, and F1-score, to address challenges posed by class imbalances. Additionally, we unravel feature importance, offering insights into the factors influencing nominations. Our work not only contributes to predicting Oscar nominations but also sheds light on the complex interplay of factors in the film industry.

Keywords: Oscar nominations, machine learning, logistic regression, neural networks, random forest, multi-label classification, award season, film studies.

1. Introduction

The Oscars highlight a film's artistic and technical merits in the movie industry. Each year, the excitement and speculation around the nominations demonstrate the importance of the awards that go beyond the screen. Campaigning for nominations is a high-stakes effort for studios and filmmakers, given the prestige, opportunities, and lucrative media coverage the award show offers. However, the journey to securing a nomination is often opaque and subject to a great deal of uncertainty. As a result, forecasting Oscar nominations can be challenging and difficult due to the weight of

Figure 1. The epitome of cinematic excellence, the iconic Oscar trophy, gleaming with aspirations of filmmakers around the globe.

cultural, economic, and political factors that contribute to them.

The consumption and recognition of cinema have long defined cultural production and reflected the zeitgeist. Unraveling the features influencing the Academy Awards offers insights into societal consciousness. Notably, age dynamics in actor and actress recognition hint at societal attitudes and power structures within award bodies. Exploring the significance of features like genre or box office in Oscar nominations raises critical questions about cultural prestige. Our machine learning-driven approach seeks to tackle these issues, shedding light on audience preferences and economic structures shaping the perceived artistic meritocracy of awards. In an era of concerns about technology misuse, our project showcases the constructive use of machine learning to diagnose cultural patterns, offering a pathway to progressive change.

Our machine learning approach aims to enhance the transparency of the Oscar nomination process. By employing algorithms to predict nominations across categories, we adopt a multi-label strategy, analyzing diverse factors contributing to a movie's likelihood of receiving recognition. This endeavor, along with a feature importance analysis, aligns with our broader goal of leveraging machine learning to gain deeper insights into the intricacies of the Oscar nomination landscape.

2. Related Work

In recent years, there has been a growing interest in employing machine learning techniques to predict the nominees and winners of prestigious awards such as the Academy Awards (Oscars). We draw insights from existing literature, where researchers have explored diverse methodologies to forecast Oscar outcomes based on various features and datasets.

Deregowski [3] takes a machine learning approach to predict Academy Award nominees and winners. Leveraging a training dataset over 5 years, 2015 to 2020, the author employs AutoML, Preferential Machine Learning, and Social Network Services to analyze features like genre, budget, reviews, and ratings. The study presents an automated model selection process, a simulation of the Academy's voting system, and an exploration of the impact of social media on predicting Oscar outcomes. The author finds that Oscar winners are usual the resultant of nominations and winners of less prestigious awards (e.g. Golden Globes, BAFTA, Critics Choice, industry awards).

Kim, Hwang, and Park [5] adopt a unique perspective by examining user-generated content on the social network service Reddit. Focusing on posts about nominated films, the authors employ natural language processing techniques to analyze sentiments from a single subreddit channel (movie) They propose an adjusted Oscar winner score, reflecting public perception based on social media buzz. The study successfully predicts the 2020 Oscar winner by considering the sentiment expressed in user-created content.

Our approach builds upon this existing body of work, contributing novel aspects to the prediction of Oscar nominations. We focus on three distinct machine learning models—logistic regression, neural network, and random forests. Unlike some studies, we adopt a temporal prediction strategy, forecasting nominations by each years rather than using a simple classification model with a threshold on the Softmax score. Furthermore, our emphasis on nominations, rather than solely on winners; we argue that this is more realistic and representative of the actual voting process, since only a few movies win Oscars each year, but many more are nominated. This also reduces data imbalance when compared to previous models which only predicts winners and non-winners. In addition to nomination predictions, our study aims to delve deeper into the features associated with Oscar success. We seek to uncover potential social and economic impacts by evaluating the feature importance from each Oscar categories, exploring how certain characteristics such as genre, budget, and critical reviews may influence the recognition of films by the Academy. We aspire to contribute valuable insights into the broader cultural and economic implications of success in the film industry.

3. Method

3.1. Data Collection

Our approach focuses on using a range of movie-related information, starting from the development stage to the final Oscar nomination announcements. We've identified key factors that we believe play a significant role in determining a movie's nomination prospects. These features include:

- Golden Globe nominations
- Screen Actors Guild (SAG) Nominations
- Release year and release month
- Movie genre (up to 3)
- Budget and revenue
- Movie run time
- Studio and production companies
- Age of actors, actresses, directors, and writers, and
- Ratings and number of votes on IMDb (Internet Movie Database)

The features are extracted by leveraging dataset sourced from reputable platforms like Kaggle and IMDb, which provide a comprehensive view of the movies, their performance metrics, and awards history. Specifically, the dataset we utilized include:

- The Movies Dataset [2]
- IMDb datasets[4]
- Golden Globe nominations dataset [8]
- Oscar nominations dataset [9]
- SAG Award nominations dataset [10]

The goal is to identify the key features that contribute to a movie's likelihood of being nominated across the 13 major Oscar categories:

- Best Picture
- Actor in a Leading Role
- Actress in a Leading Role
- Cinematography
- Directing
- Film Editing
- Actor in a Supporting Role
- Actress in a Supporting Role
- Costume Design
- Visual Effects
- Animated Feature Film
- Production Design

It is worth to note that one movie can win multiple nominations, so rather than a multi-class classification (choosing one class amongst many), it is a multi-labelling problem (each class can either be chosen or not chosen).

3.2. Train/Validation/Test Split

The dataset spans from 1927 to 2017, with movies from 2013 to 2017 isolated as the test set. The remaining data is divided into training and validation sets, as summarized in Table 1.

Data Split	Years	Movie Count
Training	1927–2010 (inclusive)	33514
Validation	2011–2012 (inclusive)	3395
Test	2013–2017 (inclusive)	7912

Table 1. Train/Validation/Test Split

Utilizing recent years for validation and test sets serves two key purposes. Firstly, it ensures the model is assessed on data that closely reflects the current landscape of the film industry. This is crucial for predicting nominations in the near future as industry trends and audience preferences evolve over time. Secondly, by training the model on historical data up to 2010 and validating and testing on subsequent years, we establish a robust evaluation framework. This approach simulates real-world scenarios where the model must generalize from past trends to make predictions for contemporary films. The focus on recent years enhances the model's adaptability to the dynamic nature of the film industry. This meticulous split strategy enables the model to learn from established patterns while fine-tuning its predictions to align with the ever-changing landscape of Oscar nominations.

3.3. Model

We propose three machine learning models to make the prediction, and we will evaluate the methods and compare the results of the three models against each other.

3.3.1 Logistic Regression Model

Logistic regression is a statistical model employed for binary classification, where the goal is to predict the probability of an instance belonging to a particular class. The logistic regression model utilizes the logistic function, also known as the sigmoid function, to transform linear combinations of input features into a probability value between 0 and 1. The logistic function is defined as[6]:

$$\sigma(z) = \frac{1}{1 + e^{-z}} \tag{1}$$

where z represents the linear combination of input features and model parameters. The logistic function ensures that the output lies within the range [0, 1], making it suitable for estimating probabilities.

For a single instance, the logistic regression model calculates the probability $P(y = 1)$ of belonging to the positive class as follows:

$$P(y = 1) = \sigma(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n) \tag{2}$$

Here, β_0 is the intercept, β_i represents the coefficients associated with each input feature x_i, and n is the number of features.

In the context of multi-label classification for predicting Oscar nominations across different categories, logistic regression can be extended to handle multiple labels. Each Oscar category is treated as a separate label, and the model is trained to predict the probability of a movie being nominated in each category independently.

The training process involves optimizing the model parameters β to maximize the likelihood of the observed nominations in the dataset. This is typically achieved through the minimization of a cost function, such as the cross-entropy loss, which penalizes deviations between predicted probabilities and actual nominations.

In summary, the logistic regression model, with its foundation in the logistic function, provides a mathematical framework for estimating probabilities and making binary decisions, making it a powerful tool for predicting Oscar nominations across diverse categories.

3.3.2 Neural Networks

A neural network, inspired by the human brain, is a computational model composed of layers of interconnected nodes. When applied to predicting Oscar nominations, neural networks are customized for multi-label classification, treating each category independently. Through training on historical data, the network learns patterns and relationships, allowing it to generalize and make predictions about the likelihood of new films receiving Oscar nominations based on various features and attributes.

The structure of a neural network involves an input layer, layers of hidden units, and an output layer. The input layer receives features denoted as X. Following the input layer, the hidden layers, containing individual hidden units (a), process information through weighted connections ($w_j^{[i]}$) and activation functions. The weighted connections capture the importance of each input feature, and the activation functions introduce non-linearity to the model. Finally, the output layer generates the final predictions ($\hat{y}$)[1]. The output layer is customized for multi-label classification. The values at these nodes represent the likelihood of a movie being nominated in each category.

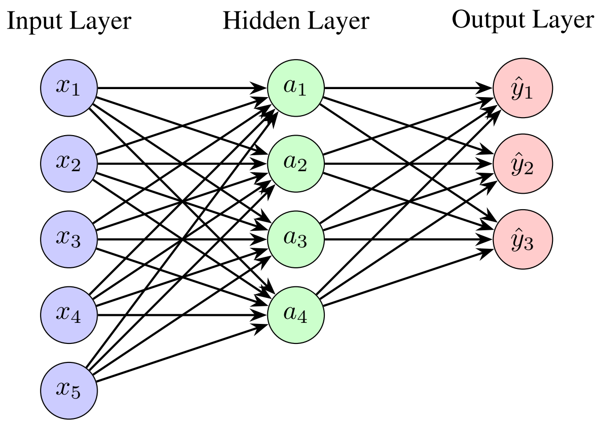

Figure 2. Neural Network Architecture

During forward propagation, inputs (x), weights, and biases (b) are used to compute the output ($z_j^{[i]}$) for each hidden unit. Here, i represents the layer number, and j represents the hidden unit within that layer:

$$z_j^{[i]} = w_j^{[i]} x + b_j^{[i]} \tag{3}$$

At the end of each hidden unit, an activation function introduces non-linearity to the model. Common activation functions include sigmoid and ReLU. After obtaining the outputs ($z_j^{[i]}$), the network moves on to the next layer, repeating the process until the final output layer is reached. The predicted values $\hat{y}$ are compared with the actual values y using a suitable loss function, such as cross-entropy for classification tasks.

The goal during training is to minimize the loss by adjusting the weights ($w_j^{[i]}$) and biases ($b_j^{[i]}$) through a process called backpropagation. The optimization algorithm updates these parameters iteratively, improving the model's ability to make accurate predictions.

In predicting Oscar nominations, this network learns from the selected film features, discerns patterns from the features through the complex network of hidden units, generalizes its understanding, and determine how likely a film is to be nominated for each category.

3.3.3 Random Forest

The Random Forest model, adept for multi-label classification tasks such as predicting Oscar nominations, integrates multiple decision trees to enhance prediction accuracy and robustness. The model's structure has:Decision Tree Classifier: A decision tree is a flowchart-like structure where each internal node represents a feature (or attribute), each branch represents a decision rule, and each leaf node represents an outcome. The decision at each node is made based on the value of the corresponding feature in the input data. Mathematically, a decision tree partitions the input space into regions $R_1, R_2, \ldots, R_n$, and for each region, it makes a prediction based on the majority label of the training instances in that region. Bagging Classifier: The Bagging Classifier, fundamental to the Random Forest, employs the concept of bootstrap aggregating. It trains N decision trees on slightly different subsets of the data (obtained by bootstrapping) and aggregates their predictions. The aggregation for classification tasks is typically performed by majority voting:

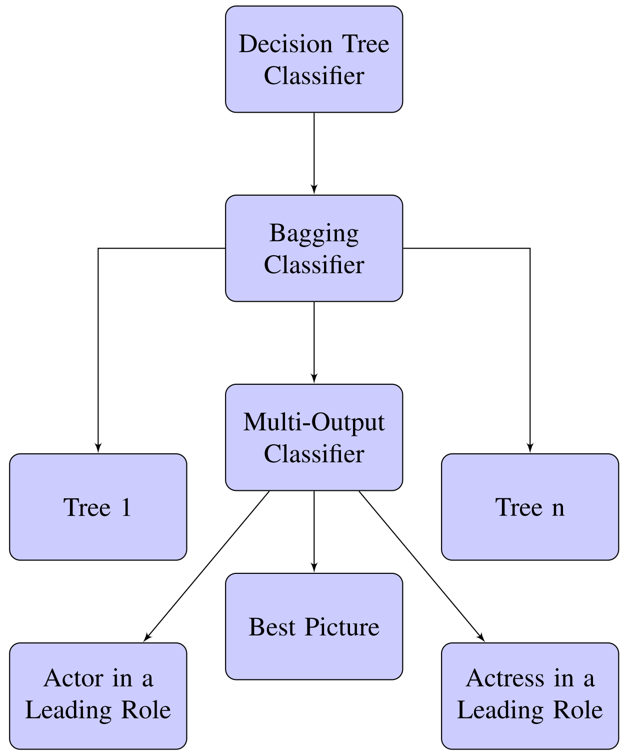

Figure 3. Random Forest Model Diagram

$$Y_{\text{RF}} = \text{mode}\{y_1, y_2, ..., y_N\} \tag{4}$$

where Y_{RF} is the final output of the Random Forest and y_i is the prediction of the i-th decision tree.

Multi-Output Classifier: In a multi-label context, the Random Forest model is extended to a Multi-Output Classifier, where it predicts across multiple labels simultaneously. If M is the number of labels, the final output for a given instance is a vector $[Y^1, Y^2, ..., Y^M]$, where each Y^m is obtained by a separate Random Forest tailored for the m-th label.

3.4. Prediction

In traditional binary classification problems, each input vector is classified based on whether its Softmax score is greater or less than a threshold. However, to ensure meaningful predictions and predict exactly 5 nominations per year per category, we adopt a unique strategy in our

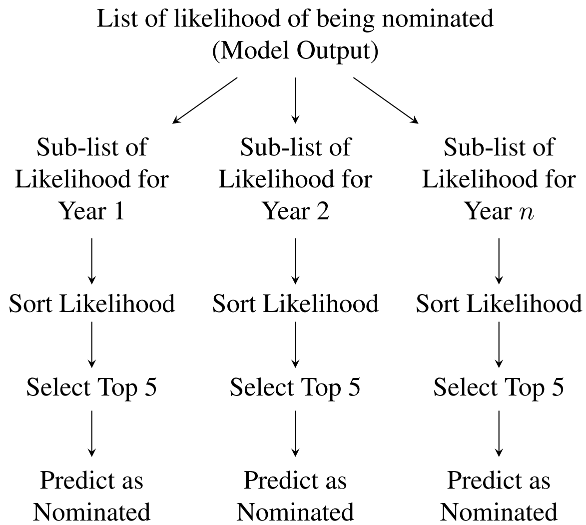

Figure 4. Prediction Flow Chart with Arrows

approach to predict Oscar nominations across the 13 major categories. After training our multi-label classification models, we obtain probability scores for each movie across the 13 Oscar categories. These probability scores represent the likelihood of each movie being nominated in a specific category, then we employ a two-step process. First, we group the movies and their associated probabilities by the release year to create year-specific data set for each category. This step allows us to analyze nominations on a yearly basis. Second, within each category for a given year, we sort the movies based on their probability scores in descending order. This sorting places the most likely nominees at the top of the list, reflecting the strength of their nomination prospects. From this ranked list, we select the top 5 movies with the highest probability scores as our predictions for Oscar nominations in that category for that year. Note that in the case of Best Pictures, the number of nomination per year varies throughout the years. The aforementioned two-step approach is adapted for the Best Picture category accordingly.

This methodology ensures that our predictions align with the historical Oscar nomination process. By considering the top 5 movies based on probability scores, we emulate the actual nominations in each category. This process is particularly valuable when evaluating our models' performance, as it allows for a direct comparison between our predictions and the official Oscar nominations. Also, this approach relies on the aforementioned process of splitting our data into training, validation, and test sets. By creating year-specific test/validation/train splits, we ensure that our predictions can be meaningfully compared and evaluated against the actual Oscar nominations.

3.5. Evaluation

The evaluation of our predictive model for Oscar nominations is a critical step in assessing its reliability and usefulness. In this section, we present a comprehensive analysis of the model's performance, feature importance, and interpretability. By employing a range of metrics and techniques, we aim to provide insights into how well the model captures the complexities of predicting nominations across diverse categories. Additionally, we delve into the significance of key features and the transparency of the model's decision-making process. Through this evaluation, we seek to demonstrate the model's effectiveness and offer valuable insights for potential users and stakeholders in the film-making industry.

Model Performance: Before delving into specific metrics, it is essential to understand the context in which the model operates. Given the inherent class imbalances in the dataset, traditional accuracy may not provide a complete picture of the model's effectiveness. We introduce a set of metrics, including balanced accuracy, precision, recall, and F1-score, to evaluate the model's performance in predicting Oscar nominations.

- **Traditional Accuracy:** Accuracy is an essential metric, measuring the proportion of the top 5 predicted nominations that match the actual nominations for a specific category and year. However, due to the highly imbalanced nature of the dataset, where non-nominations significantly outnumber nominations, traditional accuracy alone may not provide a comprehensive evaluation. Therefore, we consider accuracy as a supplementary metric. Accuracy is calculated as:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{5}$$

  where:

$$\begin{aligned} TP &= \text{True Positives} \\ TN &= \text{True Negatives} \\ FP &= \text{False Positives} \\ FN &= \text{False Negatives} \end{aligned}$$

- **Balanced Accuracy:** Recognizing the imbalances in the dataset, we incorporate balanced accuracy as a key evaluation metric. Unlike traditional accuracy, which is calculated as the ratio of correctly predicted instances to the total number of instances, balanced accuracy considers the distribution of classes. It ensures that the accuracy calculation considers the true positive rate for each class, preventing inflated accuracy scores driven by the dominance of the majority class. The balanced accuracy is calculated as:

$$\text{Accuracy} = \frac{TP}{TP + FN} + \frac{TN}{TN + FP} \tag{6}$$

- **Precision, Recall, and F1-score:** These metrics are crucial for assessing the model's performance, especially in the presence of class imbalances. Precision measures the fraction of the predicted top 5 nominations that are correct, recall evaluates the ability to capture all actual nominations, and F1-score provides a balanced measure of precision and recall. We calculate these metrics independently for each label, treating the prediction task as a multi-label classification problem. These metrics are calculated as:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{7}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{8}$$

$$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{9}$$

By incorporating a range of metrics, including balanced accuracy, precision, recall, and F1-score, we aim to provide a comprehensive evaluation of the model's effectiveness in predicting Oscar nominations across diverse categories while addressing the challenges posed by class imbalances in the dataset.

Feature Importance: Feature importance helps us understand the key features which contributing to a movie's likelihood of being nominated is crucial for model interpretability and decision-making. Feature importance analysis involves:

- Identification of Key Features: Analyzing which features significantly influence the model's predictions.
- Ranking of Feature Importance: Assigning scores or rankings to features based on their contribution to prediction outcomes.

This analysis provides valuable insights into the factors that drive the model's decision-making process.

Model Interpretability: Decisions made by a complex model is essential for building trust and understanding its behavior. Model interpretability is achieved through:

- Visualizations: Creating visual representations of model predictions, such as
confusion matrices or ROC curves.
- Decision Trees: Examining decision trees to understand the logic behind
the model's predictions.

These interpretability measures help stakeholders, including filmmakers and industry professionals, make informed decisions based on the model's predictions.

4. Experiment

4.1. Dataset Preprocessing

The dataset used in our study comprised a diverse set of Oscar nomination categories, each characterized by varying counts of true and false instances. The distribution of instances across categories is detailed in Table 2, and make note of the true/false label imbalance mentioned earlier.

Category	**False Count**	**True Count**
Cinematography	33269	245
Directing	33152	362
Film Editing	33167	347
Actor in a Supporting Role	33176	338
Actress in a Supporting Role	33199	315
Costume Design	33319	195
Sound	33311	203
Best Picture	33276	238
Actor in a Leading Role	33353	161
Actress in a Leading Role	33365	149
Visual Effects	33424	90
Animated Feature Film	33481	33
Production Design	33513	1

Table 2. Class distribution in the dataset

To ensure the dataset's suitability for machine learning, we conducted extensive preprocessing steps. Notably, some features such as revenue and budget were not available for all movies. To address this, mean imputation was employed. Moreover, considerations were made for potential biases arising from the absence of inflation adjustments. Categorical variables, including genre and release month, were transformed using one-hot encoding. The age of the first three actors and actresses from the credits were extracted, with missing values filled using the mean of other actors or actresses from the same movie. Language was split into English and non-English, and top production companies and countries were one-hot encoded. The top 10 production companies and top 5 countries were selected for encoding, with the remaining labeled as 'other' to accommodate variations in names and subsidiaries.

Notably, the dataset's cleaning process acknowledges potential nuances in film industry practices and allows for flexibility in handling diverse information across different time periods.

4.2. Software and Hardware

A high-performance personal computer will be the primary hardware resource; if the computational requirements exceed the capabilities of the personal computer, NEU discovery cluster will be utilized which provides access to high-performance computing resources including GPUs; however, at the end, this was deemed unnecessary.

The project will be coded in python. Libraries such as Pandas and NumPy for data manipulation, and Scikit-learn for machine learning will be used.

4.3. Logistic Regression Model

In the context of predicting Oscar nominations, logistic regression serves as a valuable predictive tool. By leveraging a mathematical model that analyzes the relationships between various film-related features and the likelihood of receiving an Oscar nomination, logistic regression allows us to make informed predictions based on historical data. This method enables us to discern patterns and correlations within the extensive dataset of movies from different years, aiding in the identification of key factors that contribute to the success of a film in garnering Oscar recognition. Through the meticulous examination of features such as critical acclaim, box office performance, and genre, logistic regression facilitates a quantitative understanding of the nomination process, thereby offering insights that contribute to the formulation of more accurate predictions for future Oscar contenders.

4.3.1 Model Configuration

When dealing with a problem where the task involves predicting multiple output labels simultaneously, *MultiOutputClassification*, an extension of multi-label classification, proves to be a powerful approach (Figure 5). In the specific context of predicting Oscar nominations across various categories, employing *MultiOutputClassification* allows us to build a dedicated predictive model for each nomination category, capturing the nuanced relationships between film features and the likelihood of receiving a nomination in each specific area.

In configuring the logistic regression model, the liblinear solver was employed, and regularization was implemented with both L1 and L2 penalties, denoted as [l1, l2]. A systematic exploration of the regularization parameter C was conducted, encompassing values within the range [0.001, 0.01, 0.1, 1, 10, 100]. To ensure robustness and generalizability, a 5-fold cross-validation strategy was adopted (cv = [5]). The optimization of the model was guided by the F1-score, a critical metric for assessing the balance between precision and recall in classification tasks (scoring = [F1]). These parameterization choices were made to enhance the model's predictive performance and its ability to generalize to unseen data.

4.3.2 Training and Validation

The logistic regression model was trained on a dataset comprising movies from 1927 to 2012, inclusive, with the test set consisting of movies from 2013 to 2017, inclusive. During training, the focus was on evaluating performance metrics such as F1-score, precision, and recall.

4.3.3 Feature Importance Analysis

Best Picture - Strongly Positive Features	**Scores**
gg_nominee_Best Screenplay - Motion Picture	2.07
gg_nominee_Best Motion Picture - Drama	1.86
sag_nominee_CAST IN A MOTION PICTURE	1.84
sag_winner_CAST IN A MOTION PICTURE	1.78
gg_winner_Best Performance by an Actress in a Motion Picture - Drama	1.61
gg_winner_Best Motion Picture - Animated	1.60
gg_nominee_Best Motion Picture - Musical or Comedy	1.50
sag_winner_LEADING ROLE - MALE	1.49
gg_nominee_Best Motion Picture - Animated	1.34
gg_winner_Best Motion Picture - Musical or Comedy	1.13

Table 3. Strongly positive features for Best Picture

Leading Actor - Strongly Positive Features	**Scores**
gg_nominee_Best Performance by an Actor in a Motion Picture - Drama	4.06
sag_nominee_LEADING ROLE - MALE	4.04
gg_nominee_Best Performance by an Actor in a Motion Picture - Musical or Comedy	2.31
gg_winner_Best Performance by an Actor in a Motion Picture - Musical or Comedy	1.46
productionCompanyColumbiaPictures	1.45
gg_nominee_Best Screenplay - Motion Picture	1.36
genreDrama	1.18
genreSport	0.97
gg_nominee_Best Motion Picture - Animated	0.93
releaseMonth12	0.92

Table 4. Strongly positive features for Actor in a Leading Role

To analyze feature importance, I meticulously examined the coefficients and organized tables to document the most positively and negatively correlated features for each category in the model. These tables served as critical tools for validating the model's weights, particularly as certain categories from the Screen Actors Guild (SAG) and the Golden Globes exhibited substantial influence on the model's predictions.

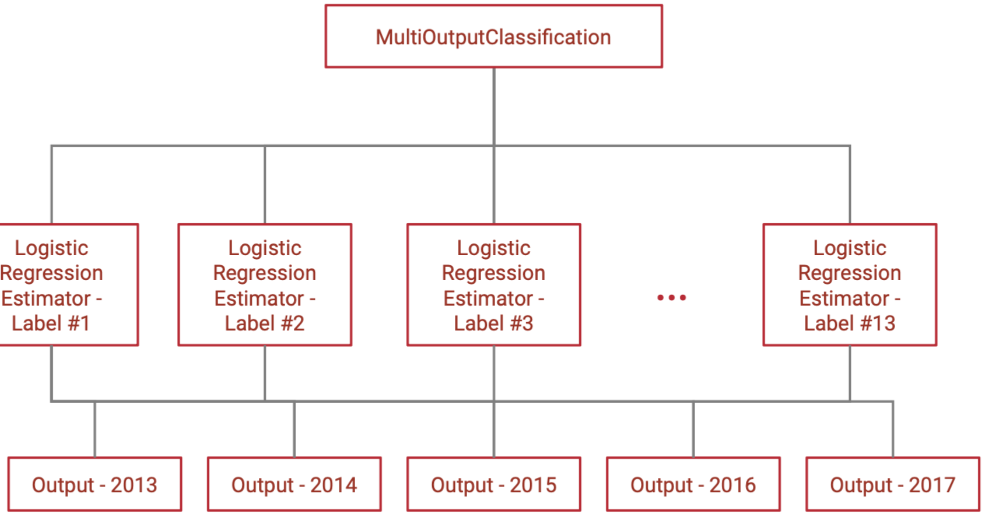

Figure 5. Diagram of Logistic Regression using MultiOutputClassification

Leading Actress - Strongly Positive Features	**Scores**
gg_nominee_Best Performance by an Actress in a Motion Picture - Drama	4.22
sag_nominee_LEADING ROLE - FEMALE	3.59
gg_nominee_Best Performance by an Actress in a Motion Picture - Musical or Comedy	2.72
gg_winner_Best Performance by an Actress in a Motion Picture - Musical or Comedy	2.43
gg_nominee_Best Screenplay - Motion Picture	1.32
productionCompanyColumbiaPictures	0.94
sag_winner_LEADING ROLE - FEMALE	0.87
genreDrama	0.76
genreSci-Fi	0.69
averageRating	0.66

Table 5. Strongly positive features for Actress in a Leading Role

Best Picture - Strongly Negative Features	**Scores**
genreFilm-Noir	-0.50
gg_winner_Best Original Song - Motion Picture	-0.55
productionCountryItaly	-0.57
genreHorror	-0.57
gg_winner_Best Performance by an Actor in a Supporting Role in any Motion Picture	-0.73
genreDocumentary	-0.79
genreSci-Fi	-0.82
productionCompanyNewLineCinema	-0.88
productionCountryGermany	-0.92
productionCountryOther	-1.07

Table 6. Strongly negative features for Best Picture

Throughout this exploration, intriguing insights surfaced for each of the three main categories. In the Best Picture category, it was noteworthy that Best Screenplay held a more substantial influence than Best Motion Picture - Drama at

Leading Actor - Strongly Negative Features	**Scores**
genreDocumentary	-0.77
productionCompanyMetroGoldwynMayerMGM	-0.80
gg_winner_Best Original Song - Motion Picture	-0.88
sag_nominee_SUPPORTING ROLE - FEMALE	-0.92
releaseMonth05	-0.98
productionCountryOther	-1.13
releaseMonth01	-1.22
sag_winner_CAST IN A MOTION PICTURE	-1.23
gg_winner_Best Original Score - Motion Picture	-1.26
genreFamily	-1.37

Table 7. Strongly negative features for Actor in a Leading Role

Leading Actress - Strongly Negative Features	**Scores**
genreWestern	-0.45
genreHorror	-0.46
genreAnimation	-0.51
genreDocumentary	-0.66
gg_winner_Best Performance by an Actor in a Motion Picture - Musical or Comedy	-0.73
genreAction	-0.77
productionCountryOther	-0.80
genreMusical	-0.89
productionCompanyColumbiaPicturesCorporation	-0.95
releaseMonth06	-0.95

Table 8. Strongly negative features for Actress in a Leading Role

the Golden Globes. Additionally, genres such as horror and sci-fi consistently displayed poor performance in this category.

An interesting pattern emerged for leading actresses, revealing that victories in Westerns or action films, traditionally associated with masculine and male-dominated genres, were uncommon. Moreover, a Leading Actor win at the Golden Globes emerged as a negative feature, suggesting a rarity in scenarios where both male and female leads secure Oscar nominations.

The temporal aspect of release months emerged as a significant factor, indicating that films released between March and June tended to underperform. Similarly, movies produced by companies based outside of the United States exhibited diminished success.

It's essential to recognize that this feature analysis only scratches the surface of the potential insights that can be derived from the model. Future work in this domain could elevate feature analysis to the forefront, making it the central focus of investigation and exploration. This could uncover deeper connections and provide more nuanced interpretations of the intricate dynamics influencing award outcomes in the film industry.

4.3.4 Performance Evaluation

In the performance evaluation phase, I deliberately avoided accuracy due to the dataset's skewness towards films not being nominated. Instead, I prioritized precision, recall, and F1-score. The evaluation involved comparing the list of predicted nominated films against the actual nominations.

4.4. Neural Network

In this section, we delve into the intricacies of employing neural networks for predicting Oscar nominations. Our approach involves model configuration, hyperparameter exploration, and a comprehensive analysis of the model's performance.

4.4.1 Model Configuration

To implement our neural network model, we utilized the *MLPClassifier* class from Scikit-Learn, with MLP standing for Multi-Label Perceptron. The model offers a choice of solvers, including Limited-memory Fletcher-Goldfarb-Shanno Algorithm (L-BFGS), Stochastic Gradient Descent, and Adam Optimization Algorithm. Given the dataset's size, L-BFGS was specifically chosen as the optimizer. L-BFGS, belonging to the family of quasi-Newton methods, eliminates the need to specify a learning rate, as the descent is computed by the algorithm [11].

Maintaining a constant number of hidden units at 100, we explored the impact of varying the number of hidden layers. Three key hyperparameters were adjusted: the arrangement of hidden layers, the activation functions, and the L2 regularization strength (α). The investigation covered the following hyperparameter ranges:

Hidden Layer Count	3 layers	2 layers	1 layer
Activation Function	Logistic	ReLU	
α	0.001	0.01	0.1

Table 9. Hyperparameter exploration

The hyperparameter exploration includes variations in the number of hidden layers, activation functions, and L2 regularization strength (α). The choice of activation function is a critical aspect of neural network design. The logistic activation function, also known as the sigmoid function, transforms input values to a range between 0 and 1. It is particularly useful in binary classification tasks, providing a smooth transition and representing probabilities. ReLU is a popular choice due to its simplicity and effectiveness. It outputs the input for positive values and zero for negative values, introducing non-linearity. ReLU is known for mitigating the vanishing gradient problem, which can enhance the learning capacity of deep neural networks.

To extend the versatility of our model, we employed the *MultiOutputClassifier* class from Scikit-Learn [7]. This class allows us to encapsulate our classifier model, effectively creating a separate model for each Oscar category. This adaptation enables our neural network to make distinct predictions for nominations across diverse categories.

Additionally, to streamline and automate the preprocessing steps, we constructed a pipeline that includes feature scaling. Feature scaling is a critical preprocessing step to ensure that all input features contribute uniformly to the model's learning process. Our pipeline employs the *StandardScaler* from Scikit-Learn, which removes the mean and scales features to unit variance for continuous values. For binary features, a transformation is applied to map the [0, 1] range to [-1, 1], ensuring that binary features also have a mean of 0.

4.4.2 Training and Validation

The dataset was divided into training and validation sets based on the temporal split outlined earlier. Cross-validation was not employed in this scenario. This decision was motivated by the desire to keep the validation set most recent, reflecting the evolving trends in Oscar nominations over time. Given the highly imbalanced nature of the dataset, where non-nominations significantly outnumber nominations, the model was evaluated using balanced accuracy instead of traditional accuracy. This metric considers the true positive rate for each class, providing a more robust evaluation in the presence of imbalances.

Hyperparameters, as explored in the previous section, were fine-tuned to achieve optimal performance. The model with the highest balanced accuracy was selected for each Oscar category. As there is a separate model for each category, the mean of the balanced accuracy across all categories was used as the overall evaluation metric. While finding the best model for each category would be ideal, for simplicity and practicality within the defined pipeline, the mean balanced accuracy served as a consolidated measure. Table 10 shows the hyperparameters and corresponding metrics on the validation set.

Models employing ReLU activation generally outperform those with logistic activation, emphasizing the importance of non-linearity in capturing complex relationships. The regularization strength (α) impacts the model's ability to generalize, with a careful balance needed to prevent overfitting. Surprisingly, more complex architectures (additional hidden layers) do not consistently yield superior performance. Considering these insights, the final model selection involves a trade-off between complexity and performance metrics. In this context, the model with the highest balanced accuracy on the validation set is chosen, providing a robust evaluation considering the dataset's class imbalances.

4.4.3 Feature Importance

In the feature importance analysis, the weights of the first layer in the neural network model were examined by summing up their absolute values. The feature with the highest weight in each Oscar category was identified as the most important. However, the interpretation of these weights is limited as it does not provide information about the direction (positive or negative) of the impact on predictions, and the distributed nature of weights in a neural network may complicate the analysis. Alternative methods for feature importance analysis were considered but not implemented.

The summarized list of the most important features for each Oscar category is presented in Table 11.

The observed most important features align with expectations to some extent. For instance, IMDb Votes and Run Time are commonly associated with a movie's success or recognition. However, the influence of certain categorical features, such as specific nominations or genres, may vary across Oscar categories. Further analysis and domain knowledge could enhance the interpretability of these findings.

4.4.4 Performance Evaluation

The performance of the neural network model was assessed using various metrics against the test set. The metrics of the best model, determined by the mean balanced accuracy across all Oscar categories, is presented in the Results section.

4.5. Random Forest

The Random Forest model, known for its robustness and effectiveness in handling diverse datasets, was selected as a key classifier for our study. Random Forest operates by constructing multiple decision trees during the training phase and outputting the mode of the classes (classification) of the individual trees. This ensemble approach enhances the model's accuracy and reduces the risk of overfitting.

4.5.1 Model Configuration

Our Random Forest model was configured with a specific number of decision trees, determined empirically via cross-validation. Key hyperparameters, including the number of trees (n_estimators), the maximum depth of trees (max_depth), and the minimum number of samples required to split an internal node (min_samples_split), were tuned to optimize performance. The model was implemented using the Scikit-learn library, which provides a comprehensive and efficient implementation of the Random Forest algorithm.

4.5.2 Training and Validation

The model was trained on a preprocessed subset of the dataset, employing a stratified K-fold cross-validation strategy to ensure comprehensive learning and generalization. This approach allowed the assessment of the model's performance across different data segments, mitigating the risk of biases in the training data. The hyperparameters were fine-tuned based on the cross-validation results, with a primary focus on maximizing the F1-score, a critical metric for imbalanced datasets.

4.5.3 Feature Importance Analysis

Post-training, we conducted a feature importance analysis using the Random Forest model's intrinsic capability to rank features based on their contribution to the prediction process. This analysis provided valuable insights into the factors most influential in predicting the target variable, aiding in the interpretation of the model's decision-making process.

4.5.4 Performance Evaluation

The final evaluation of the Random Forest model was conducted on an unseen test dataset. Performance metrics, including accuracy, precision, recall, F1-score, were computed to assess the efficacy of the model comprehensively. The results were compared with other models in the study, providing a holistic view of the model's capabilities in the context of our research objectives.

5. Results and Discussion

5.1. Logistic Regression Model

The logistic regression model exhibits diverse performances across Oscar nomination categories (Table 12), re-

Hidden Layers	Activation Function	α	Accuracy	Balanced Accuracy	Precision	Recall	F1-score
(33, 33, 33)	logistic	0.001	0.918	0.498	0.069	0.075	0.071
(33, 33, 33)	logistic	0.010	0.918	0.498	0.069	0.075	0.071
(33, 33, 33)	logistic	0.100	0.918	0.498	0.069	0.075	0.071
(33, 33, 33)	reLU	0.001	0.920	0.692	0.477	0.463	0.460
(33, 33, 33)	reLU	0.010	0.920	0.697	0.492	0.473	0.473
(33, 33, 33)	reLU	0.100	0.920	0.681	0.454	0.440	0.438
(50, 50)	logistic	0.001	0.920	0.629	0.362	0.337	0.343
(50, 50)	logistic	0.010	0.920	0.632	0.377	0.343	0.353
(50, 50)	logistic	0.100	0.920	0.640	0.392	0.358	0.368
(50, 50)	reLU	0.001	0.920	0.693	0.477	0.465	0.462
(50, 50)	reLU	0.010	0.920	0.689	0.469	0.456	0.455
(50, 50)	reLU	0.100	0.920	0.691	0.469	0.460	0.455
(100,)	logistic	0.001	0.920	0.694	0.477	0.465	0.461
(100,)	logistic	0.010	0.920	0.701	0.492	0.479	0.476
(100,)	logistic	0.100	0.921	0.715	0.515	0.507	0.502
(100,)	reLU	0.001	0.920	0.701	0.492	0.481	0.477
(100,)	reLU	0.010	0.920	0.697	0.485	0.472	0.469
(100,)	reLU	0.100	0.920	0.696	0.485	0.470	0.469

Table 10. Neural network – hyperparameter exploration results

Category	Most Important Feature
Cinematography	numVotes
Directing	runtimeMinutes
Film Editing	runtimeMinutes
Actor in a Supporting Role	runtimeMinutes
Actress in a Supporting Role	gg_nominee_Best Performance by an Supporting Actress in a Motion Picture - Drama
Costume Design	runtimeMinutes
Sound	runtimeMinutes
Best Picture	gg_nominee_Best Screenplay - Motion Picture
Actor in a Leading Role	gg_nominee_Best Performance by an Actor in a Motion Picture - Drama
Actress in a Leading Role	gg_nominee_Best Performance by an Actress in a Motion Picture - Drama
Visual Effects	Average Rating
Animated Feature Film	Genre Animation
Production Design	Genre Adventure

Table 11. Neural network – most important features by Oscar category

vealing its varying predictive capabilities. Notably, the model excels in categories such as *Directing*, *Best Picture*, *Actor in a Leading Role*, and *Actor in a Supporting Role*, where it demonstrates robust precision, recall, and F1 scores, indicating reliable predictions. These successes underscore the model's proficiency in identifying movies with high chances of receiving nominations in these key categories.

However, challenges arise in categories like *Costume Design*, *Sound*, and *Production Design*, where the model faces difficulties, particularly in precision, pointing to a higher rate of false positives. Noteworthy outliers include the *Animated Feature Film* category, where the model achieves reasonable precision, recall, and F1 scores, indicating its ability to reliably predict nominations for animated films.

Additionally, the *Actress in a Leading Role* category presents a challenge, with the model exhibiting room for improvement, especially in precision. The *Visual Effects* category demonstrates modest performance, suggesting potential areas for enhancement.

5.2. Neural Network

The selected neural network model demonstrated promising performance in predicting Oscar nominations across diverse categories. The model achieved a balanced accuracy of 73%. Precision, recall, and F1-score provide detailed insights into the each model's predictive capabilities. The metrics are reported in Table 13 by Oscar category.

In terms of predicting Oscar nominations, these metrics suggest that the model demonstrates a relatively good balance between correctly predicting nominations and avoiding false positives. The ROC AUC score of 71.46% further supports the model's discriminative ability.

Overall, the selected neural network model exhibits promising performance, considering the inherent challenges

Category	Precision	Recall	F1
Cinematography	0.40	0.48	0.43
Directing	0.64	0.70	0.67
Film Editing	0.60	0.65	0.63
Actor in a Supporting Role	0.64	0.80	0.71
Actress in a Supporting Role	0.60	0.68	0.64
Costume Design	0.28	0.33	0.30
Sound	0.04	0.50	0.07
Best Picture	0.74	0.70	0.72
Actor in a Leading Role	0.72	0.78	0.75
Actress in a Leading Role	0.48	0.75	0.59
Visual Effects	0.32	0.38	0.35
Animated Feature Film	0.52	0.72	0.60
Production Design	0.28	0.32	0.30

Table 12. Classification Report for Logistic Regression

Label	Precision	Recall	F1
Cinematography	0.20	0.24	0.22
Directing	0.48	0.52	0.50
Film Editing	0.56	0.54	0.78
Actor in a Supporting Role	0.56	0.70	0.62
Actress in a Supporting Role	0.48	0.55	0.51
Costume Design	0.20	0.24	0.2
Sound	0.00	0.00	0.00
Best Picture	0.76	0.51	0.61
Actor in a Leading Role	0.68	0.74	0.71
Actress in a Leading Role	0.52	0.81	0.63
Visual Effects	0.28	0.33	0.30
Animated Feature Film	0.48	0.67	0.56
Production Design	0.12	0.14	0.13

Table 13. Classification Report for Neural Network

posed by imbalanced datasets and the complexity of the prediction task.

The evaluation of the neural network model reveals interesting insights into its predictive performance across various Oscar categories. In the *Sound* category, all metrics—recall, precision, and F1-score—are reported as 0. This indicates that the model fails to predict any true positives for this category, possibly due to the inherent challenges or lack of discernible patterns in the provided features for this specific Oscar category.

For categories such as *Cinematography*, *Costume Design*, *Production Design*, and *Visual Effects*, relatively low metrics are observed. These lower metrics can be attributed to the limited amount of data available for these categories, making it challenging for the model to generalize effectively. In cases where there is insufficient data, the model may struggle to identify patterns and make accurate predictions. On the other hand, categories like *Best Picture*, *Actor in a Leading Role*, and *Actress in a Leading Role* exhibit higher metrics. These categories, associated with the original awards, tend to have more available data. Notably, the *Best Picture* category benefits from a higher number of nominations (more than 5) in recent years, contributing to a slight reduction in class imbalance and improving the model's predictive performance.

It's essential to acknowledge that the imbalanced nature of the dataset poses challenges for predicting categories with fewer examples, and the model may benefit from additional data to enhance its performance in these specific areas.

In summary, the neural network model demonstrates varying levels of success across different Oscar categories, with performance influenced by both the quantity and distribution of available data.

When considering feature importance, caution should be exercised when interpreting the results based on the weights between the input layer and the first hidden layer. The neural network's architecture distributes weights throughout the network, and the contribution of individual features may not be accurately represented by the weights from the input layer alone. Unfortunately, we are not confident to report the most important features from the neural network model without further analysis.

Label	Precision	Recall	F1
Cinematography	0.38	0.24	0.29
Directing	0.55	0.70	0.62
Film Editing	0.57	0.52	0.55
Actor in a Supporting Role	0.68	0.75	0.71
Actress in a Supporting Role	0.74	0.64	0.68
Costume Design	0.00	0.00	0.00
Sound	0.00	0.00	0.00
Best Picture	0.82	0.62	0.71
Actor in a Leading Role	0.65	0.48	0.55
Actress in a Leading Role	0.82	0.56	0.67
Visual Effects	0.32	0.33	0.33
Animated Feature Film	0.65	0.72	0.68
Production Design	0.00	0.00	0.00

Table 14. Classification Report for Random Forest

5.3. Random Forest

The Random Forest model's performance on the test set is summarized in the following classification report:

5.3.1 Best Picture

For the *Best Picture* category, the feature importances indicate a strong influence of industry recognition and commercial success. Key influential features are:

Notably, Golden Globe nominations for Best Director and Best Screenplay play significant roles, reflecting the im-

Feature	Importance
gg nominee - Best Director - Motion Picture	0.3266
revenue	0.0982
gg nominee - Best Screenplay - Motion Picture	0.0854
...	...

Table 15. Feature importances for the Best Picture category

portance of directorial and narrative excellence.

5.3.2 Actor in a Leading Role

The nomination for *Actor in a Leading Role* is significantly influenced by industry accolades and personal attributes of the actors:

Feature	Importance
sag nominee LEADING ROLE - MALE	0.3338
gg nominee Best Performance by an Actor	0.1313
actor1Age	0.0448
...	...

Table 16. Feature importances for the Actor in a Leading Role category

SAG and Golden Globe nominations are predominant factors, alongside the lead actor's age.

5.3.3 Actress in a Leading Role

Similar to actors, the *Actress in a Leading Role* category is influenced by both recognition in awards and the actresses' attributes:

Feature	Importance
sag nominee LEADING ROLE - FEMALE	0.3604
gg nominee Best Performance by an Actress	0.0723
actress3Age	0.0515
...	...

Table 17. Feature importances for the Actress in a Leading Role category

The influence of SAG and Golden Globe nominations is evident, along with the age factor, which might correlate with the types of roles and experience.

5.4. Discussion

The analysis of three distinct models—logistic regression, neural network, and random forest—provides valuable insights into their predictive capabilities for Oscar nominations. Each model exhibits diverse performances across various categories, shedding light on their strengths and challenges.

5.4.1 Model-Specific Performance

Logistic Regression: The logistic regression model excels in predicting nominations for categories like *Directing*, *Best Picture*, and *Actor in a Leading Role*, demonstrating robust precision, recall, and F1 scores. However, challenges are evident in categories such as *Costume Design* and *Sound*, suggesting areas for improvement.

Neural Network: The selected neural network model demonstrates promising overall performance with a balanced accuracy of 73%. It exhibits varying success across Oscar categories, with notable strengths in major awards like *Best Picture*. Challenges in categories with limited data highlight the importance of data availability for effective model generalization.

Random Forest: The Random Forest model showcases diverse performances, with high precision in categories like *Best Picture*, emphasizing the influence of industry recognition. Feature importance analysis reveals the significance of Golden Globe nominations and actor/actress ages in predicting nominations. Challenges in certain categories, such as *Costume Design* and *Animated Feature Film*, suggest the need for addressing data limitations.

5.4.2 Comparative Analysis

Comparing the performances of the logistic regression, neural network, and random forest models across common Oscar categories reveals distinct patterns. The logistic regression model excels in key categories like *Directing*, *Best Picture*, and *Actor in a Leading Role*, demonstrating robust predictive capabilities in these critical areas. In contrast, the neural network exhibits promising overall performance, particularly excelling in major awards, indicating its potential for capturing broader industry trends. The random forest model's diverse performances underscore the influence of industry recognition, as seen in its strength in predicting nominations for the prestigious *Best Picture* category. Disparities in predictions highlight areas for improvement and model refinement to achieve more consistent and accurate results.

5.4.3 Temporal Analysis

The varying success amongst some categories may demonstrate the dynamic nature of predictors over the years. While the challenges in predicting nominations for *Costume Design*, *Sound*, and *Production Design* categories could be influenced by imbalanced data, they may also indicate shifts in cinematic trends or Academy preferences. On the other hand, the relative consistency in predicting nominations for

Animated Feature Film suggests stability in certain genres over time.

5.4.4 Social and Economic Analysis

The study's focus on features related to social and economic impact provides a nuanced understanding of the broader implications of Oscar nominations, considering the influence of award prestigiousness, actor and actress age, country, language, and production company size.

Golden Globe versus SAG Awards: The logistic regression model, excelling in categories like *Best Picture*, suggests that industry recognition and commercial success significantly influence the nominations process. The prominence of features like Golden Globe nominations (over the SAG Awards) in the *Best Picture* category, as indicated by all three models, aligns with the broader industry's acknowledgment of narrative and directorial excellence.

Actor and Actress Age: The random forest models emphasize the importance of age in categories such as *Actor in a Leading Role* and *Actress in a Leading Role*. This suggests that age plays a determining role in Oscar recognition, although the direction of this influence (older or younger) requires further analysis. Conversely, the logistic regression model and the neural network models do not significantly factor in the age of actors and actresses.

Language: The logistic regression model reveals insights into how language impacts Oscar nominations. Non-English films exhibit a strong negative correlation, indicating a bias towards English-language productions. This underscores the industry's inclination towards movies with broader accessibility, potentially driven by a global audience.

Production Company Size: The logistic model highlights the impact of production company size on Oscar nominations. Smaller production companies outside the top 10 face challenges, emphasizing the industry's preference for larger, potentially more established entities. This insight reflects the industry's recognition of the resources and influence larger production companies bring to film-making.

Country of Origin: The logistic model unveils insights into how the country of origin impacts Oscar nominations. Films from countries outside the top 5 face challenges, emphasizing the industry's preference for movies from major producing nations. This bias highlights the influence of established film-making industries on the global stage.

Superhero Films: Our analysis brings forth an intriguing observation regarding superhero films, a genre often predicted to secure Oscar nominations in all three models due to its massive popularity and cultural impact. Despite these expectations reveal a nuanced reality. Features related to critical acclaim, specific awards categories, and genre dynamics play pivotal roles in the Oscar nomination process. However, this demonstrates potential biases within the Oscar nomination selection progress. Industry standards and preferences might influence the recognition of genres like superhero films, which are often associated with entertainment rather than traditional notion of "prestige".

Horror Films: The negative weights associated with horror films in certain Oscar categories, as indicated by the Logistic Regression model, also experiences a strong bias from critics. While horror films are often popular and has a dedicated fan base, this genre also may not align with traditional perceptions of award-worthy cinema. The cultural and artistic biases embedded int he film industry's recognition patterns become apparent through our machine learning model.

6. Conclusion

In this study, we conducted a comprehensive exploration of machine learning models, including Logistic Regression, Neural Networks, and Random Forest, to predict Oscar nominations across various categories. Leveraging a rich dataset that encapsulates diverse features spanning decades of cinematic history, our models were designed to uncover patterns, relationships, and influential factors contributing to the recognition of films by prestigious award committees.

The Logistic Regression model, with its ability to discern nuanced relationships, provided valuable insights into feature importance. We identified critical factors, such as nominations in specific Golden Globe categories and the impact of genres, shedding light on the dynamics influencing Best Picture nominations. The model's interpretability allowed us to uncover patterns that extend beyond simple correlations, offering a deeper understanding of the intricate dynamics at play in the film industry.

Neural Networks, known for their capacity to capture complex relationships, presented a compelling alternative. Through hyperparameter tuning, we optimized model performance. Feature importance analysis revealed the dominance of certain features, including the number of votes and runtime, emphasizing their role in predicting nominations across diverse Oscar categories. However, the actual significance of these features are ambiguous as only weights between the input layer and the first hidden layers were investigated.

The Random Forest model, with its ensemble of decision trees, showcased robustness and efficiency. Through cross-validation and feature importance analysis, we gained valuable insights into the key factors influencing the model's predictions. The Random Forest model's ability to handle diverse datasets and capture non-linear relationships further validated its utility in predicting Oscar nominations.

Comparing the performance of these models, we observed nuanced differences in their strengths and weaknesses. Logistic Regression excelled in interpretability, providing actionable insights into specific features' impacts. Neural Networks, with their capacity to capture complex relationships, offered a powerful tool for prediction. Random Forest, with its ensemble approach, showcased robustness across diverse datasets.

Our analysis of social and economic factors, including actor and actress age, language, production company size, and country of origin, enhances our understanding of the nuanced dynamics influencing Oscar nominations. The interplay of these factors reflects the complex nature of the film industry, where industry recognition, cultural context, and production dynamics collectively shape the trajectory of films within the prestigious realm of the Oscars.

This research opens avenues for further exploration, encouraging the integration of additional features, temporal dynamics, and external factors to enhance model performance. As the film industry embraces technological advancements, the combination of data science and cinema promises to yield deeper insights, enabling stakeholders to navigate the intricate landscape of award recognition with greater precision.

7. Code

Our codebase is publicly available on GitHub: https://github.com/victorzwu/oscarsprediction.

References

[1] S. Amidi. Teaching page of shervine amidi, graduate student at stanford university. `https://stanford.edu/~shervine/teaching`, 2023.

[2] R. Banik. The movies dataset. `https://www.kaggle.com/datasets/rounakbanik/the-movies-dataset`, 2023.

[3] S. Deregowski. Oscar-predictions, 2022.

[4] IMDb. Imdb data files available for download. `https://datasets.imdbws.com/`, 2023. Documentation can be found at `http://www.imdb.com/interfaces/`.

[5] J. Kim, S. Hwang, and E. Park. Can we predict the oscar winner? a machine learning approach with social network services. *Entertainment Computing*, 39:100441, 2021.

[6] C. Molnar. Interpretable machine learning. `https://christophm.github.io/interpretable-ml-book`, 2022.

[7] Scikit-learn. Machine learning in python — scikit-learn 0.24.2 documentation. `https://scikit-learn.org/stable`, 2023.

[8] UNIMAD. Golden globe awards. `https://www.kaggle.com/datasets/unanimad/golden-globe-awards`, 2023.

[9] UNIMAD. The oscar award, 1927 - 2023. `https://www.kaggle.com/datasets/unanimad/the-oscar-award`, 2023.

[10] UNIMAD. Screen actors guild awards, 1994 - 2020. `https://www.kaggle.com/datasets/unanimad/screen-actors-guild-awards`, 2023.

[11] user12075. Mlpregressor for lbfgs solver in sklearn. `https://stackoverflow.com/a/60902315/12075`, 2023.