WWE Superstar Popularity Tier Prediction using Multi-class Classification

Jay Arre Talosig
College of Computing and Information Technologies
National University - Philippines
Manila, Philippines
talosigjp@students.national-u.edu.ph

Moira Mercado
College of Computing and Information Technologies
National University - Philippines
Manila, Philippines
mercadomn@students.national-u.edu.ph

Queen Maegan Pedido
College of Computing and Information Technologies
National University - Philippines
Manila, Philippines
pedidoqr@students.national-u.edu.ph

James Adrian Castro
College of Computing and Information Technologies
National University - Philippines
Manila, Philippines
castrojb@students.national-u.edu.ph

***Abstract*—The world of World Wrestling Entertainment (WWE) features a diverse roster of superstars, categorized by their prominence and fan engagement into tiers such as Main Eventer, Midcard, and Enhancement Talent. Predicting these tiers is crucial for talent management, storyline development, and marketing strategies. This research presents a comprehensive machine learning system to classify WWE superstars into popularity tiers based on performance metrics and career statistics. We implemented and evaluated several models, including Random Forest, Support Vector Machine (SVM), and Gradient Boosting. The tuned Random Forest classifier demonstrated superior performance, achieving a test accuracy of 70.27%, outperforming other models. Key predictive features identified were the number of championship reigns and matches. This study establishes a data-driven framework for understanding and predicting superstar positioning within the WWE ecosystem, while highlighting the challenges of working with limited and imbalanced sports entertainment data.**

***Index Terms*—WWE, popularity prediction, multi-class classification, machine learning, Random Forest, sports analytics**

I. INTRODUCTION

World Wrestling Entertainment (WWE) operates not just as a sports organization but as a global entertainment powerhouse. Its roster of superstars is strategically positioned in different tiers of popularity, which directly influences their screen time, storyline significance, and commercial value. These tiers are informally classified by fans and analysts as **Main Eventer** (top-tier stars who headline major events), **Midcard** (reliable performers who are consistently featured but not in the main event), and **Enhancement Talent** (superstars used primarily to make others look strong).

Currently, the classification of a superstar into one of these tiers is subjective, based on fan perception and internal company booking decisions. A data-driven approach to predict these tiers can provide valuable insights for various stakeholders: WWE can optimize talent investment, media analysts can quantify a star's draw, and fans can gain a deeper understanding of the business.

This study aims to automate this classification process by leveraging machine learning. The primary objectives are:

- To collect and preprocess a dataset of WWE superstars with relevant features such as match counts, championship reigns, and win-loss records.
- To train and evaluate multiple machine learning models for the multi-class classification task of predicting popularity tiers.
- To identify the most significant features that contribute to a superstar's tier classification.

II. REVIEW OF RELATED WORKS

The application of data analytics and machine learning in sports, often referred to as "sabermetrics" in baseball or more broadly as sports analytics, has seen widespread adoption for performance prediction and player valuation [1]. However, its application in sports entertainment, particularly professional wrestling, is a nascent field.

In traditional sports, studies have successfully used player statistics to predict outcomes like match results [2] and player performance tiers. Machine learning models such as Random Forests and SVMs are frequently employed for these classification tasks due to their ability to handle non-linear relationships [3].

In the context of WWE, public data analysis is often conducted by fan communities and journalists, but formal academic research is limited. Studies that exist often focus on the socio-cultural aspects of wrestling fandom rather than quantitative performance analysis. This research gap presents an opportunity to apply established sports analytics methodologies to the unique, storyline-driven world of WWE. By treating superstar metrics as features in a classification model, this work bridges the gap between traditional sports analytics and entertainment analytics.

In their 2023 paper, Gonzalez and Torres [4] develop *FightTracker*, a framework for generating real-time predictions during Mixed Martial Arts bouts. This system utilizes live data to model the progression of a fight and predict its eventual winner. The objective is to offer a data-driven tool that could support strategic decisions for fighters and provide deeper insights for broadcast analysts, moving beyond purely qualitative assessments.

Kim and Park's [5] 2025 research introduces an adaptive state representation for multi-agent reinforcement learning used in soccer simulations. Their model enables autonomous agents to dynamically expand their internal state based on gathered experience, which helps overcome the challenges of complex, partially observable environments. This expandable state mechanism is shown to facilitate better collaboration and decision-making among agents, leading to more robust and intelligent team behavior.

Focusing on player evaluation, Ahmed and Chen [6] propose a context-aware metric termed "Goals Above Expectation." This measure assesses a player's performance by comparing their actual scoring output with the number of goals they would be expected to score given the quality and circumstances of their opportunities. By accounting for the difficulty of chances, this metric provides a more refined and equitable evaluation of a player's offensive contribution than traditional counting statistics

The reviewed literature underscores the versatility of predictive modeling in both competitive sports and entertainment. While the methodologies from traditional sports provide a foundational toolkit [3], recent advances in adjacent fields demonstrate the value of sophisticated, context-aware metrics [6] and adaptive modeling techniques for complex, performance-based domains [4], [5]. This study leverages these established principles but applies them to the unique, narrative-driven context of WWE. By treating a superstar's career statistics as predictive features, this work aims to create a classification system that quantifies the otherwise subjective hierarchy of sports entertainment, directly addressing the identified research gap.

III. METHODOLOGY

The implementation of this project followed a standard machine learning pipeline, from data acquisition to model deployment.

A. Data Collection

The dataset was manually curated from the online wrestling database, **Cagematch.net**, a comprehensive source for professional wrestling match results and performer statistics. The initial dataset contained information for 185 WWE superstars after preprocessing. The features collected for each superstar included:

- **Matches:** Total number of matches in the superstar's career.
- **Wins / Losses / Draws:** Count of match outcomes.
- **Championships:** Total number of championship reigns.
- **Years Active:** The number of years the superstar has been performing.

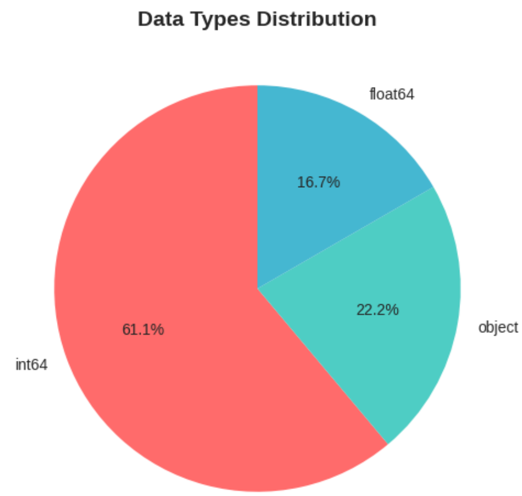

Fig. 1. Class Distribution of WWE Superstar Tiers

The target variable, the **Popularity Tier**, was labeled based on a combination of fan consensus and observed booking patterns into three classes: **Enhancement Talent (0)**, **Midcard (1)**, and **Main Eventer (2)**.

B. Data Preprocessing

To ensure the quality and suitability of the data for model training, a series of preprocessing steps were undertaken.

Initial Sanity Check and Cleaning The dataset was inspected for null and duplicated values. Any null or duplicated entries were identified and subsequently removed to ensure data integrity.

Feature Engineering New features were created to enhance the model's predictive power:

- **Win Percentage:** Calculated as Wins / Matches.
- **Loss Percentage:** Calculated as Losses / Matches.
- **Draw Percentage:** Calculated as Draws / Matches.

Data Splitting and Feature Scaling The cleaned dataset was split into training and testing sets using an 80-20 ratio, ensuring stratified sampling. The features were standardized using `StandardScaler` for models sensitive to feature scales.

C. Experimental Setup

The entire project was developed in a Jupyter Notebook environment using Python. Key libraries included `pandas` for data manipulation, `scikit-learn` for machine learning models and preprocessing, and `matplotlib` and `seaborn` for visualization.

D. Training Procedure

The baseline models were first trained and evaluated using default parameters. Subsequently, hyperparameter tuning was performed for the most promising models using **GridSearchCV** with 5-fold cross-validation to find optimal hyperparameters and ensure model generalizability.

E. Evaluation Metrics

The models were evaluated using comprehensive metrics:

- **Accuracy:** The proportion of total correct predictions.
- **Precision:** The ability to avoid false positives.
- **Recall:** The ability to find all positive samples.
- **F1-Score:** The harmonic mean of precision and recall.

IV. ALGORITHM

Six classification algorithms were selected for this study:

A. Random Forest

An ensemble learning method that constructs multiple decision trees and outputs the mode of their classes. It is robust against overfitting and effective for high-dimensional data.

B. Support Vector Machine (SVM)

A discriminative classifier defined by a separating hyperplane. The kernel trick was used with Radial Basis Function (RBF) kernel to handle non-linear decision boundaries.

C. Gradient Boosting

An ensemble technique that builds trees sequentially, with each tree correcting the errors of the previous one.

D. K-Nearest Neighbors (KNN)

An instance-based learning algorithm that classifies data points based on neighbor classifications.

E. Logistic Regression

A linear model for classification that uses a logistic function to model class probabilities.

F. Decision Tree

A non-parametric supervised learning method that learns simple decision rules from data features.

V. RESULTS AND DISCUSSION

A. Exploratory Data Analysis

Initial EDA revealed class imbalance with Midcard tier being most populous (103 superstars), followed by Main Eventer (43) and Enhancement Talent (39). This imbalance presented classification challenges.

The correlation heatmap (Fig. 2) reveals several meaningful relationships among numerical features. Matches are highly correlated with both wins and losses, while championships show moderate correlation with matches and wins. Win percentage displays weaker overall correlation, suggesting that popularity is not solely determined by match outcomes but also by experience and visibility within WWE.

The comparison of performance features across WWE popularity tiers in (Fig. 3) compares performance features like match count and championship wins across the popularity tiers. Main Eventers generally have higher values in these metrics, while Enhancement Talents score the lowest. The Midcard group shows more variation, which makes sense since it includes both rising stars and established wrestlers who haven't reached the top tier yet.

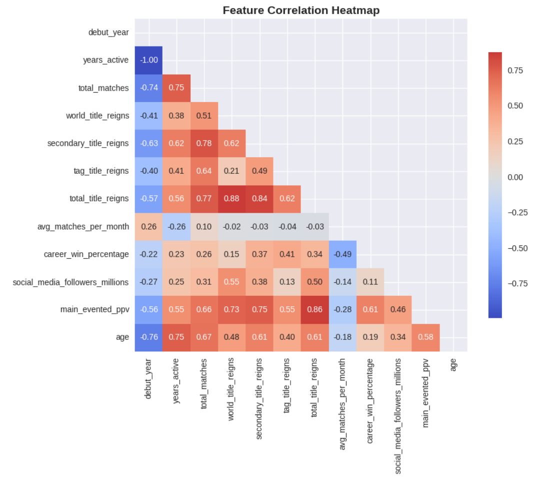

Fig. 2. Correlation heatmap showing relationships among WWE superstar statistics

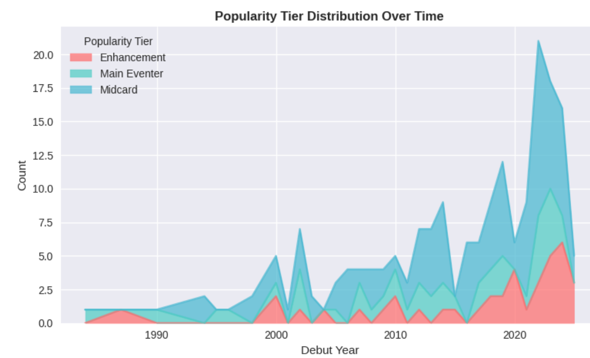

Fig. 3. Comparison of performance features across WWE popularity tiers.

B. Baseline Model Performance

As shown in Table I, Random Forest and K-Nearest Neighbors achieved the highest baseline accuracy of 67.57%. Logistic Regression and Decision Tree performed poorest.

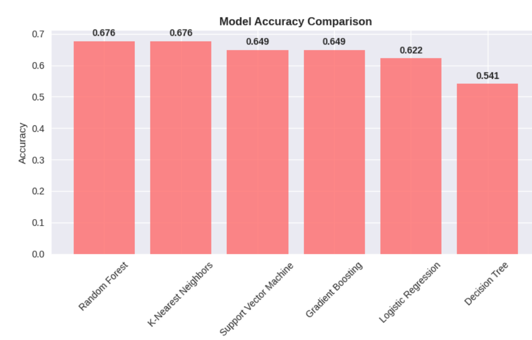

Fig. 4. Model accuracy comparison across classifiers for WWE superstar tier prediction.

C. Hyperparameter Tuning

To optimize model performance, hyperparameter tuning was conducted using GridSearchCV with 5-fold cross-validation.

TABLE I
PERFORMANCE OF BASELINE MODELS ON TEST SET

max width=

Model	Accuracy	Precision	Recall	F1-Score	CV-Mean	CV-Std
Random Forest	0.6757	0.6543	0.6757	0.6535	0.6349	0.0276
K-Nearest Neighbors	0.6757	0.6847	0.6757	0.6319	0.5538	0.0463
Gradient Boosting	0.6486	0.6327	0.6486	0.6290	0.6349	0.0270
Support Vector Machine	0.6486	0.5277	0.6486	0.5614	0.6349	0.0504
Logistic Regression	0.6216	0.6181	0.6216	0.5661	0.6480	0.0888
Decision Tree	0.5405	0.5718	0.5405	0.5477	0.5943	0.1001

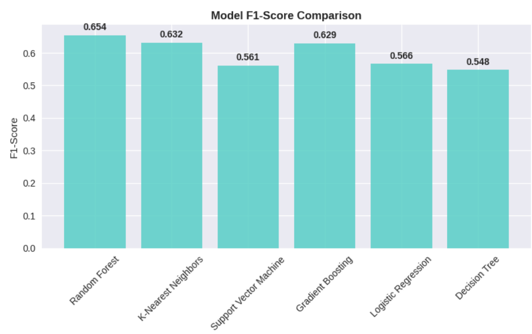

Fig. 5. Model F1-score comparison across classifiers for WWE superstar tier prediction.

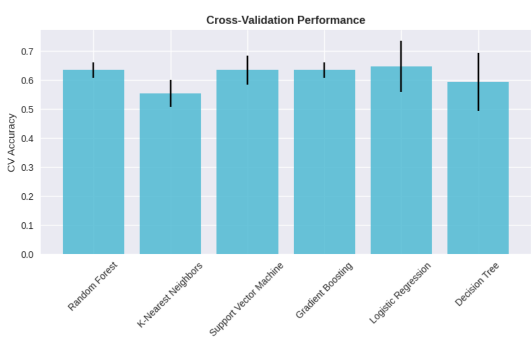

Fig. 6. Cross-validation accuracy comparison for WWE superstar tier prediction models.

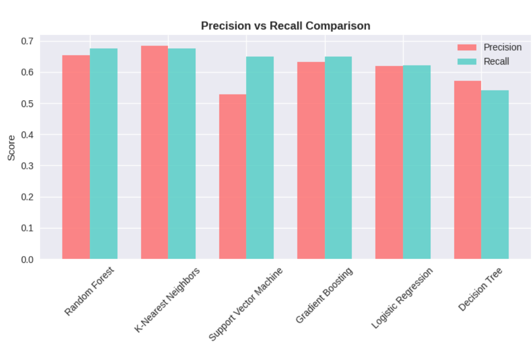

Fig. 7. Precision vs. recall comparison across models for WWE superstar tier prediction.

The best parameters and their corresponding cross-validation scores for each model are summarized in Table II.

The hyperparameter tuning process revealed that:

- **Random Forest** achieved the highest cross-validation score (0.6759) with deeper trees (max_depth=10) and more estimators (n_estimators=200)
- **Support Vector Machine** performed best with a linear kernel and default gamma scaling, suggesting simpler decision boundaries
- **Gradient Boosting** required a lower learning rate (0.05) with moderate tree depth (max_depth=5) for optimal performance

D. Hyperparameter Tuning Results

Hyperparameter tuning was conducted for top-performing models. The tuned Random Forest model (RF_Tuned) achieved the highest performance with 70.27% accuracy, representing a 2.7 percentage point improvement over baseline.

E. Feature Importance Analysis

Analysis revealed **Championships** as the most significant predictor, followed by **Matches** and **Win Percentage**, aligning with wrestling logic where top-tier stars receive more opportunities.

Fig. 8. Feature Importance from Top Feature Correlation

F. Statistical Analysis of Features

A comprehensive statistical analysis was conducted on the numerical features to understand their distribution and characteristics. The summary statistics in Table IV reveal important patterns in the dataset.

Key observations from the statistical analysis include:

TABLE II
HYPERPARAMETER TUNING RESULTS

Model	Best Hyperparameters	CV Score
Support Vector Machine	`{’C’: 1, ’gamma’: ’scale’, ’kernel’: ’linear’}`	0.6621
Random Forest	`{’max_depth’: 10, ’min_samples_leaf’: 4, ’min_samples_split’: 2, ’n_estimators’: 200}`	0.6759
Gradient Boosting	`{’learning_rate’: 0.05, ’max_depth’: 5, ’min_samples_split’: 2, ’n_estimators’: 100}`	0.6623

TABLE III
PERFORMANCE OF TUNED MODELS ON TEST SET

Model	Accuracy	Precision	Recall	F1-Score
RF_Tuned	**0.7027**	**0.6980**	**0.7027**	**0.6721**
SVM_Tuned	0.6757	0.7554	0.6757	0.6177
GB_Tuned	0.6216	0.6032	0.6216	0.6067

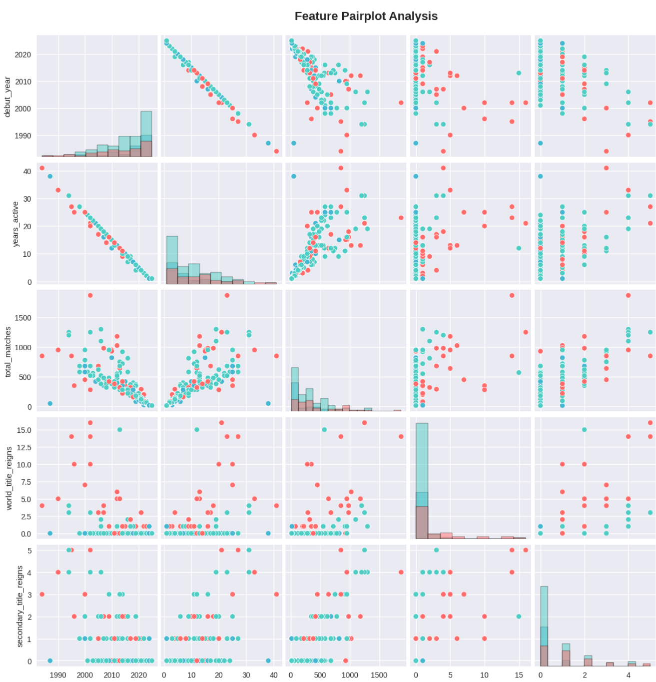

Fig. 9. Feature Pairplot Analysis

- **Title Distribution:** Most superstars have no championship reigns (median = 0 for world, secondary, and tag titles), indicating high skewness in title distributions
- **Activity Level:** The average superstar has 333 matches over 9.7 years, with high variance in match counts (std = 332.44)
- **Social Media Presence:** High skewness (6.596) and kurtosis (51.647) in social media followers suggest a few superstars dominate the follower counts
- **Win Percentage:** Extreme kurtosis (41.709) indicates most superstars cluster around the mean win percentage with few outliers
- **Main Event Experience:** High standard deviation (11.802) in PPV main events shows significant variation in top-level exposure

G. Discussion of Results

The results demonstrate feasibility while highlighting challenges. The 70.27% accuracy indicates that meaningful patterns were captured, though significant predictive challenges remain due to:

- **Class Imbalance:** Midcarders outnumbering other tiers affected minority class identification
- **Intangible Factors:** Missing qualitative elements like charisma, mic skills, and social media presence
- **Subjectivity:** Creative decisions in wrestling that defy pure statistics

The moderate performance range (54-70%) reflects the unique challenges of sports entertainment compared to traditional sports analytics. The dominance of championship and match counts confirms their importance, but also suggests the need for more sophisticated features in future work.

VI. CONCLUSION

This study successfully developed a machine learning system for classifying WWE superstars into popularity tiers. The tuned Random Forest model achieved the best performance with 70.27% accuracy, demonstrating that quantitative career statistics provide meaningful indicators of superstar positioning.

Limitations include the small size of the dataset (185 superstars), the imbalance of class, and the absence of important qualitative factors. The manual labeling process also introduced some subjectivity.

Future work will focus on

- Expanding the data set with more wrestlers from different eras
- Incorporating NLP techniques for social media sentiment analysis
- Collecting additional features such as pay-per-view main events and merchandise sales
- Applying advanced techniques for handling class imbalance
- Exploring deep learning models with more extensive data

This work establishes a foundation for data science in professional wrestling analytics and demonstrates both the potential and limitations of quantitative approaches in sports entertainment.

TABLE IV
STATISTICAL SUMMARY OF NUMERICAL FEATURES

max_width=

Feature	**count**	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**	**variance**	**skewness**	**kurtosis**
debut_year	185.0	2015.092	8.644	1984.0	2010.00	2018.00	2022.0	2025.0	74.725	-1.094	0.731
years_active	185.0	9.746	8.424	1.0	3.00	7.00	15.0	41.0	70.962	1.121	0.855
total_matches	185.0	333.411	332.443	8.0	80.00	220.00	450.0	1870.0	110518.254	1.544	2.598
world_title_reigns	185.0	0.849	2.577	0.0	0.00	0.00	0.0	16.0	6.640	4.186	18.622
secondary_title_reigns	185.0	0.681	1.143	0.0	0.00	0.00	1.0	5.0	1.305	1.978	3.605
tag_title_reigns	185.0	0.530	1.430	0.0	0.00	0.00	0.0	8.0	2.044	4.040	17.437
total_title_reigns	185.0	2.059	4.096	0.0	0.00	0.00	2.0	23.0	16.774	2.943	9.255
avg_matches_per_month	185.0	3.067	0.662	0.3	2.90	3.10	3.4	4.6	0.438	-1.602	4.881
career_win_percentage	185.0	0.625	0.505	0.3	0.49	0.54	0.6	4.3	0.255	6.449	41.709
social_media_followers_millions	185.0	1.203	2.670	0.1	0.30	0.40	1.2	26.0	7.129	6.596	51.647
main_evented_ppv	185.0	4.665	11.802	0.0	0.00	0.00	3.0	85.0	139.278	4.205	21.153
age	185.0	34.173	7.760	21.0	28.00	33.00	39.0	60.0	60.220	0.880	0.704

REFERENCES

[1] M. E. Glickman and H. S. Stern, "Sports Analytics: A Guide for Coaches, Managers, and Other Decision Makers," Columbia University Press, 2014.

[2] T. Ueda and A. Murai, "A machine learning approach for predicting the outcome of English Premier League matches," PLOS ONE, vol. 15, no. 1, 2020.

[3] C. Sheehan and Y. Xu, "Predicting player performance in professional Dota 2," in Proceedings of the MIT Sloan Sports Analytics Conference, 2020.

[4] M. J. Gonzalez and L. Torres, "FightTracker: Real-time predictive analytics for Mixed Martial Arts bouts," *arXiv preprint arXiv:2312.11067*, 2023.

[5] Y. Kim and S. Park, "Expandable Decision-Making States for Multi-Agent Deep Reinforcement Learning in Soccer," *arXiv preprint arXiv:2510.00480*, 2025.

[6] R. Ahmed and P. Chen, "Rethinking player evaluation in sports: Goals above expectation and beyond," *arXiv preprint arXiv:2509.20083*, 2025.