2024 4th International Conference on Machine Learning and Intelligent Systems Engineering (MLISE)

Data-Driven MMA Outcome Prediction Enhanced by Fighter Styles: A Machine Learning Approach

Jiajie Yin*
Faculty of Arts and Sciences, Beijing Normal University, Zhuhai 519087, China
*Corresponding author e-mail: jiajieyin@mail.bnu.edu.cn

Abstract—**Mixed Martial Arts (MMA) analysis presents unique challenges for predicting fight outcomes and clustering fighter styles. Existing research often falls short of accurately capturing the complexity of fighters' technical styles and their impact on match results. To bridge these gaps, we propose a novel approach that utilizes machine learning methods. Specifically, 1) Use factor analysis to derive high-dimensional technical style factors and applies the K-means algorithm to cluster fighters based on these factors. 2) Various machine learning models, such as Random Forest, Support Vector Machine (SVM), Extreme Gradient Boosting (XGBoost), Logistic Regression, Neural Networks are experimentally tested. 3) An ensemble learning model is proposed that employs majority voting among the mentioned models. 4) Based on data from the Ultimate Fighting Championship (UFC), the experimental results demonstrate that the ensemble learning model achieves the highest accuracy of 65.52%, significantly outperforming individual models. An ablation study further validates the importance of these factors in prediction accuracy. These findings underscore the effectiveness of incorporating technical style factors into predictive models, enhancing accuracy and interpretability in the context of MMA.**

Keywords-component; MMA; machine learning; fighter style clustering; outcome prediction; factor analysis

I. INTRODUCTION

Mixed Martial Arts (MMA) is a full-contact combat sport that incorporates techniques from various martial arts disciplines, including boxing, wrestling, judo, Brazilian jiu-jitsu, and Muay Thai. It has grown exponentially in popularity since its inception in the early 1990s, evolving from a niche sport to a mainstream phenomenon with a global following [1]. The Ultimate Fighting Championship (UFC), as the leading MMA promotion, hosts events that attract millions of viewers worldwide, showcasing the sport's dynamic nature and the diverse skill sets of its athletes.

Understanding a fighter's style in MMA is crucial for several reasons. Each fighter's approach to combat, whether they are predominantly strikers, grapplers, or versatile all-rounders, significantly impacts their performance and strategy in the cage. Clustering fighter styles in MMA presents significant challenges due to the complex and multifaceted nature of the sport. The UFC's technical system is intricate, with fighters often possessing a comprehensive blend of striking, grappling, and submission techniques. This makes it challenging to distinguish between different fighting styles using traditional methods [2]. These conventional approaches, which rely on professional experience or simplistic clustering techniques, fall short in accurately capturing the nuanced differences between fighters' styles [3].

Predicting fight outcomes is another critical aspect that holds substantial value for stakeholders across the sport. Accurate predictions can benefit betting industries, enhance fan engagement, and provide strategic insights for fighters and coaches. Traditional methods of outcome prediction often rely on expert opinions and historical performance data [4]. However, these approaches can be subjective and limited in their ability to account for the complex, multifactorial nature of MMA fights [5].

To address these challenges, data-driven and machine learning methods offer a robust solution. By leveraging large datasets, these methods can uncover patterns and relationships that are not immediately apparent through conventional analysis. Machine learning algorithms can classify fighter styles based on numerous performance metrics and predict fight outcomes with greater accuracy [6]. These techniques not only enable a more objective, comprehensive analysis, but also provide strategic insights that can drive decisions and most importantly, improve competitive performance. This is a game-changer for fighters and coaches, offering a new level of hope and motivation.

In this paper, we present a comprehensive study utilizing data-driven and machine learning methods to analyze MMA fighter styles and predict fight outcomes. We demonstrate the effectiveness of these techniques through detailed data preprocessing, feature engineering, and model evaluation processes, highlighting their potential to transform understanding and strategic planning in the sport of MMA. The key contributions of this paper can be summarized as follows:

1) We employ factor analysis for feature extraction and dimensionality reduction, identifying underlying factors that influence fighter performance. These factors are also proved to be useful in improving outcome prediction accuracy.

2) We present a comprehensive data-driven analysis of MMA fighter styles using K-means clustering, clustering fighters into distinct style categories based on performance metrics.

3) We develop and compare multiple predictive models for fight outcomes using various machine learning and ensemble learning algorithms, including Random Forest, Logistic Regression, Support Vector Machine (SVM), Extreme Gradient Boosting (XGBoost), and Neural Networks, identifying the most effective approaches for prediction.

The rest of the paper is organized as follows. First, we present the dataset in Section II, detailing the data sources, preprocessing steps, and feature extraction methods. Then, Section III describes the machine learning methodologies employed. In Section IV, we provide our experiments' results, showcasing the proposed models' performance. Finally, Section V discusses the findings and concludes the paper.

II. DATASET

A. Data Sources and Preprocessing

UFC is the premier MMA league globally, renowned for its top-tier athletes, comprehensive rules, and structured competition system. It also maintains the most complete and detailed records of competition data and fighter information. Hence, for studying mixed martial arts, UFC data serves as the most appropriate and reliable source.

Our data is derived from the statistical information available on the official UFC website [7]. We utilized web crawlers [8] to compile data on events, fighters, and matches. The final dataset encompasses a total of 690 UFC events from March 11, 1994, to May 18, 2024, covering 7,515 fight records and 4,162 fighter profiles. The data dictionary for fighter, event and fight data is provided as Table 1.

Table 1 Data Dictionary

Class	Field Name	Description	Example
Fighter	fighter_id	Primary key for ufc_fighters, unique for each fighter	1
	fighter_name	Fighter name	Tom Aaron
	fighter_nickname	Fighter nickname	"The Notorious"
	fighter_height_cm	Fighter height in centimeters	175
	fighter_weight_lbs	Fighter weight in pounds	155
	fighter_reach_cm	Fighter reach in centimeters	188
	fighter_stance	Fighter stance	"Orthodox"
	fighter_dob	Fighter date of birth	1988/7/14
	fighter_w	Number of wins	22
	fighter_l	Number of losses	4
	fighter_d	Number of draws	0
	fighter_nc_dq	Number of no contests or disqualifications	1
	SLpM	Significant Strikes Landed per Minute	5.53
	Str. Acc.	Significant Striking Accuracy (percentage)	48.5
	SApM	Significant Strikes Absorbed per Minute	4.12
	Str. Def.	Significant Strike Defence (percentage of opponents' strikes that did not land)	55.2
	TD Avg.	Average Takedowns Landed per 15 minutes	2.1
	TD Acc.	Takedown Accuracy (percentage)	45.8
	TD Def.	Takedown Defense (percentage of opponents' TD attempts that did not land)	60.3
	Sub. Avg.	Average Submissions Attempted per 15 minutes	0.8
Event	event_id	Primary key for UFC events, unique for each event	267
	event_name	Name of the event	"UFC 267"
	event_date	Date of the event	2021-10-30
Fight	fight_id	Primary key for ufc_fights, unique for each fight	12345
	event_id	Secondary key from ufc_events	67890
	fighter_id	Foreign key from ufc_fighters	54321
	knockdowns	Number of knockdowns landed	2
	total_strikes_att	Number of strikes attempted	150
	total_strikes_succ	Number of successful strikes	85
	sig_strikes_att	Number of significant strikes attempted	75
	sig_strikes_succ	Number of significant strikes successful	45
	takedown_att	Number of takedown attempts	10
	takedown_succ	Number of successful takedowns	4
	submission_att	Number of submission attempts	3
	reversals	Number of reversals	1
	ctrl_time	Control time	05:30
	fighter_age	Age of the fighter	28
	winner	Boolean for whether fighter won or lost	True

In preparation for analysis, we rigorously preprocess the dataset. We clean the data by removing duplicates, correcting errors, and standardizing formats across sources. We address missing values using mean imputation and manage outliers through statistical methods. We compute new attributes, including total matches played and winning rates for each player. Additionally, we calculate the Body Mass Index (BMI) and the "ape ratio" [9] to standardize players' physical attributes. We standardize numerical data to ensure uniformity in scale, facilitating accurate comparisons across features and enhancing the robustness of subsequent machine learning analyses.

B. Feature Extraction

Combined with professional knowledge, we select 10 technical features indicative of a fighter's style to refine the dataset. The resulting high dimensionality problem can be solved using factor analysis. The main steps and results are as follows:

1) KMO and Bartlett's Tests: We conduct the Kaiser-Meyer-Olkin (KMO) and Bartlett's tests to assess the appropriateness of our data for factor analysis. KMO values range from 0 to 1, with higher values indicating better suitability for analysis. As presented in Table 2, our KMO value exceeds 0.5, and the significance level is well below 0.05, confirming the dataset's adequacy for factor analysis.

Table 2 KMO and Bartlett's Tests

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.553
Bartlett's Test of Sphericity	*Approx. Chi-Square*	2204.477
	df	45
	Sig.	0.000

2) Eigenvalues and Scree Plot: Eigenvalues are obtained, and a scree plot is drawn to determine the optimal number of factors for dimensionality reduction. The result is demonstrated in Figure 1. Based on the criterion of eigenvalues greater than 1, we select four factors for the construction of new factors. And show in figure 2.

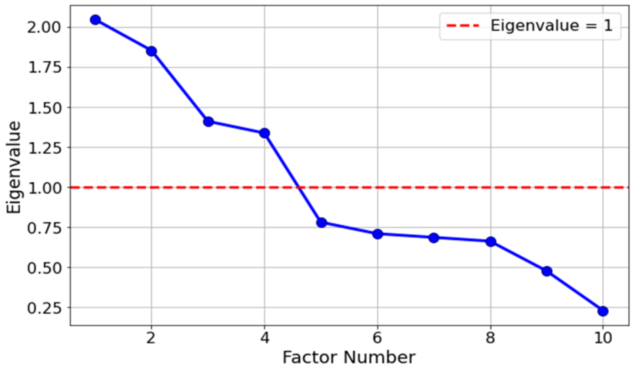

Figure 1. Scree plot of the eigenvalues

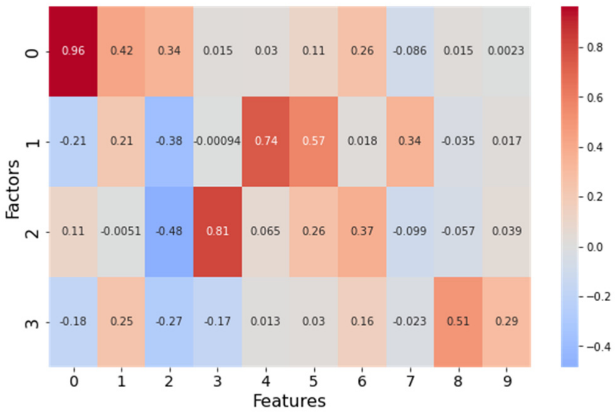

Figure 2. Heatmap plot of factor loadings

3) Construction of Factors: The factor analysis results in the identification of four distinct factors, each representing a different aspect of a fighter's technical style and physical condition. These factors are defined as follows:

Factor 1 (Prefers Stand-up Striking): This factor captures a fighter's preference and effectiveness in stand-up striking:

- SLpM: This measures the offensive capability.
- Str. Acc.: This metric reflects the precision and effectiveness in striking.
- SApM: This metric indicates defensive vulnerabilities in striking.

Factor 2 (Prefers Wrestling and Jiu-Jitsu): This factor reflects a fighter's inclination towards grappling and submission techniques. It includes:

- TD Avg.: This metric indicates the propensity for wrestling.
- TD Acc: This metric reflects proficiency in executing takedowns.
- Sub. Avg.: This metric indicates a focus on Jiu-Jitsu and grappling submissions.

Factor 3 (Defensive Ability): This factor assesses a fighter's defensive skills in both striking and grappling. It includes:

- Str. Def.: This metric indicates their ability to avoid being hit.
- TD Def.: This reflects their ability to stay on their feet and avoid grappling exchanges.

Factor 4 (Physical Condition): This factor captures key physical attributes that can influence a fighter's performance:

- BMI: This is a standardized measure indicating overall fitness and weight category suitability.
- Ape Ratio: This ratio is often used to infer reach advantage in striking exchanges.

By defining these factors, we can better understand the multifaceted nature of a fighter's abilities and preferences, allowing for more nuanced analysis and clustering.

III. METHODOLOGY

A. Fighter Style Clustering

We choose K-means clustering for this task due to its simplicity, efficiency, and scalability, which make it well-suited for handling large datasets with multiple features. Additionally, it provides clear and interpretable results, allowing us to easily identify and analyze distinct clusters of fighting styles. K-means clustering is an unsupervised machine learning algorithm used to partition a dataset into K distinct, non-overlapping subsets or clusters [10]. The primary objective is to minimize the within-cluster variance, which is achieved by assigning each data point to the cluster with the nearest mean value. The algorithm follows the steps of initialization, assignment, update and iteration.

To categorize the fighters' fighting styles, we utilize the factors derived from the factor analysis as the input features for K-means clustering. These factors encapsulate the key technical and physical attributes of the fighters, providing a concise yet comprehensive representation of their fighting styles. To ensure each factor contributes equally to the clustering process, we standardize the factor scores using z-score normalization.

We apply the K-means algorithm to the standardized factor scores. The number of clusters, K is determined based on the Elbow Method, which involves plotting the sum of squared distances from each point to its assigned centroid (within-cluster sum of squares) for different values of K and identifying the point where the rate of decrease sharply slows (the "elbow"). Once the optimal number of clusters is determined, each fighter is assigned to a cluster, representing a distinct fighting style. The centroids of these clusters are analyzed to interpret the predominant characteristics of each fighting style.

The sum of squared distances used to evaluate clustering performance is given by:

$$\text{WCSS}=\sum_{j=1}^{K}\sum_{i=1}^{n_j}\left\|x_i-c_j\right\|^2, \tag{1}$$

where n_j is the number of data points in cluster j, x_j is a data point in cluster j, and c_j is the centroid of cluster j.

By utilizing the new factors obtained from factor analysis, K-means clustering effectively group the fighters into clusters that reflect their fighting styles, providing valuable insights for further analysis and strategic planning.

B. Outcome Prediction

In this study, we employ several machine learning algorithms to predict the outcome of fights based on the factors derived from factor analysis. The algorithms include Random Forest, SVM, XGBoost, Logistic Regression, and Neural Networks. Each of these algorithms offers unique strengths in handling complex datasets and generating accurate predictions.

a) Random Forest: is an ensemble learning method that constructs multiple decision trees during training and outputs the mode of the classes or mean prediction of the individual trees. It reduces overfitting by averaging multiple trees, thus improving prediction accuracy.

b) SVM: is a supervised learning algorithm that finds the optimal hyperplane which maximizes the margin between different classes. The decision boundary is defined by support vectors, which are the data points nearest to the hyperplane. The decision function for SVM is as follows:

$$f(x) = sign\left(\sum_{i=1}^{n} \alpha_i y_i K(x_i, x) + b\right), \tag{2}$$

where α_i are the model parameters, y_i are the class labels, b is the bias term, and $K(x_i, x)$ is the kernel function.

c) XGBoost: is an optimized gradient boosting algorithm that builds decision trees sequentially, where each tree attempts to correct the errors of the previous trees. It uses a regularization term to control overfitting and enhance generalization. The objective function for XGBoost is:

$$\mathcal{L}(\theta) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k), \tag{3}$$

where l is the loss function, y_i is the predicted value, y_i is the true value, and $\Omega(f_k)$ is the regularization term for the k-th tree.

d) Logistic Regression: is a statistical model used for binary classification. It models the probability that a given input belongs to a particular class. The relationship between the input features and the probability is modeled using the logistic function.

e) Neural Networks: consist of interconnected layers of neurons, where each neuron applies a linear transformation followed by a non-linear activation function. The network learns to map inputs to outputs through backpropagation, adjusting the weights to minimize the prediction error.

f) Ensemble Learning: combines the predictions from multiple models to improve accuracy and robustness. By aggregating the strengths of various models, ensemble methods can achieve better performance than individual models. In this study, we use a simple majority voting method to combine all the models into a composite voting model. Each model votes for a class, and the class with the most votes is selected as the final prediction [11]. The prediction from the majority voting ensemble model is given by:

$$\hat{y} = \text{mod}(y_1, y_2, ..., y_n)$$

where y_1, y_2,…, y_n are the predictions from n individual models.

By following this methodology, we ensure that our models are rigorously trained, validated, and tested, leading to reliable predictions of fight outcomes based on the derived factors. We believe that incorporating fighter style factors into the prediction model enhances both the accuracy and interpretability of the predictions, providing more insightful and actionable results.

IV. EXPERIMENT

A. Fighter Style Clustering Result

Our models are trained on the UFC match data mentioned above. Given that the gender of the players in a fighting match may cause significant differences in the pattern of the match [12], we only analyze and predict matches between male players, who have more abundant match data in this experiment.

The clustering analysis results in three distinct categories: Striker, Grappler, and All-rounder. Each category represents a unique fighting style based on the factors derived from our factor analysis.

- Striker: Fighters classified as Strikers have a significantly higher Factor 1 compared to Factor 2. They also tend to have a higher Factor 4 relative to Grapplers.
- Grappler: Grapplers exhibit a significantly higher Factor 2 than Factor 1, indicating a strong preference for wrestling and jiu-jitsu techniques.
- All-rounder: All-rounders show no significant difference between Factor 1 and Factor 2, indicating a balanced proficiency in both striking and grappling. They also generally exhibit higher scores in Factor 3.

The visualization of our cluster analysis is presented in the form of a 3D scatter plot, as shown in Figure 3. The graph also highlights the positions of several famous fighters whose style classification aligns well with mainstream perceptions. Georges St-Pierre excels in wrestling, jiu-jitsu, and defensive abilities, classifying him as a Grappler. Similarly, Islam Makhachev shows a strong inclination toward wrestling and jiu-jitsu, along with notable defensive skills. Anderson Silva balances standing strikes and wrestling, positioning him near the center of the plot. This balance classifies him as a Grappler but also shows traits of an All-rounder. In the All-rounder area, Jon Jones reflects his excellent performance across multiple dimensions. Max Holloway's high preference for standing strikes clearly places him in the Striker category, as does Conor McGregor, who is also positioned in the standing strike area.

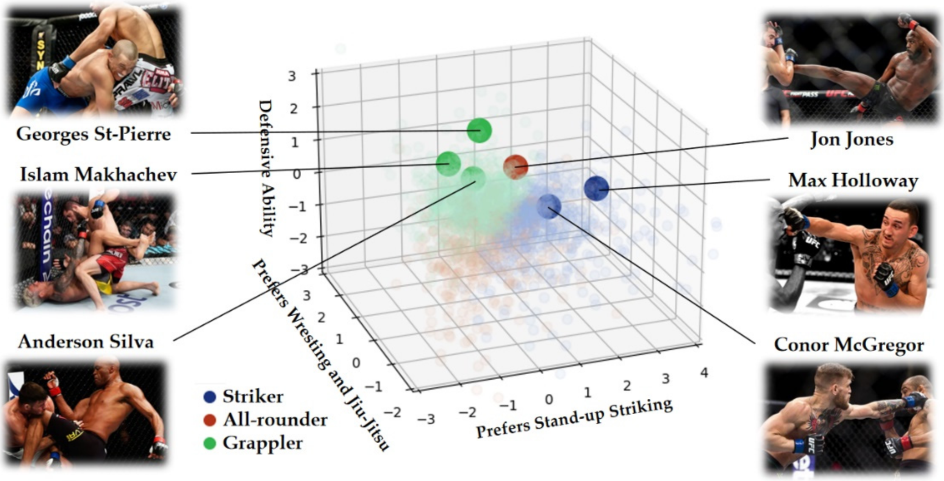

Figure 3. Clustering of MMA fighters' fighting styles

B. *Outcome Prediction Result*

In UFC matches, the outcomes are not strictly binary (win/loss) but also include "No Contest" and "Draw" results. However, these occurrences are extremely rare, constituting approximately 0.3% of all matches. Thus, it is reasonable to simplify the problem by ignoring these "unexpected" results.

Based on this assumption, our study used the difference in raw technical indicators and the difference in technical style features derived from factor analysis for both fighters in each match as inputs. The training of machine learning prediction models is conducted using Python. The details of the model training processes are as follows:

The Random Forest model had 100 decision trees, using the Gini impurity criterion for splits. Nodes expand until all leaves are pure or contained fewer than 2 samples. For the SVM model, we utilized a Radial Basis Function (RBF) kernel with a regularization coefficient set to 1. The XGBoost model is configured with 100 trees, using gbtree as the booster and a learning rate of 0.3. The Logistic Regression model is set with a maximum of 1000 iterations. The Neural Network is constructed and trained using Keras. The network's architecture includes an input layer with dimensions matching the number of features in the training data. The first hidden layer consists of 64 neurons with ReLU activation, followed by a second hidden layer with 32 neurons, also using ReLU activation. The output layer comprised a single neuron with a Sigmoid activation function. The model is compiled using the Adam optimizer and binary cross-entropy loss function, trained for 100 epochs with a batch size of 64 samples.

After independently training each model, we employ ensemble learning. By using a simple majority voting method, we combine all the models into a composite voting model. The accuracy of these models on the test set is demonstrated in Table 3, and the confusion matrix is illustrated in Figure 4. Among the individual models, Logistic Regression achieves the highest prediction accuracy at 63.99%. Furthermore, employing the ensemble majority voting strategy results in further improvement, achieving an accuracy of 65.52%.

Table 3 Prediction Accuracy for Each of Our Models

Model	Accuracy
Random Forest	62.65%
SVM	59.13%
XGBoost	60.64%
Logistic Regression	63.99%
Neural Networks	60.12%
Majority Voting	65.52%

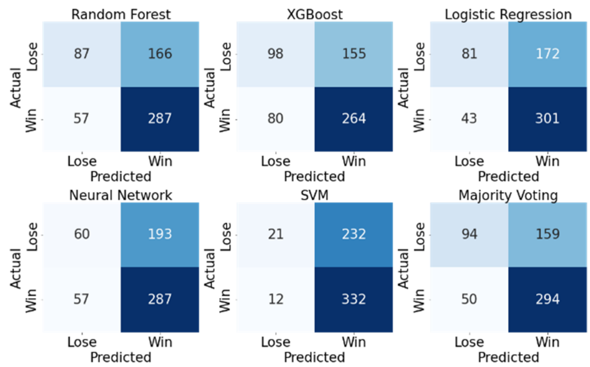

Figure 4. Confusion matrices for each model on the dataset

C. *Ablation Study*

An ablation study is a technique used in machine learning and artificial intelligence research to understand the importance and contribution of different components of a model or system. By systematically removing or "ablating" various parts of the model, researchers can observe the impact on performance, thereby identifying which components are crucial and how they affect overall functionality [13]. In this section, we aim to determine the impact of the technical style factors on model performance by removing these factors from the input features and observing any changes in accuracy.

After removing the technical style factors from the input features, we evaluate the models on the test set. The accuracy results for these models, without the technical style factors, are presented in Table 4, and the comparison of accuracy before and after ablation is demonstrated in Figure 5. It is evident from the table that the prediction accuracy of all models decreases, with reductions ranging from 2% to 4%. This decline in performance indicates that the inclusion of technical style factors indeed contributes significantly to the accuracy of the outcome predictions.

Table 4 Prediction Accuracy for Each of Our Models after Removing the Technical Style Factors from the Input Features

Model	Accuracy	Accuracy reduction
Random Forest	60.64%	3.21%↓
SVM	57.17%	3.32%↓
XGBoost	59.06%	2.61%↓
Logistic Regression	61.58%	3.77%↓
Neural Networks	58.28%	3.06%↓

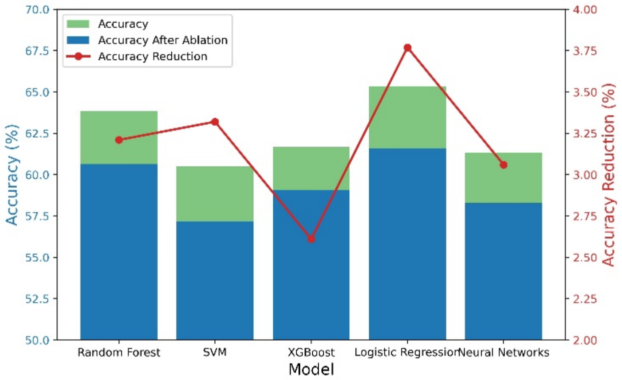

Figure 5. Model Accuracy and Accuracy Reduction After Ablation

V. CONCLUSIONS

This work focused on exploring the patterns in MMA and addressing the challenges of fighter style classification and match outcome prediction using data-driven, machine learning methods. We utilized the K-means algorithm for the first time to cluster fighters based on their technical styles, proposing a series of high-dimensional technical style factors. Through our experiments, we validated the effectiveness of various machine learning algorithms in predicting match outcomes. Additionally, our ablation study confirmed the significance of incorporating technical style factors into the prediction models. The highest accuracy achieved in our study was 65.52%, obtained through an ensemble learning model using majority voting, which significantly outperformed other related works [6]. Our findings provided valuable insights into the application of machine learning in sports analytics, specifically in the context of MMA.

REFERENCES

[1] M. Latyshev, Y. Tropin, L. Podrigalo, and N. Boychenko, "Analysis of the relative age effect in elite wrestlers," Ido movement for culture. Journal of Martial Arts Anthropology, vol. 22, no. 3, pp. 28–32, 2022.

[2] Y. Tropin, L. Podrigalo, N. Boychenko, O. Podrihalo, O. Volodchenko, D. Volskyi, and M. Roztorhui, "Analyzing predictive approaches in martial arts research," Pedagogy of Physical Culture and Sports, vol. 27, no. 4, pp. 321–330, 2023.

[3] C. Ma, "An analysis of weight and fighting styles as predictors of winning outcomes of elite mixed martial arts athletes," Sport Journal, vol. 24, 2023.

[4] B. Holmes, I. G. McHale, and K. Zychaluk, "A markov chain model for forecasting results of mixed martial arts contests," International Journal of Forecasting, vol. 39, no. 2, pp. 623–640, 2023.

[5] G. Walsh, "Predictive analysis of ufc fights: Technical report," Ph.D. dissertation, Dublin, National College of Ireland, 2022.

[6] Hitkul, K. Aggarwal, N. Yadav, and M. Dwivedy, "A comparative study of machine learning algorithms for prior prediction of ufc fights," in Harmony Search and Nature Inspired Optimization Algorithms: Theory and Applications, ICHSA 2018. Springer, 2019, pp. 67–76.

[7] "UFC Stats," [Online]. Available: http://ufcstats.com/. [Accessed: Jun. 21, 2024].

[8] "UFC-Web-Scraping," GitHub. Available: https://github.com/remypereira99/UFC-Web-Scraping. [Accessed: Jun. 21, 2024].

[9] C. Kirk, D. R. Clark, C. Langan-Evans, and J. P. Morton, "The physical demands of mixed martial arts: A narrative review using the armss model to provide a hierarchy of evidence," Journal of Sports Sciences, vol. 38, no. 24, pp. 2819–2841, 2020.

[10] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An efficient k-means clustering algorithm: Analysis and implementation," IEEE transactions on pattern analysis and machine intelligence, vol. 24, no. 7, pp. 881–892, 2002.

[11] A. Rojarath, W. Songpan, and C. Pong-inwong, "Improved ensemble learning for classification techniques based on majority voting," in 2016 7th IEEE international conference on software engineering and service science (ICSESS). IEEE, 2016, pp. 107–110.

[12] M. M. Fern´andez, C. J. Brito, B. Miarka, and A. L. D´ıaz-de Durana, "Anxiety and emotional intelligence: Comparisons between combat sports, gender and levels using the trait meta-mood scale and the inventory of situations and anxiety response," Frontiers in psychology, vol. 11, p. 505982, 2020.

[13] I. Fostiropoulos and L. Itti, "Ablator: Robust horizontal-scaling of machine learning ablation experiments," in International Conference on Automated Machine Learning. PMLR, 2023, pp. 19–1.