A Systematic Literature Review in Machine Learning in Predicting Sports Outcome

1st Athanasios Andrianis
Department of Electrical and Computer Engineering
Hellenic Mediterranean University
Heraklion, Greece
andrianis@hmu.gr

2nd Giorgos Papadourakis
Department of Electrical and Computer Engineering
Hellenic Mediterranean University
Heraklion, Greece
papadour@hmu.gr

3rd Konstantinos Karampidis
Department of Electrical and Computer Engineering
Hellenic Mediterranean University
Heraklion, Greece
karampidis@hmu.gr

4th Nuno Escudeiro
Department of Computer Engineering
Polytechnic Institute of Porto
Porto, Portugal
nfe@isep.ipp.pt

***Abstract*— This paper presents a systematic literature review on Machine Learning (ML) applications for predicting sports outcomes, following PRISMA 2020 guidelines, analyzing 30 primary studies from 2020-2025 sourced from Scopus and Google Scholar. Key findings highlight football and basketball as dominant sports (~67% of studies). Artificial Neural Networks (~30%) is the most commonly used method followed by tree-based methods (~24%), statistical & regression methods (~17%) and instant-based and kernel methods (~15%). Most methods use feature selection (~81%), over formal dimensionality reduction, and accuracy is the primary metric (80% of studies). Prediction rates vary from 55% to 93%.**

Keywords*— *Machine learning, sports outcome prediction, systematic literature review, PRISMA, neural networks, random forest, accuracy.

I. Introduction

The prediction of sports outcomes constitutes an area of considerable intellectual inquiry and commercial value, primarily fueled by the increasing amount of detailed statistical data pertaining to athletes, teams, and competitive seasons, alongside the expanding global sports betting industry[1]. The sports market is a multibillion-dollar enterprise, with leagues such as the National Basketball Association (NBA) and UEFA Champions League generating substantial revenue and data, attracting significant research interest, particularly concerning macro-level outcomes, drawing millions of viewers and billions in wagers [2], [3].

The trend in sports analytics is shifting away from traditional evaluative methods toward advanced approaches utilizing ML and Artificial Intelligence (AI)[1]. This methodological evolution is driven by the necessity to process vast datasets and capture complex, non-linear dependencies inherent in real-world sports competitions, which conventional statistical models often fail to adequately address. Consequently, ML algorithms are frequently deployed for predictive modeling and strategic decision-making across various sporting disciplines [4].

Predicting the result of a match is commonly framed mainly as a classification task searching for the two main categories of win or loss in basketball and American football, Horse Racing, Powerlifting, cricket and volleyball, while, mainly, in football and secondarily in boxing a third category, which is draw, is included. The effectiveness of these predictive models hinges upon the incorporation of rich and diverse features, including historical team performance, player-specific metrics such as defensive rebounds, three-point percentage, and assists in basketball, psychological variables, and even non-box score statistics detailing factors like audience attendance or travel distance[5].

Despite the explosion of research in this domain over the last decade, achieving consistently high prediction accuracy remains a critical challenge, with many models yielding results below seventy percent. These fluctuating and sometimes contradictory outcomes stem largely from inconsistencies in the specific datasets, features, methodologies, and evaluation metrics employed across different studies [6]. Therefore, a systematic literature review is essential to synthesize the broad spectrum of existing work, identify the most efficacious combinations of features and algorithms, and establish reliable benchmarks for subsequent efforts in ML-based sports outcome prediction.

This paper implements a systematic search strategy and adheres to the PRISMA 2020 guidelines to ensure transparent, reproducible, and reliable documentation of the review process. The contribution of this paper is to present and analyzing the literature to answer its core research questions, which include examining trends in sports-related literature, changes in algorithm usage, the most common and successful ML algorithms in sport industry, dominant strategies for dimensionality reduction (DR) and the prioritization of performance metrics in sports outcome problems.

In the following chapter 2 the insights of similar systematic literature reviews are presented. Chapter 3 describes the methodology of systematic literature review as used in this paper, as well as the results and findings. Finally in chapter 4 the paper concludes, summarizing the main findings.

II. Insights from similar Literature Reviews

The results from the conducted literature review can be subdivided into several related categories. There is a distinct agreement among researchers that ML has changed the operation of the sports industry, by providing new tools for athletes, coaching staff, and organization managers.

Reis et al. [7], presented a review focusing on publications related to ML applications in sports and proposing directions for future research. The review encompassed studies on three main areas: injury prediction, performance analysis, and personalized training. The paper established that ML is potentially useful resources for the prevention, detection, diagnosis, and treatment of various health conditions. Among others the findings showed that the Artificial Neural Network (ANN) was the most common technique used in predicting injury risk and sports performance models.

Lotfi and Rebbouj [8], evaluated how ML is used to predict match results. They looked at how data is collected and which algorithms provide the most accurate classifications. Their research noted that deep learning algorithms are particularly effective for these types of models. By comparing studies across sports like Javelin Throw, Basketball, Football, Swimming, American Football, Australian Rules Football, Rugby, Horse Racing and Baseball, they organized findings by measurement tools and sample sizes. Their results confirmed that ML is a reliable method for prediction because it can build models using specific features found in historical data. Generally, ANNs are seen as the most accurate way to predict outcomes when there is a clear relationship between variables. For example, the review cited football match predictions reaching 91% accuracy using Fuzzy Logic, while ANN models provided better predictions for javelin throwing.

Dong [9] discussed how AI fits into the management of sports events. The study analyzed the core technologies supporting applications in the fields ML, Computer Vision (CV), and Big Data Analytics (BDA). Also illustrates the necessity of algorithms of these fields for health monitoring and sports data management. The main results highlighted that AI technologies are effectively used to enhance event management, fan involvement, and the performance of athletes. A comparison showed that Deep Learning (DL) methods often perform better than traditional ML. For sports health monitoring, a hybrid model combining Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) achieved high metrics across accuracy, precision, recall, and F1-score. The review highlighted that AI systems can be used essentially for personalized training plan design and for skill enhancement.

Bunker and Susnjak [10] reviewed studies focused on predicting results in team sports. They divided these into "invasion" sports and "striking/fielding" sports, looking at which algorithms and dataset sizes worked best. While earlier studies often used ANNs, the evidence does not show that they always beat other types of ML models. A major finding from their review was that choosing and refining the right data features is more important for accuracy than simply having a larger number of instances. Predictive accuracy changed depending on the sport; Rugby Union reached 93%, while the highest accuracy for football was 78%. This suggests that sports with lower scoring totals are naturally more unpredictable.

The study by Gandhi et al. [11] addressed to expand predictive methods. Their results show that models built with ML are a promising way to predict scores and athlete performance. The paper discussed how AI could be used to generate game strategies, give feedback after a match, to generate training plans, and prediction based on post-sporting events. These methods are also used to track injury risks for elite athletes over several years. However, the study also mentioned that working with high-quality datasets is difficult because of limits on time and the amount of available data.

Summarizing the reviews, applications and algorithms improving performance analysis, designing personalized training plans, predicting injuries, monitoring athlete health, and forecasting match outcomes. Methodologically, the ANN, is identified as a fundamental technique for prediction across various sports applications. An important finding is that selecting appropriate feature sets is vital for achieving high predictive accuracy, often more important than large datasets size. Finally, all authors agree that significant challenges must be used for future progress.

III. Methodology

A systematic search strategy was implemented to identify pertinent literature on the application of ML techniques in sports outcome prediction. The PRISMA model used in this paper is a flow diagram that visually encapsulates the sequential phases of the systematic review process, adhering to the PRISMA 2020 guidelines for systematic review reporting [12]. The search was conducted across two major academic databases, Scopus and Google Scholar. The result was a total of 352 records, with 157 identified in Scopus and 195 in Google Scholar.

The initial 352 search results were imported into reference manager applications, like Mendeley and Zotero, to remove duplicate records. Then the remaining papers were screened independently against specific criteria. To be included, studies had to be written in English, published in 2020 or later, centered on sports outcomes, and built on machine learning, deep learning, or artificial intelligence methodologies. In the exclusion phase when the full text was unavailable, when the study lacked a direct link to both sports and machine learning, or when the results were unclear or of low quality the papers were excluded. Additionally, because the target was focused strictly on outcome prediction, the research on other sports applications, such as fan participation, injury risk, or betting marketing was dropped. Last selection was only peer-reviewed journal articles, conference papers, and book sections. Unpublished works, including master's theses, were left out. Finally, 30 studies were incorporated into the final systematic review.

TABLE 1 analyzes the total of 30 studies across various disciplines, revealing highly concentrated research focuses heavily prioritizing certain sports. The overwhelming majority of the research is dedicated to Football, which constitutes the largest category by far, accounting for 13 studies, or 43% of the total volume. The second largest focus is on Basketball, with seven studies, contributing 23%. Together, these two sports make up exactly two-thirds (66%) of all the research in the dataset.

TABLE 1:DISTRIBUTION OF PRIMARY STUDIES PER SPORT

Sports	Studies	Number of Studies	%
Football	[1], [2], [3], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22]	13	43%
Basketball	[4], [6], [23], [24], [25], [26], [27]	7	23%

Cricket	[28], [29], [30]	3	10%
American Football	[31], [32], [33]	3	10%
Powerlifting	[34]	1	3%
Volleyball	[35]	1	3%
Horse Racing	[36]	1	3%
Boxing	[37]	1	3%

The following section categorizes the primary algorithms used, mentioned the best result and if there were any use of DR algorithms, in the papers of the majority of the dataset meaning football and basketball.

Adebisi and Alabi in [1] predicted home wins, draws, and away wins in the Nigerian Professional Football League. They evaluated k-Nearest Neighbor (k-NN), Poisson distribution, Logistic Regression (LR), Decision Trees (DT), Random Forests (RF), Support Vector Machines (SVM), ANNs, and Naive Bayes (NB). They excluded DR algorithms and instead the input was 7 features. A hybrid model that combined k-NN and Poisson regression achieved the highest accuracy of 93%.

Ievoli et al. [2] modeled the probability of winning a UEFA Champions League match. The algorithms compared were LR, ANNs, NB, and RF. To reduce the number of features, they applied a bidirectional stepwise selection process based on the Akaike Information Criterion (AIC). The final models used 24 features. The LR model provided the highest performance achieving 76% accuracy.

Ćwiklinski et al. [3] predicted the success of professional association football player transfers. They evaluated RF, NB, and AdaBoost for this classification task. The researchers retained the full dataset without DR. The models used 11 features, including player and team metrics across technical, mental, and physical categories, alongside age, nationality, and league rules. The RF model achieved the highest performance, with an accuracy of 82%.

Anfilets et al. [13] predicted English Premier League match outcomes, categorized as wins, losses, or draws. They developed a Deep Neural Network (DNN) that incorporated Elastic Net regularization. The dataset combined historical match details to construct 108 features per game. To address dimensionality, they applied Lasso regularization (L1) as an embedded feature selector, which excluded less important features. The DNN combined with Elastic Net model performed good and achieved an accuracy of 61%.

Rahman in [14] predicted group and knockout stage results (wins, draws, or losses) for the 2018 FIFA World Cup. The author developed Long Short-Term Memory (LSTM) blocks and standard ANNs. Instead of using a separate DR algorithm, the model used deep embedding layers to process the raw input features. The final system evaluated 10 features. The LSTM network achieved the highest accuracy and correctly classified 63% of the matches.

Glowania et al. [15] predicted draws, home wins, and away wins in the German Bundesliga and the English Premier League. They constructed an ensemble model combining DT, SVM, AdaBoost, Bagging, and RF, and evaluated alternative weighted voting rules. They did not apply DR, though they noted its potential utility for addressing highly correlated features in future work. The model used 6 features. That method named as "Heterogeneous Ensemble of Classifiers" achieved an accuracy of 60%.

Eğitimi et al. [16] predicted match outcomes in the Turkish league. They evaluated LR, RF and 3 DT variants, Classification and Regression Trees (CART), Chi-square Automatic Interaction Detector (CHAID), and Quick, Unbiased, Efficient Statistical Tree (QUEST). Rather than using algorithmic DR, they manually reviewed correlation matrices to select features. Their final models used 3 features selected from an initial set of 8. The CART algorithm outperformed the other models by achieving 67% accuracy.

Govindasamy et al. [17] used a Binary version of LR which predicts the probability of a categorical outcome with the options of wins and losses in tournament matches. The authors applied forward stepwise selection to limit the feature set. They reduced an initial set of 13 features to 6. The LR model achieved an accuracy of 87%.

Obradović and Kečo [18] classified English Premier League matches into wins, losses, or draws. They compared LR, ANNs, and SVM. The researchers cleaned the dataset by removing incomplete rows and irrelevant columns, omitting mathematical DR. The final models evaluated 24 variables, which represented recent performance streaks, current points, and league standing differentials. The ANN achieved the highest accuracy at 61%.

Rebbouj and Lotfi [19] predicted the final team standings in the Moroccan national league. The researchers used Elastic Net regression as their primary methodology. Instead of reducing the feature space with a DR algorithm, they preprocessed the raw data using string casting, categorical hashing, and standard scaling. The model evaluated 9 features. The Elastic Net model achieved an R^2 score of 0.99.

Wong et al. [20] , forecasted whether a home team in the English Premier League would win or fail to win. They evaluated LR, RF, SVM, XGBoost, LightGBM, and Convolutional Neural Networks, alongside voting and stacking ensemble methods. The authors retained all 52 features rather than applying DR algorithms, but they standardized the numerical values. These features included weather conditions, player fatigue estimates, and team momentum. Convolutional Neural Networks achieved the highest overall accuracy of 63%.

Tümer et al. [21] forecasted the final league standings in the Turkish league. They evaluated ANNs, Radial Basis Function Networks, and Linear Regression (LinR). They used the ReliefF scoring method to rank feature importance but did not apply mathematical DR. The models processed 23 physical and technical statistics, including total running distance, passes per minute, and dribbles. The ANN achieved the highest accuracy, with an R^2 value of 0.60.

Gao et al. [22] predicted match outcomes home wins, draws, or away wins, using the European Soccer Database. They proposed a hybrid model combining a 1D Convolutional Neural Network (1D CNN) and a Transformer. They compared this architecture with DT, RF, k-NN, XGBoost, ANNs and LSTMs. The authors reduced data dimensions using max pooling layers within the CNN and embedding layers for categorical features. Although they did not report the total number of features, the dataset contained multiple continuous and categorical match events. The hybrid 1D

CNN and Transformer model achieved the highest performance, with an accuracy of 55%.

Lu et al. [4] explained an approach to predicting final NBA game scores relied on adaptive weighted features. The researchers set up prediction systems using CART, RF, Stochastic Gradient Boosting and XGBoost. They avoided standard DR algorithms. Instead, they designed an engineering method based on Inverse Distance Weighting to create adaptive weighted features that incorporate game-lag information. This expanded their initial fourteen basic box score features into complex feature combinations. After cross-validation, the Stochastic Gradient Boosting model slightly beat the other methods. It produced the lowest Root Mean Square Error (RMSE) with 11.55.

Kim et al. [6] presented in a paper on March Madness, the goal was to forecast matchup outcomes between any two teams in the NCAA men's basketball tournament. The researchers tested several classifiers ANNs, k-NN, SVM, LR, and RF. To reduce dataset size, they used feature scoring tools such as Information Gain, Gain ratio, Gini index, Chi-square, and ReliefF but no DR algorithms. They started with 18 features per instance and narrowed these down to the 5 most influential features. The ANN performed best and beat the other options with an accuracy rate of 67%.

Li et al. [23] focused on predicting NBA win probabilities and expected team wins. They paired a multivariate S-shape LR with a non-parametric Data Envelopment Analysis (DEA) model. They did not use a named mathematical DR algorithm, instead they relied on feature selection based on regression parameters to trim the dataset. This process reduced the feature number to 8, after they dropped statistical variables like offensive rebounds and blocks. In their testing, the combined DEA and LR setup predicted 60% wins.

Wang et al. [24] predicted an NBA team's monthly win ratio and classified whether a team would reach the post-season playoffs. The authors applied RF and ANN algorithms. For DR, they chose Principal Component Analysis (PCA) to visualize how gameplay styles changed over different eras. The raw data had many gameplay attributes, which they processed into roughly 32 metrics per team. The ANN performed best with an R^2 of 0.77 for win ratio predictions.

Horvat et al. [25] presented a machine learning model to predict NBA win-loss outcomes using a custom extended team efficiency index. They presented their model as a "data-driven machine learning algorithm" and classified as a "supervised ML model". This model compared against standard classifiers including LR, Naive Bayes, DT, ANN, k-NN and RF. When they tried feature selection based on information gain, it lowered prediction accuracy. Thus, they decided to keep the full dataset with 42 features. The extended team efficiency index model performed best by reaching a maximum accuracy of 78% and beat the standard supervised classifiers.

Patrot et al. [26] aimed to estimate individual NBA game outcomes and total season wins using past performance data. The authors built their models using LinR, SVM, and DT. For feature selection, they applied Correlation Feature Sets and Multiple Regression. They trained and tested the models on 17 features. The LinR model yielded the best results by reaching a prediction accuracy of 92%.

Walsh and Joshi [27] evaluated a study that estimated the true probability of NBA game outcomes. The analysis evaluated LR, RF, SVM, and ANN. The researchers reduced dimensionality using filter methods with the Spearman correlation coefficient, followed by a Sequential Forward Selection wrapper method. They took 24 statistics and narrowed them down 6 features. The SVM model yielded the best results with an accuracy of 66%.

Machine learning and statistical techniques show varying success when forecasting football and basketball outcomes. Researchers regularly compare new models against standard options like LR, RF, and ANNs. But basic, ordinary classifiers often struggle with the complex, volatile data typical of sports. Because of these limitations, researchers find they must customize or proceed to hybrid solutions their methods to get better results.

After concentrating on the performance in TABLE 2, the first observation is that the most accurate predictions do not come from single, basic algorithms. Instead, they rely on hybrid models and setups designed specifically for sports metrics. For example, a football forecasting model that combined k-Nearest Neighbor with Poisson regression hit an accuracy of 93% [1] and easily beats any single algorithm that performed on its own. We see the same trend in other studies. The Extended Team Efficiency Index [25] for basketball and the Heterogeneous Ensemble of Classifiers [15] for football, both beat standard classifiers. To capture how teams play and how match environments shift, researchers moved away from isolated tools and used combined, domain-specific systems.

TABLE 2: BEST RESULT AND METRIC IN EVERY PAPER

Sport	Paper	Performance	Metric	Algorithm
Football	[1]	93%	Accuracy	KNN + Poisson Distribution
	[2]	76%	Accuracy	LR
	[3]	82%	Accuracy	RF
	[13]	61%	Accuracy	DNN
	[14]	63%	Accuracy	LSTM
	[15]	60%	Accuracy	Heterogeneous Ensemble of Classifiers
	[16]	67%	Accuracy	CART (DT)
	[17]	87%	Accuracy	LR
	[18]	61%	Accuracy	ANN
	[19]	0.99	R^2	Linear Regression (Elastic Net)
	[20]	63%	Accuracy	DNN
	[21]	0.60	R^2	ANN
	[22]	55%	Accuracy	DNN
Basketball	[4]	11.55	Root Mean Square Error (RMSE)	Stochastic Gradient Boosting
	[6]	67%	Accuracy	ANN
	[23]	60%	Accuracy	LR
	[24]	0.77	R^2	ANN
	[25]	78%	Accuracy	Extended team efficiency index
	[26]	92%	Accuracy	Linear Regression
	[27]	66%	Accuracy	SVM

Looking at the broader research, prediction success depends no more on combining the models and organizing the features than on using a highly complex single model. Deep learning methods like Deep Neural Networks and LSTMs can process raw data well using embedding layers. Even so, their standalone accuracy for predicting match winners stayed low, hovering between 55% and 63%. The success of hybrid systems and custom indices shows that progress in sports forecasting requires pairing traditional statistical sports modeling with modern machine learning.

Which algorithm performed best is viewed in TABLE 3 and graphically in Figures 1 and 2 for football and basketball.

TABLE 3: ALGORITHM PER SPORT

Sports	Algorithms	Times Used	Best Result
Football	AdaBoost	2	
	Artificial Neural Network (ANN)	5	2
	Bagging	1	
	Decision Tree	4	1
	Depp Neural Networks	3	3
	k-Nearest Neighbour (k-NN)	2	1
	LightGBM	1	
	Linear Regression	2	1
	Logistic Regression	6	2
	Long Short-Term Memory Network (LSTM)	2	1
	Naive Bayes	2	
	Radial Basis Function (RBF)	1	
	Random Forest (RF)	5	1
	Support Vector Machine (SVM)	3	
	XGBoost Classifier	2	
Basketball	Artificial Neural Network (ANN)	3	2
	Decision Tree	3	
	k-Nearest Neighbour (k-NN)	2	
	Linear Regression	2	1
	Logistic Regression	1	1
	Naive Bayes	1	
	Random Forest (RF)	4	
	Support Vector Machine (SVM)	3	1
	XGBoost Classifier	1	
	Stochastic Gradient Boosting	1	1

Football research as seen both from TABLE 1 and FIGURE 1 is characterized by a wider, more exploratory approach, utilizing 41 total instances across 15 different algorithms. This translates to a lower overall efficiency, requiring researchers to test more models to achieve the same number of successes. The most used models are LR and RF with 6 and 5 instances respectively, which recorded a modest 33% and 20% success rate. Conversely, the highest efficacy is found in models designed to capture complexity. DNN has the highest success rate among frequently used models with 100% in 3 instances, with LSTM also performing well with 50%. On the contrary, simple ANNs performed 25% in 5 instances. This pattern suggests that predictive tasks in Football, which may involve more varied data sources or complex spatio-temporal dynamics, require advanced deep learning models to unlock better performance, justifying the need for broader experimentation.

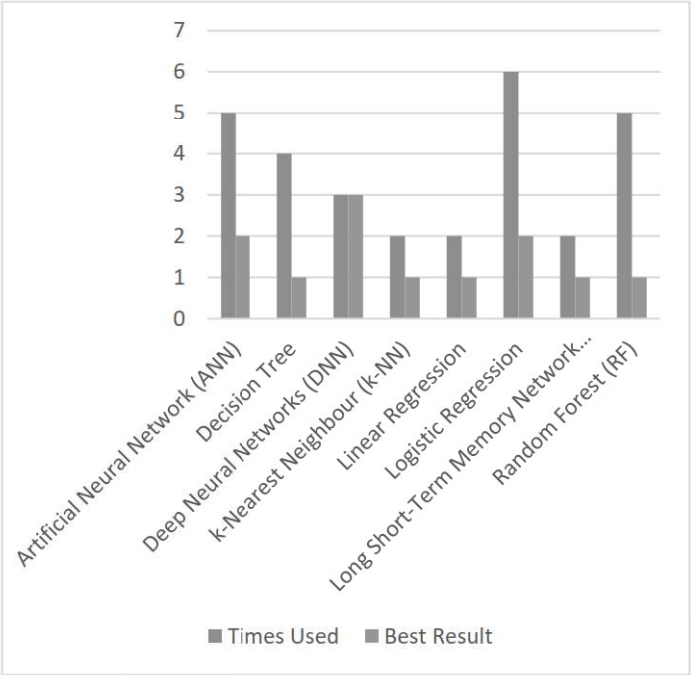


FIGURE 1:ALGORITHMS FOR FOOTBALL

In contrast, as seen in TABLE 3 and FIGURE 2, basketball research focuses on a more condensed set of 10 algorithms with 21 total instances. Highly utilized models like ANN performed well, achieving a 67% success rate. This suggests that for the types of predictive tasks documented here basketball data is highly effectively modeled by ANN methods, which exceed at capturing the non-linear feature interactions common in sports metrics.

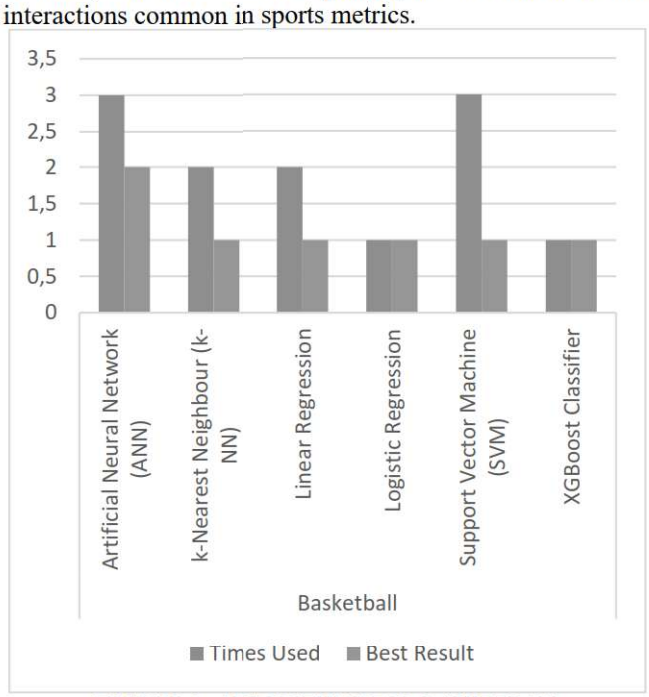


FIGURE 2: ALGORITHMS FOR BASKETBALL

One of the suggestions is that the major growth in recent sports analytics research is centered on non-deep-learning models despite their better performance. This sight likely stems from researchers favoring models that are more understandable, less computationally expensive, and sometimes highly effective for the specific predictive tasks common in sports data, despite the increasing sophistication of deep learning architectures elsewhere.

The dominant strategy in literature is to avoid explicit DR, instead relying on either no preprocessing or Feature Selection techniques. Studies did not perform any feature engineering or DR algorithm, while studies used simple or practical feature selection. Interestingly, the study that used PCA faced the largest feature spaces, with a maximum feature count of 322. Conversely, studies that did not use DR algorithms tended to deal with the smallest minimum feature. This suggests researchers only employ complex DR techniques like PCA when necessary to handle very large datasets, otherwise preferring to use simpler, rule-based Feature Selection or simply work with the original, potentially smaller, feature set. Moreover, strongly indicates a preference among these studies to avoid the complexity or computational cost of formal DR techniques.

Literature in this field demonstrates prioritization of classification tasks, with accuracy emerging as the overwhelmingly dominant performance metric. Reported in 16 out of 20 studies, accuracy far surpasses alternatives such as R^2 and RMSE. This emphasis reflects a strong orientation toward categorical prediction problems, where accuracy serves as the standard benchmark for binary and multi-class classification. The activity of high accuracy is achieved through ANNs architectures and reliable linear baselines, reinforcing the significance of classification-oriented algorithms in the selected papers.

IV. Conclusion

This systematic literature review thoroughly analyzed the application of Machine Learning (ML) techniques in sports outcome prediction across studies published from 2020 onward, adhering strictly to the PRISMA 2020 guidelines. The comprehensive screening process, starting from an initial retrieval of 352 records, culminated in the incorporation of 30 primary studies.

The collected papers show a clear concentration of interest in Football (43%) and Basketball (23%) make up two-thirds of the research studied. The rest of the papers cover fewer common sports, including Boxing, Horse Racing, Powerlifting, and Volleyball. Most of these papers focus on classification tasks, with ANNs as the main algorithm used and 80% of them use prediction accuracy as their main evaluation metric.

Prediction accuracy across models ranges from 55% to 93%. Many researchers still choose simple options like logistic regression and random forests because they are fast and easy to interpret. Standalone deep learning models struggle with sports volatility; for instance, DNNs and LSTMs posted lower accuracies between 55% [22] and 63% [20]. Hybrid and custom systems work best. A football model that combined k-Nearest Neighbors with Poisson regression achieved 93% accuracy [1], while custom basketball index reached 78% [25], outperforming standard supervised classifiers.

When preparing data, most of the studies used no preprocessing or relied only on simple feature selection. Researchers rarely used complex DR tools like PCA, saving them only for exceptionally large datasets. This choice keeps the models fast and straightforward. Researchers generally agree that selecting the right game-specific features is more useful than applying complex mathematical transformations. Finally, the literature demonstrates a strong prioritization of Accuracy serving as the overwhelmingly dominant performance metric, reported in 80% of all primary studies. This reliance on accuracy is directly correlated with the consistent high performance achieved by robust classification models such as ANN and Logistic Regression.

Even with these developments, sports forecasting faces clear physical and data limits. Gathering high-quality, complete datasets is difficult because of scheduling constraints and restricted access to team records. Prediction accuracy is also tied to the scoring style of each sport. Low-scoring sports like football are naturally more volatile and harder to predict than high-scoring sports. This volatility explains why many sports models struggle to consistently beat a 70% accuracy threshold.

Making future models more effective is essential for future progress in the field. First, future studies should focus on designing hybrid modeling solutions that pair traditional statistical sports forecasting with modern machine learning algorithms to achieve superior predictive strength. Second, predictive accuracy can be significantly enhanced by moving beyond standard box-score features to integrate richer, dynamic inputs such as player fatigue estimates, weather conditions, psychological variables, and team momentum. Lastly the primary objective can be switched, from binary win or loss results to predicting real-time match momentum swings throughout a game taking advantage of the benefits of the advanced deep learning models. This change will allow frameworks to capture the highly complex, non-linear dependencies that define real-world athletic competition.

References

[1] J. Adebisi and Alabi Damilola, "Design and Implementation of a Predictive Model for Nigeria Local Football League," International Journal of Computer Science and Security (IJCSS), vol. 15, no. 4, pp. 106–122, Aug. 2021.

[2] R. Ievoli, L. Palazzo, and G. Ragozini, "On the Use of Passing Network Indicators to Predict Football Outcomes," Knowledge-Based Systems, vol. 222, p. 106997, Jun. 2021, doi: 10.1016/j.knosys.2021.106997.

[3] B. Ćwiklinski, A. Giełczyk, and M. Choraś, "Who Will Score? A Machine Learning Approach to Supporting Football Team Building and Transfers," Entropy, vol. 23, no. 1, p. 90, Jan. 2021, doi: 10.3390/e23010090.

[4] C.-J. Lu, T.-S. Lee, C.-C. Wang, and W.-J. Chen, "Improving Sports Outcome Prediction Process Using Integrating Adaptive Weighted Features and Machine Learning Techniques," Processes, vol. 9, no. 9, p. 1563, Sep. 2021, doi: 10.3390/pr9091563.

[5] F. Sjöberg, "Football Match Prediction Using Machine Learning," PhD Thesis, Åbo Akademi University, Finland, 2023. [Online]. Available: https://www.doria.fi/handle/10024/187628

[6] J. W. Kim, M. Magnusen, and S. Jeong, "March Madness Prediction: Different Machine Learning Approaches with Non - Box Score Statistics," Manage Decis Econ, vol. 44, no. 4, pp. 2223 - 2236, Jun. 2023, doi: 10.1002/mde.3814.

[7] F. J. J. Reis, R. K. Alaiti, C. S. Vallio, and L. Hespanhol, "Artificial intelligence and Machine Learning approaches in sports: Concepts, applications, challenges, and future perspectives," Brazilian Journal of Physical Therapy, vol. 28, no. 3, p. 101083, May 2024, doi: 10.1016/j.bjpt.2024.101083.

[8] S. Lotfi and M. Rebbouj, "Machine Learning for sport results prediction using algorithms," IJITAS, vol. 3, no. 3, pp. 148–155, Aug. 2021, doi: 10.52502/ijitas.v3i3.114.

[9] H. Dong, "Overview of the Application of Artificial Intelligence in Sports Events," in Proceedings of the 2024 International Conference on Sports Technology and Performance Analysis, Guangzhou China: ACM, Dec. 2024, pp. 324–331. doi: 10.1145/3723936.3723986.

[10] R. Bunker and T. Susnjak, "The Application of Machine Learning Techniques for Predicting Match Results in Team Sport: A Review," jair, vol. 73, pp. 1285–1322, Apr. 2022, doi: 10.1613/jair.1.13509.

[11] T. Gandhi, P. Pradhan, D. Patidar, R. Kushwah, and H. Kaur, "Revolutionizing Sports: An In-depth Analysis of Artificial Intelligence in Sports," Frontiers in Health Informatics, pp. 3429–3447, Nov. 2024.

[12] M. J. Page et al., "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews," BMJ, p. n71, Mar. 2021, doi: 10.1136/bmj.n71.

[13] S. Anfilets et al., "Deep Multilayer Neural Network for Predicting the Winner of Football Matches," IJC, vol. 19, no. 1, pp. 70–77, Mar. 2020, doi: 10.31891/1727-6209/2020/19/1-70-77.

[14] Md. A. Rahman, "A Deep Learning Framework for Football Match Prediction," SN Appl. Sci., vol. 2, no. 2, p. 165, Feb. 2020, doi: 10.1007/s42452-019-1821-5.

[15] S. Głowania, J. Kozak, and P. Juszczuk, "New Voting Schemas for Heterogeneous Ensemble of Classifiers in the Problem of Football Results Prediction," Procedia Computer Science, vol. 207, pp. 3393–3402, 2022, doi: 10.1016/j.procs.2022.09.398.

[16] B. Eğitimi, S. B. Dergisi, D. Topcu, and Ö. Vupa Çilengiroğlu, "Performance Comparison of Machine Learning Methods in Turkish Super League Match Result Predictions," SPORMETRE Beden Eğitimi ve Spor Bilimleri Dergisi, vol. 22, no. 2, pp. 59–72, Jul. 2024, doi: 10.33689/SPORMETRE.1381602.

[17] K. Govindasamy et al., "Uncovering Determinants of Victory and Defeat in Men's UEFA Champions League: An Analytical Exploration Using Logistic Regression," StS, vol. 18, no. 2, Jul. 2024, doi: 10.5817/StS2024-2-5.

[18] A. Obradović and D. Kečo, "Sports Results Prediction Model Using Machine Learning," SAR Journal, pp. 184–189, Sep. 2024, doi: 10.18421/SAR73-03.

[19] M. Rebbouj and S. Lotfi, "Football League Ranking Prediction Using Machine Learning Regression Model," Annals of Applied Sport Science, 2024, [Online]. Available: https://aassjournal.com/browse.php?a_id=1226&slc_lang=en&sid=1&printcase=1&hbnr=1&hmb=1

[20] A. Wong, E. Li, H. Le, G. Bhangu, and S. Bhatia, "A Predictive Analytics Framework for Forecasting Soccer Match Outcomes Using Machine Learning Models," Decision Analytics Journal, vol. 14, Mar. 2025, doi: 10.1016/j.dajour.2024.100537.

[21] A. E. Tümer et al., "Prediction of Soccer Clubs' League Rankings by Machine Learning Methods: The Case of Turkish Super League," Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology, vol. 239, no. 3, pp. 422–436, Sep. 2025, doi: 10.1177/17543371221140492.

[22] J. Gao, Y. Cheng, and J. Gao, "Predicting sport event outcomes using deep learning," PeerJ Comput. Sci., vol. 11, p. e3011, Jul. 2025, doi: 10.7717/peerj-cs.3011.

[23] Y. Li, L. Wang, and F. Li, "A Data-Driven Prediction Approach for Sports Team Performance and Its Application to National Basketball Association," Omega, vol. 98, p. 102123, Jan. 2021, doi: 10.1016/j.omega.2019.102123.

[24] Y. Wang, W. Liu, and X. Liu, "Explainable AI Techniques with Application to NBA Gameplay Prediction," Neurocomputing, vol. 483, pp. 59–71, Apr. 2022, doi: 10.1016/j.neucom.2022.01.098.

[25] T. Horvat, J. Job, R. Logozar, and Č. Livada, "A Data-Driven Machine Learning Algorithm for Predicting the Outcomes of NBA Games," Symmetry, vol. 15, no. 4, p. 798, Mar. 2023, doi: 10.3390/sym15040798.

[26] A. Patrot, H. H, S. B, G. P L, and Sahana, "NBA Game Prediction Using Machine Learning Algorithm," in 2023 International Conference on Recent Trends in Electronics and Communication (ICRTEC), Mysore, India: IEEE, Feb. 2023, pp. 1–6. doi: 10.1109/ICRTEC56977.2023.10111906.

[27] C. Walsh and A. Joshi, "Machine Learning for Sports Betting: Should Model Selection Be Based on Accuracy or Calibration?," Machine Learning with Applications, vol. 16, p. 100539, Jun. 2024, doi: 10.1016/j.mlwa.2024.100539.

[28] U. V. Anbazhagu and R. Anandan, "Predicting the Cricket Match Outcome Using ANFIS Classifier for Viewers Opinions on Twitter Data," in Data Management, Analytics and Innovation, vol. 1174, N. Sharma, A. Chakrabarti, V. E. Balas, and J. Martinovic, Eds., in Advances in Intelligent Systems and Computing, vol. 1174. , Singapore: Springer Singapore, 2021, pp. 171–182. doi: 10.1007/978-981-15-5616-6_12.

[29] D. Abeysuriya, S. Fernando, and R. Navarathna, "Beyond the Run-rate: Forecasting Framework for First Innings Score in T20 Cricket," in 2023 Moratuwa Engineering Research Conference (MERCon), Moratuwa, Sri Lanka: IEEE, Nov. 2023, pp. 48–53. doi: 10.1109/MERCon60487.2023.10355397.

[30] K. Sharma, L. Hota, V. A. Tikkiwal, and A. Kumar, "Exploring Twitter Sentiments for Predicting Match Outcomes in the Game of Cricket," Procedia Computer Science, vol. 235, pp. 152–162, 2024, doi: 10.1016/j.procs.2024.04.018.

[31] P. Xenopoulos and C. Silva, "Graph Neural Networks to Predict Sports Outcomes," in 2021 IEEE International Conference on Big Data (Big Data), Orlando, FL, USA: IEEE, Dec. 2021, pp. 1757–1763. doi: 10.1109/BigData52589.2021.9671833.

[32] M. Gifford and T. Bayrak, "A Predictive Analytics Model for Forecasting Outcomes in the National Football League Games Using Decision Tree and Logistic Regression," Decision Analytics Journal, vol. 8, p. 100296, Sep. 2023, doi: 10.1016/j.dajour.2023.100296.

[33] N. Cliffel and R. Ramnath, "Optimizing Sports Predictions Using Model Selection: A Case Study in Reproducible Data-Science," in 2025 IEEE 49th Annual Computers, Software, and Applications Conference (COMPSAC), Jul. 2025, pp. 283–288. doi: 10.1109/COMPSAC65507.2025.00045.

[34] W. Rafajłowicz and J. Marszałek, "Intelligent Performance Prediction for Powerlifting," in Artificial Intelligence and Soft Computing, vol. 12854, L. Rutkowski, R. Scherer, M. Korytkowski, W. Pedrycz, R. Tadeusiewicz, and J. M. Zurada, Eds., Cham: Springer International Publishing, 2021, pp. 475–484. doi: 10.1007/978-3-030-87986-0_42.

[35] D. Sanghvi, P. Deshpande, S. Shanbhogue, and V. Shah, "Analyzing and Predicting NCAA Volleyball Match Outcome Using Machine Learning Techniques.," in star.informatik.rwth-aachen.de, Buenos Aires, Argentina, Oct. 2021, pp. 99–116. [Online]. Available: http://star.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-2992/icaiw_wdea_2.pdf

[36] M. Gupta and L. Singh, "Predicting Outcomes of Horse Racing using Machine Learning," IJRITCC, vol. 11, no. 9, pp. 38–47, Oct. 2023, doi: 10.17762/ijritcc.v11i9.8119.

[37] A. K. Krutikov, V. Y. Meltsov, and D. A. Strabykin, "Evaluation the Efficienty of Forecasting Sports Events Using a Cascade of Artificial Neural Networks Based on FPGA," in 2022 Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), Saint Petersburg, Russian Federation: IEEE, Jan. 2022, pp. 355–360. doi: 10.1109/ElConRus54750.2022.9755840