# CHAPTER 6

## Developing an Ensemble-Based Model for Player Performance Classification in the Indian Super League

*Swapnil Kumar Sahu* and Sushma Jaiswal***

**ABSTRACT**

In order to ensure forecast accuracy and generalizability across a variety of player positions, this study creates a machine learning ensemble framework for assessing Indian Super League (ISL) 2023–24 player performance utilizing goals and assists as key criteria. To guarantee reliable input representation, the dataset underwent pre-processing using polynomial feature engineering, uniform scaling, and rare-class management. An ensemble integrates Random Forest meta-learner with Bagging, HistGradientBoosting, Gradient Naïve Bayes, and Decision Tree classifiers The stacking ensemble provides better stability and interpretability across player categories while achieving competitive accuracy (91.11%), $R^2$ (0.9395), and RMSE (0.7888) as compared to individual models. Actionable insights for recruiting and scouting were made possible by the derivation of suitability categorization in relation to median contribution values. The results set the groundwork for future sports analytics research in developing football ecosystems by highlighting stacking ensembling as a possible methodological development for player appraisal in ISL.

**Keywords:** Ensemble learning, Player performance evaluation, Indian super league (ISL), Sports analytics.

### 1.0 Introduction

An important step in assessing an athlete's skill level and training quality is performance evaluation (Morciano *et al.*). Athletes' physical condition, past achievements, and mental health are just a few of the many variables that must be examined in order to evaluate and predict their performance (Mackenzie & Iso-Ahola). To create focused training sessions that are intended to increase athletes' individual skills and fill up any ability gaps, trainers rely on performance evaluation (Tichy).

**Corresponding author; Research Scholar, Department of Computer Science & Information Technology, Guru Ghasidas Central University, Bilaspur, Chhattisgarh, India*
*(E-mail: swapnilsahu747@gmail.com)*
***Assistant Professor, Department of Computer Science & Information Technology, Guru Ghasidas Central University, Bilaspur, Chhattisgarh, India (E-mail: jaiswalsush2@gmail.com)*

*DOI: 10.17492/JPI.DYPIU.052606* *ISBN: 978-93-49790-31-5*

Performance evaluation becomes considerably more difficult in team sports like football since it requires gathering information about several players over time while taking into consideration for their individual responsibilities on the squad (Gabbett *et al.*).

In football, scouting, training, and tactical decision-making all depend considerably on player performance evaluation. Existing evaluation techniques mostly rely on subjective judgments, which frequently result in biases and inconsistencies. Machine learning (ML), on the other hand, provides a data-driven methodology that makes it possible to analyze players performance more accurately and objectively using statistical performance labels.

Despite ML-driven player evaluation has progressed, current models sometimes have trouble generalizing to other positions, especially goalkeepers, defenders, or attackers. Evaluation criteria must be customized to each function because of the distinct duties and the skill sets required for each. As a result, one ML model might not be enough for evaluating players in various roles.

This paper investigates ensemble learning strategies, namely Stacking Classifiers, to improve prediction accuracy and resilience in order to overcome these issues. We hope to create a thorough mechanism for evaluating performance that provides a more accurate and comprehensible appraisal of football players through the integration of multiple machine learning models.

Dijkhuis et al. made an effort to predict football players' performance. Here, data on football players' distance travelled, distance in speed category, and energy expenditure were gathered by the authors using an optical tracking device (Dijkhuis et al.). The underperformance of players 15 minutes after the start of a match was then predicted using these data and two ML models. Underperformance was defined as failing to meet 100%, 95%, or 90% of each player's season average for a particular performance metric taken into consideration. If the player underperforms during the game, the trainers may obtain valuable information about the player's state, which may lead them to replace the player with another player if necessary. If the performance is above the average, additional information is obtained. For example, it can assist in the assessment of the skill level of the player in comparison with teammates, as well as specific training plans. In order for this to be achieved, sensors integrated into the vests worn by the players can be used for the acquisition of a large amount of useful data. Some sensors that can be used by the players without being cumbersome are heart rate monitors (HRM), inertial movement units (IMUs), and GPSs. Alexandre et al., Reinhardt et al. These sensors are capable of gathering a wide range of data, such as performance and biometric data.

Although there are various studies on the application of machine learning models for predicting football player performance in prominent football leagues in Europe, such as Serie A, La Liga, and the English Premier League, there remains a conspicuous absence of studies on the Indian Super League (ISL). Considering the explosive growth of the ISL and

DOI: 10.17492/JPI.DYPIU.052606 ISBN: 978-93-49790-31-5

its increased presence in the world football arena, it is surprising that there is little available investigation of the use of cutting-edge machine learning models for player performance assessment or prediction in the ISL scenario. In order to fill this research gap, this study creates and assesses data-driven models especially suited for ISL player datasets, offering fresh perspectives on performance analytics for Indian football.

The following are this work's main contributions:

- Data-driven framework for Indian football: By creating a strong machine learning framework especially suited to Indian Super League (ISL) player data, this study fills a significant research gap by offering one of the first thorough, data-driven assessments of player performance in Indian football.
- Objective player performance evaluation: It shows how machine learning methods are useful for assessing football players' performance based on position-specific and historical characteristics, allowing for precise player suitability forecasts.
- Comparative model evaluation: To identify the best approach for performance prediction, a number of machine learning models, including ensemble techniques like Random Forest (Patel *et al.*), Gradient Boosting (Pantazopoulos *et al.*), Extra Trees (Guo *et al.*)., XGBoost (Guo *et al.*), and a Stacking Ensemble, are methodically evaluated.
- Improved feature engineering and ensemble learning: To increase prediction accuracy and guarantee consistent, trustworthy player assessments, sophisticated feature engineering approaches and ensemble structures are used.

The Stacking Ensemble technique provides improved accuracy in prediction by taking advantage of the complementary characteristics of all the learning algorithms used. While the Gradient Boosting and XGBoost algorithms are good at identifying complex relationships in the data, the Random Forest and Extra Trees algorithms are good at reducing variance with the bagging approach used in them. The ensemble technique benefits from the stability and flexibility of all the algorithms used in it, as the advantages of one algorithm compensate for the disadvantage of another. In addition, the ensemble technique reduces overfitting, a common problem with all boosting algorithms. Consequently, the Stacking Ensemble improves generalization and produces a more robust, balanced, and trustworthy model for predicting football player performance.

## 2.0 Related Work

As machine learning (ML) can analyze large amounts of data in a variety of domains, such as match outcome prediction (Wong *et al.*), injury risk (Musat *et al.*), player market value (Franceschi *et al.*), athlete development (Söker *et al.*), specific skill analysis (Leddy *et al.*), and player performance prediction (Taber *et al.*), an increasing number of studies are using ML techniques in football. As stated earlier in the section, estimating and

*DOI: 10.17492/JPI.DYPIU.052606* *ISBN: 978-93-49790-31-5*

forecasting the performance of footballers can be quite challenging and demands huge amounts of information. In order to allow the audience to appreciate the progress made in this field, some insights into the main uses of machine learning in football will be provided.

An example of how to predict match outcomes is given by Bunker *et al.* This study's authors looked at 380 matches between 20 Spanish Premier League professional football teams. Two machine learning algorithms were used in making predictions on three distinct football results: win for the home side, a draw, and win for the away side. Using 33 parameters, such as crossings, accelerations, sprint speeds, and long shots, the linear support vector machine's accuracy was 73% for home wins, 75% for draws, and 71% for away wins. Rather, the logistic regression model's accuracy was 73% for home wins, 72% for draws, and 74% for road wins (Bunker *et al.)*. These results imply that ML may be utilized for predicting match outcomes, even if they fail to achieve a very high accuracy.

Higher chances of winning a match are often linked to high-intensity training, which increases the risk of injury (da Silva *et al.*). For example, Venturelli et al. note that Several factors, such as height, size of the body, jump, and previous injuries, can increase the possibility of injury to the thighs among the young footballers. Injuries can obviously affect football players, as well as their respective teams and football organizations (Venturelli et al.). As such, prediction of injuries is important. Another example of using machine learning to predict football injuries is given by Rossi et al. For this study, physical activities of 26 professional footballers were monitored for the 2013-2014 championship season. The captured attributes could be classified into three groups: kinematic attributes – measures how much the athlete moves in his training; metabolic attributes – measures the energy consumption of the athlete in his training session; and mechanical attributes – measures the mechanical stress of the athlete's muscular system in his training session. For predicting the injury to the footballers, three different models of machine learning algorithms were adopted – Decision Tree, Random Forest, and Logistic Regression. While the Random Forest predicted the injuries in 87% of the cases, the Decision Tree predicted 80%. However, using a logistic regression model, only 60% prediction was possible (Rossi et al.). The scientists came to the conclusion that certain machine learning models outperform conventional state-of-the-art methods for assessing injury risk. If injuries are precisely anticipated, football teams and players may be more careful about injuries and football clubs can better manage their resources.

Indeed, in such instances, the team can adjust the workout schedule for the athlete so as to minimize the chances of being injured again. The athletes can be helped to seek the necessary medical care as well. This would help ensure that the injury does not get any worse and even make recovery easier. Another example of identifying recurrent motion patterns is given by Zhang *et al.* The authors of this study used k-Nearest Neighbors to examine football players' most frequent movements in certain regions of the field. This

DOI: 10.17492/JPI.DYPIU.052606 ISBN: 978-93-49790-31-5

method makes it possible to evaluate how the athlete moves toward opponents and teammates as well as whether or not the athlete can move across the field to change positions as needed. The analysis may then be utilized to help trainers ascertain if the taught patterns were applied during official games or to give information on the patterns employed by other teams (Zhang *et al.*). An inventive method to enhance optimization algorithms is to modify the WOA by utilizing the Euclidean distance. Since this modification is being used for the first time, a comparison between the gridsearch algorithm and the conventional WOA was conducted to examine the variations in the prediction accuracy of the three ways. Future research projects that employ this state-of-the-art optimization method might be evaluated using this study as a standard.

## 3.0 Materials and Methods

### 3.1 Data collection and preprocessing

The study's information came from publicly accessible football statistics on FBref, an online database that offers comprehensive player-level stats for professional football leagues throughout the globe. The pertinent player statistics for the 2024–2025 season were manually retrieved from FBref webpages related to Forward player because the Indian Super League (ISL) data is not immediately accessible through an API. Relevant statistics such as games played, minutes played, goals scored, assists made, passing accuracy, etc., are some of the key performance indicators that are collected from the relevant FBref webpages pertaining to the position of the player.A series of preparation steps are undertaken before the data is used with the help of the machine learning algorithm.

### 3.2 Data preprocessing

- *Data cleaning:* All the numeric columns are converted to the correct numerical format by handling the commas and missing values, and the unnecessary textual columns like nationality, club names, and birth year are removed (Chu et al**.**).
- *Feature engineering:* To better capture player performance, additional features are created: minutes per match (Nargesian *et al.*), weighted target (a combination of PPM and minutes per game), and total contributions (the sum of goals and assists).
- *Imputation and scaling:* To ensure that the input scales are consistent across all models, the numerical features (Wu *et al.*) are normalized using the StandardScaler, and the missing values are imputed using median imputation.
- *Label binarization:* In order to distinguish between the above-average performers and the below-average performers, the target variables (Qin et al.), including the performance and suitability, were binarized by using the median values.

While the data was collected manually in this version, the authors note that this may cause human error, which could negatively influence repeatability. Future research tries to

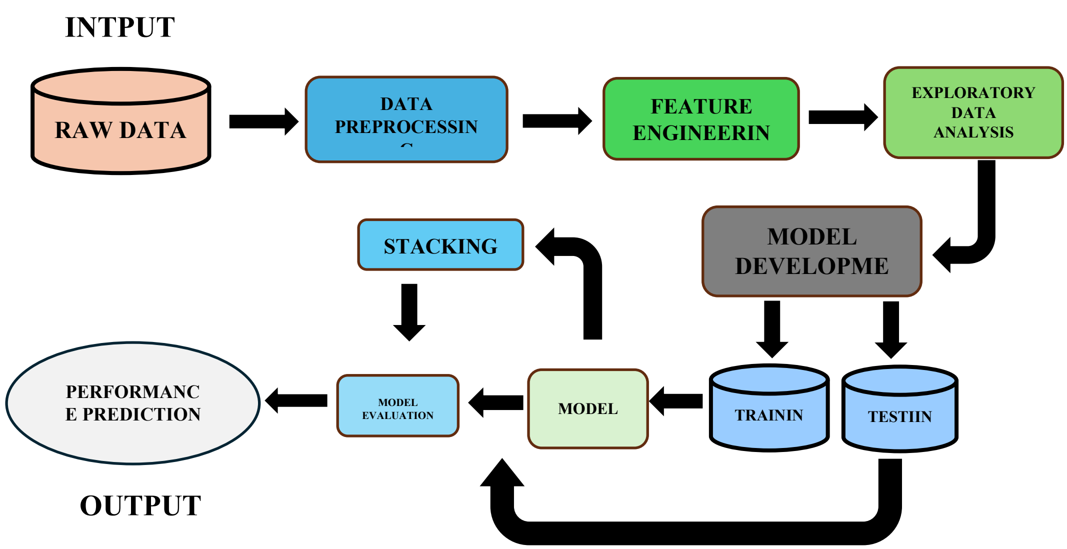
DOI: 10.17492/JPI.DYPIU.052606 ISBN: 978-93-49790-31-5

address this problem by developing a semi-automatic web scraping method that could automatically collect the data from FBref without any errors using Python tools like BeautifulSoup or Selenium.

### 3.3 Methods

To enhance the accuracy of the model, multiple ensemble classifiers are trained: Random Forest (Patel *et al.*), Gradient Boosting (Pantazopoulos A *et al.*), Extra Trees (Guo *et al.*)., XGBoost (Guo J *et al.*), These individual models are combined using a Stacking Classifier with XGBoost as the meta-learner. This is done to leverage the benefits of each model. Finally, the players are classified as "Suitable" and "Not Suitable" based on their goal contributions with respect to the median of the dataset.

**Figure 1: Proposed Framework for Football Player Performance Prediction**




### 3.4 Model architecture and contribution to the ensemble pipeline

This work employs four essential models of machine learning, namely Random Forest (Patel *et al.*), Gradient Boosting (Pantazopoulos A *et al.*), Extra Trees (Guo *et al.*), and XGBoost (Guo J *et al.*), each of which introduces special features to the ensemble pipeline that ensure the reliability and accuracy of the prediction of players' performance across different positions.

- *Random Forest (RF):* This is a model based on bagging, where many decision trees are created and their predictions are averaged (Patel *et al.*). This model is quite good at

DOI: 10.17492/JPI.DYPIU.052606 ISBN: 978-93-49790-31-5

dealing with noisy data and reducing variance. RF detects general trends in players' performance and provides robust predictions for the pipeline.

- *Gradient Boosting (GB):* This is an additive model that creates trees sequentially, where each tree aims to correct the errors of the previous trees (Pantazopoulos A *et al.*). GB is particularly good at detecting subtle connections and small patterns in the data. This model provides nuance and complexity to the ensemble.
- *Extra Trees (ET):* This algorithm is similar to Random Forest, but it has an extra level of randomness, as the threshold for splitting is selected randomly (Guo *et al.*). This results in greater generalization and minimizes the problem of overfitting. This algorithm adds diversity to the ensemble and makes it better for dealing with high-dimensional data.
- *XGBoost:* This is an efficient algorithm for gradient boosting, which uses parallel processing and regularization techniques (Guo J *et al.*). It is known for its efficiency and accuracy, especially when dealing with structured data. XGBoost is better at handling non-linear relationships and reducing error, thus acting as both a base learner and a meta-learner in the stacking ensemble.

The results of the aforementioned basic models are integrated by the stacking ensemble and input into the meta-learner (Liu *et al.*) XGBoost, which decides in the best way to combine the results of the models. The ensemble method can overcome the particular weaknesses of each model while leveraging the particular strengths of each model due to the tiered approach. For example, Gradient Boosting and XGBoost detect complex patterns, while Random Forest and Extra Trees reduce variability. The stacking model's higher performance metrics for all player positions prove that more accurate and balanced predictions are made.

### 3.5 Cross-validation

In the case of the defender and midfielder datasets, 5-fold cross-validation (Fushiki, T) was utilized during the training of the model. This ensured generalization and minimized the risk of overfitting. The average accuracy and F1 scores were utilized in assessing the model's consistency after each model was trained and tested. These predictions were utilized in the stacking ensemble method.

### 3.6 Evaluation metrices

Accuracy, Precision, Recall, and F1 Score (Yacouby, R., & Axman) were used as classification performance measures for evaluating the performance of the model. F1 Score was given more priority since it is a compromise between precision and recall, especially when there is a class imbalance.

*DOI: 10.17492/JPI.DYPIU.052606* *ISBN: 978-93-49790-31-5*

*Accuracy:* Accuracy refers to the proportion of predictions that the model gets right out of all the cases it evaluates (Yacouby, R., & Axman).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad ...1$$

For classification with confusion matrix terms (TP = true positives, TN = true negatives, FP = false positives, FN = false negatives)

*Precision:* Precision tells us how many of the cases the model labeled as positive were actually correct. It shows how trustworthy the model's positive predictions are (Yacouby, R., & Axman).

$$\text{Precision} = \frac{TP}{TP+FP} \quad ...2$$

*Recall:* Recall, or Sensitivity, shows how many of the actual positive cases the model successfully identifies. It reflects how well the model can detect the positive instances it needs to find (Yacouby, R., & Axman).

$$\text{Recall} = \frac{TP}{TP+FN} \quad ...3$$

*F1-Score:* The F1-Score is the harmonic mean of Precision and Recall, giving a single measure that balances both. It is especially helpful when the dataset is imbalanced or when it's important to consider both false positives and false negatives (Yacouby, R., & Axman).

$$\text{F1-Score} = 2 \times \frac{\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}} \quad ...4$$

## 4.0 Results and Discussion

### 4.1 Performance evaluation of individual models

Performance of machine learning algorithms including Random Forest (RF), Gradient Boosting (GB), Extra Trees (ET), XG Boost (XG B), and Stacking Ensemble was studied. In this context, these algorithms were assessed in terms of accuracy measures such as Accuracy (Yacouby, R., & Axman), Precision (Yacouby, R., & Axman), Recall (Yacouby, R., & Axman), and F1 Score (Yacouby, R., & Axman), since these measures provided an overall understanding of the performance of the algorithm. Although $R^2$ Score and Root Mean Squared Error have been used in this process, these measures are more suitable for regression than classification algorithms.

#### 4.1.1 Forward

The accuracy of the various models in forecasting the appropriateness of the forward players is presented in Table 1 below. The XGBoost model recorded the best accuracy of 0.9333 among the various models, coupled with the best F1-Score of 0.9151, recall of 0.9333, and precision of 0.9012, making it the best balance of sensitivity and

DOI: 10.17492/JPI.DYPIU.052606 ISBN: 978-93-49790-31-5

prediction accuracy. The Gradient Boosting model, with an F1-Score of 0.85, recorded the same accuracy of 0.8667 as the Stacking Ensemble model, making them the most reliable, as shown by the consistency of the two values. The Random Forest model, which recorded the best accuracy of 0.8667 and F1-Score of 0.8529, recorded the same values as the Gradient Boosting model. The Extra Trees model, however, recorded the lowest values of 0.8155 as the F1-Score and 0.8444 as the accuracy, making it the least.

Table 2 shows the suitability of the players for the stacking ensemble model where the player number is shown instead of the player's name.

**Table 1: Model Performance for Forward Player Suitability Prediction**

| Model | Accuracy (%) | F1-Score | Recall | Precision |
|---|---|---|---|---|
| Random Forest | 86.67 | 0.8529 | 0.8667 | 0.8444 |
| Gradient Boosting | 86.67 | 0.8559 | 0.8667 | 0.8456 |
| Extra Trees | 84.44 | 0.8155 | 0.8444 | 0.8067 |
| XGBoost | 93.33 | 0.9151 | 0.9333 | 0.9012 |
| Stacking Ensemble | 86.67 | 0.8593 | 0.8667 | 0.8556 |

**Table 2: Suitability of Players of Stacking Ensembling for Each Position**

| Player | Forward Suitability |
|---|---|
| 1 | Suitable |
| 2 | Suitable |
| 3 | Suitable |
| 4 | Suitable |
| 5 | Suitable |

### 4.2 Limitations and future work

The proposed paradigm also has some cons, despite the positive outcomes.

- Due to the depth and complexity of the stacking and the XGBoost models, there was a slight indication of overfitting.
- The small nature of the dataset.
- The existing methodology does not consider real-time or multi-seasonal performance data, which could make the models more robust over time; the manual extraction of the data from FBref also presents the possibility of human error.
- Future work will look to expand the dataset further and consider using hybrid models with the inclusion of deep learning techniques to make the models more flexible.

*DOI: 10.17492/JPI.DYPIU.052606* *ISBN: 978-93-49790-31-5*

## 5.0 Conclusion

Based on the findings, XGBoost and Stacking Ensemble models were found to be two of the best algorithms that can classify football players based on their appropriateness. The gradient boosting technique used by XGBoost, which maximizes bias and variance reduction through sequential learning, is responsible for the excellent results achieved by this algorithm. The ability of the Stacking Ensemble model to leverage the power of multiple learners has been found to make this model exceptionally resilient with regards to generalization.

In general, taking into account all factors involved, the findings prove the efficacy of ensemble and boosting models when dealing with the complexity of football statistics data. To include the temporal and spatial relationships in the game data, future research may expand on this model through applying deep learning techniques, including CNNs and RNNs. Deep learning techniques and ensemble models in hybrid models may improve their scalability, interpretability, and precision when forecasting several football tournaments and competitions. Such developments may help contemporary football employ a more comprehensive and data-driven approach to player recruitment and selection processes.

## References

1. Alexandre D, Da Silva CD, Hill-Haas S, Wong DP, Natali AJ, De Lima JR, Bara Filho MG, Marins JJ, Garcia ES, Karim C, et al. Heart rate monitoring in soccer: Interest and limits during competitive match play and training, practical application. *J Strength Cond Res*, 2012; 26(10):2890906. https://doi.org/10.4081/ejtm.2021.9692
2. Bunker R, Yeung C, Fujii K. Machine learning for soccer match result prediction. In: *Artificial Intelligence, Optimization, and Data Sciences in Sports*. Cham: Springer Nature Switzerland; 2024. p. 7–49. https://doi.org/10.48550/arXiv.2403.07669
3. Chu, X., Ilyas, I. F., Krishnan, S., & Wang, J. (2016, June). Data cleaning: Overview and emerging challenges. In *Proceedings of the 2016 international conference on management of data* (pp. 2201-2206). https://doi.org/10.1145/2882903.2912574
4. da Silva TM, Matos Esteves PT, Clemente FM, Ibáñez Godoy SJ. Examining the associations between weekly training load and match running demands in professional men football players: A descriptive study reflecting contextual variables. *Int J Sports Sci Coach*, 2025; 17479541251335224. https://doi.org/10.1177/17479541251335224
5. Dijkhuis TB, Kempe M, Lemmink KA. Early prediction of physical performance in elite soccer matches—a machine learning approach to support substitutions. *Entropy*, 2021; 23(8):952. https://doi.org/10.3390/e23080952

*DOI: 10.17492/JPI.DYPIU.052606* *ISBN: 978-93-49790-31-5*

6. Franceschi M, Brocard JF, Follert F, Gouguet JJ. Football players in light of economic value theory: Critical review and conceptualisation. *Manag Decis Econ*, 2024; 45(2):896–920. https://doi.org/10.1002/mde.4039
7. Fushiki, T. (2011). Estimation of prediction error by using K-fold cross-validation. *Statistics and Computing*, *21*(2), 137-146.
8. Gabbett TJ, Nassis GP, Oetter E, Pretorius J, Johnston N, Medina D, Rodas G, Myslinski T, Howells D, Beard A, et al. The athlete monitoring cycle: A practical guide to interpreting and applying training monitoring data. *Br J Sports Med*, 2017. https://doi.org/10.1136/bjsports-2016-097298
9. Guo J, Yang L, Bie R, Yu J, Gao Y, Shen Y, Kos A. An XGBoost-based physical fitness evaluation model using advanced feature selection and Bayesian hyper-parameter optimization for wearable running monitoring. *Comput Netw*, 2019; 151:166–180. https://doi.org/10.1016/j.comnet.2019.05.005
10. Indian Super League Stats | FBref.com. FBref.com. Available from: https://fbref.com/en/comps/82/Indian-Super-League-Stats
11. Iso-Ahola S. Intrapersonal and interpersonal factors in athletic performance. *Scand J Med Sci Sports*, 1995; 5(4):191–9. https://doi.org/10.1111/j.1600-0838.1995.tb00035.x
12. Leddy C, Bolger R, Byrne PJ, Kinsella S, Zambrano L. The application of machine and deep learning for technique and skill analysis in swing and team sport-specific movement: A systematic review. *Int J Comput Sci Sport*, 2024; 23(1):110–145. https://doi.org/10.2478/ijcss-2024-0007
13. Liu X, Zhang Z, Zhang X, De Meo P, Cherifi H. Multi-supervisor association network cold start recommendation based on meta-learning. *Expert Syst Appl*, 2025; 267:126204. https://doi.org/10.1016/j.eswa.2024.126204
14. Mackenzie B. *Performance evaluation tests*. London: Electr World plc; 2005; 24(25):57–158.
15. Morciano G, Zingoni A, Calabrò G. Optimization and comparison of machine learning algorithms for the prediction of the performance of football players. *Neural Comput Appl*, 2024; 36(31):19653–66. https://doi.org/10.1007/s00521-024-10260-9
16. Musat CL, Mereuta C, Nechita A, Tutunaru D, Voipan AE, Voipan D, et al. Diagnostic applications of AI in sports: A comprehensive review of injury risk prediction methods. *Diagnostics*, 2024; 14(22):2516. https://doi.org/10.3390/diagnostics14222516
17. Nargesian, F., Samulowitz, H., Khurana, U., Khalil, E. B., & Turaga, D. S. (2017, August). Learning feature engineering for classification. In *Ijcai* (Vol. 17, pp. 2529-2535). https://doi.org/10.24963/ijcai.2017/352
18. Pantazopoulos A, Maragoudakis M. Sports & nutrition data science using gradient boosting machines. In: *Proc 10th Hellenic Conf Artif Intell*; 2018. p. 1–7.

19. Patel D, Saxena A, Wang J. A machine learning-based wrapper method for feature selection. *Int J Data Warehous Min*, 2024; 20(1):1–33. https://doi.org/10.4018/IJDWM.352041
20. Qin, H., Zhang, M., Ding, Y., Li, A., Cai, Z., Liu, Z., ... & Liu, X. (2023, July). Bibench: Benchmarking and analyzing network binarization. In *International conference on machine learning* (pp. 28351-28388). PMLR. https://doi.org/10.48550/arXiv.2301.11233
21. Reinhardt L, Schwesig R, Lauenroth A, Schulze S, Kurz E. Enhanced sprint performance analysis in soccer: New insights from a GPS-based tracking system. *PLoS ONE*, 2019; 14(5):e0217782. https://doi.org/10.1371/journal.pone.0217782
22. Rossi A, Pappalardo L, Cintia P, Iaia FM, Fernández J, Medina D. Effective injury forecasting in soccer with GPS training data and machine learning. *PLoS ONE*, 2018; 13(7):e0201264. https://doi.org/10.48550/arXiv.1705.08079
23. Schulz SVW, Matits L, Schwarz E, Jerg A, Otte M, Wiedemann P, et al. Lower limb injuries and their association with anthropometrics—an observational study of 381 German elite youth football players. *BMC Sports Sci Med Rehabil*, 2025; 17(1):143. https://doi.org/10.1186/s13102-025-01190-7
24. Söker N, Schamerowski J, Zech A, Steib S, Rahlf L. Long-term athlete development in schools: A systematic review and narrative synthesis. *Front Sports Act Living*, 2025; 7:1706060. https://doi.org/10.3389/fspor.2025.1706060
25. Taber CB, Sharma S, Raval MS, Senbel S, Keefe A, Shah J, et al. A holistic approach to performance prediction in collegiate athletics: Player, team, and conference perspectives. *Sci Rep*, 2024; 14(1):1162. https://doi.org/10.1038/s41598-024-51658
26. Tichy W. Changing the game: Dr. Dave Schrader on sports analytics. *Ubiquity*, 2016; 2016(May):1–10. https://doi.org/10.1145/2933230
27. Wong A, Li E, Le H, Bhangu G, Bhatia S. A predictive analytics framework for forecasting soccer match outcomes using machine learning models. *Decis Anal J*, 2025; 14:100537. https://doi.org/10.1016/j.dajour.2024.100537
28. Wu, W., Jia, F., & Enders, C. (2015). A comparison of imputation strategies for ordinal missing data on Likert scale variables. *Multivariate behavioral research*, *50*(5), 484-503. https://doi.org/10.1080/00273171.2015.1022644
29. Yacouby, R., & Axman, D. (2020, November). Probabilistic extension of precision, recall, and f1 score for more thorough evaluation of classification models. In *Proceedings of the first workshop on evaluation and comparison of NLP systems* (pp. 79-91) https://doi.org/10.18653/v1/2020.eval4nlp-1.9
30. Zhang L, Ala'a R, Karim FK, Saidani T, Khishe M, Smerat A. Reconstruction of dance movements using evolved parallel recurrent and wavelet models. *J Franklin Inst*, 2025; 108276. https://doi.org/10.1016/j.jfranklin.2025.108276

DOI: 10.17492/JPI.DYPIU.052606 ISBN: 978-93-49790-31-5