



# Artificial Intelligence for Sports Talent Identification: A Comprehensive Review of Data Sources, Machine Learning Methods, and Decision-Support Deployment

Saba Shakil Patel, Prof. Bangar Abhishek Popat, Prof. Bhosale S.B.

Dept. of Computer Engineering  
Jaihind College of engineering, Pune, India

**Abstract.** Sports talent identification (TID) seeks to detect and select athletes with exceptional potential, but conventional practice relies heavily on coach and scout judgement and is exposed to relative age effects, differences in biological maturation and small-sample intuition. Artificial intelligence (AI), particularly machine learning (ML), is increasingly proposed as a decision-support tool that can integrate anthropometric, physical, technical, psychosocial and match-tracking data. This paper presents a focused narrative review of verified peer-reviewed studies, prioritising 2019–2026 publications, that apply AI to talent selection, performance prediction, playing-position classification, athlete profiling and scouting. The review examines three connected themes: data sources (field tests, match and tracking data, psychosocial measures, maturation indicators and video), learning methods (penalised regression, tree ensembles, support vector machines, neural networks and unsupervised clustering), and validation and deployment practice. The literature shows that models reach moderate to high internal discrimination, with area under the ROC curve (AUC) values reported from about 0.58 to 0.96 in a recent systematic review, while very high accuracies are usually reported for small or augmented datasets without external validation. Most studies are male-dominated, soccer-dominated and single-club, and few report calibration, fairness audits or prospective testing. The review contributes a comparative analysis of representative studies, a conceptual athlete-to-decision architecture that links multi-source data acquisition, age- and maturation-aware modelling, calibrated and explainable prediction, human-in-the-loop decision support and external validation, and a prioritised agenda of research gaps and future directions.

**Keywords:** Sports Talent Identification, Machine Learning, Artificial Intelligence, Athlete Selection, Explainable AI, Decision Support

## I. Introduction

Talent identification and development (TID) aims to detect, select and nurture athletes with the potential to reach elite performance. Sporting excellence is multifactorial and long term [1], and syntheses in football likewise stress its multidimensional determinants [2]. In practice, selection still depends heavily on coach and scout judgement, which is exposed to structural biases. Annual age-grouping produces relative age effects, with meta-analytic evidence showing a substantial over-representation of relatively older athletes [3], and differences in biological maturation can make early developers appear temporarily superior [4]. Predictions made early in adolescence also weaken as the follow-up period grows longer [5].



AI and ML methods can model high-dimensional, nonlinear relationships and combine heterogeneous data, and their use in sport has expanded rapidly [6], [7]. Existing syntheses address talent identification methodology broadly [8], computational models published up to 2020 [9], and ML in youth team sports [10]. A 2025 review in sport psychology adds that missing data, class imbalance for rare outcomes and the under-representation of psychological and coach-judgement information limit current models [11].

This review examines three strands together: the data sources used to describe athletes, the learning methods and validation practices applied to those data, and the way predictions are, or could be, turned into decisions. It asks how mature the evidence is for operational use and proposes a conceptual architecture that addresses the weaknesses identified.

## II. Review Methodology

A focused web search of publisher platforms, PubMed Central and institutional repositories was carried out in October 2026 with combinations of the terms talent identification, talent selection, athlete selection, scouting, machine learning, deep learning, explainable AI, playing-position classification and relative age effect, prioritising 2019–2026 publications. Studies were included when they reported sport-specific prediction, classification or profiling with a stated AI or ML method; reviews were used for context only. Reference metadata were cross-checked against publisher, DOI or repository records where available. Details of several primary studies were taken from their titles, abstracts or from the structured synthesis in [10], and the full texts were not independently re-read; this is flagged in Table I. Direct logged querying of Scopus and Web of Science was not performed, so the review is narrative rather than systematic and study counts are not reported.

## III. Data Sources for AI-Based Talent Identification

### 1. Anthropometric, Physical and Maturation Data

Field-test batteries remain the most common input. In a study of elite male youth soccer players with 13,876 athletes over seven seasons, gradient-boosted trees predicted selection versus de-selection, with speed, change of direction, countermovement jump, aerobic speed reserve and technical skill all contributing [12], [10]. Physical profiling is not always informative, however: in an elite Scottish academy followed over ten years, anthropometric and physical profiles did not predict professional contracts [13]. Because maturation confounds such tests [4], some work models it directly, for example explainable ML to predict age at peak height velocity in preadolescent team-sport athletes [14]. Smaller applications include an artificial neural network pilot for sport orientation from morphological and biodynamic parameters [15] and random forest and XGBoost models for volleyball performance from anthropometric data in female players [16].

### 2. Match, Tracking and Technical Data

Player monitoring with GPS and inertial sensors generates high-volume data that can complement traditional scouting [17]. Unsupervised ML has been used to group soccer players by match-specific running performance rather than nominal playing position [18], and clustering of shooting patterns has been used to examine early specialization in U14 versus professional basketball players [19]. Technical skill indicators have been used to classify playing position in elite junior Australian football, although some roles, such as rucks, were poorly captured [20], [10].

### 3. Psychosocial and Perceptual-Cognitive Data

Models that add non-physical variables show that selection is not explained by athletic performance alone. Psychosocial and physiological factors were combined to model selection to regional age-grade rugby union squads [21], and a multidisciplinary investigation at an English football academy used ML to study talent development [22]. The synthesis in [10] reports that psychological characteristics such as coping under pressure and emotional regulation emerged among the important predictors.



An exploratory ML study of selected and non-selected White British and British South Asian youth cricketers extends this to socio-cultural characteristics [23]. Relative age and socio-cultural effects were also found to influence selection outcomes [10], so the labels used to train models can themselves embed bias.

#### **4. Video and Scouting Tools**

Computer vision can reduce the manual cost of collecting match data: an automated system has been described that generates tactical performance statistics for individual soccer players from video [24]. Scouting-oriented tools include the YUVA-SQ cognitive scouting model [25] and WTDTool, a women's talent detection tool [26]. The synthesis in [10] describes the first as a descriptive case study without predictive validation and presents both as examples of automated data collection that may help resource-limited recruitment.

### **IV. Machine Learning and Deep Learning Methods**

#### **1. Classical Supervised Learning**

Supervised, non-deep models dominate. In the systematic review of 27 youth team-sport studies, random forests, gradient boosting, neural networks and penalised regression were the most frequent algorithms, and reported AUC values ranged from about 0.58 to 0.96 with calibration rarely reported [10]. The same review cites AUC below 0.70 in the large seven-season soccer study [12] as an example of accuracy under what practice would require [10]. A one-class support vector machine has been applied to soccer talent identification [27], with small samples and class imbalance noted as limits [10]. In Australian rules football, neural networks showed superior sensitivity to logistic regression when predicting draft outcome [28], in a study the review credits with prospective external testing [10].

#### **2. Deep Learning and Augmented Data**

Deep learning appears less often [10]. SCM-DL, a split-combine-merge model with feature selection for sport-school admission and branch allocation, reported accuracy above 97 % on a sample of 2,222 athletes, but without external validation and with imbalanced data [29], [10]. In a smaller case study, random forest and multilayer perceptron models exceeded 90 % accuracy for position assignment and lineup generation using 21 real players supplemented with synthetic data [30], [10]; the review notes that augmentation can give very high accuracy at the expense of validity, and that coach-rated inputs may overlap with the outcome [10]. Autoencoders with Gaussian mixtures have also been used to classify youth basketball fitness [31]. Apparent accuracy therefore rises with model complexity while the underlying data remain small.

#### **3. Unsupervised and Explainable Methods**

Unsupervised methods yield profiles rather than predictions [18], [19]. Explainability is applied unevenly: SHAP values were used in some studies [12], [14], simpler feature rankings in others, and many studies did not address interpretability [10]. The synthesis frames interpretability as a continuum from global feature importance, through instance-level attribution such as SHAP, to counterfactual reasoning [10].

#### **4. Validation Practice**

Small samples and inadequate validation inflate estimated performance, which makes nested cross-validation and strict separation of training and test data important [32]. Leakage, including reuse of the same individuals across folds, can overstate accuracy [33], and selection tasks are typically imbalanced, so evaluation metrics must be chosen with care [34]. Against these standards, few TID studies use nested or external validation; in the review of 27 studies, 16 (59 %) were rated at high risk of bias with the PROBAST tool [10], [35]. Structured reporting guidance such as TRIPOD+AI is available [36]. The review also distinguishes apparent validity, measured within the development sample, from transportable validity across clubs, seasons and cohorts [10].



## V. Deployment, Decision Support and Governance

In talent identification the central design question is less where inference runs than how a prediction enters a human decision. The synthesis in [10] positions ML as an exploratory decision-support aid with two pathways: an operational pathway of screening, monitoring and early flagging, and a discovery pathway that looks for interaction-based patterns. It recommends tiered protocols in which a high-sensitivity ML screen is followed by structured expert evaluation of context, adaptability and psychosocial maturity, and ongoing monitoring for drift, fairness and recalibration [10]. Wearable and video data can feed such systems [17], [24], but, as far as their abstracts and syntheses show, none of the representative studies in Table I evaluates the effect of a deployed system on selection decisions or athlete outcomes.

Governance is equally important because the data concern young people and the outcome is a high-stakes selection. A minimal fairness framework has been proposed that records birth quarter, maturation status, socio-economic background, sex and playing context, and audits model performance across these subgroups [10]. Human-machine interaction, trust in “black-box” models and the weight given to coach judgement are further open issues [11].

## VI. Comparative Analysis

Table I summarises representative primary studies. Performance values are not directly comparable because outcomes, datasets and validation schemes differ, and findings were drawn from titles, abstracts and the synthesis in [10] as described in Section II.

Table 1: Comparison of Representative AI-Based Sports Talent Identification Studies

| Author, year                 | Objective                                          | Data / method                                                                     | Major reported findings                                                                                 | Limitations                                       |
|------------------------------|----------------------------------------------------|-----------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------|---------------------------------------------------|
| Jauhiainen et al., 2019 [27] | Talent identification in soccer                    | One-class SVM                                                                     | Moderate to high predictive validity (per [10])                                                         | Small samples; class imbalance                    |
| Craig and Swinton, 2021 [13] | Predict professional contracts                     | Anthropometric and physical profiles; penalised regression; one academy, 10 years | Profiling did not predict contracts                                                                     | Weak model performance; single academy            |
| Abidin, 2021 [30]            | Position assignment and lineup                     | RF, MLP; coach-rated inputs                                                       | Over 90 % accuracy; close to coach decisions (per [10])                                                 | 21 real players plus synthetic data               |
| Owen et al., 2022 [21]       | Regional squad selection, rugby union              | ML on psychosocial and physiological data                                         | Psychological factors among important predictors (per [10])                                             | Regional age-grade setting; full text not re-read |
| Altmann et al., 2024 [12]    | Selection vs de-selection, elite male youth soccer | Gradient boosting; nested CV; SHAP; 13,876 athletes, 7 seasons                    | Speed, change of direction, jump, aerobic and technical measures contributed; AUC below 0.70 (per [10]) | No calibration or external testing (per [10])     |
| Retzepis et al., 2024 [14]   | Predict age at peak height velocity                | Explainable ML; preadolescent team-sport athletes                                 | Key predictive factors identified                                                                       | Small or single-institution data (per [10])       |
| Jennings et al., 2024 [28]   | Predict AFL draft outcome                          | Neural network vs logistic regression                                             | Neural network sensitivity superior                                                                     | Single sport; draft-outcome label                 |

|                             |                                              |                                               |                                    |                                                             |
|-----------------------------|----------------------------------------------|-----------------------------------------------|------------------------------------|-------------------------------------------------------------|
| Haan et al., 2025 [18]      | Categorise players by match running          | Unsupervised ML on match running data         | Groupings beyond nominal positions | Grouping, not outcome prediction                            |
| Abidin and Erdem, 2025 [29] | Sport-school admission and branch allocation | SCM-DL with feature selection; 2,222 athletes | Accuracy above 97 % (per [10])     | No external validation; imbalanced data; coach-rated inputs |

## VII. Proposed Conceptual Architecture

Fig. 1 presents a proposed conceptual architecture derived from the reviewed studies. Athletes are assessed within an academy, school or regional pathway (layer 1). Complementary data are collected because no single domain explains selection: field tests and anthropometry, match, GPS and video data, and psychosocial or cognitive measures [12], [17], [24], [21] (layer 2).

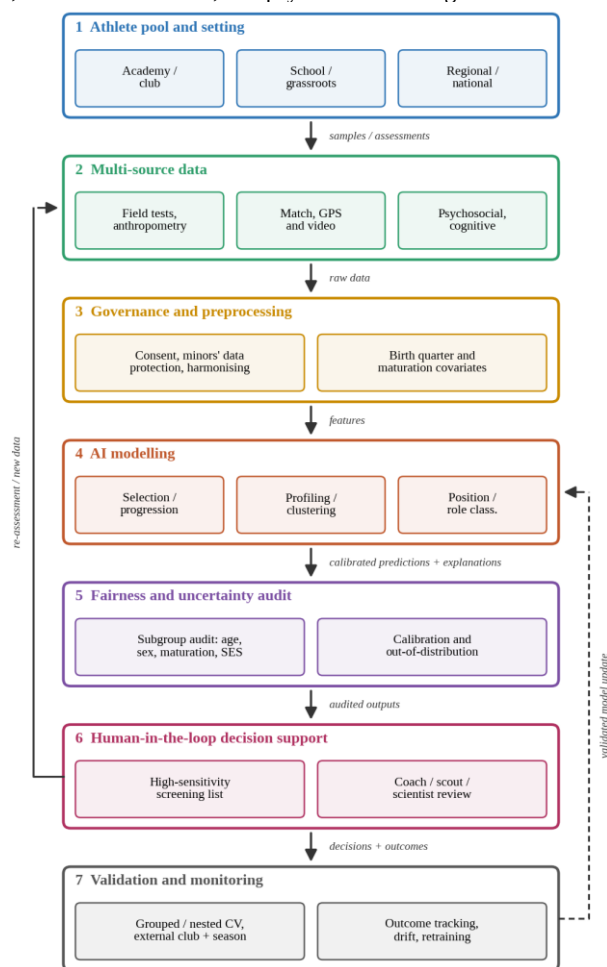


Fig. 1. Proposed conceptual architecture for AI-enabled sports talent identification from athlete data to decision (conceptual, not experimentally validated)



A governance and preprocessing layer secures consent and minors' data protection, harmonises measurements and records birth quarter and maturation status as covariates, so that relative age and maturation can be modelled or audited rather than ignored [3], [4], [10] (layer 3). Models then perform selection or progression prediction, profiling and position classification, and output calibrated probabilities with explanations such as SHAP [12], [14], [10] (layer 4). A fairness and uncertainty audit checks subgroup performance, calibration and out-of-distribution cases before results leave the system [10] (layer 5). Decision support presents a high-sensitivity screening list to coaches, scouts and sport scientists, who retain the decision [10], [11] (layer 6). Finally, grouped or nested validation, external testing on a different club and season, long-term outcome tracking and drift monitoring decide when a model may be updated [32], [33], [10] (layer 7), which addresses the lack of external validation highlighted in the literature.

## VIII. Research Gaps and Future Directions

Gaps supported by the reviewed evidence. (i) Small, male-dominated, single-club datasets: sample sizes in the synthesis range from 21 to 13,876 athletes, soccer accounts for 13 of 27 studies, and female athletes are under-represented [10]. (ii) Retrospective rather than prospective designs: outcome labels such as selection or draft position are imperfect proxies for later success [37], and predictability weakens over long follow-up [5]. (iii) Weak validation practice: nested or external testing is uncommon, calibration is rarely reported and leakage is a known risk [10], [33]. (iv) Bias in labels and inputs: relative age, maturation and socio-cultural effects shape who is selected, and coach-rated inputs may overlap with outcomes [3], [4], [23], [10]. (v) Augmented or imbalanced data inflate accuracy [30], [29]. (vi) Limited interpretability and uncertainty reporting [10].

Directions requiring investigation. Priorities include multi-club, multi-sport and longitudinal datasets with objective outcome labels; external validation across clubs, regions and seasons; reporting of grouped validation, calibration and subgroup fairness following TRIPOD+AI [36]; multimodal fusion of physical, tracking, video and psychosocial data; prospective trials that compare AI-assisted with unaided selection; human–AI interaction studies with coaches and scouts [11]; and data governance for shared datasets of young athletes.

## IX. Conclusion

AI-based talent identification has progressed from small classification case studies to large-scale selection models, explainable maturation prediction, unsupervised profiling and video-based scouting tools. Reported accuracies can be high, particularly in small or augmented datasets, but the evidence base is dominated by retrospective, male, soccer-centred and single-club studies with limited external validation and little calibration or fairness reporting. Current models are best treated as exploratory decision-support aids that complement, rather than replace, expert judgement. Reliable use will require larger and more diverse datasets, prospective and externally validated studies, age- and maturation-aware modelling, transparent reporting and continuous human oversight, as reflected in the proposed conceptual architecture.

## References

1. R. Vaeyens, M. Lenoir, A. M. Williams, and R. M. Philippaerts, "Talent identification and development programmes in sport," *Sports Med.*, vol. 38, no. 9, pp. 703–714, 2008, doi: 10.2165/00007256-200838090-00001.
2. H. Sarmento, M. T. Anguera, A. Pereira, and D. Araújo, "Talent identification and development in male football: A systematic review," *Sports Med.*, vol. 48, pp. 907–931, 2018, doi: 10.1007/s40279-017-0851-7.
3. S. Copley, J. Baker, N. Wattie, and J. McKenna, "Annual age-grouping and athlete development: A meta-analytical review of relative age effects in sport," *Sports Med.*, vol. 39, no. 3, pp. 235–256, 2009, doi: 10.2165/00007256-200939030-00005.



4. R. M. Malina, A. D. Rogol, S. P. Cumming, M. J. Coelho e Silva, and A. J. Figueiredo, "Biological maturation of youth athletes: Assessment and implications," *Br. J. Sports Med.*, vol. 49, pp. 852–859, 2015, doi: 10.1136/bjsports-2015-094623.
5. D. Deprez, M. Buchheit, J. Fransen, J. Pion, M. Lenoir, R. M. Philippaerts, and R. Vaeyens, "A longitudinal study investigating the stability of anthropometry and soccer-specific endurance in pubertal high-level youth soccer players," *J. Sports Sci. Med.*, vol. 14, pp. 418–426, 2015.
6. N. Chmait and H. Westerbeek, "Artificial intelligence and machine learning in sport research: An introduction for non-data scientists," *Front. Sports Act. Living*, vol. 3, Art. no. 682287, 2021, doi: 10.3389/fspor.2021.682287.
7. M. Rico-González, J. Pino-Ortega, A. Méndez, F. Clemente, and A. Baca, "Machine learning application in soccer: A systematic review," *Biol. Sport*, vol. 40, pp. 249–263, 2023, doi: 10.5114/biolSport.2023.112970.
8. S. Barraclough, K. Till, A. Kerr, and S. Emmonds, "Methodological approaches to talent identification in team sports: A narrative review," *Sports*, vol. 10, no. 6, Art. no. 81, 2022, doi: 10.3390/sports10060081.
9. N. J. Khan, G. Ahamad, M. Naseem, and S. S. Sohail, "Computational efficiency in sports talent identification – a systematic review," *Int. J. Appl. Decis. Sci.*, vol. 16, no. 3, 2023.
10. Q. Tang, X. Wei, and B. Tan, "The role of machine learning in talent identification for team sports: A systematic review," *J. Sports Sci. Med.*, vol. 25, no. 1, pp. 58–83, 2026, doi: 10.52082/jssm.2026.58.
11. F. Sartor, S. A. Rosito, M. Raimo, M. Beato, and M. Bertollo, "Artificial intelligence in sport psychology: Implications for the identification and development of talent," *Psychol. Sport Exerc.*, advance online publication, Sep. 15, 2025.
12. S. Altmann, L. Ruf, S. Thiem, T. Beckmann, O. Wohak, C. Romeike, and S. Härtel, "Prediction of talent selection in elite male youth soccer across 7 seasons: A machine-learning approach," *J. Sports Sci.*, pp. 1–14, 2024, doi: 10.1080/02640414.2024.2442850.
13. T. P. Craig and P. Swinton, "Anthropometric and physical performance profiling does not predict professional contracts awarded in an elite Scottish soccer academy over a 10-year period," *Eur. J. Sport Sci.*, vol. 21, pp. 1101–1110, 2021, doi: 10.1080/17461391.2020.1808079.
14. N.-O. Retzepis et al., "Identifying key factors for predicting the age at peak height velocity in preadolescent team sports athletes using explainable machine learning," *Sports*, vol. 12, no. 11, Art. no. 287, 2024, doi: 10.3390/sports12110287.
15. P. F. de Almeida-Neto et al., "Using artificial neural networks to help in the process of sports selection and orientation through morphological and biodynamic parameters: A pilot study," *Sport Sci. Health*, vol. 19, pp. 929–937, 2023, doi: 10.1007/s11332-022-00986-1.
16. S. Sanjaykumar et al., "Machine learning analysis for predicting performance in female volleyball players in India," *J. Hum. Sport Exerc.*, vol. 20, pp. 207–215, 2024, doi: 10.55860/cn2vdj44.
17. G. Ravé, U. Granacher, D. Boullousa, and A. C. Hackney, "How to use global positioning systems (GPS) data to monitor training load in the 'real world' of elite soccer," *Front. Physiol.*, vol. 11, Art. no. 944, 2020, doi: 10.3389/fphys.2020.00944.
18. M. Haan, S. de van der Zwaard, J. Sanders, P. J. Beek, and R. T. Jaspers, "Beyond playing positions: Categorizing soccer players based on match-specific running performance using machine learning," *J. Sports Sci. Med.*, pp. 565–577, 2025, doi: 10.52082/jssm.2025.565.
19. J. M. Contreras-García, J. Courel-Ibáñez, M. I. Piñar-López, and S. J. Ibáñez, "Early specialization in formative basketball: A machine learning analysis of shooting patterns in U14 and professional players," *J. Sports Sci.*, pp. 1–8, 2024, doi: 10.1080/02640414.2024.2436809.
20. C. T. Woods, J. Veale, J. Fransen, S. Robertson, and N. F. Collier, "Classification of playing position in elite junior Australian football using technical skill indicators," *J. Sports Sci.*, vol. 36, pp. 97–103, 2018, doi: 10.1080/02640414.2017.1282621.



21. J. Owen, R. Owen, J. Hughes, J. Leach, D. Anderson, and E. Jones, "Psychosocial and physiological factors affecting selection to regional age-grade rugby union squads: A machine learning approach," *Sports*, vol. 10, no. 3, Art. no. 35, 2022, doi: 10.3390/sports10030035.
22. A. L. Kelly, C. A. Williams, R. Cook, S. L. J. Sáiz, and M. R. Wilson, "A multidisciplinary investigation into the talent development processes at an English football academy: A machine learning approach," *Sports*, vol. 10, no. 10, Art. no. 159, 2022, doi: 10.3390/sports10100159.
23. T. Brown, R. Cook, L. A. Gough, I. Khawaja, A. B. T. McAuley, and A. L. Kelly, "Exploring the multidimensional characteristics of selected and non-selected White British and British South Asian youth cricketers: An exploratory machine learning approach," *Youth*, vol. 4, pp. 718–734, 2024, doi: 10.3390/youth4020048.
24. R. Theagarajan and B. Bhanu, "An automated system for generating tactical performance statistics for individual soccer players from videos," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, pp. 632–646, 2021, doi: 10.1109/TCSVT.2020.2982580.
25. S. Venkataraman, K. Sundharakumar, A. Bharathi Malakreddy, and S. Natarajan, "YUVA-SQ: A cognitive scouting model for the beautiful game," in *Proc. 5th Int. Conf. Innov. Trends Inf. Technol. (ICITIIT)*, 2024, pp. 1–6, doi: 10.1109/ICITIIT61487.2024.10580784.
26. J. López-De-Armentia, "WTDTool: Women's talent detection tool," in *Proc. IEEE Int. Workshop Sport, Technol. Res. (STAR)*, 2024, pp. 144–149, doi: 10.1109/STAR62027.2024.10635969.
27. S. Jauhiainen, S. Äyrämö, H. Forsman, and J.-P. Kauppi, "Talent identification in soccer using a one-class support vector machine," *Int. J. Comput. Sci. Sport*, vol. 18, pp. 125–136, 2019, doi: 10.2478/ijcss-2019-0021.
28. J. Jennings, J. C. Perrett, D. W. Wundersitz, C. J. Sullivan, S. D. Cousins, and M. I. Kingsley, "Predicting successful draft outcome in Australian rules football: Model sensitivity is superior in neural networks when compared to logistic regression," *PLOS ONE*, vol. 19, Art. no. e0298743, 2024, doi: 10.1371/journal.pone.0298743.
29. D. Abidin and M. G. Erdem, "SCM-DL: Split-combine-merge deep learning model integrated with feature selection in sports for talent identification," *IEEE Access*, vol. 13, pp. 71148–71172, 2025, doi: 10.1109/ACCESS.2025.3562551.
30. D. Abidin, "A case study on player selection and team formation in football with machine learning," *Turk. J. Elect. Eng. Comput. Sci.*, vol. 29, pp. 1672–1691, 2021, doi: 10.3906/elk-2005-27.
31. C. Ge, "Optimization study of a dynamic assessment model of physical fitness for youth basketball training," *Appl. Math. Nonlinear Sci.*, vol. 9, 2024, doi: 10.2478/amns-2024-3396.
32. A. Vabalas, E. Gowen, E. Poliakoff, and A. J. Casson, "Machine learning algorithm validation with a limited sample size," *PLOS ONE*, vol. 14, Art. no. e0224365, 2019, doi: 10.1371/journal.pone.0224365.
33. S. Kapoor and A. Narayanan, "Leakage and the reproducibility crisis in machine-learning-based science," *Patterns*, vol. 4, no. 9, Art. no. 100804, 2023, doi: 10.1016/j.patter.2023.100804.
34. E. Richardson, R. Trevizani, J. A. Greenbaum, H. Carter, M. Nielsen, and B. Peters, "The receiver operating characteristic curve accurately assesses imbalanced datasets," *Patterns*, vol. 5, Art. no. 100994, 2024, doi: 10.1016/j.patter.2024.100994.
35. R. F. Wolff et al., "PROBAST: A tool to assess the risk of bias and applicability of prediction model studies," *Ann. Intern. Med.*, vol. 170, pp. 51–58, 2019, doi: 10.7326/M18-1376.
36. G. S. Collins et al., "TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods," *BMJ*, Art. no. e078378, 2024, doi: 10.1136/bmj-2023-078378.
37. B. J. Gogos, P. Larkin, J. A. Z. Haycraft, N. F. Collier, and S. Robertson, "Combine performance, draft position and playing position are poor predictors of player career outcomes in the Australian Football League," *PLOS ONE*, vol. 15, Art. no. e0234400, 2020, doi: 10.1371/journal.pone.0234400.