Intelligent Classifiers for Football Player Performance Based on Machine Learning Models

Original Scientific Paper

Baydaa M. Merzah
Department of Computer Science,
College of Computer Science and Information Technology,
University of Anbar, Ramadi, Iraq
Al-Nahrain University, Baghdad, Iraq
bai21c1007@uoanbar.edu.iq

Muayad S. Croock
Department of Control and Systems
University of Technology,
Baghdad, Iraq
muayad.s.croock@uotechnology.edu.iq

Ahmed N. Rashid
Department of Computer Networks Systems,
College of Computer Science and Information Technology,
University of Anbar, Ramadi, Iraq
rashidisgr@uoanbar.edu.iq

Abstract *– The remarkable effectiveness of Machine Learning (ML) methodologies has led to a significant increase in their application across various academic domains, particularly in diverse sports sectors. Over the past decade, scholars have utilized Machine Learning (ML) algorithms in football for varied objectives, encompassing the analysis of football players' performances, injury prediction, market value forecasting, and action recognition. Nevertheless, there has been a scarcity of research addressing the evaluation of football players' performance, which is a noteworthy concern for coaches. Hence, the objective of this work is to categorize the performance of football players into active, normal, or weak based on activity features. This will be achieved through the utilization of the Performance Evaluation Machine Learning Model (PEMLM), employing two novel datasets that cover both training and match sessions. To attain this goal, seven machine learning methods are applied, namely Random Forest, Decision Tree, Logistic Regression, Support Vector Machine, Gaussian Naïve Bayes, Multi-Layer Perceptron, and K-Nearest Neighbor. The findings indicate that in the dataset corresponding to match sessions, the Decision Tree classifier attains the highest accuracy (100%) and the shortest test time. In contrast, the K-Nearest Neighbor demonstrates the best accuracy (96%) and a reasonable test time for the training dataset. These reported metrics underscore the reliability and validity of the proposed assessment approach in evaluating the performance of football players in online games. The results are verified and the models are assessed for overfitting through a k-fold cross-validation process.*

Keywords*: Dataset structuring, Football, Machine learning, Player performance*

1. INTRODUCTION

Machine Learning (ML) has emerged as a powerful catalyst, transforming various fields by effectively extracting valuable insights from extensive and complex datasets. Its significance goes beyond technological limitations, profoundly impacting a wide range of industries, including healthcare [1-3], wireless sensor networks [4, 5], sports [6-9], and various other domains [10-12]. In the realm of football, the applications of ML can be categorized into distinct groups, as depicted in Fig. 1.

The prevention and anticipation of injuries are extremely important in the sports industry, significantly affecting the financial stability of sports clubs and team performance. The absence of essential players from games and training sessions due to injury has a significant financial impact, costing the team a total of EUR 188 million annually. The financial burden incorporates various factors, encompassing expenses associated with player recuperation, efforts in rehabilitation, and the salaries of players [13]. Recent empirical studies have emphasized the effectiveness of ML techniques in the injury prediction domain. Additionally, these techniques have demonstrated excellent outcomes for predicting injuries in adult handball and football players [14].

Fig. 1. ML applications in football

Furthermore, ML has proven to be more sensitive in forecasting injuries among young football players who are at an elite level [15].

Furthermore, the game analysis data plays a pivotal role as an essential indicator in action recognition. This data empowers sports physiologists and coaches to enhance the monitoring, evaluation, and design of training plans [16]. As a result, sports analytics has received significant attention in the field of artificial intelligence (AI). Several sources demonstrate how the field of action recognition has significantly advanced due to the amount of precise data [17,18].

Conversely, a player's market value represents an estimation of the sum that a team could be able to get for the sale of their contract to another team. It bears substantial significance in the negotiations between football clubs and the agents representing players. The method used in [19] can provide an objective and quantitative way to estimate the transfer fees and salaries of players, which are usually determined by subjective and non-transparent expert judgments. The paper also analyzes the most important factors that affect the market value of players, such as age, position, skills, and potential. Consequently, in appraising players' market worth, it becomes imperative to take into account their specific skills contingent upon their positional role on the field. The work by [20] introduces an effective approach to forecast the market values of football players by leveraging the FIFA 20 dataset. After implementing the clustering according to the playing area, the model was trained and evaluated each cluster using the regression technique. This method is highly effective in identifying relevant characteristics and simplifies the determination of the market value for the player.

In addition, assessing individual and team performances helps to understand the strategic and successful approach of team sports and is consistent with a fundamental objective of sports science [21]. As a result, ML has become imperative to build professional teams [18, 22, 23].

This work aims to assess the physical performance of football players in training and real match sessions and to introduce intelligent classification methods using ML algorithms. The utilized algorithms are the state of art algorithms used in the sports field and they are commonly used in previous studies. The algorithms are Random Forest (RF) [24], Gaussian Naïve Bayes (GNB) [25], Decision Tree (DT) [26], Logistic Regression (LR) [27], Support Vector Machine (SVM) with linear and radial base function kernels [28], Multi-Layer Perceptron (MLP) [29], and the K-Nearest Neighbor (KNN) [30]. The model implementation used Anaconda 2022.10, Jupyter Notebook 1.0, Scikit-learn library 1.0.2, and Python 3.11. Two separate datasets were constructed: the initial set comprised activity features extracted from player footage during actual match sessions, and the second set employed sensor data to replicate player activity features during training sessions. These datasets were employed for model testing to identify the most effective one for assessing player performance and validating categorization choices. The classifiers demonstrated high accuracies, utilizing important measurable activity features.

The contributions of this work are:

- The proposed method uses ML techniques along with specialized datasets to classify the physical performance of football players during the training and match sessions.
- The proposed models present new datasets one for players metrics during the match session and the second for training session.
- The efficacy of the proposed approach has been evaluated based on specific parameters, including accuracy, precision, recall, and F1-score.

The remainder of the paper is structured as follows: Section 2 provides an overview of the literature, Section 3 outlines the methodology employed in this work, and Sections 4 and 5 present the findings and conclusions, respectively.

2. RELATED WORK

Football is now acknowledged to encompass a wealth of statistical information on seasons, games, clubs, and players. While traditionally considered the domain of specialists and analysts, sports organizations are increasingly recognizing the scientific insights concealed within their data. Accordingly, these organizations have adopted ML technologies to unlock this knowledge and maximize strategic advantages. ML demonstrates the diversity in handling various types of sports data, including match statistics, players' metrics, videos, and time series. The outcomes derived from these technologies contribute significantly to the capabilities of trainers and coaches, aiding in the prediction of match results and player injuries, performance assessment of players, identification of sports talents, recognition of match actions, and estimation of players' market value.

Several studies employing ML techniques to assess player performance have been identified in the domain of football. Notably, Tindaro et al. [31] conducted a study aiming to forecast the physical performance of elite football players, utilizing their anthropometric characteristics. To achieve this, the researchers developed a regression model that incorporated both upper and lower body features to forecast the physical prowess of these players. The investigation involved enrolling sixteen male soccer players under the age of 15, all belonging to the elite group, the study administered the Yo-Yo test to gather the requisite data for the proposed model. The study's findings underscored the significance of the selected characteristics as pivotal indicators of both sprint efficiency and aerobic fitness. It is worth noting, however, that the model's predictive capability exhibited a subpar level. This suggests that factors beyond anthropometric characteristics may play a role in predicting variations in performance among these athletes.

The primary objective of the PlayeRank data-driven framework, as developed and implemented by Luca et al. [32], is to assess and rank soccer players' performances using an extensive dataset of match events. The paper delves into the PlayeRank assessments, revealing significant patterns that illuminate the attributes of exceptional performances and the factors that differentiate elite players. The dataset employed in this study includes millions of match events for 18 major football leagues. Despite the numerous advantages of the paper, it acknowledges certain inherent limitations of the PlayeRank framework. Notably, factors such as tactical arrangements, opponent strength, match results, and other contextual aspects are conspicuously overlooked.

The objective outlined by Ahmet et al. [33] is to employ a ML model for utilizing football players' attributes and performance data to determine their market value. The study employs a dataset encompassing 18,000 players sourced from the Football Manager 2018 game, coupled with their corresponding values from transfermarkt.com. The researchers undertake various preprocessing techniques to appropriately condition the data for modeling purposes. Subsequently, they conduct a comparative analysis of different ML algorithms, including linear regression, decision tree, random forest, and artificial neural networks, to identify the most suitable one for the task at hand. The results indicate that the random forest algorithm surpasses its counterparts, demonstrating superior performance in both mean absolute error (MAE) and coefficient of determination (R2) metrics.

Bartosz et al. [22] explored the use of ML techniques to predict the success of player transfers in professional football. The study defines various parameters for player evaluation, including age, position, goals, assists, passes, and tackles. Three definitions of a successful transfer are proposed, based on player, team, and transfer fee performance. Employing Random Forest, Naive Bayes, and AdaBoost algorithms on data from Transfermarkt, the authors report promising results, suggesting ML's potential in team building and player transfer planning. They propose further development for application as a tool of professional utility for scouts specializing in football talent. However, the study's limitation lies in its reliance on data solely from the English Premier League 2018/2019 season, potentially restricting the generalizability of findings to other contexts and periods.

The primary objective of Mikael et al. [34] was to implement and compare various ML algorithms for the classification of professional goalkeepers' performance levels based on their technical data. The researchers utilize a dataset comprising 14,671 player-match observations from the elite divisions of England, Spain, Germany, and France. Three ML algorithms, namely Logistic Regression, Random Forest Classifier, and Gradient Boosting Classifier are applied in the study. Recursive feature elimination is employed to identify the most crucial features for the classification task. The results suggest that the ability of goalkeepers with their feet is more important than their ability with their hands to distinguish elite and sub-elite performance. In addition, the essential features for predicting goalkeepers' performance levels encompass factors such as passes received, successful passes, short distribution, and clean sheets. However, the lack of information on physical and psychological parameters was the main limitation of this work.

In her study, Didem [35] employs seven distinct ML algorithms to identify the optimal player combination for the U13 team at Altınordu Football Academy. The article integrates data derived from player training sessions and coach evaluations, supplementing the analysis with synthetically generated data to enhance classification accuracy. The findings of the study highlight the utility of ML algorithms in the player selection and team formation processes. Notably, the random forest algorithm emerges as particularly effective, achieving a 93.93% reliability in player selections. Furthermore, the lineup suggestions generated by these algorithms exhibit a remarkable 97.16% similarity to the coach's ideal team. These findings underscore the precision of ML algorithms in addressing player classification challenges. Incorporating additional input data, such as coach assessments and quantitative measurements detailing player talents and position-specific capabilities, could enhance the overall reliability of these results.

These studies offer valuable insights into possible areas for development and propose future avenues for sports-related research, making a crucial contribution to the field of football analytics. Table 1 presents a summary of the preceding research.

Table 1. Preceding research summary

Author	Aim	Limitations
Tindaro et al. [31],2021	To predict elite football players' physical performance	Required specific tests and requirements. It is not suitable for use during training or match sessions. Limited data samples
Luca et al. [32], 2019	To rank the players	The (ball-touches) represent one part of football match actions
Ahmet et al. [33], 2020	To assess the value of football player	-
Bartosz et al. [22], 2021	To predict the transfer success of the players	Data only concerned the male players/leagues.
Mikael et al. [34], 2021	To classify the performance levels	Absent data on physical/ psychological parameters
Didem [35], 2021	To find the best combination of players.	Quantitative data to represent players' skills was missed

3. METHOD: PERFORMANCE EVALUATION ML MODEL (PEMLM)

The assessment of players' performance holds paramount importance, whether during training to optimize future workloads or in matches to empower coaches in strategic decision-making and substitutions. Consequently, this work incorporates both training and match session datasets in the Performance Evaluation Machine Learning Model (PEMLM). Unlike prior works limited by the number of tested samples, this research overcomes such constraints by providing a sufficiently ample sample size, ensuring the generalizability of ML results. Notably, the datasets employed both male and female samples and incorporated quantitative metrics reflective of players' physical skills. An additional aspect often overlooked in previous studies is the consideration of the time necessary for conducting ML tests. In online applications, time emerges as a crucial factor, prompting this work to incorporate it as a key evaluation metric.

3. 1. GENERAL BLOCK DIAGRAM

Various supervised ML models are employed to evaluate the physical performance of football players utilizing recorded features extracted from both match and training sessions. The PEMLM system employs three classes: weak, normal, and active performance for both training and match datasets. The recorded datasets undergo clustering and labeling before the pre-processing step, responsible for dataset reformation to prepare it for training and testing. Subsequently, the datasets are partitioned into training and testing groups, with an 80% and 20% ratio, respectively. The test samples are then fed into the trained ML models to be classified into the three designated classes. Figure 2 provides the complete system as a block diagram.

3. 2. DATASETS STRUCTURING

The dataset in this work comprises two sources: the match session dataset (MSDS) and the training session dataset (TSDS), each containing 2040 samples. The records within these datasets represent the characteristics of an individual player. They serve as inputs to the proposed (ML) model to determine the performance level of the player. This classification is considered a guide for making decisions related to player replacements or game strategy adjustments during matches and for optimizing the training workload.

The datasets utilized in this work undergo various pre-processing techniques to make them suitable for use in the ML models, as illustrated in Figure 3. The two datasets are sourced from different origins, as detailed in the subsequent subsections. Initially, area-based clustering is employed for dataset labeling, utilizing K-Nearest Neighbors (KNN) where n=3 to group players occupying the same field areas (midfielders, forwarders, defenders). Each instance in the group is then associated with one performance level (active or normal or weak). The subsequent step involves removing outliers to identify and eliminate extreme values, reducing the match dataset to 1949 samples, while the number of samples in the training dataset remains unchanged. The normalization phase follows, employing quantile normalization, which proves effective in transforming features into a normal distribution format. Subsequently, stratified sampling ensures the representation of each subgroup in the final sample, enhancing precision and reliability in analyzing the entire population.

3.2.1. MSDS

The data for this work is sourced from actual match videos captured by tactical cameras. Six features are extracted from the adopted videos: ID, Gender, Area, CD, Sp, and AC, along with the Class. The methodology for obtaining the dataset is illustrated in Figure 4 [36]. The definitions of these features are as follows:

ID: Every player in the dataset has a distinctive code, which functions as an identifying record.

G: Using a binary gender classification approach, which assigns a value of 1 to male teams and a value of 0 to female teams.

Area: There are three main categories of playing positions on the field, denoted by numbers: forwards (1), defenders (2), and midfielders (3).

CD: This measure represents the estimated cross-distance traveled by every player while playing.

Sp: This feature gives information about each player's speed and agility.

AC: The activity count metric is represented by a numerical value, it reflects the player's level of engagement in ball touches during the match and interactions with other players.

The CSRT tracker was used in this work to implement the tracking module, utilizing the OpenCV Python library [37]. The KNN unsupervised learning algorithm accomplishes the dataset labeling process where k is set to 3. Accordingly, three discrete classes were generated: active performance denoted by Class 0; weak performance indicating Class 1; and normal performance signifying Class 2. Table 2 views the head from the total of 2040 collected samples of the MSDS. Additionally, the general statistics of the collected dataset are presented in Table 3, while position-based statistical metrics are illustrated in Table 4, where Max and Min represent the maximum and minimum values, respectively, Mean is the mean value for each feature, and STD is the standard deviation.

Table 2. MSDS samples

ID	G	Area	CD	Sp	AC	Class
1	0	2	0.855	5.263	23	2
2	0	2	0.85	5.232	26	0
3	0	2	0.794	4.891	32	0
4	0	2	0.711	4.377	4	1
5	0	3	0.631	3.888	0	1
6	0	3	0.678	4.177	0	1
7	0	3	0.627	3.86	28	0
8	0	3	0.52	3.206	0	1
9	0	1	0.623	3.835	16	2
10	0	1	0.75	4.619	18	2

Table 3. Original MSDS statistics

Statistics	CD	Sp	AC
Max	1.16	7.144	33
Min	0.032	0.2	0
Mean	0.556	3.425	17.62
STD	0.209	1.285	12.28

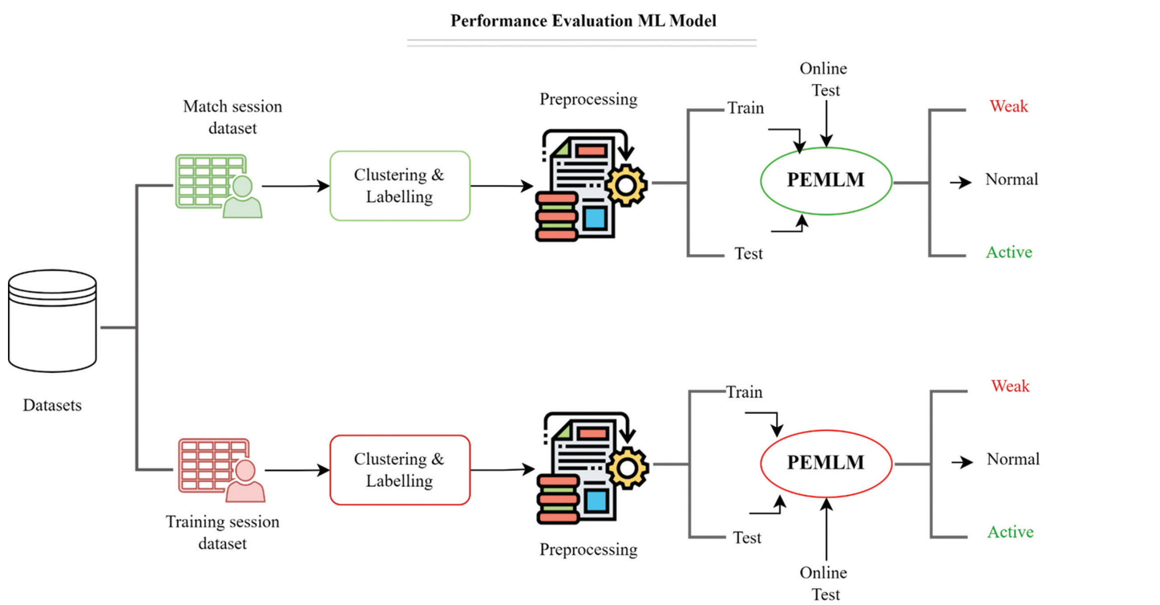

Fig. 2. General block diagram

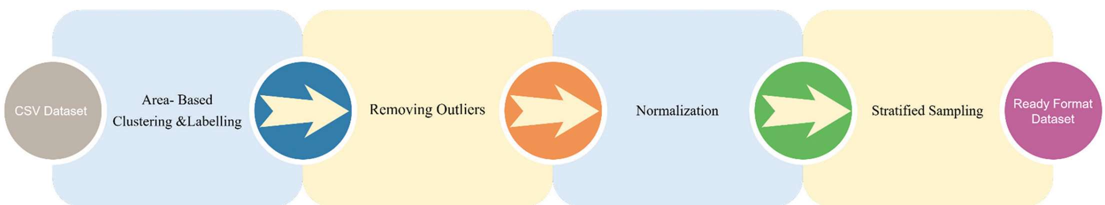

Fig. 3. Dataset pre-processing workflow

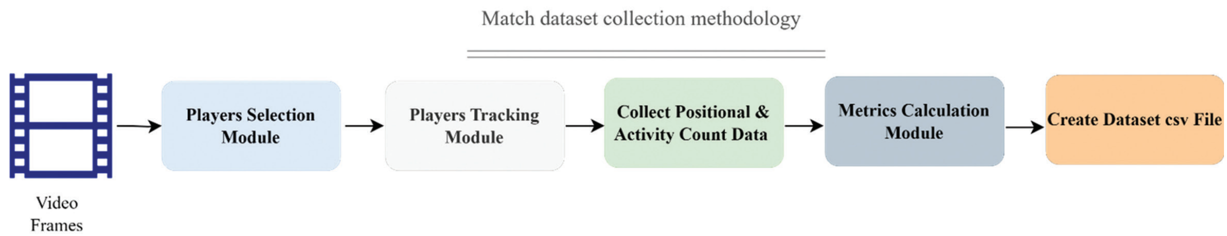

Fig. 4. MSDS collection methodology

Table 4. Area-based statistics

Statistics	CD	Sp	AC
Area 1 = forwarders			
Max	1.16	7.144	33
Min	0.081	0.5	0
Mean	0.556	3.426	16.824
STD	0.208	1.283	12.147
Area 2 = defenders			
Max	1.155	7.114	33
Min	0.032	0.2	0
Mean	0.553	3.409	19.614
STD	0.212	1.308	11.816
Area 3 = midfielders			
Max	1.159	7.135	33
Min	0.032	0.2	0
Mean	0.559	3.442	16.017
STD	0.205	1.261	12.52

3.2.2. TSDS

Precision in gathering data concerning the physiological attributes of soccer athletes holds paramount importance for coaches and trainers, enabling them to make well-informed decisions regarding team selection and the formulation of effective training strategies. Due to the high privacy of football players' data and the unavailability of the necessary data in the literature, synthetic data for 2040 samples are generated to simulate the real values. Tables 5, 6, and 7 showcase examples from the dataset, overall statistics, and playing region statistics, respectively. The MSDS included features such as ID, Gender, and Area, with additional features being incorporated, namely:

Heart Rate (HR): The heart rate is measured by bpm, which indicates the number of beats per minute. Football players' heart rates vary depending on fitness level, age, and playing position. In general, the resting heart rate is 40-60 (bpm), maximum heart rate is about 193.85 ± 5.2 [38], [39].

Oxygen Level (O2): Oxygen level is usually measured by %SpO2, which represents a percentage of oxygen-bound hemoglobin in the blood. Hemoglobin is a blood protein responsible for transporting oxygen to the body's tissues. The standard SpO2 range indicated by pulse oximeters is 95% to 100% [40].

Steps: The tactical effectiveness, work rate, and durability assessment of football players during the match is usually measured by their step count. This feature is used by the coaches to gain insights into the positioning, physical condition, and overall contribution to the game. Continuous monitoring enables the patterns' identification, helping optimize training strategies and playing. Maintaining an optimal step count indicates sustained performance and efficient field coverage. The count increases gradually from zero every ten seconds.

Energy: The vitality of football players significantly influences the team's on-field performance. Adequate energy levels are imperative for optimal running, decision-making, and passing, concurrently mitigating the risk of injuries. Coaches meticulously observe and manage players' energy expenditure, develop a customized training plan, and emphasize appropriate hydration and nutrition to optimize performance. Energy-effective management is essential to reach optimum performance in professional matches. It is important to note that each athlete's energy needs mainly depend on physical activity [41].

Table 5. TSDS sample

ID	G	Area	HR	O2	Steps	Energy	Class
1	0	2	76	92	3	90	1
2	0	2	78	90	14	96	1
3	0	2	76	93	10	92	1
4	0	2	74	92	8	93	1
5	0	3	80	94	8	93	0
6	0	3	74	91	12	91	0
7	0	3	85	80	28	75	2
8	0	3	78	91	18	91	0
9	0	1	96	84	38	84	2
10	0	1	102	80	70	61	0

Table 6. TSDS statistics

Statistics	HR	O2	Steps	Energy
Max	113	94	106	98
Min	73	80	2	48
Mean	88.335	85.313	39.63	79.162
STD	8.919	4.605	22.658	11.327

Table 7. Area-based statistics

Statistics	HR	O2	Steps	Energy
Area 1 = forwarders				
Max	109	94	106	98
Min	74	80	4	50
Mean	88.863	85.196	42.637	79.302
STD	8.76	4.559	23.346	11.368
Area 2 = defenders				
Max	110	94	106	98
Min	73	80	2	48
Mean	87.110	85.191	36.933	79.244
STD	8.677	4.582	22.443	11.16
Area 3 = midfielders				
Max	113	94	100	98
Min	74	80	4	48
Mean	89.295	85.494	40.824	79.011
STD	9.093	4.645	22.226	11.469

3.3. PEMLM ALGORITHM

The PEMLM algorithm provides a comprehensive procedural overview of the proposed models, with a detailed breakdown of the steps elucidated below:

1. Data Pre-processing: The utilized datasets undergo pre-processing, as detailed in Section 3.2.
2. Testing Module: This step involves determining the training or match module.
3. Dataset Selection Module: The type (Training or Match) of the dataset is selected based on the cho-

sen testing module, contingent on the evaluation period.
4. ML Model Training: The ML models are trained on the mentioned datasets to identify the right model for each scenario.
5. Model Testing: All employed algorithms undergo testing using the test sets and validation through a 5-fold validation.
6. Output: Three performance levels are the output of the model (active, normal, and weak).

Recommendations are then adopted based on the session state. For weak performance during a training session, it is suggested to intensify the training workload for the evaluated player. Conversely, in a match session, replacement is recommended. Normal results suggest maintaining or altering the workload, depending on whether it is a training or match session, respectively. Active players are advised to continue playing and training at the existing level.

PEMLM algorithm: Pseudo Code

1. Data Pre-processing
2. Testing module (Training or Match) session
3. Dataset selection module
4. Model Training
5. Model Testing
6. Output

IF (MSDS): THEN
- **IF** (Weak Performance): THEN Replacement Recommended
- **ELSE IF** (Normal Performance):
  - **THEN** Playing Strategy Change Recommended
- **ELSE** (Active Performance): THEN Continue

IF (TSDS) THEN
- **IF** (Weak Performance):
  - **THEN** Enhance Workload Recommended
- **ELSE IF** (Normal Performance):
  - **THEN** Training Workload Justification Recommended
- **ELSE** (Active Performance): THEN Continue

7. END

4. RESULTS

In the preceding section, we applied the PEMLM models to assess the performance of football players using both MSDS and TSDS datasets. The obtained results are categorized into two primary groups based on the dataset type. We employed five metrics to assess the outcomes: accuracy, precision, recall, F1-score, and test time. These metrics are defined by equations (1-4), wherein TP denotes true positives, FP signifies false positives, TN represents true negatives, and FN stands for false negatives. The time denotes the total seconds required to evaluate the model's performance on the test set.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$precesion = \frac{TP}{TP + FP} \tag{2}$$

$$recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1 - score = 2.\frac{precision\,.recall}{precision + recall} \tag{4}$$

4.1. MSDS RESULTS

The analysis of various ML models applied to MSDS reveals noteworthy performance metrics. The DT model in Table. 8 stands out with impeccable scores of 100% in accuracy, precision, recall, and F1-score, indicating precise and accurate predictions, coupled with an efficient classification time of 0.0032 seconds.

Table 8. MSDS results

Model	Acc	Pre	Rec	F1	Time
DT	100	100	100	100	0.0032 s
RF	99	99	99	99	0.0167 s
LR	97	97	97	97	0.0016 s
GNB	87	89	87	86	0.0022 s
SVM (Linear)	97	97	97	97	0.0022 s
SVM (RBF)	97	97	97	97	0.0019 s
KNN	95	95	95	95	0.0046s
MLP	99	99	99	99	0.0026 s

The RF model exhibits robust performance, securing a commendable 99% across all metrics, albeit with a slightly longer classification time of 0.0167 seconds due to its ensemble approach. LR delivers solid performance with a 97% score across all metrics, accompanied by a swift classification time of 0.0016 seconds. GNB shows respectable performance, although it falls short in comparison, particularly in accuracy and recall, achieving 86%- 89% in these metrics, respectively, with a classification time of 0.0022 seconds. The linear and (RBF) SVM both achieve an impressive 97% in all measures; the linear SVM takes 0.0022 seconds, while the RBF SVM takes 0.0019 seconds, which is a little faster. KNN performs admirably, scoring 95% on every metric, however, it takes 0.0046 seconds longer to classify data than other algorithms. The MLP model has a reasonable classification time of 0.0026 seconds and scores 99% on accuracy, precision, recall, and F1-score, matching the top-performing models across all criteria. Finally, the models exhibit varied levels of performance, with DT, RF, LR, and MLP emerging as top performers, stressing the significance of evaluating both accuracy and time based on online application needs. The confusion matrix for the highest accuracy model is shown in Fig. 5.

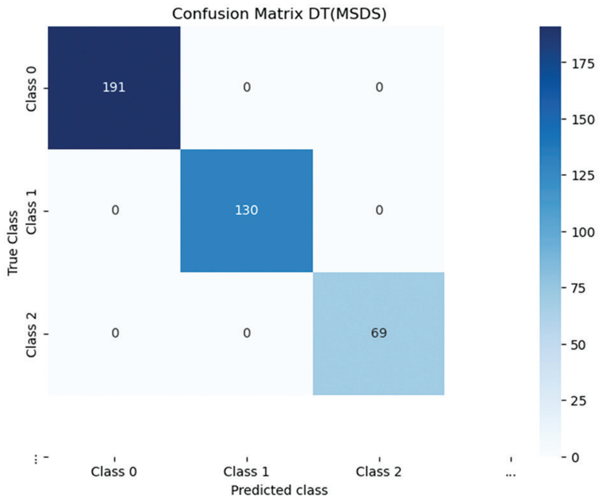

Fig. 5. Confusion matrix for DT model

4.2. TSDS RESULTS

The examination of the performance metrics for diverse ML models applied to the training session dataset reveals distinctive outcomes. DT, RF, GNB, and MLP demonstrate a relatively consistent accuracy range of 68-71%, reflecting balanced performance across metrics such as precision, recall, and F1-score. These models exhibit a comparable processing time, with DT and RF requiring 0.0029 seconds and 0.0272 seconds, respectively, showcasing efficiency in their computations. In contrast, LR and (SVM) (both Linear and RBF) display lower accuracy ranging from 50-58%, suggesting less optimal performance in classification tasks. KNN, however, stands out with an impressive accuracy of 96%, indicating strong predictive capabilities. KNN is good when the basic patterns in the data are clearly represented by the distances between samples. These results underscore the importance of considering both accuracy and computational time, with certain models showcasing more balanced performance across the evaluated metrics. The KNN and MLP are preferred for online applications. The classification performance metrics for the proposed models are listed in Table 9 and the confusion matrix for the best-fit model is presented in Fig. 6.

Table 9. TSDS results

Model	Acc	Pre	Rec	F1	Time
DT	68	68	68	68	0.0029 s
RF	68	68	68	68	0.0272 s
LR	50	47	50	48	0.0011 s
GNB	70	69	70	70	0.0025 s
SVM (Linear)	58	55	58	54	0.00167 s
SVM (RBF)	50	47	50	41	0.0015 s
KNN	96	96	96	96	0.0033 s
MLP	71	70	71	70	0.0037 s

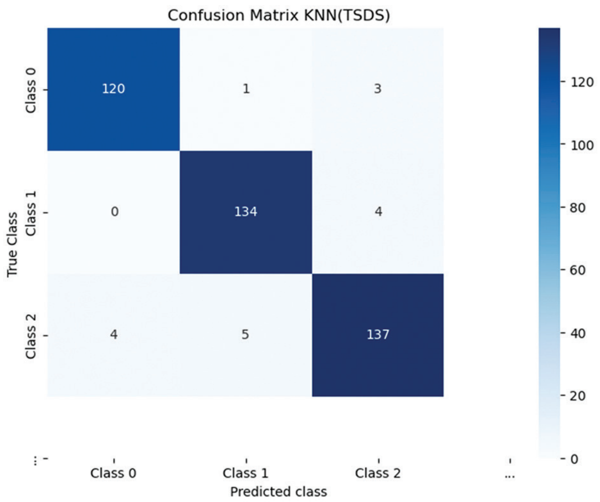

Fig. 6. Confusion matrix for KNN model

4.3. VALIDATION AND COMPARISONS WITH PREVIOUS WORK

The k-fold cross-validation process is used to validate the results and ensure that the models don't have an overfitting problem. The outcomes as presented in Table 10 for both datasets, reveal consistent accuracy, precision, recall, and F1-score across different models. Specifically, for the MSDS, both DT and MLP exhibit identical accuracy levels. Conversely, when assessed on the testing data within the training dataset, KNN demonstrates the best performance in correct classification compared to the other models. The overfitting problem is not observed in all models. The 5-fold performance for all models is illustrated in Figures 7 and 8 for the MSDS, while Figures 9,10, and 11 presented the models' performance validation of the training dataset.

The proposed models operate on both training and match sessions, providing a comprehensive assessment and classification of football player performance. Focusing on match session data, we scrutinized four crucial physical activity metrics: playing position, total distance covered during play, speed during play, and a novel metric gauging player activities during a match. In the case of the training session dataset, synthetic data was employed to simulate the authentic metrics due to the unavailability of an actual dataset, driven by privacy concerns. The selected features have been substantiated in the existing literature as reliable indicators for gauging the physical performance of players. These datasets are utilized in supervised ML models to evaluate players' physical performance.

Table 10. 5-fold cross-validation results

	Model	Avg Acc	Avg Pre	Avg Rec	Avg F1	Avg Time
Match session dataset	DT	100	100	100	100	0.0032 s
	MLP	100	100	100	100	0.0026 s
	GNB	68	66	68	67	0.0024 s
Training session dataset	MLP	71	70	71	71	0.0037 s
	KNN	94	94	94	94	0.0033 s

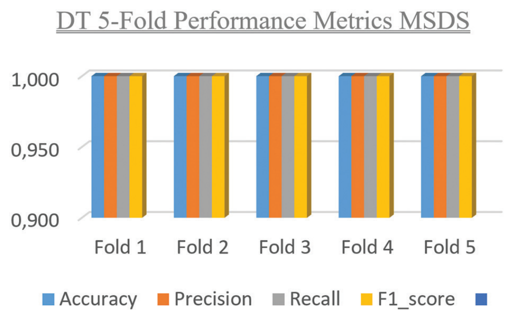

Fig. 7. DT 5-fold cross-validation (MSDS)

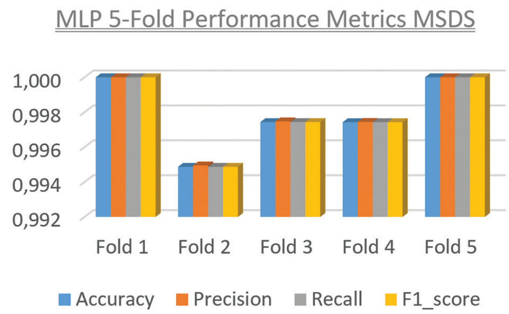

Fig. 8. MLP 5-fold cross-validation (MSDS)

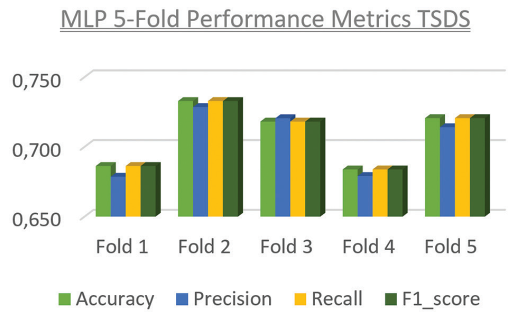

Fig. 9. MLP 5-fold cross-validation (TSDS)

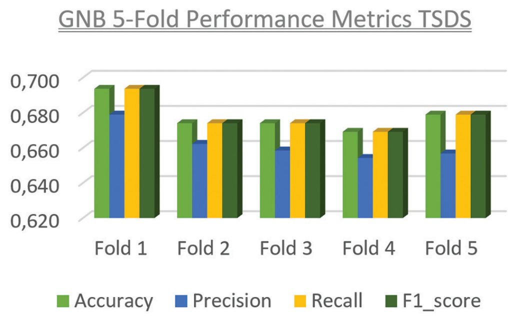

Fig. 10. GNB 5-fold cross-validation (TSDS)

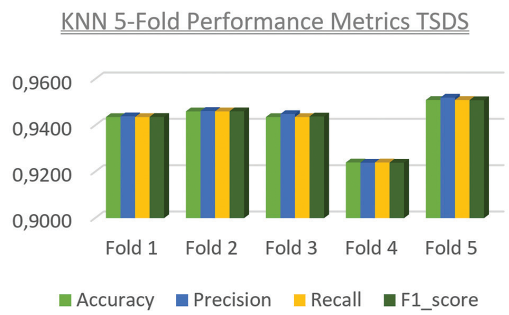

Fig. 11. KNN 5-fold cross-validation (TSDS)

The findings underscore the resilience and effectiveness of the proposed methodology in terms of time efficiency, accuracy, and suitability for integration into online applications. In contrast to the studies presented by [22], [24-27], the PEMLM handles substantial datasets derived from both training and match sessions. The inclusivity of a diverse set of studied features ensures a robust evaluation of players, distinguishing it from the approach in [25], where the prediction of physical performance relies on ball-touch descriptions. Notably, the introduced ML models account for dataset samples from both male and female players, offering a more comprehensive perspective compared to prior studies that exclusively analyzed male samples. Additionally, the considered features in PEMLM encompass the physical skills of players, addressing a gap apparent in the works of [26] and [27].

5. CONCLUSION

In this paper, we present an innovative approach to constructing datasets for the evaluation of football player performance in training and match sessions. The method leverages various ML models and features to train these models effectively. The proposed methodology undergoes a comprehensive evaluation, considering key metrics such as accuracy, precision, recall, and F1 score. Additionally, the validation process employs a robust 5-fold cross-validation technique. Our experimental results reveal that the proposed DT model attains an exceptional classification accuracy of 100% when applied to match session datasets. Conversely, the KNN model demonstrates superior performance when applied to training session datasets. We are confident that the implementation of these techniques will prove invaluable to coaches and trainers, streamlining the task of evaluating physical player performance. This, in turn, will enhance strategic decision-making during matches and optimize training session workloads. Consequently, both models exhibit potential for integration into online systems. We advocate exploring the feasibility of implementing the proposed PEMLM for real-time predictions during live matches. It is crucial to acknowledge and address the challenges associated with real-time data processing and model deployment, paving the way for future research endeavors. Additionally, a more nuanced analysis of gender-based performance patterns could reveal potential differences in play styles or strategies between them.

6. REFERENCES:

[1] L. Abdullah, A. Mazin, N. Ahmed, K. Seifedine, H. Karrar, N. Jan, M. Radek, R. Imran, "Restricted Boltzmann machine assisted secure serverless edge system for internet of medical things", IEEE Journal of Biomedical and Health Informatics, Vol. 27, No. 2, 2022, pp. 673-683.

[2] A. M. Abdulrahman, A. Belal, A.M. Mazin, "Breast cancer images Classification using a new transfer

learning technique", Iraqi Journal for Computer Science and Mathematics, Vol. 4, No. 1, 2023, pp. 167-180.

[3] A. M. Mazin, L. Abdullah, A. Z. Dilovan, Karrar A. Hameed, N. Jan, M. Radek, T. Usman, A. Majed, T. Prayag, "Adaptive secure malware efficient machine learning algorithm for healthcare data", CAAI Transactions on Intelligence Technology, 2023. (in press)

[4] H. S. Sama, S. C. Muayad, "Increasing WSN Lifetime using Clustering and Fault Tolerance Methods", Iraqi Journal for Electrical and Electronic Engineering, Vol. 17, No. 1, 2021, pp. 94-99.

[5] H. A. Sahar, N. Ahmed, "Prediction of Single Object Tracking Based on Learning Approach in Wireless Sensor Networks", Proceedings of the 14th International Conference on Developments in eSystems Engineering, Sharjah, United Arab Emirates, 7-10 December 2021, pp. 352-357.

[6] J. Rong, "Research on Basketball Shooting Action Based on Image Feature Extraction and Machine Learning", IEEE Access, Vol. 8, 2020, pp. 138743-138751.

[7] G. Brandon, P. Peter, K. Stephanie, R. Machar, "Differentiating movement styles in professional tennis: A machine learning and hierarchical clustering approach", European Journal of Sport Science, Vol. 23, No.1, 2021, pp. 44-53.

[8] O. Musa, T. Cevdet, S. Boran, A. Caner, U. Y. Hasan, "Performance Prediction and Evaluation in Female Handball Players Using Machine Learning Models", IEEE Access, Vol 8, 2020, pp. 116321-116335.

[9] B. Daniel, L. Phillippe, D. Werner, "Incorporating domain knowledge in machine learning for soccer outcome prediction", Machine Learning, Vol. 108, No. 1, 2019, pp. 97-126.

[10] T. A. Saja, A. Rafah, S. C. Muayad, "Enhancement of student performance prediction using modified K-nearest neighbor", TELKOMNIKA, Vol. 18, No. 4, 2020, pp. 1777-1783.

[11] A. U. Masar, S. C. Muayad, "Multiclassification of license plate based on deep convolution neural networks", International Journal of Electrical and Computer Engineering, Vol. 11, No. 6, 2021, pp. 5266-5276.

[12] K. S. Abrar, N. R. Ahmed, "Applications of machine learning for earthquake prediction: A review", Proceedings of the Al-Kadhum 2nd International Conference on Modern Applications of Information and Communication Technology, Baghdad, Iraq, 8-9 December 2021.

[13] R. Alessio, P. Luca, "A Narrative Review for a Machine Learning Application in Sports: An Example Based on Injury Forecasting in Soccer", Sports, Vol. 10(1), No. 5, 2022, pp. 1-16.

[14] L. O. Jon, A. Francisco, D. S. C. Mark, L. S. Rhodri, D. M. Greg, J. R. Paul, "Using machine learning to improve our understanding of injury risk and prediction in elite male youth football players", Journal of Science and Medicine in Sport, Vol. 23, No. 11, 2020, pp. 1044-1048.

[15] V. Emmanuel, S. Nicolas, I. Abdelhak, M. Jacky, P. Stephane, "Combining Internal- and External-Training-Loads to Predict Non-Contact Injuries in Soccer", Applied Sciences, Vol. 10, No. 15, 2020, p. 5261.

[16] G. Abraham, M. Moisés, C. Javier, R. Asier, L. Sergio, "In-game behavior analysis of football players using machine learning techniques based on player statistics", International Journal of Sports Science & Coaching, Vol. 16, No. 1, 2021, pp. 148-157.

[17] G. Yaparla, S. T. Allaparthi, K. M. Sai, R. M. Garimella, "A Novel Framework for Fine-Grained Action Recognition in Soccer", Proceedings of the International Work-Conference on Artificial Neural Networks, Gran Canaria, Spain, 12-14 June 2019, pp. 137-150.

[18] L. Guiliang, L. Yudong, S. Oliver, K. Tarak, "Deep soccer analytics: learning an action-value function for evaluating soccer players", Data Mining and Knowledge Discovery, Vol. 34, 2020, pp. 1531-1559.

[19] A. A. Mustafa, T. Sakir, "Predict the Value of Football Players Using FIFA Video Game Data and Machine Learning Techniques", IEEE Access, Vol. 10, 2022, pp. 22631-22645.

[20] B. Iman, M. R. Seyed, "A novel machine learning method for estimating football players' value in the transfer market", Soft Computing, Vol. 3, No. 1, 2020, pp. 2499-2511.

[21] K. Shitanshu, S. Sergiy, Y. Zhu, G. Paul, A. Maya, "Machine Learning Enabled Team Performance

Analysis in the Dynamical Environment of Soccer", IEEE Access, Vol. 8, 2020, pp. 90266-90279.

[22] Ć. Bartosz, G. Agata, C. Michal, "Who Will Score? A Machine Learning Approach to Supporting Football Team Building and Transfers", Entropy, Vol. 23, No. 1, 2021, p. 90.

[23] L. G. B. Tom, G. P. F. Wouter, M. N. A. Susan, R. M. Rob, J. R. Ruud, "How soccer scouts identify talented players", European Journal of Sport Science, Vol. 22, No. 7, 2022, pp. 994-1004.

[24] K. H. Tin, "Random Decision Forests", Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, Canada, 14-16 August 1995, pp. 278-282.

[25] B. Thomas, "An Essay Towards solving problem in the doctrine of chances", Royal. Society, Vol. 53, 1997, pp. 370-418.

[26] R. Q. John, "Induction of Decision Trees", Machine Learning, Vol. 1, 1986, pp. 81-106.

[27] A. Žilinskas, "Reviewed Work(s): Practical Mathematical Optimization: An Introduction to Basic Optimization Theory and Classical and New Gradient-Based Algorithms", INFORMS, Vol. 36, No. 6, 2006, pp. 613-615.

[28] C. Corinna, V. Vladimir, "Support-Vector Networks", Machine Learning, Vol. 297, 1995, pp. 273-297.

[29] M. T. Hagan, "Neural network design, Second Edition", 2nd Edition, PWS Publishing Co., 2003.

[30] G. Jacob, E. H. Geoffrey, R. Sam, R. S. Russ, "Neighborhood Components Analysis", Proceedings of the 17th International Conference on Neural Information Processing Systems, Vancouver, BC, Canada, 1 December 2004, pp. 513-520.

[31] B. Tindaro, T. Athos, C. Luca, R. Alessio, P. Enrico, P. Giulio, F. M. Laia, A. Giampietro, "Importance of anthropometric features to predict physical performance in elite youth soccer: a machine learning approach", Research in Sports Medicine, Vol. 29, No. 3, 2021, pp. 213-224.

[32] P. Luca, C. Paolo, F. Paolo, M. Emanuele, P. Dino, G. Fosca, "PageRank: Data-driven Performance Evaluation and Player Ranking in Soccer via a Machine Learning Approach", ACM Transactions on Intelligent Systems and Technology, Vol. 10, No. 5, 2019, pp. 1-27.

[33] T. Y. Ahmet, S. Barış, K. Tolga, "Football Player Value Assessment Using Machine Learning Techniques", Proceedings of the Intelligent and Fuzzy Techniques in Big Data Analytics and Decision Making: Proceedings of the INFUS 2019 Conference, Istanbul, Turkey, 23-25 July 2020, pp. 289–297.

[34] J. Mikael, P. Ashwin, M. Saumya, B. Marco, M. Daniel, C. Mark, "Using multiple machine learning algorithms to classify elite and sub - elite goalkeepers in professional men' s football", Scientific Reports, Vol. 11, 2021, p. 22703.

[35] A. Didem, "A case study on player selection and team formation in football with machine learning", Turkish Journal of Electrical Engineering and Computer Sciences, Vol. 29, No. 3, 2021, pp. 1672-1691.

[36] M. M. Baydaa, S. C. Muayad, N.R. Ahmed, "Football player tracking and performance analysis using the OpenCV library", Mathematical Modelling of Engineering Problems, Vol.11, No. 1, 2024, pp. 123-132.

[37] L. Alan, V. Tomas, C. Z. Luca, M. Jiri, K. Matej, "Discriminative Correlation Filter Tracker with Channel and Spatial Reliability", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21-26 July 2017, pp. 671-688.

[38] R. Vincenzo, K. Peter, M. Rafael, R. Antonio, M. Magni, "Training load and submaximal heart rate testing throughout a competitive period in a top-level male football team", Journal of Sports Sciences, Vol. 38, No. 11-12, 2020.

[39] A. Aşçı, "Heart Rate Responses during Small Sided Games and Official Match-Play in Soccer", Sports, Vol. 4(2), No. 31, 2016, pp. 11-14.

[40] G. Chiara, C. Emanuele, C. Andrea, N. Stefano, B. Carmine, C. Carlo, G. Michele, "COVID-19 disease in professional football players: symptoms and impact on pulmonary function and metabolic power during matches", Physiological Reports, Vol. 10, No. 11, 2022, pp. 1-11.

[41] D. Hubert, K. Aleksandra, W. Dariusz, "Nutrition for Female Soccer Players Recommendations", Medicina, Vol. 56, No. 1, 2020, pp. 28-44.