J. Electrical Systems 20-6s (2024): 64-74

[1]Lin Leng

Research on Evaluation and Optimization Algorithm of Athletes' Sports Skills Based on Trajectory Analysis and Machine Learning

Abstract: - Evaluating athletes' sports skills is significant for optimizing their training strategies and improving competitive outcomes. Although various studies are done to assess athletes' skills, they face challenges because of their limitations in capturing the complex, intricate patterns in athletes' movements. Hence, we proposed an innovative approach utilizing trajectory analysis and machine learning algorithms to evaluate and optimize athletes' sports skills. Initially, the trajectory data is collected using advanced motion tracking technology during either training or competitive sessions of the athletes. The collected data was pre-processed, and then feature extraction was performed to extract the most important and relevant features within the data. Further, the Gradient Boosting Decision Tree (GBDT) was trained using the extracted features to predict the athletes' performances. The GBDT is a powerful machine learning algorithm that can handle complex, non-linear interconnections among variables, enabling it to accurately predict the athletes' performances. Finally, we applied the Whale Optimization Algorithm (WOA) to refine the training process of the GBDT, enabling precise training and accurate predictions. The proposed methodology was implemented in the MATLAB software, and it is validated using the real-world athlete data collected during training sessions. The experimental results are validated in terms of parameters such as accuracy, precision, recall, and f-measure. Furthermore, we made a comparative study with the existing methods to validate the effectiveness and robustness of the proposed technique.

Keywords: Gradient Boosting Decision Tree, Whale Optimization Algorithm, Athlete Sport Evaluation

1. INTRODUCTION

The most important pillars of competitive sports are sports competition and sports training. Although these two pillars are closely interconnected, they have certain variations. The major objective of sports training is to transform the physical and psychological qualities of the players to improve their sports outcomes [1]. In sports training, the trainers aim to maximize the potential of all players to their fullest to make them ready for competition. On the other hand, sports competition quantifies the efficiency of sports training by allowing all players allowing all players to compete on the same field. Both sports training and sports competitions promote physical fitness, skill development, personal growth, and teamwork among athletes [2]. Although sports training is measured through competitions, high-level sports training is the basic and significant for high-quality sports competitions. Achieving competitive success is the primary goal of sports training. Therefore, assessing the athletes' sports skills is important to improve their performances. Generally, the trainers or coaches use quantitative assessment approaches to measure and evaluate the athletes' sports skills [3].

These traditional assessment techniques evaluate athletes' outcomes, including physical and technical features like speed, agility, mental resilience, etc. This assessment enables the trainers to identify the athletes' strengths and weaknesses. However, this assessment was made based on the athletes' feedback; hence, providing a detailed understanding of athletes' skills is impossible [4]. The advancement of technologies such as artificial intelligence (AI) and machine learning (ML) offers an innovative tool to improve the accuracy and robustness of the assessment of athletes' sports skills. These advanced technologies offered a more accurate evaluation of athletes' performances than traditional models. The AI and ML methods are trained with the large real-time sports database collected using wearable sensors and motion-tracking tools, leading to more precise prediction of athletes' performances [5]. These techniques can help understand and learn the complex, intricate patterns

[1] HeNan College of Transportation, HeNan, Zheng Zhou, 450000, China

*Corresponding author's e-mail:lenglin2001@163.com https://email.163.com/

within data, making it effective in assessment compared to conventional models. Despite these advantages, some challenges exist in incorporating ML techniques in sports assessment. The most significant drawback of ML algorithms is the limitation of adaptability.

Typically, the ML algorithms get trained on historical data to predict the athletes' performances; hence, they cannot adapt to evolving changes in sports and athletes' progress. Sometimes, the ML algorithms get overtrained on training data, leading to overfitting and limited generalization. In addition, the ML approaches often face interpretability challenges, which limits their practical utility [6]. To address these issues, we proposed an innovative athlete sports skill evaluation algorithm based on trajectory analysis and an optimized machine learning algorithm. By analyzing the trajectory of the athletes' movements, the designed algorithm aims to predict the performance of athletes more accurately through continuous learning of the evolving dynamics of sports and athletes' progress.

The presented algorithm is organized as follows: section 2 provides the recent literature related to the proposed framework, section 3 provides the developed methodology, section 4 presents the performance analysis of the proposed technique, and section 5 provides the research conclusion.

2. RELATED WORKS

A few recent studies associated with athletes' sports skill assessment are reviewed below,

Predicting athletes' performance in sports enables them to improve their training process, leading to high-competition success. Haixia Zhao et al. [7] developed a predictive framework using genetic algorithm-assisted backpropagation neural networks to assess athletes' sports performances. The primary concern of this work is to improve sports skills and performances by creating a robust and effective prediction algorithm. This framework was validated using the real-time sports data collected from China, and the experimental results depict that this predictive algorithm achieved 97.6% accuracy. However, this methodology acquired a high false positive rate.

Jinjuan Wang et al. [8] presented a hybrid predictive algorithm to assist coaches in formulating training strategies for improving athletes' sports performances. This work constructed a genetic algorithm-assisted backpropagation neural network to predict sports performances. The implementation of this algorithm showed that it achieved faster convergence than the conventional approaches. Also, it illustrated that this algorithm can identify each athlete's strengths and weaknesses. In addition, it achieved better stability, high accuracy, and greater application value than the traditional methods. However, this methodology is limited to generalizability.

Yu-Hung Tsai et al. [9] introduced an innovative model for predicting the performances of tennis athletes by accumulating their dynamic brain waves. This study designed a deep neural network (DNN) algorithm to predict the athletes' performances. This study commences with the collection of dynamic brain waves of the table tennis players, and the collected data was transformed from the time domain into the frequency domain, which is processed using the DNN algorithm to predict the athletes' performances. The implementation results show 96.70% accuracy for detecting dynamic brain waves interconnected with good elite sports performance. However, this methodology requires more computational resources for training and is limited to scalability.

Musa Oytun et al. [10] presented a study to examine various ML algorithms for predicting the kinds of athletes' performances in female handball players. This study utilized ML models such as support vector machine (SVM), Linear regression, Radial-Basis Function Neural Network (RBFNN), Decision Tree, and Backpropagation neural network. These algorithms are executed to identify the performance of female handball players. The implementation results validated that the RBFNN outperformed other algorithms with an accuracy of 0.97. Although the RBFNN offers greater accuracy, it faces limited scalability and adaptability issues.

Duan Hongyun et al. [11] presented an improved neural network-based prediction model for identifying sports performances. This framework utilizes the data mining algorithm for analyzing the athletes' grades. In addition, it utilized the firefly optimization algorithm to optimize the calculation of the designed improved neural system. This collaborative framework aims to determine the interconnections between physical education teaching algorithms and grades. The result of this study demonstrates that teaching methodologies influence sports performance. However, the collaboration of multiple techniques leads to increased algorithm complexity.

2.1 Challenges

Evaluating athletes' sports skills is one of the important tasks for improving their competition success rates. This evaluation enables the trainers and coaches to optimize their training strategies, enhancing athletes' sports performances and physical strengths. Typically, the trainers assess sports performances by collecting the athletes' feedback. However, this methodology is time-consuming and may not accurately assess the athletes' strengths and weaknesses. Recent studies have concentrated on developing automatic sports performance assessment tools using artificial intelligence algorithms such as machine learning and deep learning to resolve this issue. These techniques use real-time sports data for training in which it understands the interrelations and patterns within the data for predicting sports performances. Although it is less time-consuming and more accurate than the traditional approach, it has certain challenges described below.

❖ The traditional way of evaluating the athletes' sports skills depends on the quantitative assessment by coaches and trainers, which is prone to bias and inconsistency.

❖ The conventional ML-based methods are often computationally intensive, limiting their applicability for evaluating the sports skills across multiple athletes.

❖ The machine learning algorithms are limited to scalability, so they cannot perform well on large-scale datasets, reducing their practical utility in real-world scenarios.

❖ The ML methods are typically data-dependent and cannot generalize well across diverse sports fields. Moreover, they require extensive data for optimal training, which increases their computational costs.

To resolve these issues, we proposed a hybrid predictive algorithm using an optimized machine learning algorithm. The proposed model integrates the WOA into the GBDT to predict athletes' performances precisely.

3. PROPOSED METHODOLOGY FOR ATHLETES' SPORTS SKILLS EVALUATION

A hybrid and innovative prediction algorithm was proposed to evaluate and assess the athletes' sports skills based on trajectory analysis and machine learning. The proposed framework consists of five modules: data collection, pre-processing, feature engineering, sports skill evaluation, and optimization. The sports trajectory database was collected using the motion tracking technology in the first module. Secondly, the collected database was pre-processed to improve the dataset's quality.

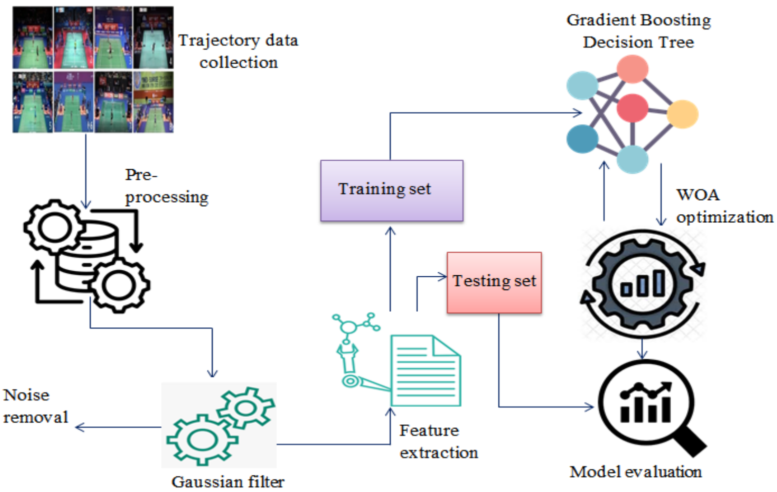


Figure 1: Proposed Algorithm Architecture

The pre-processing module follows steps like noise removal and normalization. Thirdly, the feature engineering was performed using principal component analysis (PCA) to extract the most important and relevant features from the pre-processed data. This extracted feature sequence was used to train the GBDT model, which learns the patterns within the data for predicting the athletes' sports performances. Finally, WOA was incorporated into the GBDT module to optimize its training process by refining its parameters to its optimal range. Figure 1 depicts the architecture of the developed framework.

3.1 Data collection

The developed work begins with collecting trajectory data from the sports field during training or competition sessions. The presented study utilized the shuttlecock trajectory database. This dataset contains 26 broadcast videos with a resolution and frame rate of 1280x720 and 30 fps, respectively. There are 78,200 video frames in the database, of which 63,675 frames (23 videos) are collected during professional games, and the remaining 9525 frames (3 videos) are recorded during fun playing. The samples of the dataset are presented in Figure 2. This input database was split in the ratio of 80:20 for model training and testing.

Figure 2: Sample dataset images

3.2 Pre-processing

Pre-processing is one of the essential steps that improves the data quality and makes the dataset into an appropriate format for further analysis. Here, we performed two pre-processing steps: noise removal and normalization. Noise removal indicates the process of removing or discarding unwanted irregularities from the trajectory data. These irregularities or noises distort the actual patterns of the athletes' movements, which may reduce the efficiency of subsequent processes. We applied a Gaussian filter to remove these irregularities from the trajectory data in the presented framework. The Gaussian filter is mathematically represented in Eqn. (1).

$$f_i = \frac{1}{2\pi\sigma} \sum_{j=-\infty}^{\infty} x_j . e^{-\frac{(i-j)^2}{2\sigma^2}} \quad (1)$$

Where f_i defines the filtered value at position i, σ indicates the standard deviation of the Gaussian distribution, and x_j denotes the original value at position j. This filter smoothens the trajectory data by discarding the high-frequency noise values while preserving the important features. Secondly, normalization was applied, where the filtered data was scaled to a common range. This step reduces the influence of changing measurements in trajectory analysis by ensuring that all attributes contribute equally to the trajectory analysis. Here, we used a min-max scaling algorithm for performing normalization, and it is represented in Eqn. (2).

$$x_s = \frac{x - x_{mi}}{x_{mi} - x_{mx}} \quad (2)$$

Where x_s defines the scaled value of the data point, x_{mi} represents the minimum value in the database, and x_{mx} refers to the maximum value. The min-max scaling algorithm retains the interconnections between all data points in the database while ensuring they fall within a specified range. These steps standardize the raw dataset and transform it into a suitable format for further trajectory analysis.

3.3 Feature extraction

Feature extraction defines extracting the most relevant and significant features from the pre-processed database by converting them into a sequence of attributes. The extracted feature sequence represents the important characteristics of the data, and it serves as the input to the created prediction model, enabling it to understand the patterns and make accurate sports skill predictions. In the developed study, we used the PCA technique for feature extraction. The PCA is a feature engineering approach which identifies the principal elements (important features) in the pre-processed data. This process minimizes the data dimensionality by discarding the irrelevant attributes in the database and highlights the dataset's most relevant attributes. The PCA approach follows certain steps: data centering, covariance matrix calculation, eigenvalue decomposition, and selection of principal elements. In the data-centering step, the mean of each attribute is subtracted from its corresponding feature value. Secondly, the covariance matrix is calculated using Eqn. (3).

$$C_m = \frac{1}{n}\left(\bar{c}^T . \bar{c}\right) \quad (3)$$

Where $\bar{c}$ indicates the centered data matrix, $\bar{c}^T$ denotes the transpose of the centered data matrix, C_m represents the covariance matrix, and *n* defines the number of samples in the pre-processed database. Then, the eigenvalue decomposition was performed on the estimated covariance matrix to determine eigenvalues and vectors formulated in Eqn. (4).

$$C_m . e_i = \kappa_i . e_i \quad (4)$$

Where e_i and κ_i indicate the i^{th} eigenvalue and its corresponding eigenvector. Finally, select the eigenvalues with maximum variance to form principal components. These principal components represent the most relevant and informative attributes present in the dataset.

3.4 Gradient Boosting Decision Tree

The GBDT is a supervised machine learning approach called a gradient boosting regression tree (GBRT). This algorithm integrates the regression tree using the gradient boosting approach, and it is widely used in tasks such as prediction, classification, fault assessment, etc. It uses the least-squares function minimization preceded by a single parameter optimization to resolve the difficult function minimization problem. This unique feature of GBDT enables it to achieve high accuracy in athletes' sports performance evaluation. The GBDT must be trained using the extracted features to make accurate predictions. Before the training process, we split the dataset in the ratios of 80:20 for training and testing purposes. Each sample in the training sequence is expressed in Eqn. (5).

$$T_s = \{(u_1, v_1), (u_2, v_2), \ldots\ldots\ldots, (u_h, v_h)\} \quad (5)$$

Where T_s indicates the training set, and h denotes the total number of samples in the training set. During training, the GBDT aims to reduce the expected value of the loss function $l_s(v_i, f(u_i))$ by creating a weak learner $h_t(u_i; w)$ iteratively. This weak learner defines a classification tree where the parameter *w* denotes the splitting variable, split locations, and the terminal nodes of the individual tree. The square loss function utilized in the designed model is represented in Eqn. (6).

$$l_s(v_i, f(u_i)) = \frac{1}{2}(v_i - f(u_i))^2 \quad (6)$$

The prediction of athletes' sports performances using DBDT follows certain steps. Firstly, initialize the weak learner for the training sequence, and it is mathematically presented in Eqn. (7).

$$f_0(u) = \arg\min_{\eta} \sum_{i=1}^{h} l_s(v_i, \eta) \quad (7)$$

Where $f_0(u)$ defines the regression tree containing a single root node. Since the loss function used in the designed methodology is square loss, the weak learner becomes $f_0(u) = \bar{v}$. Then, determine the negative gradient using the expression in Eqn. (8).

$$\tilde{v}_i = -\left[\frac{\partial l_s(v_i, f(u_i))}{\partial f(u_i)}\right]_{f(u)=f_{r-1}(u)} \quad (8)$$

Where *r* defines the iteration number ranging from 1 to *R*. On the other hand, $\tilde{v}_i$ indicates the negative gradient of the loss function concerning the current model. Further, replace the label of the training sequence with the determined gradient to create a new database, represented in Eqn. (9).

$$T_r = \{(u_1, \tilde{v}_1), (u_2, \tilde{v}_2), \ldots\ldots\ldots, (u_m, \tilde{v}_m)\} \quad (9)$$

Where represents the new dataset T_r obtained by replacing the labels v_i with $\tilde{v}_i$. This replacement creates a new regression tree by training the newly created database mathematically expressed in Eqn. (10).

$$w_t = \arg\min_{w} \sum_{i=1}^{h} (\tilde{v}_i - h_r(u_i; w))^2 \quad (10)$$

Where $h_r(u_i; w)$ denotes the new regression tree trained on the new dataset and w_t represents the parameters of the tree, which minimize the squared loss between the predictions. Then, update the strong learner using the Eqn. (11).

$$f_r(x) = f_{r-1}(x) + \gamma h_r(u_i; w_r) \quad (11)$$

Where γ defines the learning rate. This process is continued for each iteration, and the final output of the GBDT is represented in Eqn. (12).

$$f_M(x) = f_0(x) + \sum_{r=1}^{M} \gamma h_r(u; w_r) \quad (12)$$

Where $f_M(x)$ is employed for predicting the sports performances in an unknown data sample. This final prediction output of GBDT provides the ranking of each athlete in the particular sport. This ranking enables the coaches and trainers to optimize their training progress to improve their competition performance. Although the DBDT offers high accuracy, training the system is quite complex. In the developed work, we applied WOA to optimize the training process of GBDT, which refines the DBDT parameters, such as maximum iteration, negative gradient, tree parameters, etc., to minimize the loss function.

3.5 Optimization

The Whale Optimization Algorithm (WOA) was a meta-heuristic optimization algorithm developed based on the characteristics of humpback whales. In the developed work, we used the WOA algorithm to optimize the GBDT by refining its parameters, such as a number of trees, learning rate, tree depth, etc. Fine-tuning these parameters to their optimal value at each iteration enhances the training process and improves the overall prediction efficiency. Typically, the WOA approach follows three major steps: encircling prey, spiral updating position, and searching phase (exploration). In this study, we used the searching phase of the WOA to find the optimal value of GBDT parameters. The optimization process begins with the initialization of the whale population. Here, the population defines the parameter sequences, which are randomly initialized, and it is represented in Eqn. (13).

$$P_p = \{a_{p1}, a_{p2}, a_{p3}, \ldots\ldots, a_{pz}\} \quad (13)$$

Where P_p indicates the parameter population, a_{p1} indicates parameter set, and z denotes the population size. Then, the fitness value of each parameter solution is determined based on the pre-defined objective function. Here, the objective function is to reduce the square loss function of GBDT. If the loss value is minimum for a parameter solution, its fitness value will be high and vice versa. After fitness evaluation, the searching phase begins. In this phase, the parameter solution explores the search space to find its optimal value by iteratively updating its values. The exploration phase is mathematically represented in Eqn. (14).

$$a_p(t+1) = a_p(t).\vec{a}_p - \vec{A}.\vec{D} \quad (14)$$

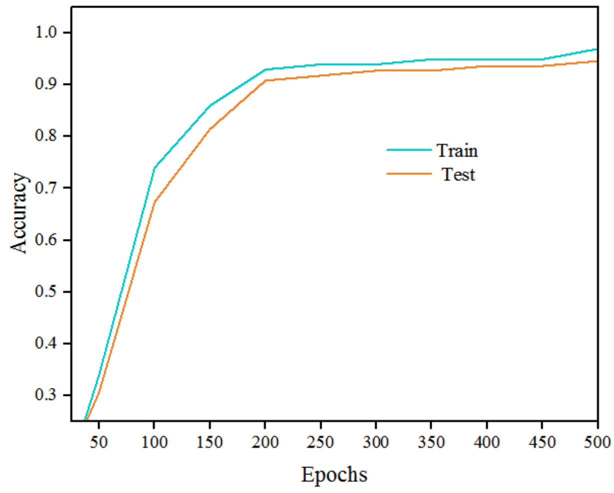
Where $a_p(t+1)$ defines the updated parameter set, $\vec{a}_p$ defines the randomly selected parameter set, $\vec{A}$ and $\vec{D}$ denotes the coefficient vector. After parameter updation, the fitness value was calculated for the updated parameter set. Finally, the parameter solution with a high fitness value was selected for GBDT training. The parameter optimization is a continuous process, and it refines the parameter solution at each iteration, which ensures adaptability to the system (the system produces outputs considering the changes in each iteration).

4. RESULTS AND DISCUSSION

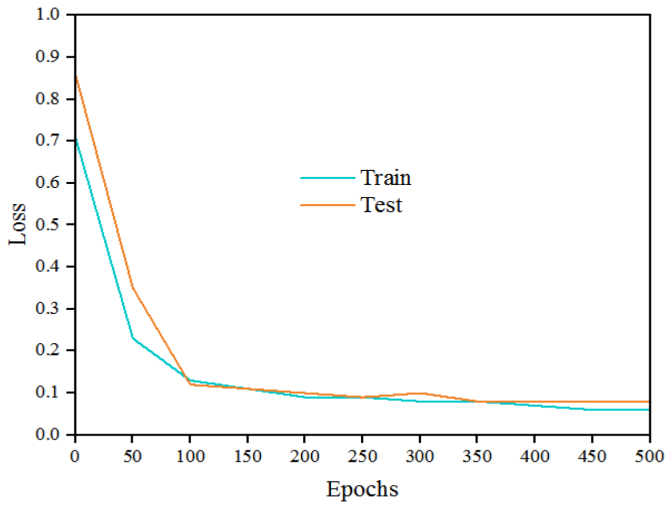
In this study, we developed an innovative predictive model combining the efficiency of machine learning, meta-heuristic optimization, and trajectory analysis for predicting athletes' sports performances. The developed methodology uses a whale optimization algorithm and the gradient boosting decision tree algorithm for evaluating the athletes' skills, which aids the coaches in improving their training progress and competition success. The presented methodology was designed and executed in the MATLAB software, version R2020a, running on a 64-bit Windows Operating System. The study's results are examined in terms of accuracy, precision, recall, and f-measure.

4.1 Assessment of training and testing performances

This section discusses the proposed algorithm's training and testing performances. Initially, the collected database was split into the ratios of 80:20 for training and testing processes, and the performances were evaluated in terms of accuracy and loss. The training accuracy measures the developed model's proficiency in learning and understanding the complex, intricate patterns within the data. It also measures how quickly the designed framework extracts the patterns, and it is evaluated by increasing the epoch count from 0 to 500. The performance evaluation suggests that the designed algorithm acquired a high training accuracy of 0.98, highlighting its efficiency in predicting the sports skills on the train dataset. Consequently, we determined the testing accuracy of the model. The testing accuracy evaluates how precisely the proposed model applies the learned patterns in the unknown samples and predicts the sports skills. The developed methodology achieved a high testing accuracy of 0.96, which depicts that it generalizes well on the unseen data samples.



(a)



(b)

Figure 3: Training and testing outcomes: (a) accuracy, (b) loss

On the other hand, we evaluated the loss to determine the false prediction made by the proposed algorithm. The training loss evaluates the difference between the real and predicted values on the training samples. It measures how effectively the developed algorithm fits the training samples and resolves the overfitting issue. The training loss evaluation manifests that this algorithm incurred a minimum loss of 0.06. Subsequently, we determined the testing loss over increasing epochs from 0 to 500. The testing loss measures the deviation between the actual and predicted values on unseen data samples, which determines the model's generalization ability to the real-time sports data. This methodology acquired a reduced testing loss of 0.07. Figure 3(a, b) depicts the graphical representation of the training and testing outcomes. This intensive evaluation shows that the proposed methodology achieved high accuracy and minimum loss in both train and test phases, highlighting its efficiency in precisely predicting the athletes' sports performances.

5.2 Performance comparison

The outcomes achieved by the developed algorithm were evaluated with the traditional models in this section. Here, the performance parameters such as accuracy, precision, recall, and f-measure are compared with conventional techniques such as Support Vector Machine (SVM), Deep Neural Network (DNN), Genetic Algorithm-based Backpropagation Neural Network (GA-BNN), Radial Basis Functional Neural Network (RBFNN), and Decision Tree (DT).

Accuracy measures the ratio of correct predictions made by the developed algorithm among all the predictions made. It also quantifies how correctly the designed methodology predicts the athletes' performances. Figure 4(a) provides the comparative evaluation of accuracy. Here, the accuracy earned by the designed strategy is compared with the conventional models such as SVM, DNN, GA-BNN, RBFNN, and DT. Their performances are determined by increasing data volume from 500 to 3000. When the data volume is 3000, these existing techniques and the proposed methodology acquired accuracies of 84.67%, 95.76%, 95.57%, 95.45%, 90.43%, and 98.84%, respectively. This evaluation clearly shows that the proposed strategy achieved more accuracy than the conventional techniques. Moreover, the designed algorithm maintained consistent accuracy over increasing data volume, highlighting its scalability and adaptability over real-time data.

Consequently, we compared the precision performance of the designed methodology with the above-stated existing techniques. The precision metric determines the proportion of true positive predictions out of all instances detected as positive by the algorithm. It measures the model's capacity to avoid false positives. Figure 4(b) depicts the comparative assessment of precision. The comparative study depicts that the existing algorithms, such as SVM, DNN, GA-BNN, RBFNN, and DT, achieved precision values of 84.31%, 93.47%, 95.22%, 95.34%, and 90.82%, respectively, for 3000 data volume. On the other hand, the developed algorithm achieved an improved precision of 99.03% at 3000 data volumes, illustrating that it is effective in predicting the correct positive instances. The comparison of recall performance of the developed system with the existing models is graphically presented in Figure 4(c). The recall metric, also known as sensitivity, quantifies the ratio of correct positive predictions made by the developed system out of all positive cases. It also measures the system's capacity to detect all positive cases correctly. The recall performance of the developed algorithm is evaluated with conventional models such as SVM, DNN, GA-BNN, RBFNN, and DT over increasing data volumes. When the data volume is 3000, these methodologies obtained recall rates of 91.24%, 96.43%, 95.75%, 94.11%, and 89.12%, respectively, while the developed algorithm earned 98.3%. This validates that the proposed model achieved an enhanced recall rate compared to the existing techniques.

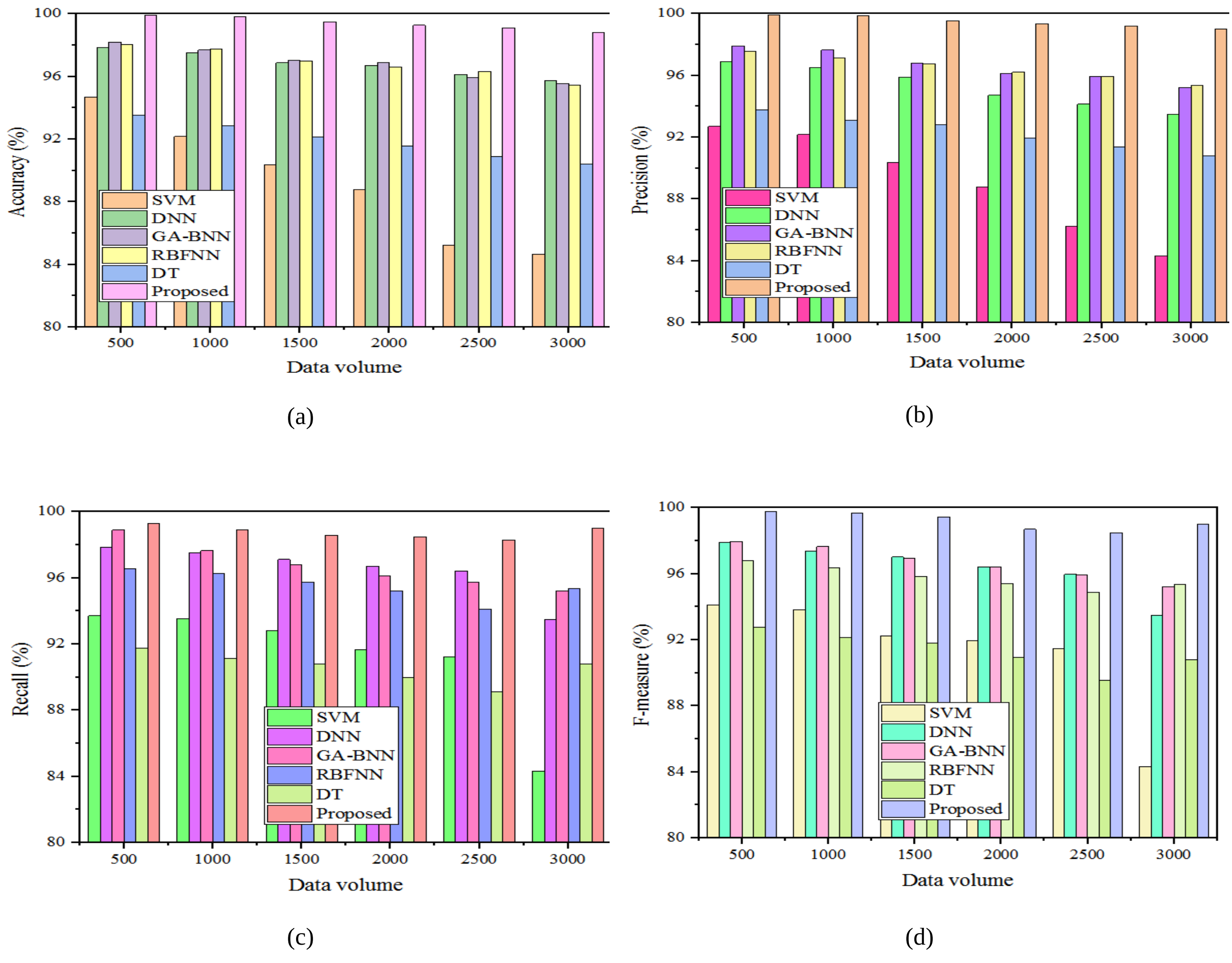


Figure 4: Performance comparison: (a) accuracy, (b) precision, (c) recall, and (d) f-measure

Subsequently, we evaluated the f-measure performance of the developed algorithm with the traditional algorithms such as SVM, DNN, GA-BNN, RBFNN, and DT, presented in Figure 4(d). The f-measure metric combines recall and precision metrics into a single value, offering a balanced evaluation of the system's performance. It indicates the harmonic mean of recall and precision. The above-mentioned existing algorithms obtained f-measures of 91.47%, 95.98%, 95.95%, 94.87%, and 89.56%, respectively, while the proposed methodology achieved f-measures of 98.50%. The improved f-measure manifests that the designed algorithm accurately predicts the athletes' sports performances better than conventional algorithms.

Through this intensive evaluation of the system's performances with the conventional models, it is validated that the trajectory analysis using the combination of GBDT and WOA increases the prediction performances. Furthermore, evaluating the model's outcomes over increasing data volumes suggests that the designed algorithm obtained high scalability, making it effective and robust for real-time sports performance prediction.

5.3 Discussion

The research introduces a hybrid predictive model for assessing the athletes' sports performances. The proposed algorithm incorporates the efficiency of WOA into the GBDT to accurately predict the athletes' sports outcomes. Through this research, we aim to assist the coaches and trainers enhance the athletes' training progress to achieve high-competition success. The presented study was trained and tested with the shuttlecock trajectory data. The experimental results of the study show that the developed algorithm achieved a high accuracy of 99.49%, a greater precision value of 99.56%, an improved recall rate of 98.6%, and an increased 99.46%. Furthermore, a comparative assessment was made with the existing techniques, including SVM, DNN, GA-BNN, RBFNN, and DT, and it validated that the designed methodology achieved higher performances in accuracy, precision, recall, and f-measure. Also, the consistent performance of the developed algorithm over increasing data volumes highlights its scalability and adaptability to changing environments. These enhanced performances make the proposed approach suitable for predicting the sports performances of athletes in real-world scenarios.

6. CONCLUSION

This work introduced a hybrid predictive model using the combined efficiency of Whale optimization and the gradient-boosting decision tree algorithm. The objective of the proposed algorithm is to evaluate and predict the athletes' sports skills to improve their training progress and competition success. The developed methodology involves collecting trajectory data from sports fields during training and competition. Then, the database was pre-processed, and the most significant features were extracted. Further, the GBDT was trained using the extracted features to learn the intricate patterns within the data, enabling it to predict the sports performances of the athletes. Consequently, we applied the WOA to refine the GBDT training by fine-tuning its parameters to its optimal value. This collaborative training process reduces the computational time and enhances the prediction efficiency. The developed algorithm was validated using the shuttlecock trajectory data. Further, a comparative analysis was made with existing models like SVM, DNN, GA-BNN, RBFNN, and DT, and it is observed that the parameters such as accuracy, precision, recall, and f-measure are improved by 2.47%, 2.74%, 1.48%, and 2.44%, respectively. These improved outcomes of the designed algorithm make it an optimal and effective real-time-based athlete sports performance analysis.

Although the developed methodology achieved higher results, it faces significant challenges, such as a lack of interpretability and adaptation to diverse sports environments. To resolve these challenges, future studies should concentrate on designing a simple predictive algorithm suitable for assessing the performances of athletes in different sports domains.

REFERENCES

1. Charest, Jonathan, and Michael A. Grandner. "Sleep and athletic performance: impacts on physical performance, mental performance, injury risk and recovery, and mental health: an update." Sleep medicine clinics 17.2 (2022): 263-282.

2. Gu, Song, and Lan Xue. "Relationships among sports group cohesion, psychological collectivism, mental toughness and athlete engagement in chinese team sports athletes." International Journal of Environmental Research and Public Health 19.9 (2022): 4987.

3. Liu, Chao, et al. "Analysis of Physical Education Classroom Teaching after Implementation of the Chinese Health Physical Education Curriculum Model: A Video-Based Assessment." Behavioral Sciences 13.3 (2023): 251.

4. Deng, N., Soh, K. G., Abdullah, B., Huang, D., Xiao, W., & Liu, H. (2023). Effects of plyometric training on technical skill performance among athletes: A systematic review and meta-analysis. Plos one, 18(7), e0288340.

5. Liu, J., Huang, G., Hyyppä, J., Li, J., Gong, X., & Jiang, X. (2023). A Survey on Location and Motion Tracking Technologies, Methodologies and Applications in Precision Sports. Expert Systems with Applications, 120492.

6. Xu, D., Zhou, H., Quan, W., Ugbolue, U. C., Gusztav, F., & Gu, Y. (2024). A new method applied for explaining the landing patterns: Interpretability analysis of machine learning. Heliyon.

7. Zhao, H., Li, W., Gan, L., & Wang, S. (2023). Designing a prediction model for athlete's sports performance using neural network. Soft Computing, 27(19), 14379-14395.

8. Wang, Jinjuan. "Analysis of sports performance prediction model based on GA-BP neural network algorithm." Computational Intelligence and Neuroscience 2021 (2021).

9. Tsai, Y. H., Wu, S. K., Yu, S. S., & Tsai, M. H. (2023). A Novel Hybrid Deep Neural Network for Predicting Athlete Performance Using Dynamic Brain Waves. Mathematics, 11(4), 903.

10. Oytun, M., Tinazci, C., Sekeroglu, B., Acikada, C., & Yavuz, H. U. (2020). Performance prediction and evaluation in female handball players using machine learning models. IEEE Access, 8, 116321-116335.

11. Hongyun, Duan. "Application of data mining technology based on improved neural network algorithm in sports performance prediction." Soft Computing (2023): 1-11.