

Chapter 6

Prediction and Performance Assessment in Woman Handball Athletes by Employing Machine Learning Methods



Kolla Vivek , J. Harikiran, Kunda Suresh Babu, and Ch. Anil Carie

Introduction

Handball is recognized as a game needing endurance, agility, endurance, and an intermittent pace, with sporadic features of the game containing rapid defense and attack [1]. In addition, handball is performance-oriented, i.e., winning a game depends on performance of the athletic and incorporates items that are mechanical, strategic, and psychological.

As a consequence of enforcing modish regulations, the abilities required and relevant to players have increased considerably in the past few years. Handball seems to have become a quick and intense sport in which players are supposed to do better with a disciplined run, drive, leap, shoot, move, and block capabilities. Consequently, the information from multiple workout experiments for the evaluation of actual game expertise in the field is very essential to examine. Sport performance analysis and methods permit sports managers, selection committees, and trainers to evaluate player's capability objectively. Thus, the capability and potential of the players have become a key factor of selection and training. Researchers have implemented several

K. Vivek (✉)

Computer Science and Engineering, QIS College of Engineering and Technology,
Ongole 523001, India

e-mail: kollavivek@gmail.com

J. Harikiran · K. Suresh Babu · Ch. Anil Carie

School of Computer Science and Engineering, VIT-AP University, Amaravati, India

e-mail: harikiran.j@vitap.ac.in

K. Suresh Babu

e-mail: suresh.20phd7019@vitap.ac.in

Ch. Anil Carie

e-mail: anil.carie@vitap.ac.in

techniques and approaches to assess the most critical attributes of sport and exercise biomechanics, the science of life, and psychobiology [2]. Because of the low error rate and high accuracy rate, artificial intelligence (AI) and its techniques are widely used to develop robust solutions to complex problems. Deep learning as a subgroup of ML and ML which is a field of AI is also used in different fields to develop solutions to complex problems. ML always utilizes experience to predict the results, and the prediction rate always depends on the attributes and features of the dataset. A very little research was done on handball sport using ML techniques. The remaining parts of the work are presented as follows. Section 2 discussed related studies. It is in Sect. 3 that the proposed method is presented. The findings were addressed in detail in Sect. 4. Section 5 concludes with findings and recommendations for the future.

Related Works

Many experts and practitioners developed prediction models using machine learning for analyzing the player's performance in different sports sciences and games like volleyball, basketball, handball, cricket, swimming, football, etc. Jovanovic et al. [1] conducted experiments on 20 female respondents aged between 16 and 25 years. The results had proved an R-squared value of 0.876. Tests conducted on players included SAMO reaction-agility test and the Illinois test. Gomez-Lopez et al. [2] considered the expectation capacity over the convictions that causes the achievement in handball sport. The example was included 444 elite competitors (233 young men and 211 young ladies). The relation between self-regulation techniques, mindfulness practice, and success was discussed by Popa et al. [3]. The study is comprised of 288 handball representatives from Romania. The participants were 30% male and 70% female, varying in age from 12.01 to 14 years. In a hierarchical multiple regression, the parameters (state consciousness of the individual, self-monitoring, and self-efficacy) clarified 87% of the variation in sports success.

Hermassi et al. [4] conducted experiments on 72 young handball players whose age group lies between 15.2 and 16.4 years, and performance measurements of lower limbs in young handball players were taken into consideration. They assessed the player's performance twice, on separate days. Tests like squat jumps, counter-movement jumps, 5m sprint, 10-m sprint, and handball sport skill test were conducted and recorded finishing times using electronic timing gates. The model forecasted the player's performance with R-squared values between 0.52 and 0.68.

Sekulic et al. [5] conducted experiments on 32 male futsal players whose age group lies between 26.22 and 31.22 years, body height lies between 182.13 and 187.99 cm, and body mass lies between 77.43 and 85.00 kg. Tests like reactive agility (RAG), change of direction speed (CODS), and 10 m sprint were conducted and recorded finishing times of players using power timer 300. Soslu et al. [6] conducted experiments on 23 male basketball players whose age group lies between 23.2 and 26.7 years, body height lies between 197.1 and 206.1, and body mass lies between 95.3 and 105.3 kg; anthropometric performance measurements,

vertical jump measurements, sprint performance evaluation, anaerobic performance measurements of isokinetic knee strength, muscle strength in basketball players were taken into consideration. The investigation was led over 1 week, during which the players did not take an interest in some other preparation or matches. Tests like Wingate anaerobic test, T-drill were conducted and recorded finishing times.

All of the above-mentioned experiments are intended to execute ML models to establish standards of player exhibitions of particular skills in different areas of the game, as well as to help mentors decide on reasonable decisions in terms of squad or individual player determination [7–10]. Only a certain number of research on the use of AI models in sports science have been published. In addition, each game has a remarkable structure, physical abilities, and prerequisites. For certain activities, games, and sports to be settled, the diverse and dynamic layout of athletic exhibitions requires an appropriate ML model [11–14]. As far as we could know, this is the main investigation to utilize ML models to anticipate explicit exhibitions in handball players. Further investigations may be required for various games, age gatherings, and genders.

Four ML models are used in this article to predict the performances of women handball players and allow mentors to accurately assess the performance of players before games. Furthermore, huge elements impacting the thought about exhibition abilities were resolved to improve the player execution. The intension and considerations of this analysis can be defined as follows, because of the previously mentioned data and a literature survey: To actualize a few ML techniques to determine the optimal standard for the competencies considered and to carry out a related test with different measurements.

- (i) To consider numerous ability skills of players for the prognosis.
- (ii) To anticipate player's ability in four competencies and to help mentors with the effective choice of players in games.
- (iii) To decide the elements that influence the considered competencies and help coaches to concentrate on critical variables to improve player performance of specific competencies.

Proposed Methodology

Dataset

Data related to 52 players whose age group lies between 20.2 and 26.1 years, height lies between 164.1 and 170.7 cm, bodyweight lies between 62.4 and 72.1 kg, and body mass index (BMI) lies between 23.4 and 25.7 kg m² were collected from Kaggle repository. For two intervals, socioeconomic data and psychometrics were collected. For the ML models, a total of 23 attributes (Table 6.1) and attribute metrics were captured, and 117 occurrences were captured for training samples. Each metric was captured for every player in parallel during two separate intervals, and a dataset

Table 6.1 Dataset features

Age (years)	Backside middle upper arm double skin density (ST) (mm)
Supraspinal double skin density in millimeters	Wingate average power (AP) in kilograms
Athletic's body weight in kilograms	Front side middle upper arm double skin density (mm)
Midriff double skin density in millimeters	Wingate comparative mean power (cMP) (kg W^{-1})
Athletic's body length in meters	Medial calf double skin density in millimeters
Thigh double skin density in millimeters	Wingate peak power (PP) in kilograms
Quetelet index (kg/m^2)	lower shoulder blade double skin density (mm)
Midaxillary double skin density (mm)	Wingate relative peak power (rPP) (kg W^{-1})
Ten-meter sprint race (SP10) in seconds	Squat jump (SJ) (cm)
Thorax double skin density (mm)	Handball sport-skill test (HSST) in seconds
Twenty-meter sprint race(SP20) in seconds	Squat jump on roes (SJT) in centimeters
Shuttle run (SR) in kilometers per hour	–

involving two or more variable quantities was created. Any time-related information was not recorded in the dataset. Nine skin fold points have been chosen (upper arm middle front, upper arm middle back, lower shoulder blade, supraspinal, thorax, midriff, lineaxillaris, upper leg, and anterior calf), and these points have been quantified in compliance with the recommendations in [15]. Quetelet index (QI) can be formulated by taking measurements of player's weight and height into consideration. It is expressed in Eq. (6.1) by its formula.

$$\text{QI (or) BMI} = \frac{m}{h^2} \quad (6.1)$$

where ' m ' is the weight of the players measured in kilograms (kg), and ' h ' is the height of the players measured in meters (m).

Using the Wingate test, Bar-Or [16] registered normal force, relative normal force, top force, and relative pinnacle strength. On a 20 m straight track, the speed was measured, and information was registered at 10 and 20 m. For any athlete, experiments were rehashed numerous times for time intervals of minutes. For the test times of 10 m (SP10) and 20 m (SP20), the fastest time was recorded [17]. Following the guidelines of Iacono et al. [17], the handball sport-skill test (HSST) was conducted. For each athlete with a 60s gap between tests, squat jumps (SJ) and squat jumps on toes (SJT) measurements were carried out on a piezoelectric force plate 3 times, and the best jump was recorded as per Moir [18].

Machine Learning Models

The four simplest standard ML models considered in this research for the expectation and investigation of the abilities of players are SLR, CT, SVR, and RBFNN.

Simple Linear Regression (SLR)

One of the critical and fundamental models of ML for prognostication is SLR. It is mainly used for the prognostication of information that generally has a linear association between its features and occurrences. For a 'Z' tagged dataset $(p_i, q_i)_{i=1}^Z$, where 'Z' is the total data size, p_i denotes input vector, and q_i denotes output vector. Equation (6.2) represents the general linear regression model.

$$f_{v,a}(p) = vp + a \quad (6.2)$$

$f_{v,a}(p)$ is a linear mixture of attributes of instance p , v is an N-directional tensor, and a is an actual numerical integer.

Classification Tree (CT)

This model has a tree form that is used for both prediction and classification problems. The CT starts at the root node and ends at the terminal nodes. It is possible to classify each terminal node based on its content. The computational time is reduced using decision trees. Multiple classification trees can be obtained from the training data and set of features. ID3, entropy, and Gini are several approaches to obtain optimal classification trees for classification problems that can give better and precise outcomes than others. In prognostication domains, the SE is utilized to find out the most critical and useful feature, which supplies the level of defects, where a lower level of defects denotes a more powerful and competent node. SE can be formulated as in Eq. (6.3).

$$SE = \frac{1}{M} \sum_{b=1}^M (z_u - \mu)^2 \quad (6.3)$$

M is the total occurrences, z_u is the tagged occurrences, and μ is the average of all tagged occurrences.

Support Vector Regression (SVR)

To conclude prediction problems, SVR is an altered form of support vector machine (SVM) that uses inputs instead of simply binary outputs. Input attributes are connected to a greater level by SVR, and the prognosis of nonlinear information is made possible. It generates a subgroup of the data points that are closer to the hyperplane from the input data and diminishes the margin of the classifier. The general SVR equation is denoted in Eq. (6.4).

$$S(v) = \sum_{g=1}^Z (\alpha_g^* - \alpha) n \cdot (v_g, v) \quad (6.4)$$

α_g^* and α are multipliers for finding local maxima and minima of function, and n is the kernel method.

Neural Networks that utilizes Radial Basis Function (RBFNN)

The main intention of this model is extracted from the concept of function approximation. This model differs from other models in the evaluation process done in the hidden layers. In this model, weights are calculated by applying the Euclidean distance to the radial basis function (RBF). Equation (6.5) represents Euclidean distance, and Eq. (6.6) formulates radial basis function, respectively.

$$d_q = \sqrt{\sum_{p=1}^Z f_p - u_{pq}^2} \quad (6.5)$$

where 'f' represents input data, and 'u' indicates hidden neuron's weight, consequently.

$$\phi = e^{\frac{-d^2}{2\sigma^2}} \quad (6.6)$$

The radius of the Gaussian curve is denoted by $\sigma > 0$, and d is the radial distance mentioned in Eq. (6.5).

The output calculation in this model is given by Eq. (6.7).

$$O(f) = \sum_{p=1}^M u_p \phi \quad (6.7)$$

$O(f)$ represents the model's output, 'M' represents radial basis function's count, and u_p depicts weights, respectively.

Results and Discussion

Two separate methods were taken into consideration to carry out observations, firstly discovering and getting the best and efficient prognosis model and grades for the woman players, and then to indicate the most critical attributes influencing the efficiency of the outperformed model, i.e., the ability of the athlete for the competency considered. To forecast the abilities of woman handball players, four separate fitness events, particularly a squat jump (SJ), squat jump on toes (SJT), sprint over a 10-m

distance (SP10), and a *handball sport-skill test (HSST)* were considered. To minimize the intricacy of the data and to improve the prognosis accuracy of the ML models, all occurrences were scaled by min–max scaling. In Eq. (6.8), the min–max scaling formula is specified.

$$N_p = \frac{M_p - \min(M)}{\max(M) - \min(M)} \quad (6.8)$$

The scaled value is denoted by N_p , the data point is given by M_p , and lowest and highest values are given by $\min(M)$ and $\max(M)$ for the respective features.

The four ML models referred to in section 3(B) were independently trained in each fitness-related case using 80% of the cumulative occurrences of the remaining 23 features. After the adjustment of the attributes, end-most measurements were made by tenfold cross-validation, and 20% of the untrained instances of the dataset were used for testing. Three primary parameters were used to analyze the performance of all the models in this experiment: R^2 score, SE, and mean AE. Using mean squared error, the R-squared score is a mathematical method employed to assess the association between examined and estimated data, which is the quantification of how far a data item appears to diverge from its expected value. In Eq. (6.9), the simple R-squared score formula is given.

$$R^2 = 1 - \frac{\sum(e_p - \hat{e}_p)}{\sum(e_p - \bar{e}_p)} \quad (6.9)$$

Examined data is denoted by e_p , the expected value is given by \hat{e} , and the average value of all observed data is given by \bar{e} , consequently.

SE is the square average of the errors that have been accumulated from the residual of the values observed and the values predicted. In Eq. (6.3), the SE formula is already given. AE is the average of absolute errors that can be quantified based on the residue of the data detected and predicted. In Eq. (6.10), the AE formula is given.

$$AE = \frac{1}{n} \sum_{i=1}^n |x_i - x| \quad (6.10)$$

where n denotes total defects, and $|x_i - x|$ represents the AE between the data observed and expected. Inefficiency evaluation of four features, namely an SJ, SJT, SP10, and HSST, four observations were conducted individually to figure out the superior variant of the considered ML models. The performance trait of woman handball players, namely SJ, was considered in the first study to forecast the abilities of players. CT had given the worst results with an R-squared score of 0.10. SLR and SVR gave very nearer outcomes of 0.707 and 0.660 for the R-squared score. Furthermore, SVR minimized the deviation. The performance trait of woman handball players, namely SJT, was considered in the second study to forecast the abilities of players. The RBFNN gave efficient results of 0.96, 0.0042 and 0.0075 for R-squared score, SE,

Table 6.2 Results of all four ML models

Skill	Metric	RBFNN	CT	SLR	SVR
SJ	SE	0.0013	0.033	0.0865	0.011
	AE	0.0033	0.1458	0.107	0.0891
	R ² score	0.97	0.10	0.707	0.66
SJT	SE	0.0042	0.030	0.018	0.020
	AE	0.0075	0.1153	0.1131	0.1111
	R ² score	0.96	0.65	0.61	0.63
SP10	SE	0.0034	0.024	0.0094	0.012
	AE	0.0316	0.1297	0.0692	0.0925
	R ² score	0.86	0.286	0.80	0.68
HSST	SE	0.0033	NA	NA	NA
	AE	0.0674	NA	NA	NA
	R ² score	0.93	NA	NA	NA

and AE. SP10 was predicted in the third analysis, and CT (0.286, 0.024 and 0.1297) received the lowest R-squared score and the best SE and AE results.

The largest prognostication was indeed achieved by the RBFNN in the last study, which was the HSST. Table 6.2 shows the overall results of 4 ML models by considering 3 performance metrics (R² score, SE, and AE) against 4 skills (SJ, SJT, SP10, and HSST) of female handball players. Figures 6.1, 6.2 and 6.3 represent the graphical comparison of the four ML models against performance metrics R² score, SE, and AE results, respectively.

Fig. 6.1 Comparison of 4 ML models against R² score

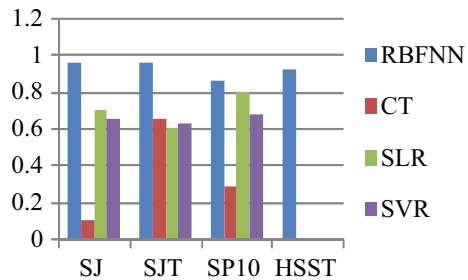


Fig. 6.2 Comparison of 4 ML models against SE values

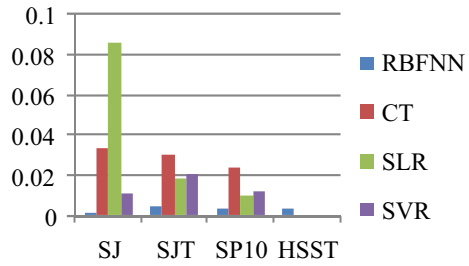
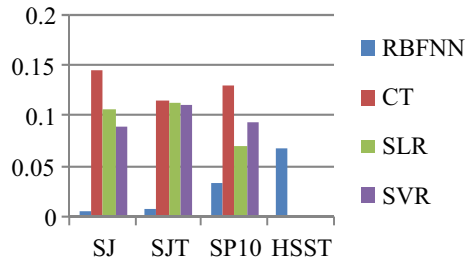


Fig. 6.3 Comparison of 4 ML models against AE values



Conclusion

Analyzing the abilities and performance of athletes before selecting a game is crucial. Athlete’s performance depends on several factors and predicting which factors are important and trivial is highly impossible. Identifying the trivial parameters can help coaches to get the pros and cons of the players so that the coaches can arrange extra training sessions for weak performers. This will also help the coaches to select the best players for the final team preparation. ML models as predictive procedures are likely the most significant methods of performing expectations for complex and sophisticated tasks. Outcomes from this study demonstrated that in women handball athletes with an ML model, in particular a radial base function neural network, it is conceivable to set up nonlinear connections for different body-related and exercise limits.

References

1. Jovanovic, S., Markovic, S., Kleva, M.: Prediction of successful defense movement of female handball players. *Sport Scientific and Practical Aspects*. **17**(1) (2020)
2. Gomez-Lopez, M., Manzano-Sanchez, D., Merino-Barrero, J.A., Valero-Valenzuela, A.: Causes of success in handball through the beliefs about ability. *Revista Internacional de Medicina y Ciencias de la Actividad Fisica y del Deporte*. **20**(77) (2020)
3. Popa, D., Mindrescu, V., Iconomescu, T.M., Talaghir, L.G.: Mindfulness and self-regulation strategies predict performance of romanian handball players. *Sustainability* **12**(9), 3667 (2020)

4. Hermassi, S., Souhail, C.M., Fieseler, G., Schulze, S., Irlenbusch, L., Delank, K.S., Schwesig, R., Hoffmeyer, B.: Validity of new handball agility test: association with specific skills and muscular explosive determinants of lower limbs in young handball players. *DRASSA J. Developm. Res. Sport Sci. Activities*. **3**(1), 79–103 (2017)
5. Sekulic, D., Foretic, N., Gilic, B., Esco, M.R., Hammami, R., Uljevic, O., Versic, S., Spasic, M.: Importance of agility performance in professional futsal players; reliability and applicability of newly developed testing protocols. *Int. J. Environ. Res. Public Health* **16**(18), 3246 (2019)
6. Soslu, R., Ozkan, A., Goktepe, M.: The relationship between anaerobic performances, muscle strength, hamstring/quadriceps ratio, agility, sprint ability and vertical jump in professional basketball players. *BedenEgitimiveSporBilimleriDergisi* **10**(2), 164–73 (2016)
7. Csato, L.: Optimal tournament design: lessons from the men’s handball Champions League. *J. Sports Econ.* **21**(8), 848–868 (2020)
8. Trejo Silva, A., Camacho Cardenosa, A., Camacho Cardenosa, M., Gonzalez Ramirez, A., BrazoSayavera, J.: Offensive performance under numerical inequality during exclusions in female handball. *RICYDE. RevistaInternacional de Ciencias del Deporte*. **62**(16), 396–409 (2020)
9. Pobar, M., Ivasic-Kos, M.: Active player detection in handball scenes based on activity measures. *Sensors* **20**(5), 1475 (2020)
10. Martinez-Rodriguez, A., Martinez-Olcina, M., Hernandez-Garcia, M., Rubio-Arias, J.A., Sanchez-Sanchez, J., Sanchez-Saez, J.A.: Body composition characteristics of handball players: systematic. In: *Camp 2020 Jan*, pp. 52.
11. Bonnet, G., Debanne, T., Laffaye, G.: Toward a better theoretical and practical understanding of field player’s decision-making in handball: A systematic review. *Movement Sport Sci.* **4**, 1–9 (2020)
12. Groll, A., Heiner, J., Schauburger, G., Uhrmeister, J.: Prediction of the 2019 IHF World men’s handball championship—a sparse Gaussian approximation model. *J. Sports Anal.* **6**(3), 187–197 (2020)
13. Niksic, E., Beganovic, E., Joksimovic, M.: The impact of the program of basketball, volleyball and handball on the situation-motorized capability of the first classes of the elementary school. *Pedagogy of Physical Culture and Sports* **24**(2) (2020)
14. Trebinjac, S., Nair, M.K.: Injury mechanisms in sports. In: *Regenerative Injections in Sports Medicine*, pp. 7–16. Springer, Singapore (2020)
15. Pescatello, L.S., Riebe, D., Thompson, P.D., (eds).: In: *ACSM’s Guidelines For Exercise Testing and Prescription*. Lippincott Williams & Wilkins (2014)
16. Bar-Or, O.: The Wingate anaerobic test an update on methodology, reliability and validity. *Sports Med.* **4**(6), 381–394 (1987)
17. Iacono, A.D., Eliakim, A., Meckel, Y.: Improving fitness of elite handball players: small-sided games vs. high-intensity intermittent training. *J. Strength and Conditioning Res.* **29**(3), 835–843 (2015)
18. Moir, G.L.: Three different methods of calculating vertical jump height from force platform data in men and women. *Measuram. Phys. Educ. Exercise Sci.* **12**(4), 207–218 (2008)