OTJHS
Online Turkish Journal of Health Sciences

e-ISSN: 2459-1467

OTSBD
Online Türk Sağlık Bilimleri Dergisi

Online Turkish Journal of Health Sciences 2024;9(1):14-20 *Online Türk Sağlık Bilimleri Dergisi 2024;9(1):14-20*

Determining the Landing Error Scoring System after a Jump by Artificial Intelligence

Sıçramadan Sonra Yere İniş Hata Puanlama Sistemi'nin Yapay Zeka ile Belirlenmesi

[1]Sabriye ERCAN, [2]Ahmet Ali SÜZEN, [3]Ferdi BAŞKURT, [3]Zeliha BAŞKURT

[1]Suleyman Demirel University, Medicine Faculty, Department of Sports Medicine, Isparta, Türkiye
[2]Isparta University of Applied Sciences, Department of Information Security Technology, Isparta, Türkiye
[3]Suleyman Demirel University, Health Sciences Faculty, Department of Physiotherapy and Rehabilitation, Isparta, Türkiye

Sabriye Ercan: https://orcid.org/0000-0001-9500-698X
Ahmet Ali Süzen: https://orcid.org/0000-0002-5871-1652
Ferdi Başkurt: https://orcid.org/0000-0002-8997-4172
Zeliha Başkurt: https://orcid.org/0000-0001-7488-9242

ABSTRACT

Objective: The study aims to examine the predictability of the Landing Error Scoring System (LESS) results after the jump with the Adaptive Boosting (AdaBoost) algorithm.

Materials and Methods: A model has been developed by artificial intelligence to shorten the scoring system significantly. In the data preprocessing stage, 17 different items contained in the original dataset were reduced to 13. A total of 3790 data items were included in the dataset used in the study, and the dataset was divided into 4 different sub-datasets. AdaBoost was chosen to give the highest accuracy tested in five different machine learning used for regression. The model's reliability was evaluated by testing the proposed AdaBoost model with performance metrics.

Results: The error score given by the clinician in the LESS was in the range of 0-86.6%. Recommended AdaBoost model for Sub_1, Sub_2, Sub_3, and Sub_4 respectively 98%, 87%, 88%, 89% accuracy has been achieved.

Conclusions: The score given to the LESS's 8^{th}, 10^{th}, 16^{th}, and 17^{th} items can be predicted with high accuracy, and the total score can be reached through the model proposed in the research.

Keywords: AdaBoost model, artificial intelligence, dataset, jump, Landing Error Scoring System

ÖZ

Amaç: Çalışmada, Adaptive Boosting (AdaBoost) algoritması ile Sıçramadan Sonra Yere İniş Hata Puanlama Sistemi (SSYİ-HPS) sonuçlarının öngörülebilirliğinin incelenmesi amaçlanmıştır.

Materyal ve Metot: Puanlama sistemini daha da kısaltmak için yapay zeka yardımıyla bir model geliştirilmiştir. Veri ön işleme aşamasında, orijinal veri setinde yer alan 17 farklı madde 13'e düşürülmüştür.
Çalışmada kullanılan veri setinde toplam 3790 veri yer almış ve veri seti 4 farklı alt veri setine ayrılmıştır. Regresyon için kullanılan beş farklı makine öğrenim modelinden en yüksek doğruluğu veren AdaBoost seçilmiştir. Modelin başarısı, önerilen AdaBoost modelinin performans metrikleri ile test edilmesiyle değerlendirilmiştir.

Bulgular: SSYİ-HPS'de klinisyen tarafından verilen hata puanı %0-86,6 aralığındaydı. Önerilen AdaBoost modelinde sırasıyla Sub_1, Sub_2, Sub_3 ve Sub_4 için %98, %87, %88, %89 doğruluk sağlanmıştır.

Sonuç: Araştırmada önerilen model ile SSYİ-HPS'nin 8., 10., 16. ve 17. maddelerine verilen puan yüksek doğrulukla tahmin edilebilmekte ve toplam puana ulaşılabilmektedir.

Anahtar Kelimeler: AdaBoost modeli, Sıçramadan Sonra Yere İniş Hata Puanlama Sistemi, veri seti, yapay zeka, sıçrama

Sorumlu Yazar / Corresponding Author:
Sabriye ERCAN
Suleyman Demirel University, Medicine Faculty, Department of Sports Medicine, Isparta, Türkiye
Tel: +90 246 211 92 48
E-mail: sabriyeercan@gmail.com

Yayın Bilgisi / Article Info:
Gönderi Tarihi/ Received: 21/06/2023
Kabul Tarihi/ Accepted: 21/12/2023
Online Yayın Tarihi/ Published: 11/03/2024

Atıf / Cited: Ercan S and et al. Determining the Landing Error Scoring System after a Jump by Artificial Intelligence. *Online Türk Sağlık Bilimleri Dergisi* 2024;9(1):14-20. doi: 10.26453/otjhs.1296565

INTRODUCTION

Assessment of biomechanical risk factors plays a key role in protecting against sports injuries.[1-3] Although three-dimensional (3D) motion analysis systems are shown as the gold standard the development of 2D motion analysis systems has been brought to the agenda. The widespread use of digital video cameras and software has also popularized the use of 2D motion analysis systems.[1,3-5] In addition, Padua et al.[6] has found that the results obtained in 2D motion analysis systems are valid and reliable with 3D motion analysis systems, which also increases confidence in these systems.[7,8]

Following the Landing Error Scoring System (LESS) protocol, the test sequence is asked to land on the ground by making a bilateral 'drop vertical jump' at the determined length.[6,9-11] From the images at the front and side camera angles where the landing on the ground is recorded after the jump, the error status of movements can be scored.[6]

The LESS: users risk analysis, neuromuscular training, post-development monitoring, etc.[7] in conjunction with the offering, this system for motion analysis in the analysis of each athlete in the image of an experienced evaluator, and there is a need for at least 30 minutes. On the other hand, it is predicted that this scoring process can be achieved in a much shorter time and independent of experience with artificial intelligence (AI) techniques. It is thought that AI methods[12-15] in the field of health and sports can be used to make this system more practical.

According to the information we have obtained from the previous research studies, AI methods are not used to estimate the LESS scoring. The purpose of this study is to examine the predictability of the LESS score with AI methods.

MATERIALS AND METHODS

Ethics Committee Approval: The study was approved by the Isparta University of Applied Sciences Ethics Committee (Date: 23.03.2021, decision no: 3). The study was planned under the Helsinki Principles. The results of 112 people (21.7±1.2 years, 54.5% male, 45.5% female) were evaluated. To evaluate the results of the LESS with AI techniques and to develop a model, they were applied.

Data Preprocessing: Seventeen different items contained in the original dataset[6] were reduced to 13. This inference on the dataset is determined by the following inference.

- S7. and S8. substance affects the response to each other.
- S9. and S10. substance affects the response to each other.
- The outcome of substance S12., S13. and S14. determines the outcome of substance 16.
- The outcome of substance S5. and S16. determines the outcome of substance S17.

Thirteen input and 4 output parameters were determined in the dataset (3790 items) with feature extraction. Since the number of items affected by the determined inferences is different, the dataset is divided into 4 different sub-datasets. Sub_1 dataset was 224 counts. Sub_1 dataset's classification was Substance 8, and the classification type was 0-1-Null. The Sub_2 dataset was 224 counts. Sub_2 dataset's classification was Substance 10, and the classification type was 0-1-Null. Sub_3 dataset was 502 counts. Sub_3 dataset's classification was Substance 16, and the classification type was 0-1-2. The Sub_4 dataset was 336 counts. The Sub_4 dataset classified Substance 17, and the classification type was 0-1-2. As a result of this partitioning, 1286 data items were extracted for training and testing the model. Of these four sub-models, 80% of the dataset was used for training, and 20% was used for testing.

Development of the Model: According to two statistical concepts, model selection begins with predicting the performance of different models to choose the best model. According to the results, the generalization error is estimated, and the best model is evaluated.[16,17] Adaptive Boosting (AdaBoost) from ensemble learning algorithms was used in the proposed model (Figure 1). Four subsets of data are sent to the model separately. The AdaBoost model is trained

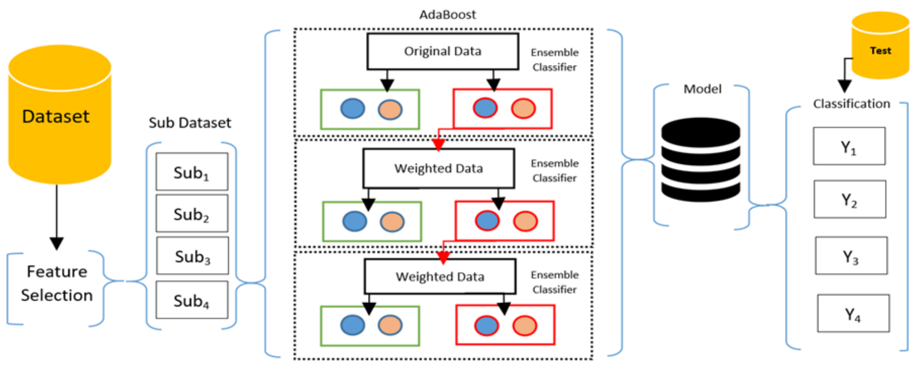

Figure 1. The proposed model.

and classified with initial training data. It then transfers the relative weight of misclassified training data to the next training. The second classifier model is trained with increased weights and classified again. In the third step, the weight is updated this way, and the consequences are created for the final model. In the last stage, the classification is completed by giving the model test data.[18]

First of all, in the mathematical structure of the model, the dataset is represented as . Where N is the size of the real numbers or the number of attributes in the dataset. X is the set of scoring data. Y is a target variable of 0, 1, or 2 because it is a triple classification problem. The same weights are used to train all data in the initial training phase of the model. The addition of weighted samples is always 1, as shown in Equation 1. For this reason, the value of each weight is between 0 and 1 in the first stage.

(1) $$w = \frac{1}{n} \in [0,1]$$

In the second step, using Equation 2 for this classifier, its actual effect on the classification of the scoring data is calculated. ε_t is the numerical value of how effective this step will be in the final classification. is the total number of incorrect classifications for the current training set divided by the training set size.

(2) $$\varepsilon_t = \frac{1}{2}\ln\frac{(1-\sum error)}{\sum error)}$$

After entering the actual values for each classification step, the weights, initially taken as 1/N for each data point, are updated according to Equation 3. Here, two cases occur for ε as plus and minus. The ε is positive when the predicted score and actual output match. In this case, the weight update does not occur. The ε value is negative when the predicted output does not match the actual score. In this case, the sample weight should be increased so that the same incorrect classification is not repeated in the next training. This process is repeated until the error function changes or the maximum limit of the classifier number is reached. The classification steps of the proposed model are shown in the rough code (Table 1).

(3) $$w_i = w_{i-1}e^{\pm\varepsilon}$$

Performance metrics for machine learning are used to evaluate the developed model. Performance metrics are used to evaluate training and test data estimation results. The ratio of correctly identified sam-

Table 1. Pseudocode of the classification algorithm of the model.

Input: Initial training dataset $N = \{x_i, y_i\}, y_i \in \{0,1,2\}$

(1) Initialize the sample weights $w_i = \frac{1}{m}$

(2) for u=1,2, 3,…., U do

$h_t = L(N, N_t)$ Train a sample from D with Dt $\varepsilon_t = p_x(h_t(x) \neq f(x))$ Evaluate the errors of ht

if $\varepsilon_t > 0.5$ go to (2)

(3) $\alpha_t = \frac{1}{2}\ln(\frac{(1-\varepsilon_t)}{\varepsilon_t}$ determine the weight of ht

end

(4) $$\begin{Bmatrix} \exp(-\alpha_t) & h_t(x) = f(x) \\ \exp(\alpha_t) & h_t(x) \neq f(x) \end{Bmatrix}$$

Output:

$$H(x) = sign\sum_{t=1}^{T} 1(y = h_t(x))$$

ples to total samples is considered by many academics to be the most plausible performance metric. By definition, accuracy (ACC) also functions in situations when there are more than two labels.[19-22] However, accuracy loses its reliability when the dataset is unbalanced, leading to an overly optimistic estimate of the classifier's performance on the majority class. The Matthews correlation coefficient (MCC) offers a useful remedy for the class imbalance problem.[19,21,22]

For the performance evaluation of the proposed model, ACC (Equation 1), Precision (Equation 2), Recall (Equation 3), and F1-score (Equation 4) are measured. Pseudocode of the classification algorithm of the model is below:

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

$$(2) \quad Precision = \frac{TP}{TP + FP}$$

$$(3) \quad Recall = \frac{TP}{TP + FN}$$

$$(4) \quad F1_{score} = 2 \times \left(\frac{Precision \times Recall}{Precision + Recall}\right)$$

Abbreviation in the formulas above: TP: True Positives; FP: False Positives; FN: False Negatives; TN: True Negatives.

Statistical Analysis: The SPSS v.23 package program was used for the analysis. Clinician' data were presented as frequency (n), percentile (%), mean±standard deviation.

RESULTS

The score of the LESS determined by the clinician was calculated as 6.8±2.1. The error score rate of item 1 (knee flexion angle at initial contact) was 86.6%. The error score rate of item 2 (hip flexion angle at initial contact) was 0%. The error score rate of item 3 (trunk flexion angle at initial contact) was 48.2%. The error score rate of item 4 (ankle plantarflexion angle at initial contact) was 9.8%. The error score rate of item 5 (knee valgus angle at initial contact) was 14.3%. The error score rate of item 6 (lateral trunk flexion angle at initial contact) was 7.1%. The error score rate of item 7 (stance width–wide) was 0%. The error score rate of item 8 (stance width–narrow) was 70.5%. The error score rate of item 9 (foot position–toe in) was 0%. The error score rate of item 10 (foot position–toe out) was 17%. The error score rate of item 11 (symmetric initial foot contact) was 25.9%. The error score rate of item 12 (knee flexion displacement) was 33%. The error score rate of item 13 (hip flexion at max knee) was 0%. The error score rate of item 14 (trunk flexion at max knee flexion) was 31.3%. The error score rate of item 15 (knee valgus displacement) was 69.6%. The error score rate of item 16 (joint displacement) was 84.8% (35.7%: 1 point, 49.1%: 2 points). The error score rate of item 17 (overall impression) was 98.2% (60.7%: 1 point, 37.5%: 2 points).

The model was developed in Spyder software with Python language. The training and testing of the model were completed on an AI machine with an I9 processor and a 24 GB video card. The confusion matrix of the classification of 4 different scores in different intervals in the dataset is shown in Figures 2a, b, c, and d. The scoring result density in the Sub_1 dataset is 1, so 98% of the model has correctly classified the result 1 (Figure 2a). It is seen that the classification results are close to each other (0-88%, 1-84%) as the scoring result density in the Sub_2 dataset is approximately equal (Figure 2b). The classification success was similar due to the equal distribution of the scoring result density in the Sub_3 data (Figure 2c). It is seen that the scoring result density in the Sub_4 data is almost all 1 and 2, so the result is classified as 1 and 2 (Figure 2d).

ACC and MCC performance criteria were used to evaluate the performance of the classification model.[19,21-23] Accuracy, Precision, Recall, and F1-score values were calculated with TP, TN, FP, and FN values in the confusion matrix shown in Figure 2. Accordingly, Accuracy, Precision, Recall, and F1-score values obtained in each dataset and the average success of the model are given in Table 2.

After the model's training and testing process, test software was developed with the C # programming language. The trained file of the AdaBoost model was saved in Keras software with the h5 format. Then, the model was run by loading it into the test

Table 2. Metric values from scoring classification and comparison.

Dataset	AdaBoost Accuracy	AdaBoost Precision	AdaBoost Recall	AdaBoost F1score	K-Nearest Neighbors Accuracy	Support Vector Machine Accuracy	Decision Trees Accuracy	Gaussian process regression Accuracy
Sub_1	0.98	0.97	0.92	0.95	0.91	0.85	0.95	0.92
Sub_2	0.87	0.86	0.85	0.86	0.85	0.87	0.86	0.83
Sub_3	0.88	0.88	0.87	0.88	0.87	0.88	0.84	0.83
Sub_4	0.89	0.89	0.86	0.87	0.88	0.82	0.87	0.84
Avg	0.90	0.89	0.87	0.89	0.87	0.85	0.88	0.85

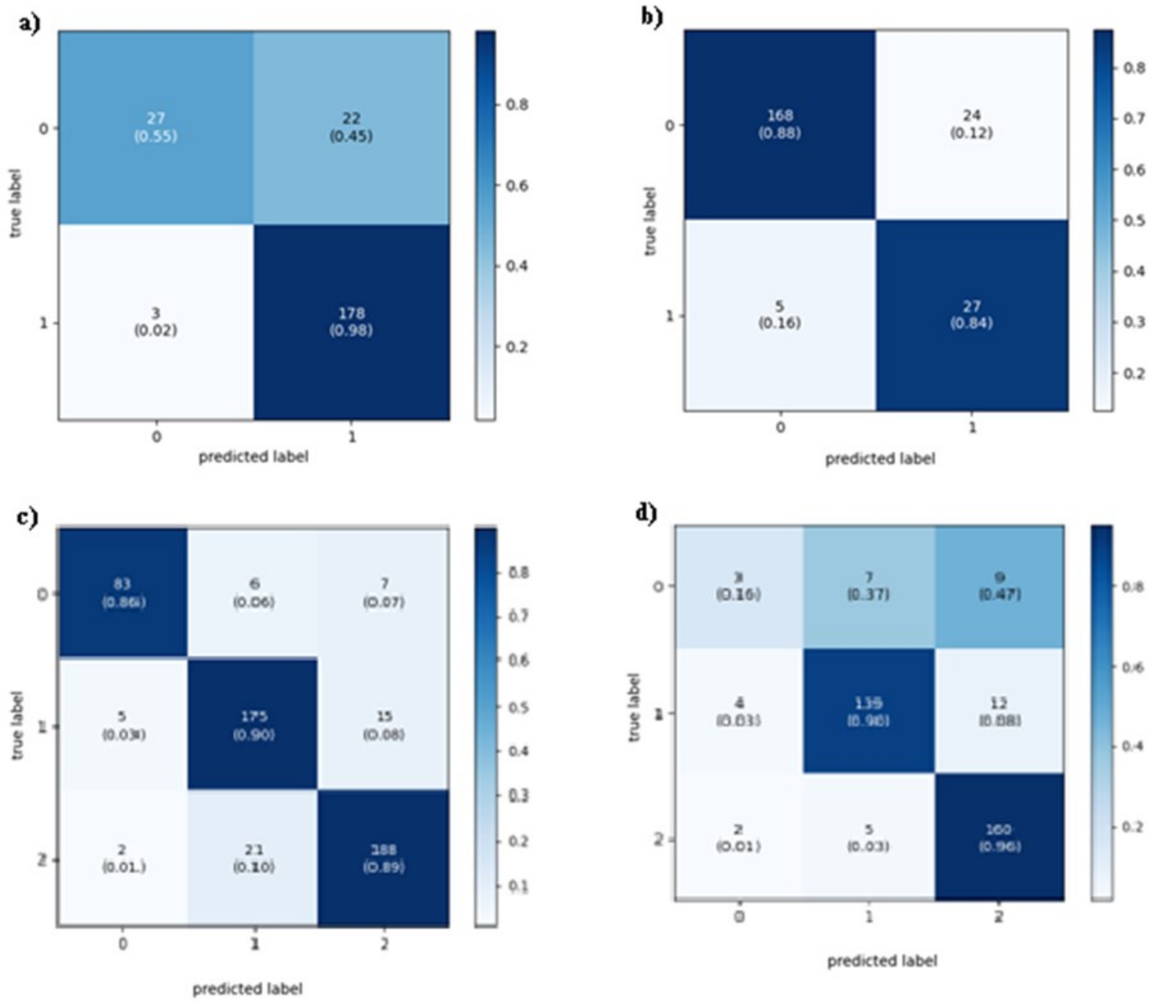

Figure 2. Confusion matrix of the score classification model for 4 datasets.

software. After the data entry of 13 items from the test results, the model estimates for 4 items. After the model estimates, it also calculates the total score for expert evaluation.

DISCUSSION AND CONCLUSION

The original scoring system of 17 items could be shortened to 13 items using AI methods. It was ensured that items 16th and 17th, whose scoring may vary depending on experience, could be scored easily and accurately using AI methods. The score to be given to the LESS's 8th, 10th, 16th, and 17th items can be predicted with high accuracy, and the total score can be reached with the proposed model.

It was observed that an attempt was made to easily develop evaluation methods/tools with the help of automated systems, such as the markerless motion-capture system, to score the LESS.[11,24] But after the jump with automated systems, the 17th item of the LESS (Overall impression item) was excluded from the analysis because it could not be evaluated.[11,24] In our research, the predictability of substances shortened by the model we proposed without any original substances being excluded from the analysis was high. The ability to predict the substances (items 16th and 17th) that experience will come into play with our proposed model has created an advantage.

Technology usage areas of the sports industry cover a wide spectrum, such as health, education, and tourism.[25] Another fact that technological progress has brought into our lives is AI.[26] AI is a system capability that will help to shorten the LESS with its feature of helping motion analysis[27] and supporting decision -making processes[26] without compromising its reliability. As demonstrated in our study, the fact that the motion analysis processes of AI systems provide convenience to the rater in the decision-making process will make the motion analysis systems more common and user-friendly.

In conclusion, the score given to the 8th, 10th, 16th, and 17th items of LESS can be estimated with a high accuracy rate, and the total score can be reached. In this way, in addition to providing ease of use to researchers who will use the LESS, 16th and 17th items can be scored easily and with significant accuracy using AI methods. In addition, the fact that the error scores in the dataset studied were relatively high (3-11 points) was considered a limitation of the study.

Ethics Committee Approval: Our study was approved by the Isparta University of Applied Sciences (Date: 23.03.2021, decision no: 3). The study was carried out following the international declaration and guidelines.
Conflict of Interest: No conflict of interest was declared by the authors.
Author Contributions: Concept – SE, AAS, FB, ZB; Supervision – FB, ZB; Materials – AAS; Data Collection and/or Processing – SE; Analysis and/or Interpretation – AAS; Writing – SE, AAS.
Peer-review: Externally peer-reviewed.

REFERENCES

1. Alahmari A, Herrington L, Jones R. Concurrent validity of two-dimensional video analysis of lower-extremity frontal plane of movement during multidirectional single-leg landing. Phys Ther Sport. 2020;42:40-45. doi:10.1016/j.ptsp.2019.12.009
2. Padua DA, DiStefano LJ, Beutler AI, De La Motte SJ, DiStefano MJ, Marshall SW. The landing error scoring system as a screening tool for an anterior cruciate ligament injury–prevention program in elite-youth soccer athletes. J Athl Train. 2015;50(6):589-595. doi:10.4085/1062-6050-50.1.10
3. James J, Ambegaonkar JP, Caswell SV, Onate J, Cortes N. Analyses of landing mechanics in division I athletes using the landing error scoring system. Sports Health. 2016;8(2):182-186. doi:10.1177/1941738115624891
4. Peebles AT, Arena SL, Queen RM. A new method for assessing landing kinematics in non-laboratory settings. Phys Ther Sport. 2021;49:21-30. doi:10.1016/j.ptsp.2021.01.012
5. Rabin A, Einstein O, Kozol Z. Agreement between visual assessment and 2-dimensional analysis during jump landing among healthy female athletes. J Athl Train. 2018;53(4):386-394. doi:10.4085/1062-6050-237-16
6. Padua DA, Marshall SW, Boling MC, Thigpen CA, Garrett JrWE, Beutler AI. The landing error scoring system (LESS) is a valid and reliable clinical assessment tool of jump-landing biomechanics: the JUMP-ACL study. Am J Sports Med. 2009;37(10):1996-2002. doi:10.1177/0363546509343200
7. Hanzlíková I, Athens J, Hébert-Losier K. Factors influencing the landing error scoring system: Systematic review with meta-analysis. J Sci Med Sport. 2021;24(3):269-280. doi:10.1016/j.jsams.2020.08.013
8. Hanzlíková I, Hébert-Losier K. Is the landing error scoring system reliable and valid? A systematic review. Sports Health. 2020;12(2):181-188. doi:10.1177/1941738119886593
9. Beese ME, Joy E, Switzler CL, Hicks-Little CA. Landing error scoring system differences between single-sport and multi-sport female high school–aged athletes. J Athl Train. 2015;50(8):806-811. doi:10.4085/1062-6050-50.7.01
10. Smith HC, Johnson RJ, Shultz SJ, et al. A prospective evaluation of the landing error scoring system (LESS) as a screening tool for anterior cruciate ligament injury risk. Am J Sports Med. 2012;40(3):521-6. doi:10.1177/0363546511429776
11. Dar G, Yehiel A, Cale'Benzoor M. Concurrent criterion validity of a novel portable motion analysis system for assessing the landing error scoring system (LESS) test. Sports Biomech. 2019;18(4):426-436. doi:10.1080/14763141.2017.1412495
12. Fister I, Fister D, Deb S, Mlakar U, Brest J. Post hoc analysis of sport performance with differential evolution. Neural Comput & Applic. 2020;32:10799-10808
13. Rajšp A, Fister I. A systematic literature review of intelligent data analysis methods for smart sport training. Appl Sci. 2020;10(9):3013. doi:10.3390/app10093013
14. Rigamonti L, Albrecht UV, Lutter C, Tempel M, Wolfarth B, Back DA. Potentials of digitalization in sports medicine: a narrative review. Curr Sports Med Rep. 2020;19(4):157-163. doi:10.1249/JSR.0000000000000704
15. Schmidt SL. 21st Century Sports: How Technologies Will Change Sports in the Digital Age. 1st ed. Cham, Switzerland: Springer Nature; 2020.
16. Emmert-Streib F, Dehmer M. Evaluation of regression models: Model assessment, model selection and generalization error. Mach Learn Knowl Extr. 2019; 1(1): 521-551. doi:10.3390/make1010032
17. Haste T, Tibshirani R, Friedman J. The elements of statistical learning: Data mining, inference and prediction. New York, USA: Springer; 2009.
18. Wang R. AdaBoost for feature selection, classification, and its relation with SVM, a review. Physics Procedia. 2012;25:800-807. doi:10.1016/j.phpro.2012.03.160
19. Chicco D, Jurman G. The advantages of the Matthews correlation coefficient (MCC) over F1

score and accuracy in binary classification evaluation. BMC genomics. 2020; 21(6): 1-13. doi: 10.1186/s12864-019-6413-7

20. Phasinam K, Mondal T, Novaliendry D, Yang CH, Dutta C, Shabaz M. Analyzing the performance of machine learning techniques in disease prediction. J Food Qual. 2022; 2022: 1-9. doi.org/10.1155/2022/7529472
21. Akosa JS. Predictive accuracy: a misleading performance measure for highly imbalanced data. In: Proceedings of the SAS Global Forum 2017 Conference. Cary, North Carolina: SAS Institute Inc.; 2017: 942–2017.
22. Baldi P, Brunak S, Chauvin Y, Andersen C, Nielsen H. Assessing the accuracy of prediction algorithms for classification: an overview. Bioinformatics. 2000; 16(5): 412–424.
23. Gensler A, Sick B. Novel criteria to measure performance of time series segmentation techniques. In LWA. 2014: 193-204.
24. Mauntel TC, Padua DA, Stanley LE, et al. Automated quantification of the landing error scoring system with a markerless motion-capture system. J Athl Train. 2017;52(11):1002-1009. doi:10.4085/1062-6050-52.10.12
25. Ratten V. Sport technology: A commentary. J High Technol Manag Res. 2020;31(1):100383. doi:10.1016/j.hitech.2020.100383
26. Farrokhi A, Farahbakhsh R, Rezazadeh J, Minerva R. Application of internet of things and artificial intelligence for smart fitness: A survey. Computer Networks. 2021;107859. doi:10.1016/j.comnet.2021.107859
27. Taborri J, Molinaro L, Santospagnuolo A, Vetrano M, Vulpiani MC, Rossi S. A machine-learning approach to measure the anterior cruciate ligament injury risk in female basketball players. Sensors. 2021; 21(9):3141. doi:10.3390/s21093141