Prediction of Marathon Performance using Artificial Intelligence

Authors
Lucie Lerebourg[1], Damien Saboul[2], Michel Clémençon[1], Jérémy Bernard Coquart[1, 3]

Affiliations
1 Centre d'Etudes des Transformations des Activités Physiques et Sportives Normandie Univ, UNIROUEN, CETAPS, 76000 Rouen, France
2 Research and Innovation, Be-ys-research, Argonay, France
3 Unité de Recherche Pluridisciplinaire Sport, Santé, Société Eurasport, 413 avenue Eugène Avinée, 59 120 Loos, France

Key words
artificial neural networks, k-nearest neighbor, machine learning, endurance running, modeling

accepted 05.12.2022
published online 17.02.2023

Bibliography
Int J Sports Med 2023; 44: 352–360
DOI 10.1055/a-1993-2371
ISSN 0172-4622

Georg Thieme Verlag, Rüdigerstraße 14,
70469 Stuttgart, Germany

Correspondence
Miss Lucie Lerebourg
Université de Rouen
STAPS, CETAPS
Boulevard Siegfried
76821 Mont-Saint-Aignan
France
Tel.: 02 32 10 77 93, Fax: 02 32 10 77 93
lucie.lerebourg@univ-rouen.fr

ABSTRACT

Although studies used machine learning algorithms to predict performances in sports activities, none, to the best of our knowledge, have used and validated two artificial intelligence techniques: artificial neural network (ANN) and k-nearest neighbor (KNN) in the running discipline of marathon and compared the accuracy or precision of the predicted performances. Official French rankings for the 10-km road and marathon events in 2019 were scrutinized over a dataset of 820 athletes (aged 21, having run 10 km and a marathon in the same year that was run slower, etc.). For the KNN and ANN the same inputs (10-km race time, body mass index, age and sex) were used to solve a linear regression problem to estimate the marathon race time. No difference was found between the actual and predicted marathon performances for either method ($p > 0{,}05$). All predicted performances were significantly correlated with the actual ones, with very high correlation coefficients ($r > 0{,}90$; $p < 0{,}001$). KNN outperformed ANN with a mean absolute error of 2,4 vs 5,6 %. The study confirms the validity of both algorithms, with better accuracy for KNN in predicting marathon performance. Consequently, the predictions from these artificial intelligence methods may be used in training programs and competitions.

Introduction

The marathon, an athletic endurance event of 42,195 km, was created for the first modern Olympic Games in Athens in 1896. Since the first "urban tour" marathon in New York City in 1976 [1], the marathon has been gaining in popularity and has evolved from an Olympic event to a worldwide social phenomenon [2]. As enthusiasm for this event has increased [1–3], race times have steadily improved for the best runners (e. g., Top 100 world best performers in Boston Marathon between 1990 and 2010 in the study by Marc et al., and Top 100, Top 10 and winners from 1897 to 2017 in Knechtle at al. study) [3, 4]. From recreational runners of all ages to elite athletes, the objectives may differ widely, from being a finisher, to running the race as fast as possible, to winning it and/or breaking records (e. g., personal, national, world records) to win money (i. e., economic reasons) [5, 6]. Although long-distance performances, as in the marathon, can be influenced by factors beyond the athlete's control (e. g., climate conditions, seasonal characteristics like temperature, humidity and barometric pressure, etc.) [7, 8], they mainly depend on personal characteristics (age, sex, physical qualities, psychological traits and states, etc.) and training variables (tactics, pacing strategy, etc.) [4, 9–13]. For example, Weiss et al. [7] showed that temperature and humidity affect pac-

ing in age group marathoners differently (i. e., slowing down for runners of both sexes aged 20–59 with increasing temperature, and slowing down for runners aged under 20 and over 80 with increasing humidity). Other studies [12, 13] indicated that pacing strategy which may also be dependent on the profile of the runners in relation to age (e. g., pace changing is more prominent in younger and older marathoners compared to the other age groups of marathoners) [12] or sex (e. g., men tend to opt more for a "risk" strategy by starting out at fast speeds and then modulating or slowing down afterwards, whereas women tend to err on the side of caution) [13]. Athletes and coaches need to be aware of these parameters and should focus on developing appropriate training programs, with particular emphasis on setting speeds for tempo runs and building competitive or optimal pace strategies to optimize performance [14–18]. For these reasons, the ability to predict marathon performance can be of great interest in the calibration of training sessions and the definition of the athlete's potential speed limits in order to achieve the best performance.

The relationships between running distance (or speed) and time have long been used for this purpose [19–21]. Several studies have sought to predict long-distance running performances for events like the marathon with mathematical models (e. g., logarithmic, hyperbolic, exponential, multiple regression models, etc.) [18, 22–24], including concepts of critical speed [25, 26] or power laws [25, 27], and machine learning algorithms (i. e., artificial intelligence: AI) [14, 16, 17].

Recently, there has been growing interest in machine learning algorithms, notably with supervised learning, one of the intelligent methodologies that have shown promising results in the prediction of continuous variables in many areas such as weather [28], health [29] and sports [30]. The literature indicates that sport is one of the expanding areas requiring good predictive accuracy [16, 30, 31]. However, although machine learning regression models like artificial neural networks (ANN) [29, 31–33] or k-nearest neighbors (KNN) [14] have been used to predict performances in some sports activities, the validity and accurate prediction of individual or team performances using AI merits further exploration [34–37].

ANN is a powerful black-box supervised learning algorithm capable of producing nonlinear input-output mapping [34, 38, 39]. The model consists of one input layer, one or more hidden layers and one output layer. The interconnected components (i. e., neurons) transform a set of inputs into a desired output [38, 39]. The accuracy of this type of model is typically improved by using additional data (i. e., weights associated with interconnected components are continuously changing) during the ANN training process [30, 38]. On the other hand, the KNN model uses one of the simplest types of supervised machine learning algorithms based on learning by analogy, that is, by comparing a given test example with training examples that are similar to it [29, 38]. The basic KNN algorithm has two steps: find the k training examples that are closest ("closeness" is defined in terms of a distance metric, such as the Euclidean distance) to the unseen example and take the average of these k label values [34, 38]. This machine learning model is also noted for not requiring learning (i. e., the computation of the algorithm occurs during runtime) as it memorizes the training dataset [34, 36, 38]. Moreover, it seems that compared to ANN, KNN tends to perform better on datasets with a small number of samples and has less risk of overfitting [40].

Although studies have focused on the use of machine learning algorithms (bagging, local matrix completion, etc.) to predict marathon performance [14, 17, 22] and future slowdowns during the race [16], no study to the best of our knowledge has used and validated ANN and KNN supervised machine learning in this running discipline and compared the accuracy (i. e., the nearness of the actual performance to the predicted performance, and thus a lower mean absolute error or a lower bias, meaning higher accuracy) or precision (i. e., the closeness of the predicted performances, and thus a smaller distance between limits of agreement, meaning higher precision).

The objectives of the current study were therefore to test the validity of two supervised machine learning methods (ANN and KNN) and to compare the accuracy and precision of the marathon performance predictions to determine which one performed best. Based on the literature and our data, we believe that both artificial intelligence techniques will be valid, and that KNN will be the better performing method.

Materials and Methods

Experimental approach

All French official rankings of the French Athletics Federation (FFA for *Fédération Française d'Athlétisme*) for the 10-km road race (n = 217,669) and the marathon (n = 92,813), both performed in 2019, were retrospectively analyzed. In France, the marathon is not open to younger categories of athletes, so only athletes over the age of 21 years were selected for both races (n = 201,990 on the 10-km and n = 92,813 on the marathon). If the athletes had not self-reported their body mass and/or height, they were removed from the analysis. Thus, 7,716 athletes with a 10-km performance, and 4,130 in the marathon were included. Then, only those athletes (women and men) who performed the 10-km and the marathon in the same year (i. e., 2019) were retained. Thus, 1,728 performances were collected. However, as the aim was to predict marathon performance based on 10-km road performance, athletes who ran a marathon before their 10-km were removed (n = 833). Moreover, athletes who maintained a higher speed in the marathon than in the 10-km race were also eliminated (n = 11). Finally, among the 884 remaining athletes, those with a performance in the 10-km below the lowest ranking of the FFA were eliminated (i. e., performance > 50 and 60 min, respectively, for men and women).

This study was approved by the National Ethics Committee for Research in Sports Sciences (CERSTAPS2019220231) [41]. The protocol for this study was legally declared, in accordance with the European General Data Protection Regulations.

Participants

The analysis was thus performed with a dataset of 820 athletes. For each athlete, the sex (i. e., female vs male), date of birth (to calculate age), body mass and height (to calculate the body mass index: BMI), and race times (i. e., the performances on the 10-km and marathon) were recorded.

Procedures & data treatment

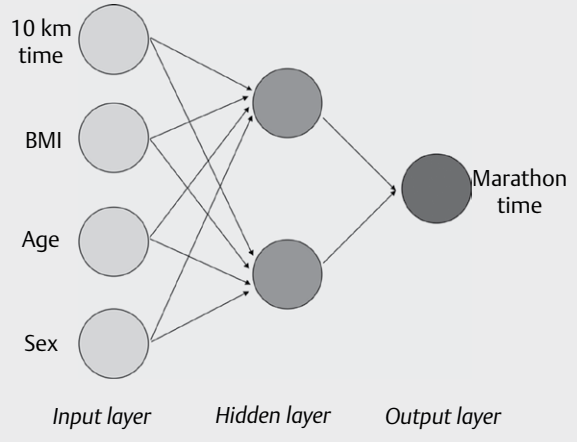
Two supervised machines learning regression algorithms were used: ANN and KNN [30, 31, 37–39]. Both algorithms were implemented in R language. R software (version R X64 3,6,1 – R Development Core Team, Vienna, Austria) was used for our analysis, and the following R packages for machine learning approaches were used: dplyr version 0,8,3, neural net version 1,44,2 [42, 43].

All data were normalized to meet the requirement of the sigmoid transfer function (for ANN) and to remove the scale differences between the input variables (for KNN).

The data from the 820 athletes were randomly separated into a random train/test split for training and testing processes, respectively. The 90:10 ratio was used, meaning that 90 % of the data was randomly selected for the training process (i. e., 738 out of the 820 performances), while 10 % of the data was randomly selected for the testing process (i. e., 82 performances) [44, 45].

For both supervised machine learning algorithms, the same inputs (10-km race time, BMI, age and sex) were used to solve the linear regression problem, which consisted in estimating the value of the same continuous output (marathon race time). We also specified that exactly the same training (n = 738) and testing (n = 82) data were used for the two algorithms.

In ANN, a multilayer perceptron was used with four inputs and one output (▶ **Fig. 1**) [37]. In this network, the computing units are arranged into three layers, which are conveniently ordered. The information flows forward from the four neurons of the input layer to the two connecting neurons of the hidden layer and, finally, to the single neuron of the output layer using no backward connection. The first layer (the input layer) corresponds to the independent variables (i. e., performance on 10-km, BMI, age and sex), while the third layer (the output layer) corresponds to the dependent variable score (marathon performance). The intermediate layer, which is the hidden layer, consists of all possible connections between the input and output layers and allows for the combined impact of a multiple set of independent variables on the output layer. This ANN makes use of Rprop, which is short for resilient back propagation, a training technique without weight backtracking for supervised learning, [43]. The training stopping point (i. e., threshold) was set at 0,01.



▶ **Fig. 1** Neural network architecture.; *Note*: BMI: body mass index.

To compare with the neural network algorithm, KNN was applied using the same four input variables and the same output variable for comparison with ANN. In this study, the KNN algorithm was tested by selecting the closest neighbors (k = 3). In other words, for each athlete of the testing dataset, we retained the three athletes of the training dataset having the smallest Euclidean distance (sum of the differences between their four respective inputs). The estimated output (marathon time) was calculated using the average of the marathon times of the three closest neighbors (athletes) weighted by the inverse of their respective Euclidean distances to the testing athlete.

Statistical analysis

Mean values and standard deviation (SD) of variables were calculated.

The Shapiro-Wilk test was used to test whether the data followed a normal (Gaussian) distribution. A Student paired samples t-test was used for normally distributed data to compare the actual and predicted marathon performances for each machine learning algorithm. When these data did not pass the test for normality, a Wilcoxon signed-rank test was used. The magnitude of the differences was assessed by the effect size (ES), which was classified using the Cohen scale [46].

The association between the actual and predicted performances was tested with Pearson's product-moment or Spearman's rank order correlations, depending on whether the data followed a normal Gaussian distribution. We considered a correlation of $r = 0{,}90$ or more as very high, between 0,70 and 0,89 as high, between 0,50 and 0,69 as moderate, and between 0,26 and 0,49 as low [47].

The coefficient of determination (r^2) and the mean absolute error (MAE) criteria based on a common 90:10 training/test data split were chosen to evaluate the numerical fit of the output from the ANN and KNN models. MAE is the average of the absolute errors, with lower error values typically meaning the model is more accurate and the predictions closely match the actual values [14, 29]. The r^2 value determines the precision of the predictions and how well the model fits the data [48]. This also makes it easier to compare and evaluate the results [31, 45].

Moreover, the bias (i. e., difference between actual and predicted performances, to access accuracy) and 95 % limits of agreement (95 % LoA, i. e., ± 1,96 SD, to access precision) were computed according to the Bland-Altman method.

Finally, the KNN and ANN models were compared to determine which one performed better. The outputs of the predicted marathon performances were verified with actual marathon performances. The model was considered valid if the MAE was less than 5 % and if the biases and approval limits were acceptable. The best model was selected on that which had the lowest MAE and the highest accuracy (from bias, i. e., the closeness of the actual performance to the predicted performance, and thus a lower average absolute error or bias, thus means a higher accuracy) and precision (from LoA, i. e., the proximity of the predicted performance, and thus a smaller distance between LoAs, means a higher accuracy) [49].

▶ **Table 1** Mean values and standard deviation (SD) of actual and predicted performances on marathon from each algorithm (*i. e.*, artificial neuronal network (ANN) and k-nearest neighbors (KNN)) in 82 performances athletes (*i. e.*, 10 % of the data is selected for testing process), difference between actual and predicted performances (*p*), magnitude of the difference, Pearson's product-moment with actual performance (*r*), bias and SD (min), bias and 95 % LoA (min and %) and mean absolute error (MAE) (%).

Distance		Mean performance (min) ± SD	p	Magnitude of difference		Correlation		Bias (min) ± 95 % LoA	Bias (%) ± 95 % LoA	MAE (%)
				ES	Interpretation	r	Interpretation			
Marathon	Actual	199,75 ± 36,01								
	Predicted from the ANN	202,77 ± 30,24	0,063	−0,091	Trivial	0,918*	Very high	3,023 ± 28,492	1,5 ± 14,1	5,6
	Predicted from the KNN	198,96 ± 32,74	0,333	−0,024	Trivial	0,982*	Very high	-0,79 ± 14,35	-0,4 ± 7,2	2,4

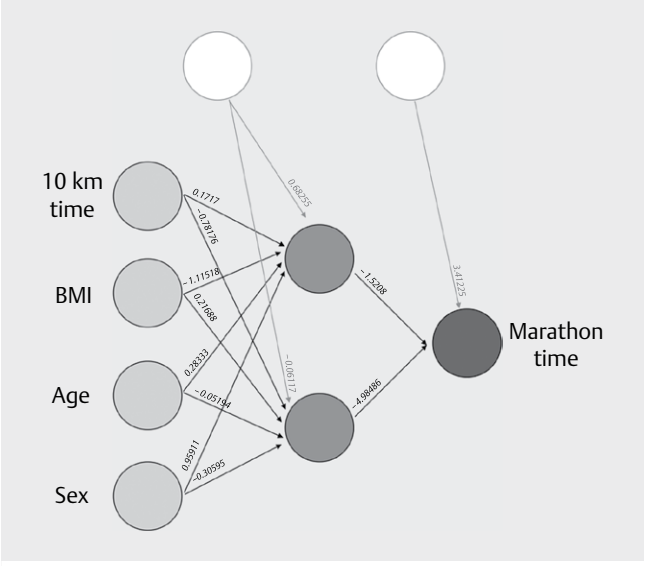
Note: Bias = difference between actual and predicted performances; *r* = coefficient of correlation; LoA = limits of agreement; ES = effect size. *Significantly correlated at $p < 0{,}001$.

▶ **Table 2** Syntax (Excel spreadsheet) of the artificial neural network-based equation to estimate marathon performance (min) from performance on 10-km, BMI, age and sex. Mean values and standard deviation (SD) of input variables from artificial neuronal network (ANN) algorithm.

Marathon Performance = (((1/(1 + EXP(-((((C3–40,734675)/5,886275)*(0,1717)) + (((D3–21,518491)/2,061137)*(-1,11518)) + (((E3–43,432927)/9,513137)*(0,28333)) + (((F3–1,8304878)/0,3754327)*(0,95911)) + (0,68255)))))*(-1,5208)) + ((1/(1 + EXP(-((((C3–40,734675)/5,886275)*(-0,78176)) + (((D3–21,518491)/2,061137)*(0,21688)) + (((E3–43,432927)/9,513137)*(-0,05194)) + (((F3–1,8304878)/0,3754327)*(-0,30595)) + (-0,06117)))))*(-4,98486)) + (3,41225)*37,77462 + 204,8372

Input variables	Mean	SD
10 -km time (min)	40,734675	5,886275
BMI (kg.m^{-2})	21,518491	2,061137
Age (years)	43,432927	9,513137
Sex	1,8304878	0,3754327
Marathon time (min)	204,83720	37,77462

Note: C3 = performance on 10-km; D3 = BMI (body mass index); E3 = age in years; F3 = sex (girls = 1; boys = 2).



▶ **Fig. 2** Neural network architecture with computational details.; *Note*: BMI: body mass index.

The level of statistical significance was set at $p < 0{,}05$, and all analyses were performed with the Statistical Package for the Social Sciences (SPSS, release 20,0, Chicago, IL, USA).

Results

Means and SD of actual and predicted marathon performances are presented in ▶ **Table 1**.

The ANN-equation to estimate marathon performance from performance on 10-km, BMI, age and sex is depicted in ▶ **Table 2** and ▶ **Fig. 2**.

No statistically significant difference was found between the actual and predicted performances for either algorithm ($p > 0{,}05$, ▶ **Table 1**). Moreover, the magnitude of the bias in these predicted performances was systematically trivial (ES ≤ -0,091; ▶ **Table 1**).

All predicted running performances were correlated with the actual ones, with a very high correlation coefficient ($p < 0{,}001$, $r \geq 0{,}918$, ▶ **Table 1**).

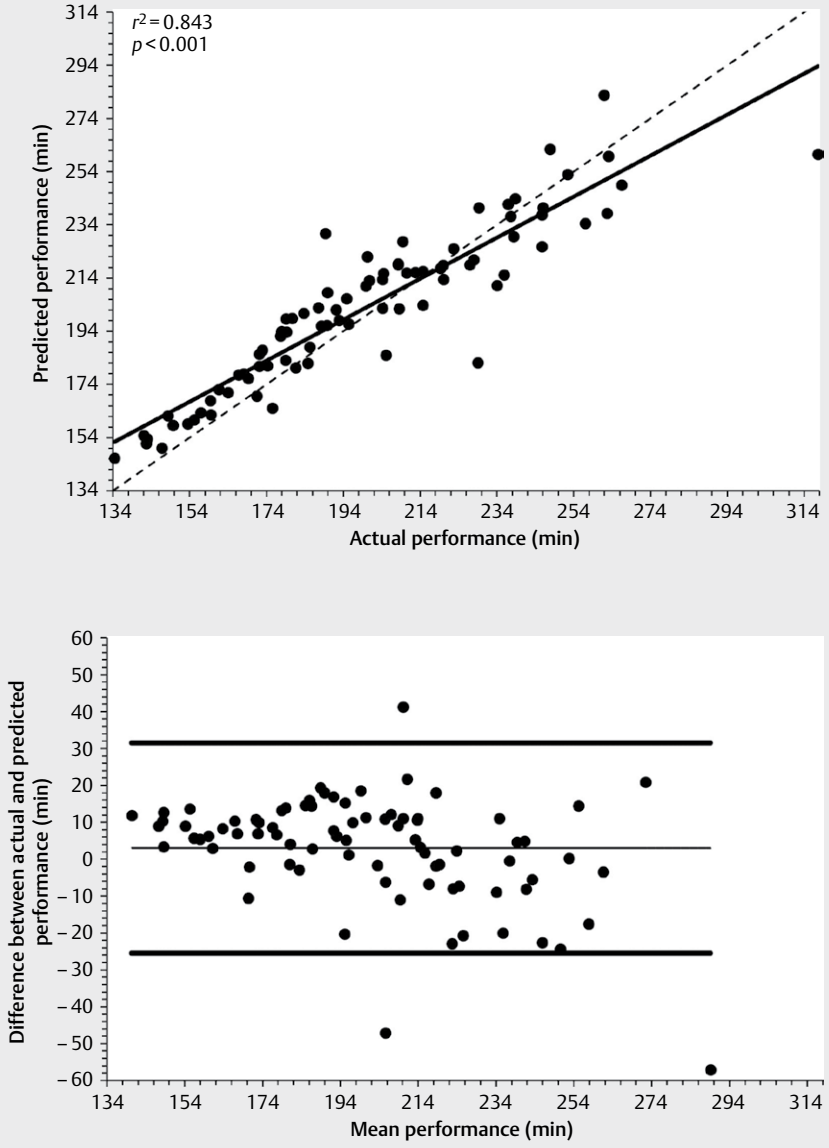
The MAE, presented in ▶ **Table 1**, were 11 min 16 s (i. e., 5,6 %) and 4 min 48 s (i. e., 2,4 %) for the ANN and KNN, respectively.

The bias ± 95 % LoA are shown in ▶ **Figs. 3**–▶ **4**. The bias ± 95 % LoA of the ANN and KNN were 3 min 14 s ± 28 min 30 s (i. e., 1,5 ± 14,1 %) and -47 s ± 14 min 21 s (i. e., -0,4 ± 7,2 %), respectively (▶ **Table 1** and ▶ **Figs. 3,4**).

Discussion

The objectives of this study were to test the validity of two supervised machine learning algorithms (ANN and KNN) and to compare them to determine which one was better at predicting marathon performances in terms of accuracy (from the MAE and the bias) and precision (from 95 % LoA). We have hypothesized that both techniques would be valid, and that KNN would be the better performing method. In view of the results obtained and our dataset, this hypothesis appears to be confirmed.

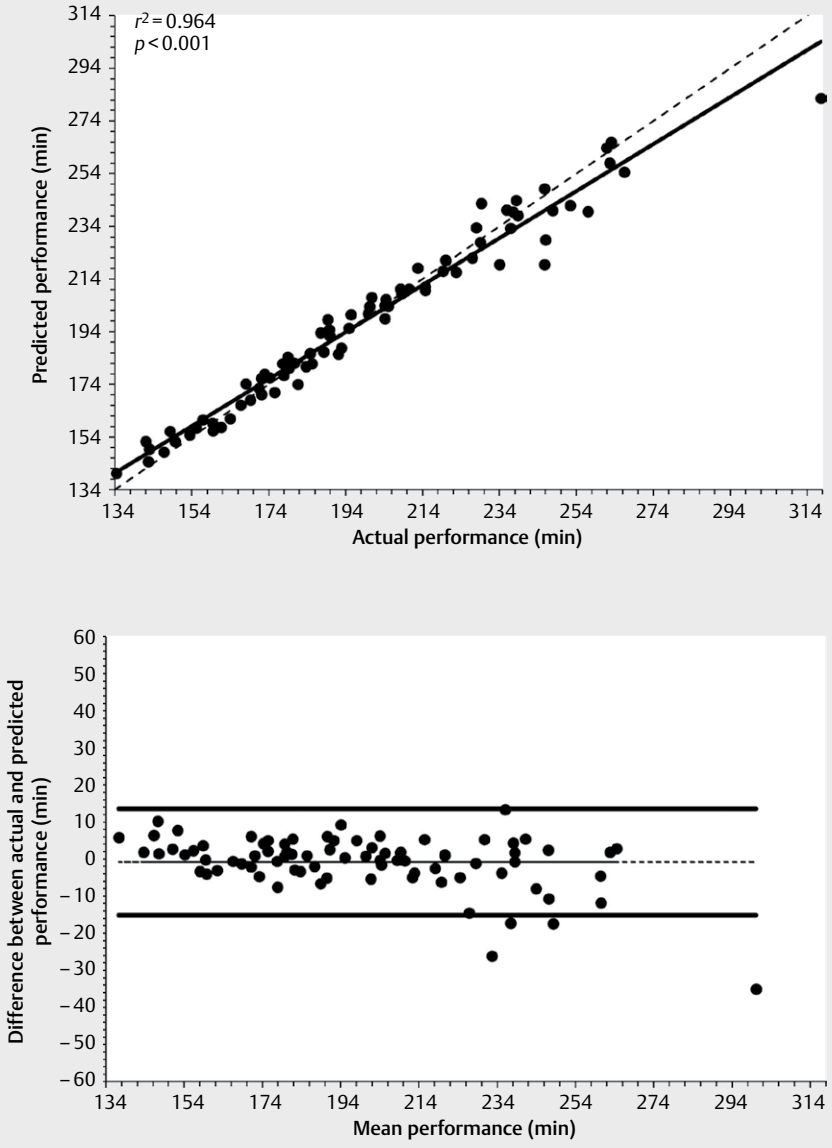
One of the main findings was that the two algorithms can indeed be considered valid, accurate (i. e., a lower bias means a high-



▶ **Fig. 3** Validity of measurements with the ANN algorithm to predict performance.; *Top panel:* association between actual and predicted performance from the ANN algorithm in the marathon in 82 athletes. The solid line is the linear regression. r^2 is the coefficient of determination. *Bottom panel:* Bland and Altman plots for the comparison between actual and predicted performance in the marathon in 82 athletes. Dashed line is the bias, solid lines are the 95 % limits of agreement.

er accuracy) and precise (i. e., a lower distance between limits of agreement means a higher precision) for predicting marathon performances, as all the results confirmed their validity for predicting performances from independent variables (i. e., performance on 10-km, body mass index, age and sex), with both of them demonstrating a prediction accuracy above 94 %. These results were comparable to those of another study [14] that showed 97 % accuracy in elite runners with a local matrix completion machine learning technique.

The MAE was lower with KNN than with ANN, meaning that KNN was more accurate (up to 98 % instead of 94 %) and the predictions matched more closely to actual performances. The fields of application differed (i. e., classification algorithm), and it should be noted that the results were in accordance with Mustafa et al. [50] and Tamilarasi and Porkodi [51], who showed that KNN demonstrated better accuracy than ANN, but in disagreement with those of Peace et al. [38], Musa et al. [52] and Anyama et al. [53, 54], who obtained better predictions with ANN than KNN. Regardless of the domain, there is no real consensus on which algorithm/model is best in terms of regression or classification, and each model has advantages and limitations [29, 32, 55, 56]. Indeed, compared to ANN, KNN tends to perform better on datasets with a small number of samples and has less risk of overfitting. Therefore, even though the data sample size (i. e., 820 athletes) was relatively large



▶ **Fig. 4** Validity of measurements with the KNN algorithm to predict performance.; *Top panel*: association between actual and predicted performance from the KNN algorithm in the marathon in 82 athletes. The solid line is the linear regression. r^2 is the coefficient of determination. *Bottom panel*: Bland and Altman plots for the comparison between actual and predicted performance in the marathon in 82 athletes. Dashed line is the bias, solid lines are the 95 % limits of agreement.

in this study, the ANN might have exhibited some limitations *(*i. e., generalizability, or the risk of overfitting the data), which would explain the greater accuracy of KNN in predicting marathon performances. In other words, the prediction results may be partly explained not only by the breadth of the data (e. g., size of the dataset [40]) but also by the model parameters (e. g., ratio used for training and testing datasets, number of hidden layers or learning rate for training a neural network, k in KNN, distance type in KNN, etc.) and the practitioner's method (i. e., the modeling procedure [30, 38, 39, 57], which defines the algorithms).

Moreover, the results revealed that the respective coefficients of correlation were, respectively, 0,918 and 0,982 for ANN and KNN and also that in 95 % of the dependent variable score (i. e., marathon performance), the biases from ANN and KNN were 3 min 14 s ± 28 min 30 s (i. e., 1,5 ± 14,1 %) and − 47 s ± 14 min 21 s (i. e., − 0,4 ± 7,2 %). Therefore, the coefficient of correlation was slightly higher (i. e., higher prediction precision) for KNN than ANN, and the bias and 95 % LoA were lower with KNN (*i. e.,* higher accuracy).

Coquart et al. [15] showed that marathon performance could be predicted from the nomogram of Mercier et al. [58] with acceptable levels of accuracy and precision. Indeed, the authors found a bias and 95 % LoA of −1 min 25 s ± 27 min 3 s (i. e., − 0,7 ± 13,2 %) [15], which is comparable to ANN in the current study. However, for the same level of accuracy and precision, Coquart et al. [15] had the athletes perform two long-distance maximal performances (*i. e.,* 10-km and 20-km), while the predictions from ANN (and KNN,

which is more accurate and precise) were obtained from only one performance (i. e., 10-km). More recently, Vickers and Vertosick [18] explored several methods (i. e., the Riegel formula [59] and two models based on one or two prior races, respectively) for predicting the marathon race times of recreational runners. Between the predicted and observed marathon times, they found a mean square error (MSE) of 6 min 21 s for the Riegel formula [59], 3 min 48 s for the model based on one prior race, and 3 min 28 s for the model based on two prior races. It thus seems that in addition to the number of prior performances a model includes to predict performance, the association of certain variables (e. g., sex, age, BMI) with race velocity also improves the accuracy of predictions. Moreover, it is especially interesting to note that in the current study, which integrated similar factors in the algorithms (i. e., sex, age, BMI), the MSEs were lower, with an MSE of 2 min 39 s for ANN and an MSE of 39 s for KNN, by taking only one performance (i. e., 10-km time). Therefore, it might be interesting to add another performance (e. g., on the half-marathon) to ANN and KNN to see whether prediction accuracy/precision is significantly better than prediction methods not relying on AI. To limit athlete fatigue, ANN and especially KNN should nevertheless be preferred over other prediction methods that require two performances.

The main potential limitations of this study were the sample size and the content of our dataset. As mentioned in the discussion, KNN has the advantage of performing better on small data samples (less risk of overfitting) than the ANN algorithm. Moreover, the choice and inclusion of input variables (limited, with only four variables, i. e., 10-km performance, BMI, age and sex) determined the algorithms and influenced the prediction results. In other words, the possibility of adjusting the algorithm to better model the problem domain will always be conceivable. Indeed, although marathon performance can be affected by a multitude of factors since the run involves several performance elements, physiological (e. g., maximal oxygen uptake, running economy, anaerobic threshold, etc.) [22], psychological (e. g., motivation, stress) and environmental variables can also be determinant in long-distance running performance [4, 9, 10]. Thus, the other limitation of this study concerned the quality of the recovered data. Height and body mass were not measured, with the athletes only reporting their anthropometric data (i. e., self-declaration of body height and mass to calculate the athletes' BMI). However, runners are known to self-report accurately for this type of data [60]. Moreover, it is likely that the performance data (i. e., 10-km time) was influenced by the race profile (e. g., course profile: uphill, downhill, etc.), meteorological conditions (e. g., temperature, wind, rain, etc.), opposition field or race strategy (e. g., run to win or run for a time), but this information was not available.

Practical applications

The applications from the current study would be further extended by performing validation studies in race (i. e., middle- or long-distance running on track or road), with other inputs that influence running performance (e. g., maximal oxygen uptake, running economy and anaerobic threshold) in specific levels of runners (i. e., sub-regional, regional, inter-regional, national and international). It would also be interesting to use machine learning techniques to identify the determinant factors of the marathon (i. e., the use of classification techniques like random forest or naive Bayes) in order to help athletes, staff and professional sport analysts to design training programs and detect athletic talent, in addition to predicting results (marathon time).

Conclusions

Few studies on the prediction of running performance, especially for the marathon, have used artificial intelligence with ANN and/or KNN algorithms. The results of the current study demonstrated that both KNN and ANN were able to predict the performances of marathon runners with an acceptable level of accuracy. Both models were valid and able to attain a prediction accuracy above 94 %, although KNN appears to be superior to ANN as it accurately predicted marathon performance above 98 %. These approaches can therefore be used to predict performances over the course of appropriate training programs, with particular emphasis on prescribing speeds for tempo runs and determining competitive strategies. Future studies should be directed toward the use of machine learning techniques to gain insight into other parameters that impact marathon performance by means of classification techniques in order to detect talented athletes, for example, and not only to predict marathon performance.

Acknowledgements

The authors are grateful to the French Athletics Federation (Fédération Française d'Athlétisme) for data collection, and the Orthodynamica Center at Mathilde Hospital 2.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

[1] Burfoot A. The history of the marathon. Sports Med 2007; 37: 284–287

[2] Vitti A, Nikolaidis PT, Villiger E et al. The "New York City Marathon": participation and performance trends of 1,2M runners during half-century. Res Sports Med 2020; 28: 121–137

[3] Knechtle B, Di Gangi S, Rüst CA et al. Men's participation and performance in the Boston marathon from 1897 to 2017. Int J Sports Med 2018; 39: 1018–1027

[4] Marc A, Sedeaud A, Guillaume M et al. Marathon progress: demography, morphology and environment. J Sports Sci 2014; 32: 524–532

[5] Knechtle B, Nikolaidis PT, Onywera VO et al. Male and female Ethiopian and Kenyan runners are the fastest and the youngest in both half and full marathon. Springerplus 2016; 5: 223

[6] Waśkiewicz Z, Nikolaidis PT, Gerasimuk D et al. What motivates successful marathon runners? The role of sex, age, education, and training experience in Polish runners. Front Psychol 2019; 10: 1671

[7] Weiss K, Valero D, Villiger E et al. The influence of environmental conditions on pacing in age group marathoners competing in the "new york city marathon". Front Physiol 2022; 13: 842935

[8] Weiss K, Valero D, Villiger E et al. Temperature and barometric pressure are related to running speed and pacing of the fastest runners in the 'Berlin Marathon'. Eur Rev Med Pharmacol Sci 2022; 26: 4177–4287

[9] El Helou N, Tafflet M, Berthelot G et al. Impact of environmental parameters on marathon running performance. PloS One 2012; 7: e37407

[10] Nikolaidis PT, Di Gangi S, Chtourou H et al. The role of environmental conditions on marathon running performance in men competing in Boston marathon from 1897 to 2018. Int J Environ Res Public Health 2019; 16: 614

[11] Boullosa D, Esteve-Lanao J, Casado A et al. Factors affecting training and physical performance in recreational endurance runners. Sports (Basel) 2020; 8: 35

[12] Cuk I, Nikolaidis PT, Villiger E et al. Pacing in Long-Distance Running: Sex and Age Differences in 10-km Race and Marathon. Medicina (Kaunas) 2021; 57: 389

[13] Deaner RO, Carter RE, Joyner MJ et al. Men are more likely than women to slow in the marathon. Med Sci Sports Exerc 2015; 47: 607–616

[14] Blythe DAJ, Király FJ. Prediction and quantification of individual athletic performance of runners. PloS One 2016; 11: e0157257

[15] Coquart J, Alberty M, Bosquet L. Validity of a nomogram to predict long distance running performance. J Strength Cond Res 2009; 23: 2119–2123

[16] Berndsen J, Smyth B, Lawlor A. Pace my race: recommendations for marathon running. In: Proceedings of the 13th ACM Conference on Recommender Systems. 2019: 246–250

[17] Ruiz-Mayo D, Pulido E, Martınoz G. Marathon performance prediction of amateur runners based on training session data. Proc Mach Learn Data Min Sports Anal 2016; 8

[18] Vickers AJ, Vertosick EA. An empirical study of race times in recreational endurance runners. BMC Sports Sci Med Rehabil 2016; 8: 26

[19] Billat LV, Koralsztein JP, Morton RH. Time in human endurance models. From empirical models to physiological models. Sports Med Auckl NZ 1999; 27: 359–379

[20] Ettema JH. Limits of human performance and energy-production. Int Z Für Angew Physiol Einschließlich Arbeitsphysiologie 1966; 22: 45–54

[21] Scherrer J, Monod H. Le travail musculaire local et la fatigue chez lhomme. Presse Med 1960; 68: 1717–1717

[22] Alvero-Cruz JR, Carnero EA, García MAG et al. Predictive performance models in long-distance runners: a narrative review. Int J Environ Res Public Health 2020; 17: 8289

[23] Péronnet F, Thibault G. [Physiological analysis of running performance: revision of the hyperbolic model]. J Physiol (Paris) 1987; 82: 52–60

[24] Tanda G. Prediction of marathon performance time on the basis of training indices. J Hum Sport Exerc 2011; 6: 521–520

[25] Vandewalle H. Modelling of running performances: comparisons of power-law, hyperbolic, logarithmic, and exponential models in elite endurance runners. BioMed Res Int 2018; 2018: 8203062

[26] Florence SL, Weir JP. Relationship of critical velocity to marathon running performance. Eur J Appl Physiol Occup Physiol 1997; 75: 274–278

[27] Mulligan M, Adam G, Emig T. A minimal power model for human running performance. PloS One 2018; 13: e0206645

[28] Huang Z-Q, Chen Y-C, Wen C-Y. Real-time weather monitoring and prediction using city buses and machine learning. Sensors (Basel) 2020; 20: 5173

[29] Uddin S, Khan A, Hossain ME et al. Comparing different supervised machine learning algorithms for disease prediction. BMC Med Inform Decis Mak 2019; 19: 281

[30] Bunker RP, Thabtah F. A machine learning framework for sport result prediction. Appl Comput Inform 2019; 15: 27–33

[31] Wiseman O. Using Machine Learning to Predict the Winning Score of Professional Golf Events on the PGA Tour. National College of Ireland 2016

[32] Bunker R, Susnjak T. The application of machine learning techniques for predicting results in team sport: a review. Prepr Submitt ArXiv 2019; 1–48. doi: 10.1613/jair.1.13509

[33] McCabe A, Trevathan J. Artificial intelligence in sports prediction. In: Fifth International Conference on Information Technology: New Generations (itng 2008). IEEE. 2008: 1194–1197

[34] Claudino JG, de Oliveira Capanema D, de Souza TV et al. Current approaches to the use of artificial intelligence for injury risk assessment and performance prediction in team sports: a systematic review. Sports Med Open 2019; 5: 28

[35] Velichkov B, Koychev I, Boytcheva S. Deep learning contextual models for prediction of sport event outcome from sportsman's interviews. In: Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2019). 2019: 1240–1246

[36] Dhanabal S, Chandramathi S. A review of various k-nearest neighbor query processing techniques. Int J Comput Appl 2011; 31: 14–22

[37] Maszczyk A, Gołaś A, Pietraszewski P et al. Application of neural and regression models in sports results prediction. Procedia Soc Behav Sci 2014; 117: 482–487

[38] Peace IC, Uzoma AO, Ita SA et al. A comparative analysis of K-NN and ANN techniques in machine learning. Int J Eng Res Technol 2015; 4: 420–425

[39] Ruiz JR, Ramirez-Lechuga J, Ortega FB et al. Artificial neural network-based equation for estimating VO2max from the 20 m shuttle run test in adolescents. Artif Intell Med 2008; 44: 233–245

[40] Ajiboye AR, Abdullah-Arshah R, Hongwu Q. Evaluating the effect of dataset size on predictive model using supervised learning technique. International Journal of Software Engineering & Computer Sciences 2015; 1: 75–84

[41] Harriss DJ, Jones C , MacSween A. Ethical standards in sport and exercise science research: 2022 update. Int J Sports Med 2022; 43: 1065–1070

[42] Renganathan V. Overview of artificial neural network models in the biomedical domain. Bratisl Lek Listy 2019; 120: 536–540

[43] Fritsch S, Guenther F, Suling M et al. Package 'Neuralnet': Training of Neural Networks 2016

[44] Malone C, Fennell L, Folliard T et al. Using a neural network to predict deviations in mean heart dose during the treatment of left-sided deep inspiration breath hold patients. Phys Med 2019; 65: 137–142

[45] Smith PF, Ganesh S, Liu P. A comparison of random forest regression and multiple linear regression for prediction in neuroscience. J Neurosci Methods 2013; 220: 85–91

[46] Cohen J. Statistical power analysis. Curr Dir Psychol Sci 1992; 1: 98–101

[47] Munro BH. Statistical Methods for Health Care Research. lippincott williams & wilkins; 2005

[48] Lin L, Torbeck LD. Coefficient of accuracy and concordance correlation coefficient: new statistics for methods comparison. PDA J Pharm Sci Technol 1998; 52: 55–59

[49] Freund Y. Concordance entre deux méthodes de mesure d'une même variable: diagramme de Bland et Altman. Ann Fr Médecine Urgence 2016; 6: 143–146

[50] Mustafa M, Taib MN, Murat Z et al. Comparison between KNN and ANN classification in brain balancing application via spectrogram image. J Comput Sci Comput Math 2012; 2: 17–22

[51] Tamilarasi R, Porkodi DR. A study and analysis of disease prediction techniques in data mining for healthcare. Int J Emerg Res Manag Technoly ISSN 2015; 1: 2278–9359

[52] Musa RM, Majeed AA, Taha Z et al. The application of Artificial Neural Network and k-Nearest Neighbour classification models in the scouting of high-performance archers from a selected fitness and motor skill performance parameters. Sci Sports 2019; 34: e241–e249.

[53] Anyama OU, Nwachukwu EO. A hybrid prediction system for american NFL results. Int J Comput Appl Technol Res 2015; 4: 42–47

[54] Anyama OU, Igiri CP. An application of linear regression & artificial neural network model in the NFL result prediction. Int J Eng Res Technol 2015; 4: 457–461

[55] Domingos P. The Master Algorithm: How the Quest for the Ultimate Learning Machine will Remake our Uorld. Basic Books 2015

[56] Kök H, Acilar AM, İzgi MS. Usage and comparison of artificial intelligence algorithms for determination of growth and development by cervical vertebrae stages in orthodontics. Prog Orthod 2019; 20: 41

[57] Yang L, Shami A. On hyperparameter optimization of machine learning algorithms: Theory and practice. Neurocomputing 2020; 415: 295–316

[58] Mercier D, Léger L, Desjardins M. Nomogramme pour prédire la performance, le VO2max et l'endurance relative en course de fond. Médecine du Sport 1984; 58: 181–187

[59] Riegel PS. Athletic records and human endurance: A time-vs.-distance equation describing world-record performances may be used to compare the relative endurance capabilities of various groups of people. Am Sci 1981; 69: 285–290

[60] Nikolaidis PT, Knechtle B. Validity of recreational marathon runners' self-reported anthropometric data. Percept Mot Skills 2020; 127: 1068–1078