applied sciences

MDPI

Article

Classification of Motor Competence in Schoolchildren Using Wearable Technology and Machine Learning with Hyperparameter Optimization

José Sulla-Torres *, Alexander Calla Gamboa, Christopher Avendaño Llanque, Javier Angulo Osorio and Manuel Zúñiga Carnero

Escuela Profesional de Ingeniería de Sistemas, Universidad Católica de Santa María, Arequipa 04000, Peru; christopher.avendano@ucsm.edu.pe (C.A.L.); jangulo@ucsm.edu.pe (J.A.O.)
* Correspondence: jsullato@ucsm.edu.pe; Tel.: +51-9599-50467

Abstract: Determining the classification of motor competence is an essential aspect of physical activity that must be carried out during school years. The objective is to evaluate motor competence in schoolchildren using smart bands, generate percentiles of the evaluation metrics, and classify motor performance through machine learning with hyperparameter optimization. A cross-sectional descriptive study was carried out on 764 schoolchildren (451 males and 313 females) aged 6 to 17 years. Five state schools in the city of Arequipa, Peru were evaluated. Weight, height, and waist circumference were assessed, and body mass index (BMI) was calculated. The tests evaluated in the schoolchildren measured walking and running for 6 minutes. These tests were carried out using smart bands, capturing cadence, number of steps, calories consumed, speed, stride, and heart rate. As a result, the percentiles were created through the LMS method [L (asymmetry: lambda), M (median: mu), and S (coefficient of variation: sigma)]. The cut-off points considered were <P25 (below average), p25 to p75 (average), and >p75 (above average). For classification, the machine-learning algorithms random forest, decision tree, support vector machine, naive Bayes, logistic regression, k-nearest neighbor, neural network, gradient boosting, XGBboost, LightGBM, and CatBoost were used, and the hyperparameters of the models were optimized using the RandomizedSearchCV technique. In conclusion, it was possible to classify motor competence with the tests carried out on schoolchildren, significantly improving the accuracy of the machine-learning algorithms through the selected hyperparameters, with the gradient boosting classifier being the best result at 0.95 accuracy and in the ROC-AUC curves with a 0.98. The reference values proposed in this study can be used to classify the walking motor competence of schoolchildren. Finally, the mobile software product built based on the proposed model was validated using the prototype of the Software Quality Systemic Model (SQSM) based on three specific categories: functionality, reliability, and usability, obtaining 77.09%. The results obtained can be used in educational centers to achieve the suggested recommendations for physical activity in schoolchildren.

Keywords: machine learning; classification; motor competence; schoolchildren; wearable; hyperparameters

check for updates

Citation: Sulla-Torres, J.; Calla Gamboa, A.; Avendaño Llanque, C.; Angulo Osorio, J.; Zúñiga Carnero, M. Classification of Motor Competence in Schoolchildren Using Wearable Technology and Machine Learning with Hyperparameter Optimization. *Appl. Sci.* **2024**, *14*, 707. https://doi.org/10.3390/app14020707

Academic Editor: Grzegorz Dudek

Received: 19 November 2023
Revised: 14 December 2023
Accepted: 15 December 2023
Published: 14 January 2024

1. Introduction

1.1. Background

Machine-learning classification is training a computer model to recognize and categorize data based on specific characteristics. It is typically classified as supervised, unsupervised, semi-supervised, or reinforced [1]. In supervised classification, the model is trained using labeled data to identify patterns and predict new data [2]. This technique has many applications, including healthcare, education, and technology. One such application is using smart bands in schoolchildren to monitor their physical activity and health status [3].

Physical activity, such as motor competence, is crucial for developing children and adolescents, as it promotes overall health and well-being. Unfortunately, sedentary lifestyles and lack of physical activity have become increasingly common among school students. Smart bands have emerged as a tool to combat this problem and measure physical activity levels in students [4].

Smart bands are wearable devices that can monitor a person's vital health statistics, including heart rate, using a combination of sensors and machine-learning algorithms [3]. In the context of schoolchildren, smart bands can be used to track their physical activity levels, sleep patterns, and overall health status. Machine-learning classification can be used to analyze the data collected with smart bands and identify patterns that can help improve the health and well-being of schoolchildren [1].

The importance of machine-learning classification for schoolchildren lies in its ability to provide personalized and adaptive learning experiences. By analyzing data collected from smart bands, machine-learning algorithms can create profiles for individual students and design learning paths tailored to their specific needs and abilities [5,6]. Additionally, machine-learning classification can identify students at risk for health issues and provide early interventions to prevent or mitigate these issues [7]. Overall, the use of machine-learning classification with smart bands in school children has the potential to improve their physical and mental health and academic performance significantly. Machine-learning algorithms can be optimized by tuning with hyperparameters, obtaining better results [8].

This research aims to explore the use of smart bands and the application of machine learning to promote physical activity and motor competence in educational centers and analyze the potential benefits of this approach. We will examine how smart bands measure levels of motor competence, how machine learning with hyperparameter optimization is used, and the benefits of using them in educational centers. In doing so, we hope to shed light on the potential of these technologies to improve levels of motor competence among students and ultimately contribute to improved health and academic outcomes.

1.2. Related Work

Smart bands are wearable devices that have become increasingly popular for tracking and monitoring various physical and health activities, including motor competence. The main works in this regard are presented.

The work of Kounoudes, Kapitsaki, and Katakis [1] indicates that smart bands and wearable fitness devices can measure physical activity levels through various sensors and tracking mechanisms. One of the most common ways smart bands can measure physical activity levels is by monitoring the number of steps the user takes throughout the day. Pedometer readings are commonly used to describe physical activity levels in adults, and fitness trackers like Fitbit record the number of steps the user takes each day. Additionally, VO2max measurements are often used to determine whether the user has increased or decreased fitness level, and variations in VO2max are used as an indicator of overall health. Garmin activity data include VO2max measurements, which can help determine fitness levels. Daily step counts can show people's habits and lifestyle as well as their risk of mortality that may occur to them.

Amor and James [2] suggest that activity monitoring (AM) is a well-established method for assessing an individual's physical activity. The swift emergence of smartwatch technology provides the capability to monitor activity and seamlessly engage with other healthcare systems. Al-Janabi and Hamza [3] propose an intelligent data analysis model to find optimal patterns in human activities based on biometric characteristics obtained from smartwatches and smartphones. The forward and backward rule-based pattern finder generates the optimal patterns that help humans organize their activities. The results show that good patterns are generated for human activities. In addition, Weiss, Yoneda, and Hayajneh [5] suggest that wearable devices include sensors that provide a platform to implement and deploy motion-based mobile behavioral biometrics using the smartwatch accelerometer by basically investigating the physical activity of walking. Therefore, they used different sensors to evaluate physical activity better. The results show that motion-based biometrics using smartwatches yield good results for the activities assessed.

Ali et al. [7] concentrated on introducing innovative techniques for identifying and recording physical activities using machine-learning methods and wearable sensors. Everyday physical activities tend to be unstructured or unplanned, with specific activities or actions (such as sitting or standing) occurring more frequently than others (such as walking or going up and down stairs). Existing activity classification systems have not explored the impact of such class imbalances on the effectiveness of machine-learning classifiers. Therefore, the study's primary aim was to examine the influence of class imbalance on the performance of machine-learning classifiers and to identify which classifiers are more sensitive to class imbalance than others. The research utilized motion sensor data from 30 participants recorded during the performance of various activities of daily living.

Wang, Lizardo, and Hachen [9] demonstrate that social, psychological, and environmental characteristics are related to physical activity among students by using Fitbit devices to collect data on daily movement activities, number of calories, and heart rate, among other measurements from a sample of 692 students, taking four indicators; they showed that there is a strong relationship between the growth of group and individual physical activity.

According to Sabry et al. [10], the application of machine learning to promote physical activity includes fall detection, seizure detection, vital sign monitoring and prediction, and activity recognition. Machine-learning techniques are also being explored for health monitoring, elder care, and fitness tracking. Additionally, machine-learning techniques have been used to link physical activity to obesity by examining the relationship between physical activity and weight status in a large-scale dataset. The study found that the weighted SVM algorithm with a penalized approach offered the best classification performance, followed by ADA(RF).

Zhou et al. [11] indicate that machine learning can also automate and personalize physical activity promotion programs by tracking people's activity patterns and developing therapy and exercise plans to reduce obesity. Additionally, machine-learning methods predict exercise relapse and improve physical activity interventions by identifying unlikely individuals to adhere to a physical exercise regimen.

Creaser et al. [12] note that smart bands and machine learning can promote the health, well-being, or understanding of children or adolescents in schools. However, more research is needed to explore their full impact. Wearables can also be used beyond encouraging physical activity, such as studying and teaching health concepts. However, research indicates when and how wearable devices with the most frequently used functions can be used in schools. It suggests that they are acceptable instruments in the school environment to monitor students' physical activity levels or educate them about the importance of physical activity.

Site, Nurmi, and Lohan [13] reviewed machine-learning algorithms to analyze eHealth data collected from wearable devices, emphasizing the significant potential for enhancing healthcare quality and customer satisfaction through machine learning (ML). The ML algorithms were applied to both time and frequency domain healthcare data derived from wearable devices and sensors. The authors explored how ML techniques can effectively process and analyze health sensor data, noting that accelerometers, gyroscopes, ECG (electrocardiogram), EEG (electroencephalogram) monitors, and blood glucose sensors are the primary sources of eHealth data. The study delved into various aspects, including types of features and methods for feature extraction and ML algorithms commonly used in eHealth data analysis. Notably, the authors concluded that, among the ML algorithms studied in the literature, neural network (NN) algorithms and support vector machines (SVMs) had demonstrated the most promising performance for analyzing healthcare data.

Himi et al. [14] introduce a predictive system named "MedAi", which is based on a smartwatch and employs machine-learning algorithms to predict multiple diseases. The system consists of three main components: a "Sense O'Clock" smartwatch prototype equipped with eleven sensors to gather body statistics, a machine-learning model for analyzing the collected data and making predictions, and a mobile application to display the prediction results. The researchers obtained a dataset of body statistics from patients at a local hospital. Several machine-learning algorithms were utilized in the study, including support vector machine (SVM), support vector regression (SVR), k-nearest neighbor (KNN), extreme gradient boost (XGBoost), long short-term memory (LSTM), and random forest (RF). The goal was to identify the most effective algorithm. The experimental results on the dataset revealed that the random forest (RF) algorithm outperformed the other machine-learning algorithms tested.

Machine-learning algorithms commonly rely on a set of hyperparameters, the values of which need to be chosen thoughtfully, and these choices often have a substantial impact on the algorithm's performance [15]. Hyperparameters are settings that are not directly learned from the dataset but especially impact model performance. The most used search strategies are grid search, manual search, and random search [14]. Several works are related to optimization with hyperparameters, such as the one developed by Yagin et al. [16], who used neural networks with hyperparameter optimization to predict obesity based on physical activity. Rivera, Avilés, and Castillo-Castaneda [17] classified the physical activity indicator using machine learning, and after feature, importance selection, and hyperparameter were tuned. There are also works regarding health in general with the optimization of hyperparameters [18].

2. Materials and Methods

2.1. Methodology

Machine-learning classification with smart bands in schoolchildren involves several steps, including data collection and analysis. Smart bands are equipped with sensors that collect data on physiological parameters, such as heart rate and physical activity levels. The data are then analyzed to identify patterns and trends that can be used to develop machine-learning models for classification.

Feature selection and preprocessing are crucial steps in the machine-learning classification process. Feature selection involves identifying the most relevant features from the data that will be used to train the model. Preprocessing involves cleaning and transforming the data to ensure they are suitable for analysis. These steps aim to provide the machine-learning model access to high-quality data that will enable accurate classification.

The machine-learning model is trained and tested once the data have been collected, analyzed, and preprocessed. Supervised machine-learning techniques are commonly used for classification tasks, where the model is trained on labeled data. The model's performance is evaluated using accuracy, precision, f1-score, and recall metrics. The goal is to develop a model that can accurately classify schoolchildren based on their physiological parameters, which can be used to monitor their health and well-being.

The methodology used was CRISP-DM [19]. It comprises six phases: business understanding, data understanding, data preparation, modeling, evaluation, and deployment.

2.1.1. Business Understanding

This initial phase will identify the problem caused by traditional methods of evaluating motor competence in educational centers, and a solution will be projected to resolve the present issue. In this phase, there are different activities for understanding the business, each of which will be described below.

1. Determination of objectives: The main goal focuses on exploring the use of smart bands and the application of machine learning optimized to promote physical activity and motor competence in schoolchildren and analyzing the potential benefits of this approach.
2. Evaluation of the situation: A descriptive cross-sectional study was conducted on 764 schoolchildren (451 males and 313 females) aged 6 to 17. The sample selection was non-probabilistic by convenience. Five state schools in the city of Arequipa, Peru, were evaluated. The schoolchildren attended physical education classes twice a week. Permission was requested from each school's administration to conduct the study in both schools. Then, parents were informed about the objective of the project. Parents

who agreed to participate in the study signed the informed consent form to authorize their children's participation.

3. Determination of the goal of machine learning: At this stage, the determination will be made to apply a correct supervised machine-learning technique to determine the best algorithm that finds the best accuracy, precision, f1-score, and recall in students' motor competence classification. A classifier is a function f that takes as input a set of features $x \in X$, where X is the feature space, and outputs a class label $y \in \{1, \ldots, C\}$, where C is the class space.

2.1.2. Data Understanding

Anthropometric measurements and the utilization of the smart band followed the recommendations outlined by the local ethics committee (UCSM-096-2022) and adhered to the principles of the Declaration of Helsinki (World Medical Association) concerning ethical standards for human research.

1. Collection of initial data: Anthropometric measurements were conducted on-site at each school. The evaluation team comprised professional physical education teachers and research assistants. Weight and height were measured using Ross and Marfell Jones's standardized method. To determine body weight (kg), a BC 730 (Tanita Corporation) electronic scale was used, with a scale from 0 to 150 kg. Standing height was measured using a portable stadiometer (Seca 216, Seca Gmbh and Co., Hamburg, Germany), accurate within 0.1 mm. Waist circumference (WC) was measured using a tape measure (Seca) to the nearest 1 mm. The body mass index (BMI) was calculated by dividing the kilograms of weight by the square of the height in meters: BMI = weight (kg)/height^2 (m).
   According to the BMI Z-score, patients were classified as underweight/normal weight with Z-scores between -2 and +0.99, overweight from 1 to 1.99, obese from 2 to 2.99, and very obese $\geq$3 [20]. To categorize abdominal adiposity (WC) by age and sex, the suggestions described by Fernández et al. [21] were used. It was categorized into two groups (without risk < p75 and with risk > p75).
   The motor competence tests that were evaluated were the quantification of the number of steps during school recess, and the 6-minute walk test was performed using a smart band (Huawei band 7) with an AMOLED screen of 194 $\times$ 368 and with 1.47 inches. This smart band has been used in other similar research [22]. The smart band was placed on the wrist of each student's hand, and its use was explained.
2. Describe and explore the data: The Kolmogorov–Smirnov test verified the dataset's normality. Descriptive statistics (mean, standard deviation, min, and max) were calculated. Table 1 shows the description of the data of the schoolchildren.

Table 1. Description of schoolchildren's data.

	Age	Weight	Height	Waist	BMI	Cadence	Steps	Speed	Stride
mean	12.48	47.57	1.49	71.93	20.74	59.58	910.64	2.99	79.73
std	2.75	16.18	0.16	11.37	4.27	34.46	504.13	1.93	10.35
min	6	18.00	1.140	50.00	11.71	10.00	126.00	0.310	53.00
25%	10	34.00	1.370	63.00	17.66	30.00	499.00	1.315	72.50
50%	13	48.00	1.520	70.50	20.30	53.00	833.00	2.570	78.00
75%	15	58.10	1.630	78.50	23.35	82.50	1241.50	4.365	86.00
max	17	107.70	1.810	114.60	39.95	163.00	2189.00	9.760	111.00

t-test was used to calculate differences between both sexes in independent samples. Differences between BMI and waist values were determined via one-way ANOVA and Tukey's specificity test. In all cases, $p < 0.05$ was significant.

2.1.3. Data Preparation

In this phase, the data are selected according to the most critical attributes to train and test the algorithms that will be chosen for the study.

1. Data selection: Through data selection, it became feasible to identify and emphasize those fields that would provide valuable contributions to the analysis of physical activity for motor skill tests. Each data record has the following attributes within the database:
    a. Anthropometric data: age (years), weight (kg), height (m), sitting height (cm), and waist circumference (cm).
    b. Average pace: the time the person can walk a kilometer; they are a number in minutes and seconds format.
    c. Average cadence: these are the steps per minute you can do; they are raw numbers.
    d. Steps: these are all the steps the person has taken during the activity; they are raw numbers.
    e. Calories: the calories the person has burned during the activity; they are numbers without formats.
    f. Average speed: the average speed at which the person has moved during the activity in kilometers/hour; it is in number format with decimals without arrangements.
    g. Average stride: it is the average distance taken by each step; they are numbers without formats.
    h. Heart rate: these are the beats per minute the heart has given during the activity; they are numbers without formats.
    i. Maximum heart rate: this is the maximum number of beats per minute the individual has given in the activity; they are numbers without formats.
2. Data cleaning: Data cleaning tasks allowed us to discover correct and sometimes eliminate erroneous data records or outliers and convert and standardize the data types necessary for processing in machine-learning algorithms. The Jupyter dashboard [23] was used with the Python 3 programming language, with its Pandas library; it is a rapid, robust, adaptable, and user-friendly open-source tool for data analysis and manipulation. The Seaborn library was used with its boxplot function to visualize the classes, as shown in Figure 1, where the high class has the highest proportion for both sexes. Points outside a boxplot are visual indicators of values that may be unusual or outliers compared to the rest of the data in the set.

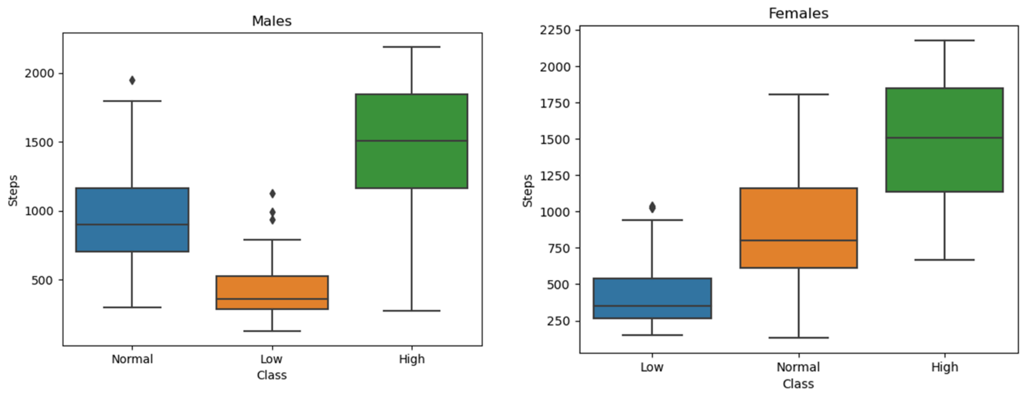

Figure 1. Types of motor competence according to the tests the male and female schoolchildren carried out.

2.1.4. Modeling

When conducting a detailed analysis of the data source, it was determined that the classification output resulting from the motor competence tests is labeled as high, normal, and low. For this reason, the decision was made that the most appropriate type of prediction would be classification.

To do this, a comparison of popular supervised machine-learning techniques for a classification model will be conducted. According to the literature, the most used algorithms and optimizers are:

1. Decision tree: A non-parametric supervised technique that constructs a classification model as a tree structure, applicable for classification and regression tasks [24].
2. Random forest: It generates a set of decision trees by employing random resampling on the training set [25].
3. Support vector machine: Creates effective boundaries to separate datasets by solving a constrained quadratic optimization problem [26].
4. Naive Bayes: It is a probabilistic classifier based on Bayes' theorem, assuming strong independence within attributes of an instance [27].
5. Logistic regression: This type of regression analysis is used to predict the outcome of a categorical variable based on the independent or predictor variables [28]. While commonly recognized as a classifier, logistic regression can also be employed as a regressor to predict numeric values. Its adaptability allows it to address classification and regression problems, depending on the nature of the data and the analysis objectives [28].
6. Neuronal network: Most current neural network applications are concerned with pattern recognition problems. Artificial neural networks consist of assemblies of perceptrons designed for multi-layer feedforward networks [29].
7. K-nearest neighbors: It seeks to predict outputs by computing the distance between the test data and training points, subsequently selecting the K number of points closest to the test data [30].
8. Gradient boosted: This ensemble learning technique builds and combines several weak learning models to form a more robust model. The main idea is to correct the errors of the previous model by iteratively adding soft models. It focuses on fitting the residuals of the previous model using a gradient-based approach [31].
9. XGBoost: short for "eXtreme Gradient Boosting", is a specific implementation of gradient boosting. It was developed to be fast and efficient in terms of resource usage. It includes regularization, missing value handling, and a custom cost function [32].
10. LightGBM: Gradient boosting machines build sequential decision trees, with each tree constructed based on the errors of the preceding tree. In the end, predictions are made by summing the contributions of all these trees.
11. CatBoost: CatBoost stands for "Category" and "Boost"; it handles categorical, numeric, and text features. The CatBoost algorithm employs a symmetric tree or an oblivious tree structure [33].

Anaconda Navigator Software 2.5.1 was used with its Jupyter Notebook 6.5.2 with the Python 3 programming language and its Scikit-learn optimization library to compare supervised machine-learning techniques for classification.

With careful preparation of the input data, it will be imported in a specific CSV format, representing it as a table with the attributes selected in the "Data selection" section. Figure 2 shows the modeling developed to classify motor competence data, using information generated by smart bands. This encompasses data processing, modeling, comparisons with machine-learning algorithms, and achieving classification with optimization using hyperparameters in the study.

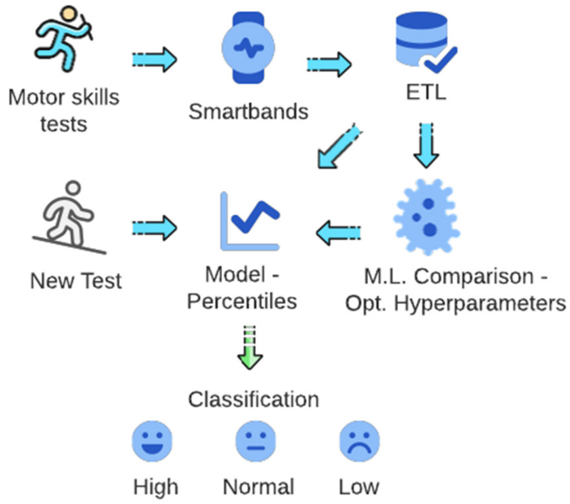

Figure 2. Modeling of data from the proposed study.

The percentiles were constructed using the LMS method [34]. The curves L represent skewness (lambda), M represents the median (mu), and S represents the coefficient of variation (sigma). The LMS method uses the Box–Cox transformation to fit the data distribution to a normal distribution by minimizing the effects of skewness.

For this purpose, the P25, P50, and P75 percentiles of males (Table 2) and females (Table 3) were created about the motor competence metrics: cadence (Figure 3), steps (Figure 4), speed (Figure 5), and stride (Figure 6) for schoolchildren males and females.

Table 2. Percentiles of motor competence metrics of schoolchildren males.

Metrics	Age					
Cadence	6–7	8–9	10–11	12–13	14–15	16–17
L	0.94	−0.22	0.22	0.33	0.85	1.52
M	28	96	60	63	49	29
S	0.72	0.42	0.46	0.50	0.55	0.53
P25	20	55	44	38.5	29	21
P50	28	96	60	63	49	29
P75	65.5	120	88	86.8	68	44
Steps						
L	1.33	−0.16	0.42	0.30	0.67	1.09
M	455	1437	884	1000	732	496
S	0.70	0.42	0.48	0.45	0.54	0.55
P25	299	772	628	690	470	309
P50	455	1437	884	1000	732	496
P75	796	1747	1179	1299	1097	726
Velocity						
L	0.85	0.19	0.80	0.33	0.93	1.85
M	1.62	4.69	3.08	3.15	2.28	1.16
S	0.70	0.47	0.53	0.52	0.63	0.61
P25	1.11	2.73	2.10	1.84	1.17	0.90
P50	1.62	4.69	3.08	3.15	2.28	1.16
P75	3.48	6.00	4.30	4.51	3.69	1.92
Stride						
L	−0.02	0.65	0.91	0.41	0.27	0.66
M	88	82	80	82	76	71
S	0.14	0.12	0.12	0.12	0.106	0.12
P25	74.5	76	74	76	71	66
P50	88	82	80	82	76	71
P75	97.5	90	86	87	83	78

Legend: P: percentile, L: (skewness, lambda), M: (median, mu), S: (coefficient of variation, sigma).

Table 3. Percentiles of motor competence metrics of schoolchildren females.

Metrics	Age					
Cadence	6–7	8–9	10–11	12–13	14–15	16–17
L	−0.06	−0.41	0.25	0.41	1.04	1.48
M	75	94	69	60	36	29
S	0.52	0.34	0.47	0.49	0.67	0.53
P25	40	68.75	33	35	20	20
P50	75	94	69	60	36	29
P75	108	111.3	89	84	60.5	44
Steps						
L	0.21	−0.24	0.58	0.54	0.99	1.29
M	941	1391	802	784	613	529
S	0.52	0.37	0.57	0.49	0.61	0.52
P25	615	989	462	578	426	359
P50	941	1391	802	784	613	529
P75	1584	1608	1233	1226	937	762
Velocity						
L	−0.06	−0.29	0.79	7.97	1.11	1.84
M	3.91	4.75	2.91	3.04	1.58	1.27
S	0.52	0.36	0.62	5.50	0.74	0.62
P25	2.15	3.33	1.52	1.77	0.94	0.92
P50	3.91	4.75	2.91	3.04	1.58	1.27
P75	5.41	5.69	4.16	4.26	3.09	1.78
Stride						
L	−2.97	1.47	1.11	0.79	0.27	0.57
M	84	82	78	77	76	71
S	0.24	0.10	0.11	0.11	0.09	0.12
P25	81	79	75	73	72	65.5
P50	84	82	78	77	76	71
P75	89.5	86.3	86.5	86	81.5	77.5

Legend: P: percentile, L: (skewness, lambda), M: (median, mu), S: (coefficient of variation, sigma).

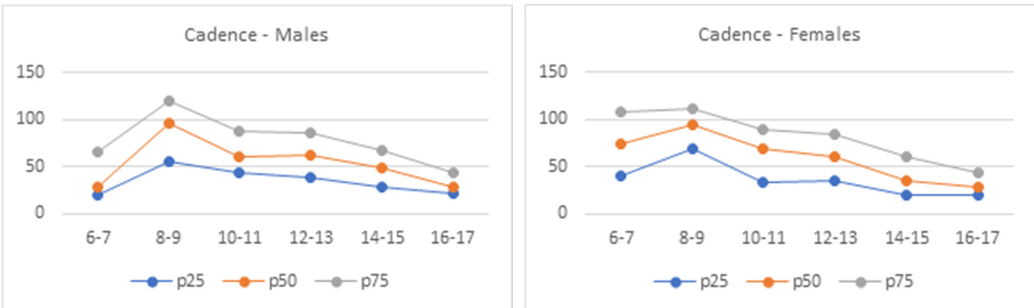

Figure 3. Distribution of percentiles for the cadence in schoolchildren in both sexes.

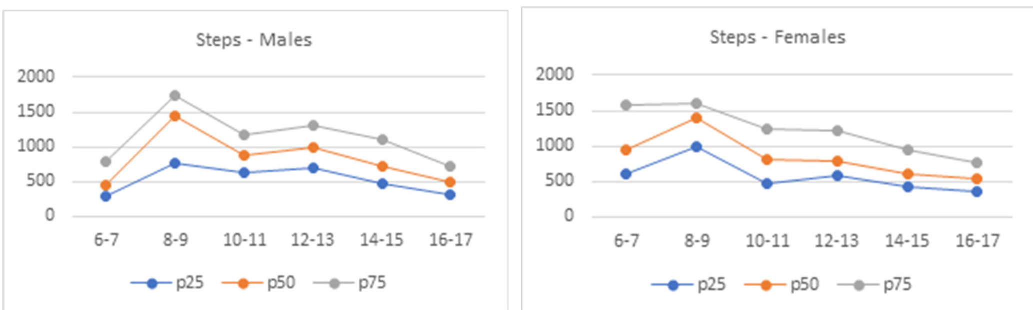

Figure 4. Distribution of percentiles for the number of steps in schoolchildren in both sexes.

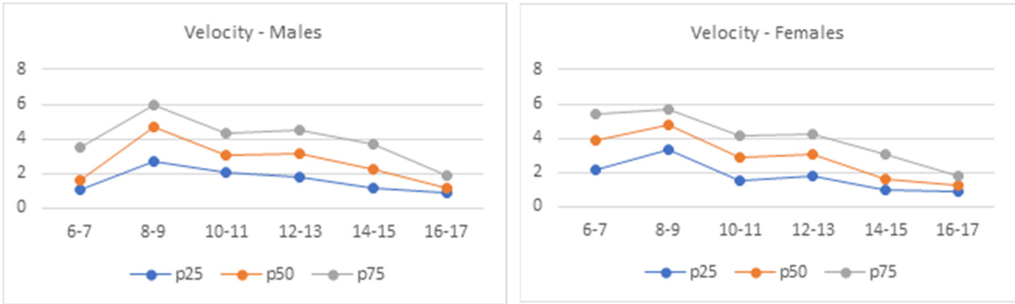

Figure 5. Distribution of percentiles for the velocity in schoolchildren in both sexes.

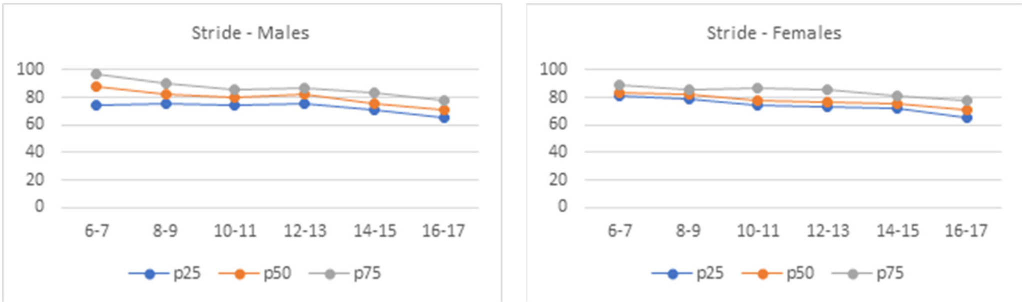

Figure 6. Distribution of percentiles for the stride in schoolchildren in both sexes.

3. Evaluation and Results

For the Anaconda Navigator platform with Jupyter Notebook, different machine-learning techniques used in further research were modeled [35], which were decision tree, support vector machine, random forest, naive Bayes, logistic regression, k-nearest neighbors, neuronal network, gradient boosted, and smart bands. Of the data to model, 80% was used for training and 20% for testing.

Jupyter Notebook is a widely used tool in the machine-learning community, where you can import a set of libraries, train a dataset classification classifier, and evaluate the model by just using a few lines of code in Python with its library Scikit-learn.

Likewise, the optimization of hyperparameters with the Scikit-learn optimization library was used. Hyperparameters are parameters that are not directly learned by the learning algorithm. The basic hyperparameter tuning models are manual search, grid search, and random search. Random search was used for training since it allows us to find equal or better models in computing time.

The configuration used for the hyperparameters in the case of gradient boosted was the loss function to optimize with the "log_loss" option, the criterion with the process to measure the quality of a division with the "friedman_mse" option, the "learning_rate" with 0.1, "max_depth": 3, "min_samples_leaf": 1, "min_samples_split": 2, "n_estimators": 100, "random_state": none, "subsample": 1.0, "tol": 0.0001, "validation_fraction": 0.1, and with the other default parameters.

Next, the tests were carried out, and the results of the modeling that was executed for the different techniques that were previously chosen were obtained using the CSV file as input data and made up of the motor competence variables.

To evaluate the model [36], accuracy in Equation (1), recall in Equation (2), precision in Equation (3), and the f1-score in Equation (4) were used. For many classes C_i, fp_i represents false positive, tp_i true positive, fn_i false negative, and tn_i true negative.

$$Accuracy = \frac{tp_i + tn_i}{tp_i + tn_i + fp_i + fn_i} \tag{1}$$

$$Recall = \frac{\sum_{i=1}^{l} \frac{tp_i}{tp_i + fn_i}}{l} \tag{2}$$

$$Precision = \frac{\sum_{i=1}^{l} \frac{tp_i}{tp_i + fp_i}}{l} \quad (3)$$

$$F1 - score = 2 \times \frac{(Precision) \times (Recall)}{(Precision) + (Recall)} \quad (4)$$

Tables 4 and 5 compare the classical and optimized machine-learning techniques results with selected hyperparameters and the accuracy, f1-score, recall, and precision metrics for both schoolchildren males and females.

Table 4. Comparison of results of supervised machine-learning techniques in males.

Algorithm	DT	SVM	RF	NB	LR	KNN	MLP	GB	XGB	LGBM	CB
Accuracy	0.88	0.69	0.92	0.74	0.78	0.81	0.68	0.93	0.92	0.94	0.93
Accuracy optimized hyperparameter	0.87	0.82	0.92	0.75	0.79	0.85	0.70	0.95	0.91	0.94	0.93
f1-score	0.87	0.74	0.93	0.77	0.82	0.88	0.72	0.92	0.91	0.93	0.92
Recall	0.86	0.67	0.94	0.73	0.84	0.88	0.67	0.92	0.90	0.94	0.92
Precision	0.88	0.82	0.92	0.82	0.80	0.88	0.77	0.92	0.92	0.92	0.92

Legend: DT: decision tree, SVM: support vector machine, RF: random forest, NB: naïve Bayes, LR: Logistic Regression; KNN: k-nearest neighbors, MLP: multilayer perceptron, GB: gradientboosted, XGB: extreme gradient boosting, LGBM: light gradient boosting machine, CB: CatBoost.

Table 5. Comparison of results of supervised machine-learning techniques in females.

Algorithm	DT	SVM	RF	NB	LR	KNN	MLP	GB	XGB	LGBM	CB
Accuracy	0.86	0.68	0.88	0.71	0.78	0.72	0.72	0.84	0.87	0.88	0.90
Accuracy optimized hyperparameter	0.86	0.80	0.87	0.72	0.83	0.81	0.74	0.89	0.84	0.88	0.88
f1-score	0.91	0.68	0.92	0.76	0.89	0.88	0.82	0.93	0.92	0.90	0.93
Recall	0.97	0.61	0.94	0.69	0.94	0.92	0.86	0.92	0.86	0.91	0.94
Precision	0.86	0.76	0.89	0.83	0.85	0.82	0.78	0.94	0.94	0.89	0.92

Legend: DT: decision tree, SVM: support vector machine, RF: random forest, NB: naïve Bayes, LR: Logistic Regression; KNN: k-nearest neighbors, MLP: multilayer perceptron, GB: gradient boosted, XGB: extreme gradient boosting, LGBM: light gradient boosting machine, CB: CatBoost.

In the results of the previous tables, it can be identified that, for the classification techniques, those that gave the best results concerning accuracy for the case of males and females were gradient boosting, whose values were the highest, indicating a better adjustment to the estimated prediction with a value of 0.95. For the f1-score metrics, the algorithms gave similar values of 0.92 between males and females. In the case of *recall*, it was 0.92. The confusion matrix of the gradient boosting machine-learning algorithm with the highest score found is shown in Figure 7.

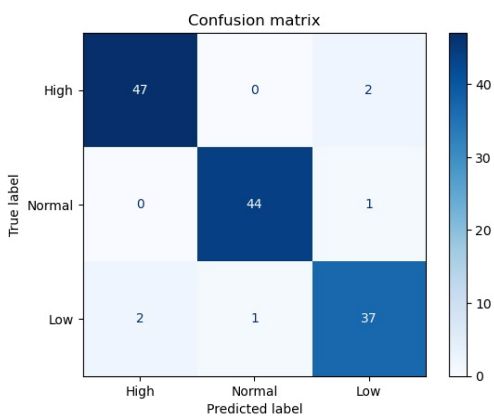

Figure 7. The confusion matrix of the gradient boosting machine-learning algorithm.

Figures 8 and 9 show a graph of the ROC-AUC curves of gradient boosting for males and females.

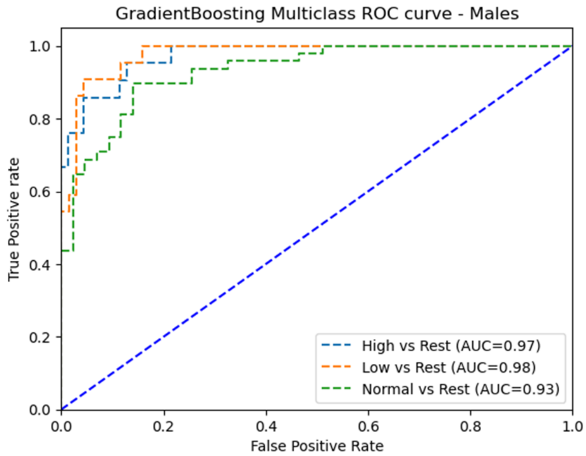

Figure 8. ROC-AUC curves of gradient boosting in males.

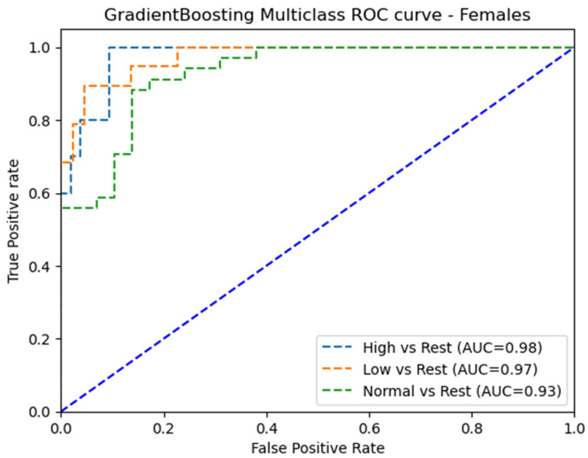

Figure 9. ROC-AUC curves of gradient boosting in females.

Observing the results of the ROC curves, in the case of male schoolchildren, the "Low" class shows an intense elevation towards the upper left corner of the graph and has a high area under the curve (AUC) of 0.98; this indicates that the model has a high sensitivity to detect the "Low" classification of motor competence. The false positive rate is low, suggesting the model does not misidentify many schoolchildren with the "Low" motor competence classification.

In the case of the female schoolchildren, the "High" class shows an intense elevation towards the upper left corner of the graph and has a high area under the curve (AUC) of 0.98; this indicates that the model has a high sensitivity to detect the classification "High" motor competence. The false positive rate is low, suggesting the model does not misidentify many schoolchildren with the "High" motor competence classification. On the other hand, the dark blue dotted line refers to the fact that the test has been bad, which is not the case of the study.

Deployment

A mobile app was developed in Android Studio using the Flutter framework. The code structure was organized by the best mobile application development practices and following the design pattern recommended by Flutter, which is the model–view–controller (MVC) design pattern. Firebase real-time database was used for storage.

The primary interfaces are shown in Figure 10, where you have the login screen and the options menu.

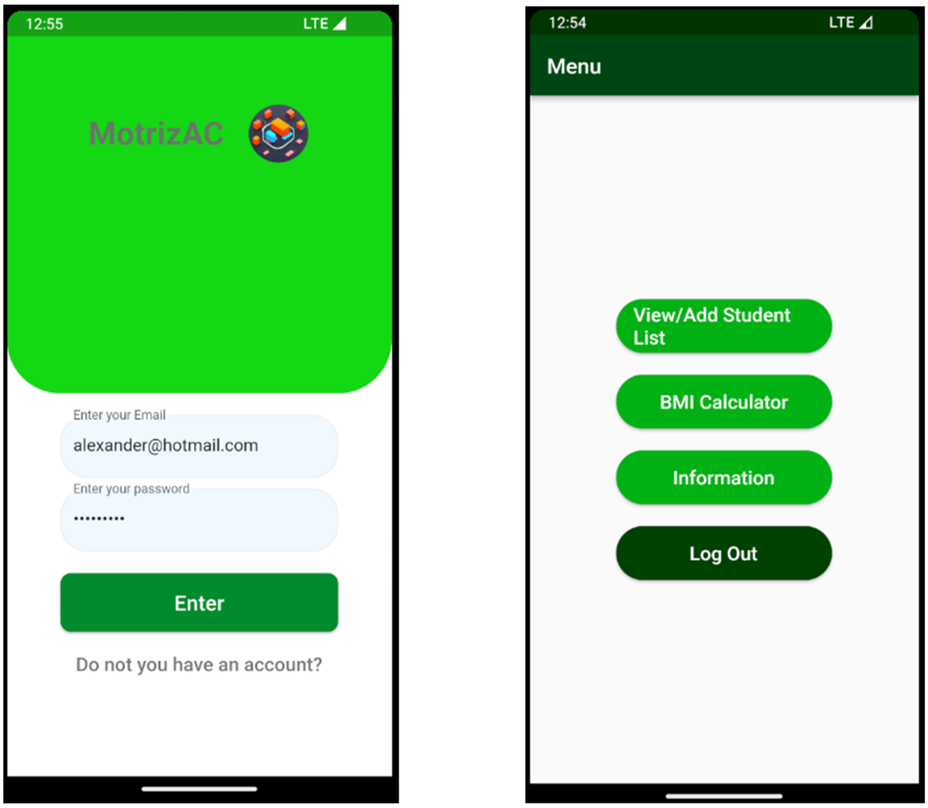

Figure 10. Initial interfaces of the mobile app.

Figure 11 shows the interface that allows entering the data corresponding to the student; the body mass index (BMI) will be able to be calculated automatically. Additionally, table is provided to identify the student's BMI classification. Likewise, the interface that allows entering the data captured with the smart band, such as speed, cadence, stride, and steps of the schoolchildren, is shown. It contains an Artificial Intelligence function that allows you to locate your position within the percentiles generated from the sample of schoolchildren.

The process was evaluated through the prototype of the Software Quality Systemic Model (SQSM). This model is planned based on six standardized international quality characteristics through a set of characteristic categories and metrics, which measure and evaluate the software quality of a product; this model is made into an instrument for measuring great value covering essential aspects of software quality.

It was carried out in the selected educational centers to evaluate the application's functionalities. The information collected meets expectations as it explains the population's needs.

Once the various interviewees were integrated, we proceeded with the evaluation method based on the Systemic Quality Model, which includes 11 categories. We have selected three specific classes related to the software: functionality, reliability, and usability because these apply to the case study. Table 6 shows a detailed description of each category with the interviews.

Considering the levels of satisfaction achieved in the categories of functionality, reliability, and usability, by calculating the arithmetic mean, it is observed that the degree of satisfaction of the application for the specialists of the educational centers is 77.09%. By contrasting this result with the quality level, it is evident that the application meets significant standards.

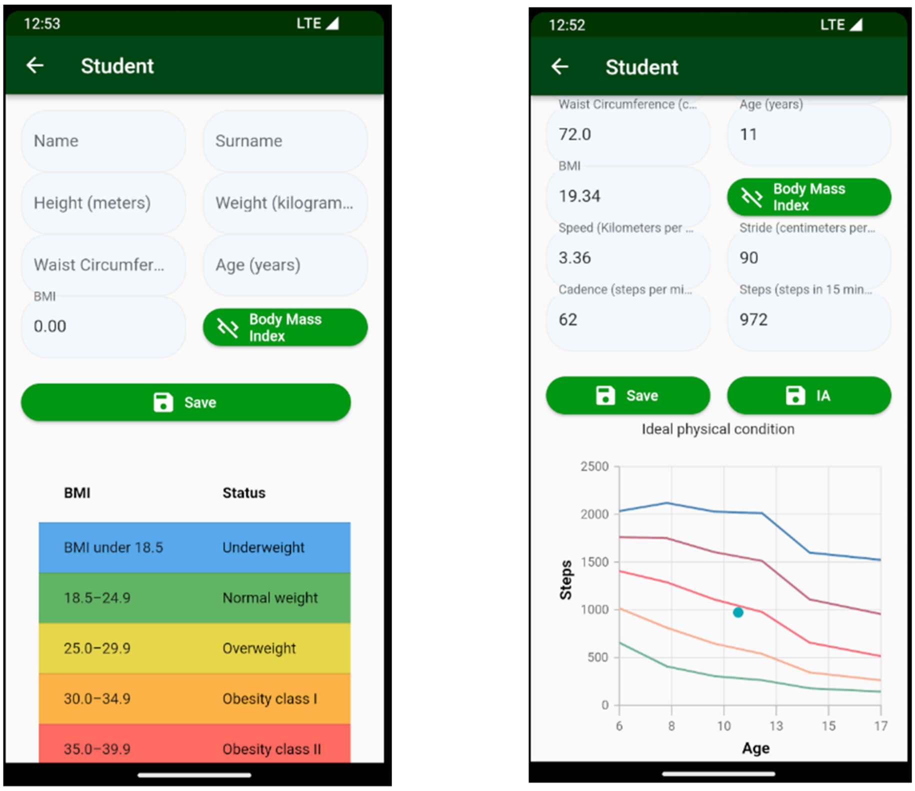

Figure 11. Calculation interfaces and use of AI to generate the value of your motor competence within the percentiles.

Table 6. Categories and characteristics according to the prototype of Software Quality Systemic Model (SQSM).

Characteristics	Subcharacteristics	Metrics	Percentage of Compliance	Quality Level
Functionality (FUN)	FUN. 1 Fit for purpose	20	84.00%	Satisfied
	FUN. 2 Accuracy	6	86.67%	Satisfied
	FUN. 3 Interoperability	4	86.67%	Satisfied
	FUN. 4 Security	3	53.33%	Does not satisfy
	Subtotal	33	81.88%	Satisfied
Usability (USA)	USA.1 Ease of Compression	5	90.00%	Satisfied
	USA. 2 Learning Capacity	8	92.00%	Satisfied
	USA. 3 Graphical Interface	4	95.00%	Satisfied
	USA. 4 Operability	3	52.00%	Does not satisfy
	Subtotal	20	82.73%	Satisfied
Reliability (RIA)	RIA. 1 Maturity	5	60.00%	Does not satisfy
	RIA. 2 Fault tolerance	4	90.00%	Satisfied
	RIA. 3 Recovery	2	50.00%	Does not satisfy
	Subtotal	11	69.09%	Does not satisfy
	Total	64	77.09%	Satisfied

4. Discussion

The study's objective was to classify the motor competence of schoolchildren performed in a school using smart bands according to age range and sex, using machine-learning techniques optimized with hyperparameters acceptably for the classification indicated through the passing results in the tests carried out.

The results of the study have shown that according to the motor competence tests carried out, step cadence values obtained range from 96 to 29 steps in both sexes and the number of steps ranges from 1437 to 455, with speed from 4.75 to 1.16 and the stride from

88 to 71. Furthermore, it was verified that the measurements obtained decreased rapidly with age, being more pronounced in females than in males.

In this context, the cut-off points suggested in other recent research [37] were taken as a basis, where the study proposed percentiles of the number of steps in a day. The percentiles indicate <p25 below average, p25 to p75 average, and >p75 above average. In essence, percentiles, regardless of the method used, can be applied to establish improvement goals, especially for schoolchildren below the 25th percentile [38]. For example, the schoolchildren in this study classified below the p25th percentile in the tests performed showed higher BMI values.

Precisely, these results coincide with those found in [4], when it states that the measurements obtained decrease rapidly with age. This information could help promote strategies for developing physical activity in schoolchildren in educational centers to improve their performance. However, a limitation found in the study was the need for a broader set of tests to be performed.

Furthermore, the results of the study have shown that using machine-learning techniques such as those evaluated by Himi et al. [14,39], then optimized with hyperparameters like work carried out by Yadav et al. [38] and Yang et al. [40], they achieve reasonable classifications for schoolchildren according to their motor competence. It can also be seen that the naïve Bayes algorithm obtains the lowest results [41,42].

It was determined that the most suitable supervised machine-learning technique would be the gradient-boosted model, which has obtained the best accuracy of 0.95 with an f-score of 0.93, recall of 0.92, and precision of 0.94, which was used in works such as [43,44]. This has been corroborated by the ROC-AUC curves, where the "Low" classes for males and "High" for females show a substantial elevation towards the upper left corner of the graph and have a high area under the curve of 0.98; this indicates that the model has a high sensitivity to detect motor competence classification.

This study represents a significant contribution with the use of wearable devices for data capture and classification using machine-learning algorithms optimized with hyperparameters for motor competence in schoolchildren that can support people interested in finding more precise ways to evaluate motor competence with the support of technology.

5. Conclusions

In conclusion, smart bands have shown great potential to improve the motor competence of schoolchildren through machine learning and hyperparameters. The gradient-boosted algorithm is a good model for classifying schoolchildren. Motor competence is a crucial aspect of the development of children and adolescents, and smart bands can facilitate this process by providing personalized feedback. By harnessing the power of technology, smart bands can help children develop essential motor skills, leading to better physical health and overall well-being. As more research is conducted in this field, we expect to see more advancements in smart band technology, which will undoubtedly positively impact child development.

Finally, the mobile software product built based on the proposed model was validated using the prototype of the Software Quality Systemic Model (SQSM) based on three specific categories: functionality, reliability, and usability, obtaining 77.09%

Expanding the determining attributes for classifying motor competence in schoolchildren is recommended in future work. Additionally, constructing the application of other classification algorithms is essential to compare results and assess efficiency, as well as deep learning. Different approaches, such as transfer learning, can be used to adapt domain data to train at high fidelity.

Author Contributions: Methodology, J.A.O.; Software, A.C.G. and C.A.L.; Validation, A.C.G. and C.A.L.; Formal analysis, J.S.-T.; Investigation, J.S.-T.; Writing—original draft, J.S.-T.; Writing—review & editing, M.Z.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Catholic University of Santa María grant number 7995-CU-2022.

Institutional Review Board Statement: Anthropometric measurements and the utilization of the smart band followed the recommendations outlined by the local ethics committee (UCSM-096-2022) and adhered to the principles of the Declaration of Helsinki (World Medical Association) concerning ethical standards for human research (Universidad Católica de Santa Maria (096-07/10/2022)).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data is not publicly available due to the privacy of handling schoolchildren's data.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Dini Kounoudes, A.; Kapitsaki, G.M.; Katakis, I. Enhancing user awareness on inferences obtained from fitness trackers data. *User Model. User Adapt. Interact.* **2023**, *33*, 967–1014. [CrossRef] [PubMed]
2. Amor, J.D.; James, C.J. Validation of a Commercial Android Smartwatch as an Activity Monitoring Platform. *IEEE J. Biomed. Health Inform.* **2018**, *22*, 968–978. [CrossRef] [PubMed]
3. Al-Janabi, S.; Salman, A.H. Sensitive integration of multilevel optimization model in human activity recognition for smartphone and smartwatch applications. *Big Data Min. Anal.* **2021**, *4*, 124–138. [CrossRef]
4. Sulla-Torres, J.; Vidal-Espinoza, R.; Avendaño Llanque, C.; Calla Gamboa, A.; Zúñiga Carnero, M.; Cossio-Bolaños, M.; Gomez-Campos, R. Quantification of the Number of Steps in a School Recess by Means of Smart Bands: Proposal of Referential Values for Children and Adolescents. *Children* **2023**, *10*, 915. [CrossRef] [PubMed]
5. Weiss, G.M.; Yoneda, K.; Hayajneh, T. Smartphone and Smartwatch-Based Biometrics Using Activities of Daily Living. *IEEE Access* **2019**, *7*, 133190–133202. [CrossRef]
6. Cosoli, G.; Antognoli, L.; Scalise, L. Wearable Electrocardiography for Physical Activity Monitoring: Definition of Validation Protocol and Automatic Classification. *Biosensors* **2023**, *13*, 154. [CrossRef] [PubMed]
7. Alsareii, S.A.; Awais, M.; Alamri, A.M.; AlAsmari, M.Y.; Irfan, M.; Aslam, N.; Raza, M. Physical Activity Monitoring and Classification Using Machine Learning Techniques. *Life* **2022**, *12*, 1103. [CrossRef]
8. Ali, Y.A.; Awwad, E.M.; Al-Razgan, M.; Maarouf, A. Hyperparameter Search for Machine Learning Algorithms for Optimizing the Computational Complexity. *Processes* **2023**, *11*, 349. [CrossRef]
9. Wang, C.; Lizardo, O.; Hachen, D.S. Using Fitbit data to examine factors that affect daily activity levels of college students. *PLoS ONE* **2021**, *16*, e0244747. [CrossRef]
10. Sabry, F.; Eltaras, T.; Labda, W.; Alzoubi, K.; Malluhi, Q. Machine Learning for Healthcare Wearable Devices: The Big Picture. *J. Healthc. Eng.* **2022**, *2022*, 4653923. [CrossRef]
11. Zhou, M.; Fukuoka, Y.; Goldberg, K.; Vittinghoff, E.; Aswani, A. Applying machine learning to predict future adherence to physical activity programs. *BMC Med. Inform. Decis. Mak.* **2019**, *19*, 169. [CrossRef] [PubMed]
12. Creaser, A.V.; Frazer, M.T.; Costa, S.; Bingham, D.D.; Clemes, S.A. The Use of Wearable Activity Trackers in Schools to Promote Child and Adolescent Physical Activity: A Descriptive Content Analysis of School Staff's Perspectives. *Int. J. Environ. Res. Public Health* **2022**, *19*, 14067. [CrossRef] [PubMed]
13. Site, A.; Nurmi, J.; Lohan, E.S. Systematic Review on Machine-Learning Algorithms Used in Wearable-Based eHealth Data Analysis. *IEEE Access* **2021**, *9*, 112221–112235. [CrossRef]
14. Himi, S.T.; Monalisa, N.T.; Whaiduzzaman, M.; Barros, A.; Uddin, M.S. MedAi: A Smartwatch-Based Application Framework for the Prediction of Common Diseases Using Machine Learning. *IEEE Access* **2023**, *11*, 12342–12359. [CrossRef]
15. Bischl, B.; Binder, M.; Lang, M.; Pielok, T.; Richter, J.; Coors, S.; Thomas, J.; Ullmann, T.; Becker, M.; Boulesteix, A.; et al. Hyperparameter optimization: Foundations, algorithms, best practices, and open challenges. *WIREs Data Min. Knowl. Discov.* **2023**, *13*, e1484. [CrossRef]
16. Yagin, F.H.; Gülü, M.; Gormez, Y.; Castañeda-Babarro, A.; Colak, C.; Greco, G.; Fischetti, F.; Cataldi, S. Estimation of Obesity Levels with a Trained Neural Network Approach optimized by the Bayesian Technique. *Appl. Sci.* **2023**, *13*, 3875. [CrossRef]
17. Rivera, O.; Avilés, O.F.; Castillo-Castaneda, E. Classifying the physical activity indicator using machine learning and direct measurements: A feasibility study. *Acta Sci. Technol.* **2023**, *45*, e61317. [CrossRef]
18. Triwiyanto, T.; Pawana, I.P.A.; Purnomo, M.H. An Improved Performance of Deep Learning Based on Convolution Neural Network to Classify the Hand Motion by Evaluating Hyper Parameter. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2020**, *28*, 1678–1688. [CrossRef]
19. Chapman, P.; Clinton, J.; Kerber, R.; Khabaza, T.; Reinartz, T.P.; Shearer, C.; Wirth, R. CRISP-DM 1.0: Step-by-step data mining guide. *PSS Inc.* **2000**, *9*, 1–73.
20. Freedman, D.S.; Berenson, G.S. Tracking of BMI z Scores for Severe Obesity. *Pediatrics* **2017**, *140*, e20171072. [CrossRef]
21. Fernández, J.R.; Redden, D.T.; Pietrobelli, A.; Allison, D.B. Waist circumference percentiles in nationally representative samples of African-American, European-American, and Mexican-American children and adolescents. *J. Pediatr.* **2004**, *145*, 439–444. [CrossRef] [PubMed]

22. Bayoumy, K.; Gaber, M.; Elshafeey, A.; Mhaimeed, O.; Dineen, E.H.; Marvel, F.A.; Martin, S.S.; Muse, E.D.; Turakhia, M.P.; Tarakji, K.G.; et al. Smart wearable devices in cardiovascular care: Where we are and how to move forward. *Nat. Rev. Cardiol.* **2021**, *18*, 581–599. [CrossRef] [PubMed]
23. Piazentin Ono, J.; Freire, J.; Silva, C.T. Interactive data visualization in Jupyter notebooks. *Comput. Sci. Eng.* **2021**, *23*, 99–106. [CrossRef]
24. Safavian, S.R.; Landgrebe, D. A Survey of Decision Tree Classifier Methodology. *Electr. Eng.* **1991**, *21*, 660–674. [CrossRef]
25. Qi, Y. Random Forest for Bioinformatics. In *Ensemble Machine Learning: Methods and Applications*; Springer: Berlin/Heidelberg, Germany, 2012.
26. Cristianini, N.; Shawe-Taylor, J. *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*; Cambridge University Press: Cambridge, UK, 2000. [CrossRef]
27. Zhang, H.; Su, J. Naive Bayes for optimal ranking. *J. Exp. Theor. Artif. Intell.* **2008**, *20*, 79–93. [CrossRef]
28. Cheng, X.; Lin, S.Y.; Liu, J.; Liu, S.; Zhang, J.; Nie, P.; Fuemmeler, B.F.; Wang, Y.; Xue, H. Does physical activity predict obesity—A machine learning and statistical method-based analysis. *Int. J. Environ. Res. Public Health* **2021**, *18*, 3966. [CrossRef] [PubMed]
29. Bishop, C.M. *Neural Networks: A Pattern Recognition Perspective. Neural Networks*; Oxford University Press: Oxford, UK, 1996.
30. Cunningham, P.; Delany, S.J. k-Nearest Neighbour Classifiers—A Tutorial. *ACM Comput. Surv.* **2022**, *54*, 1–25. [CrossRef]
31. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [CrossRef]
32. Liu, M.; Guo, C.; Guo, S. An explainable knowledge distillation method with XGBoost for ICU mortality prediction. *Comput. Biol. Med.* **2023**, *152*, 106466. [CrossRef]
33. Prokhorenkova, L.; Gusev, G.; Vorobev, A.; Dorogush, A.V.; Gulin, A. Catboost: Unbiased boosting with categorical features. In Proceedings of the 32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montréal, QC, Canada, 3–8 December 2018.
34. Cole, T.J. Establishing a standard definition for child overweight and obesity worldwide: International survey. *BMJ* **2000**, *320*, 1240–1243. [CrossRef]
35. Dai, R.; Lu, C.; Yun, L.; Lenze, E.; Avidan, M.; Kannampallil, T. Comparing stress prediction models using smartwatch physiological signals and participant self-reports. *Comput. Methods Programs Biomed.* **2021**, *208*, 106207. [CrossRef] [PubMed]
36. Sokolova, M.; Lapalme, G. A systematic analysis of performance measures for classification tasks. *Inf. Process. Manag.* **2009**, *45*, 427–437. [CrossRef]
37. Lammers, A.E.; Romanowski, A.L.; Baumgartner, H.; Diller, G.P.; Uebing, A. Reference Values for Wristband Accelerometry Data in Children Aged 6–11 Years of Age. *Front. Pediatr.* **2022**, *10*, 808372. [CrossRef] [PubMed]
38. Yadav, P.; Sharma, S.C.; Mahadeva, R.; Patole, S.P. Exploring Hyper-Parameters and Feature Selection for Predicting Non-Communicable Chronic Disease Using Stacking Classifier. *IEEE Access* **2023**, *11*, 80030–80055. [CrossRef]
39. Wang, J.; Xie, Z.; Li, Y.; Song, Y.; Yan, J.; Bai, W.; Zhou, T.; Qin, J. Relationship between Health Status and Physical Fitness of College Students from South China: An Empirical Study by Data Mining Approach. *IEEE Access* **2020**, *8*, 67466–67473. [CrossRef]
40. Yang, H.; Chen, Z.; Yang, H.; Tian, M. Predicting Coronary Heart Disease Using an Improved LightGBM Model: Performance Analysis and Comparison. *IEEE Access* **2023**, *11*, 23366–23380. [CrossRef]
41. Mannini, A.; Sabatini, A.M. Machine Learning Methods for Classifying Human Physical Activity from On-Body Accelerometers. *Sensors* **2010**, *10*, 1154–1175. [CrossRef]
42. Saez, Y.; Baldominos, A.; Isasi, P. A Comparison Study of Classifier Algorithms for Cross-Person Physical Activity Recognition. *Sensors* **2016**, *17*, 66. [CrossRef]
43. Zhang, B.; Ren, J.; Cheng, Y.; Wang, B.; Wei, Z. Health data driven on continuous blood pressure prediction based on gradient boosting decision tree algorithm. *IEEE Access* **2019**, *7*, 32423. [CrossRef]
44. Zhao, L.; Zhang, F.; Zhang, H.; Liang, Y.; Zhou, A.; Ma, H. Robust Respiratory Rate Monitoring Using Smartwatch Photoplethysmography. *IEEE Internet Things J.* **2023**, *10*, 4830–4844. [CrossRef]