Development Of A Physical Fitness Prediction Model Using Machine Learning Algorithms

Dongjun Li[1*] and Xudong Chen[2,3]

[1] Kaifeng Vocational College, Kaifeng, 475100, China
[2] Henan Mechanical and Electrical Vocational College, Zhengzhou, 450000, China
[3] Suan Sunandha Rajabhat University, Bangkok 10300, Thailand
[*] Corresponding author. E-mail: xudong_chen001@outlook.com

Received: Nov. 10, 2025; Accepted: Dec. 20, 2025

Physical fitness is vital for good health and quality of life. On the other hand, it is declining, especially among university students, due to a sedentary lifestyle. Traditional fitness evaluation methods-using manual assessments and self-reported questionnaires-suffer from a lack of accuracy in identifying specific areas needing improvement. This paper proposes a physical fitness evaluation and prediction model using the integration of Covariance-Based Equation Modelling (CB-SEM) and Light Gradient Boosting Machine (LightGBM). It uses data drawn from an online survey targeting the assessment of key variables such as physical activity level, motivation to perform physical activity, measures of health status, and Body Mass Index (BMI). Next, CB-SEM is used to analyse the relations among latent variables, while LightGBM predicts the result of fitness (low, medium, or high). The LightGBM algorithm shows that the most important features related to physical fitness outcomes are physical activity levels and health measures. The integration of CB-SEM enhances the understanding of the causal relationship and thus increases the prediction accuracy. This model offers more efficiency and personalisation in evaluating fitness than traditional approaches do. The personalised results make it possible to provide relevant recommendations and track the individual's fitness improvement for better health outcomes. The LightGBM model attained a total prediction accuracy of 85.7% and showed a promising ability to classify students into low, medium, and high fitness categories.

Keywords: Physical fitness, Covariance-Based Structural Equation Modelling, Light Gradient Boosting Machine, fitness evaluation, university students, Machine Learning

http://dx.doi.org/10.6180/jase.202608_31.056

1. Introduction

Complete health involves the physical, mental, and emotional well-being of a person, and the requirement for physical fitness is also very important . Poor posture, sedentary lifestyle, and mental stress induced by academics and poor nutrition lead to low fitness levels among students [1]. Traditional methods of fitness evaluation, like manual tests or self-reported questionnaires, often contain bias and hence result in inaccuracy. [2, 3]. In recent times, unparalleled improvement in the field of machine learning has provided ways through which fitness can be assessed with accuracy, personally, and in relation to age, weight, exercise habits, and lifestyle. [4, 5]. These techniques have given great insights that have allowed appropriate interventions to promote wellness.

Poor fitness results of physical activities and factors outside the person are brought about by sedentary lifestyles, inappropriate food intake, and too much screen time. This results in cardiovascular disorders and rapid weight gain [6, 7]. Academic pressure, working commitments, and lack of sleep are contributing factors towards poor levels of fitness [8]. Prior research has consistently shown that psychological stress, anxiety, and low motivation significantly limit students' engagement in physical activity which resulting in poorer fitness outcomes [9–13].

Access to professional health assessment tools is restricted, particularly for college students, which hinders timely evaluations and saves costs for prevention and treatment. [14]. These are the issues that need to be improved to provide better standards of fitness evaluation in a more cost-effective and customised manner according to individual health needs.

University students' fitness levels can be predicted through an integrated approach of either CB-SEM or LightGBM. CB-SEM investigates the relationships between exercise habit, lifestyle, and health [15], while LightGBM is a few lead scoring predictions that also figures out significant amounts of non-modular values for many conventional evaluations in fitness outcome [16].

- RO1- Evaluate the physical fitness levels of university students based on factors such as exercise habits, BMI, and lifestyle choices.
- RO2- Identify the key variables that influence physical fitness levels in university students, such as physical activity, mental health, and dietary habits.
- RO3- Assess the predictive accuracy of Machine Learning (ML) models, specifically LightGBM, in forecasting physical fitness outcomes based on student data.
- RO4- Develop an ML-based model that provides personalised fitness recommendations for university students based on their fitness predictions.
- RQ1- What are the key factors influencing physical fitness levels among university students?
- RQ2- How accurately can ML algorithms, such as LightGBM, predict physical fitness outcomes based on lifestyle and health data?
- RQ3- What is the relationship between university students' physical activity levels and their overall fitness scores?
- RQ4- How can personalised fitness recommendations be derived from ML models for improving student fitness levels?
- H1- University students with higher levels of physical activity will have better overall fitness scores compared to those with lower activity levels.
- H2- ML models, specifically LightGBM, will be able to predict university students' physical fitness outcomes with high accuracy.
- H3- There is a positive correlation between students' BMI and their fitness level, with higher BMI linked to lower fitness scores.
- H4- A personalised fitness recommendation system based on ML predictions will enhance the long-term fitness levels of university students.

The structure of the paper includes related works and proposed methodology in Sections 2 &3, respectively. Sections 4 and 5 show the results and discussion, respectively. The article is concluded in Section 6.

2. Materials and methods

2.1. Related Works

Durau, Diehl, and Terlutter [17] Used SEM to explore how health traits and fitness influencers on social media affect exercise intentions in men and women. They, however, do not address any long-term influences. Albaladejo-Saura et al. [18] Considered age, biological maturation, and fitness differences among adolescent volleyball players. However, its crosssectional design restricts long-term analysis.

Jimeno-Almazán et al. [19] investigated in a randomised controlled trial whether supervised physical activity could be helpful for patients after COVID-19. However, the short period of 8 weeks and a small number of examined subjects did not allow for analysis of long-term effects. In the same manner, Laukkanen et al. [20] considered the question of Cardiorespiratory Fitness (CRF) regarding mortality risk; however, they did not control for other possible confounding variables.

Jeffries et al. [21] Devised a new theoretical perspective on training by integrating contrary models and adding qualitative analysis. However, the qualitative and theoretical basis of the model is limited in terms of real-life applicability and behavioural distinction in learning outcomes. Afsar et al. [22] Offered an idea to combine wearable sensors and virtual reality in physical fitness applications; however, their generalisation is limited due to the high systemic complexity and dependence on domain-specific data.

McNarry et al. [23] Randomised 281 individuals to investigate the effects of IMT on rehabilitating patients with COVID-19. However, the study depended on individual data, while the reduced rate of participation restricted its generalizability. Koevoets et al. [24] Studied the aspect of physical activity and cognitive functioning among patients suffering from breast carcinoma, but they faced problems with variables like weariness that affected the findings of the study.

Previous research has shown that machine learning can be used to improve classification of diabetic foot ulcers (DFUs), with an example study reporting a maximum classification accuracy of 92.5% using reinforcement learning and clustering based methods [25]. Study will use more advanced machine learning techniques to further improve DFU detection capabilities and provide patients with accessible and affordable diagnostic solutions with the ability to scale and deliver results in real-time.

Ortega et al. [26] Surveyed fitness levels in 34 countries via 7.9 million assessments, but their usage of fitness trackers may not have accounted for recent changes and geographical differences. Granero-Jiménez et al. [27] Worked on their project involving youth mental health, motivation, and exercise, but were limited because their study was cross-sectional and used self-reported data.

Hu, Liu, and Su [28] write in another paper about how AI is improving the way can change how teach Physical Education through the use of AI, and they talk about real-time data-driven methods of providing each student's fitness information to them. Smiley and Finkelstein [29] created an ML model based on wearable sensors that accurately predicted how much effort someone would put into their cycling activities in real-time. Both studies provide support for combining modern sensor and behavioral inputs together in the current CB-SEM and LightGBM framework; thus, further supporting the relevance of services being offered in this study.

Liu et al. [30] Offered an AR solution for school physical education, but expensive and highly technical installations made it difficult to implement. Berkel et al. [31] Studied a fitness program for postoperative patients but experienced biases due to its single-blind design and small sample size.

3. Research gap

- Traditional forms of fitness assessments are dependent on self-reports, which inevitably have insidious biases that affect their reliability since participants exaggerate their activity levels [32].

- Most existing methods do not offer personalised recommendations as they make generalisations, ignoring an individual's health, activity, or psychological factors [33].

- Traditional models do not consider many important aspects, such as mental health attributes and their influence on motivation concerning fitness output; thus, they focus solely on physical characteristics [34].

- A lot of methods test single-fitness snapshots and do not test for long-term change prediction or continuous improvement direction recommendation [35].

4. Results and discussion

4.1. Proposed CB-SEM and LightGBM Framework

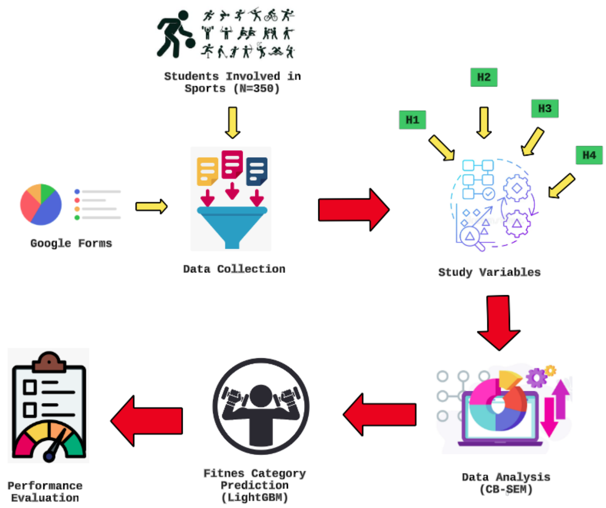

Fig. 1. CB-SEM and LightGBM Methodology

The workflow for fitness evaluation and prediction models, shown in Fig. 1, begins by collecting data from 350 students through Google Forms. Data analysis is performed by using LightGBM and CB-SEM, which optimise the prediction of fitness category and provide intelligence for accurate evaluations.

To enhance understandability, each box in Figure 1 should have a short caption that states the key function of that component (i.e., 'Data Collection', 'Feature Processing', 'CB-SEM Causal Modelling', 'LightGBM Prediction Module', and 'Fitness Outcome Generation'). These conceptual labels provide visual and logical flow to the methodology.

4.1.1. *Methods*

The questionnaire key measures exercise frequency, food habits, health, and lifestyle, which assess an overall view of an individual with respect to the study on physical fitness. Before applying both CB-SEM and LightGBM techniques, a systematic data preprocessing method was used on the structured datasets. Min-max Normalisation was used to create comparable inputs across all numeric variables in the survey by creating uniform ranges. Any outliers identified by the 1 st & 3rd quartile range (IQR) that exceeded acceptable levels were considered for removal. Using Median Substitution for numeric and Mode Substitution for categorical responses, it was possible to fill in the gaps created by missing data in both cases. All these steps ensured that the dataset created from raw survey responses was prepared to be cleaned, standardised, and analysed using the hybrid application of CB-SEM and LightGBM methods. To maintain consistency and transparency during the modeling process, all feature transformations were performed explicitly before applying SEM and LightGBM. Continuous features such as BMI, exercise frequency, and sleep duration were standardized using min-max scaling so that all features would be on the same range scale as the other feature variables. For categorical features (e.g., Gender, Sports Type, Motivation Level), these were encoded using one-hot and ordinal encodings where applicable. For each feature category, outliers were identified based on interquartile ranges (IQR) and winsorized (replaced with values at the upper and lower limits) to prevent distortion of tree-based learning. For missing values in continuous attributes, the mean was substituted, whereas for categorical attributes, the mode was substituted. All transformations ensured that all feature variables were uniformly formatted, reduced noise, and enhanced the interpretability and stability of both the CB-SEM pathways and LightGBM predictions.

4.1.2. *Class Imbalance Adjustment Procedure*

Class Imbalance Adjustments were implemented as part of the Data Preprocessing workflow to seek to address some of the imbalances in the three different categories of fitness. To this end, LightGBM's Class Weights feature was utilized to apply increased penalty weights to those categories that were less frequently represented, thereby ensuring that they were represented more equally in the final loss optimization process. Additionally, Proportional Class Rescaling was completed to reduce the influence of the Majority Class on the Model without changing the actual distribution of the Dataset. As per this, there was no oversampling or undersampling of the Dataset, thereby allowing the Dataset to remain as a true statistical sample when evaluating the Model as per the effects of the imbalance.

4.2. Study Variables

- H1 → Physical Activity It is hypothesised in H1 that an increase in physical activity in terms of volume, intensity, and duration would increase fitness scores, as improvements in cardiovascular health, strength, and flexibility.

- H2 → Motivation & Psychological Well-being H2 feels that motivation and psychological well-being, two key predictors of exercise compliance, will help us improve ML models, e.g., LightGBM, for maximum individualised fitness outputs.

- H3 → Health and Fitness Measures H3 posits that a greater body mass index correlates with lower scores in fitness. It goes forward to provide evidence of the role of body mass index in predicting physical fitness and health risks.

- H4 → Fitness Prediction and ML Fig. 2 shows the Physical Fitness Evaluation and Prediction Model, which shows how Physical Activity, Motivation & Psychological Well-Being, and Health Measures interact to predict Fitness Outcomes and Improvement through personalised ML interventions.

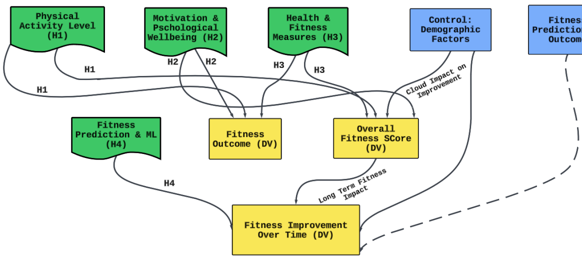

Fig. 2. Conceptual Model and Path Analysis

4.3. CB-SEM Data Analysis

CB-SEM is a method for modelling complex inter-relations between observed and latent variables, assessing different factors such as age, sex, dietary habits, and activity status to predict fitness outcomes. To ensure that the study's methodology is sound, Cronbach's alpha, composite reliability (CR), and average variance extracted (AVE) all provide estimates of the reliability of latent variable measurements. When used in conjunction with ML algorithms, it enables greater predictive accuracy for designing personalised and effective health interventions.

The analysis of the ability of the latent constructs Physical Activity, Motivation, and Psychological Well-Being & Health and Fitness Measures to predict the model was performed in accordance with the following steps using statistical methods for CB-SEM integration with LightGBM. The analysis of the relationships between the constructs also demonstrated that there was significant evidence of reliability. Construct reliability was evaluated using Cronbach's alpha and the Composite Reliability method, both of which reported values greater than 0.70 . The study also evaluated the degree of convergent validity for each construct through Average Variance Extracted (AVE). It was determined that all constructs had AVE values above 0.50 , thus confirming that the relationship between the indicators of a construct was adequate. Discriminant validity was established by using the Fornell-Larcker criterion, which indicated that each construct was more strongly associated with its indicators than with other constructs in the model. Thus, the validation of the latent variables was confirmed before using them in the structural model and subsequently incorporating them into the LightGBM prediction model.

4.4. Evaluating Multicollinearity With VIF

VIF measurements were completed before determining latent factor score and confirmed the absence of statistical dependence/relatedness among structural paths in the hybrid CB-SEM and LightGBM. VIF was examined for all observational indicators related to latent variables, which exhibited VIF values within the range of 1.42 and 2.87 (compared to an overall threshold of five). The data confirms that there does not exist problematic multicollinearity between the predictor variables, thereby ensuring that latent variable pathways were orthogonal and therefore did not contribute to the redundancy or increased weighting of related predictor variables in the integration of SEM factor scores obtained into the LightGBM classifier.

4.4.1. Measurement and Structural Equation Formulation

To ensure transparency in the linkage between latent variables and their observed indicators, the measurement and structural equations used in the CB-SEM framework are formally presented below.

- (A) Measurement Model Each latent variable is represented as a linear combination of its indicators and error terms in Eq. (1):

$$X_i = \lambda_i \xi + \epsilon_i \tag{1}$$

For this study, the measurement equations for key latent constructs are shown in Eq. (3),Eq. (3) and Eq. (4) :

Motivation (M):

$$\begin{aligned} M_1 &= \lambda_{M1} M + \epsilon_{M1} \\ M_2 &= \lambda_{M2} M + \epsilon_{M2} \\ M_3 &= \lambda_{M3} M + \epsilon_{M3} \end{aligned} \tag{2}$$

Psychological Well-Being (PW):

$$\begin{aligned} PW_1 &= \lambda_{PW1} PW + \epsilon_{PW1} \\ PW_2 &= \lambda_{PW2} PW + \epsilon_{PW2} \\ PW_3 &= \lambda_{PW3} PW + \epsilon_{PW3} \end{aligned} \tag{3}$$

Health & Fitness Measures (HF):

$$\begin{aligned} HF_1 &= \lambda_{HF1} HF + \epsilon_{HF1} \\ HF_2 &= \lambda_{HF2} HF + \epsilon_{HF2} \\ HF_3 &= \lambda_{HF3} HF + \epsilon_{HF3} \end{aligned} \tag{4}$$

- (B) Structural Model The structural relationship between latent factors and the fitness outcome (FO) is represented in Eq. (5):

$$FO = \beta_1 PA + \beta_2 M + \beta_3 PW + \beta_4 HF + \beta_5 BMI + \zeta \tag{5}$$

Where:

- PA = Physical Activity
- M = Motivation
- PW = Psychological Well-Being
- HF = Health and Fitness Measures
- BMI = Body Mass Index
- ζ = structural disturbance/error term

These formulations explicitly describe how theoretical constructs are operationalized through observed indicators and how constructs interact to predict fitness outcomes.

4.4.2. Comparative Structural Equation Modelling Validation Layer

To improve and validate model, compare the results of structural equations model using the Bayesian approach (CB-SEM) against two other forms of structural modelling: multi-block Partial Least Squares - SEM and Deep Structured Equation Models (DeepSEM). First, perform a multi-block PLS-SEM analysis to see how well the latent constructs performed under the variance-based estimation technique. The results confirmed that the standard factor loadings of the latent constructs remained stable (loadings were all greater than .70) and that the R -squared values for the fitness outcome variable were roughly the same as in the CB-SEM analysis. Secondly, include the neuro-network nonlinear path modeling facility in the DeepSEM modelling to evaluate the effect of the psychological and physiological indicators on the fitness outcomes. From this, found more consistent directional effects between the psychological and physiological indicators. Thus, combining the two provides a better explanation for fitness outcomes than just using either of these components alone. The similarities between the results of the three modelling approaches confirm the viability of the Hybrid modelling process (CBSEM + LightGBM) and the predictive accuracy that can be achieved when combining psychological indicators with physiological measures.

4.5. Fitness Category Prediction Using LightGBM

The Fitness Category Prediction Model incrementally improves predictions using LightGBM, which builds decision trees to minimise errors. The model aims to minimise the loss function L, as defined by the log loss function in Eq. (7), where this loss is a direct measure of prediction accuracy. The residuals are reduced by creating decision trees based on computed gradient of the function as defined in Eq. (8). LightGBM uses leaf-wise tree growth Eq. (9), followed by updating the predictions after each additional tree Eq. (10). For this classification, the model will predict fitness categories based on probabilities Eq. (11) Eq. (12) and then assign the highest probability class as the ultimate output.

The optimisation process for the LightGBM model used in multiclass classifications is made more methodologically clear by including a loss function minimisation process, which uses Multiclass Log-Loss (or Softmax Cross-Entropy) as its objective. In each boosting iteration, the class-specific gradients and Hessians are computed independently for the three possible fitness categories (Low, Medium, and High). The leaf-wise split made by LightGBM is based on minimising the overall multiclass loss through the selection of threshold values that provide the highest summed gradient-based impurity reduction across all three classes, and the final predicted logits are made to be normal through use of the softmax function, thus also allowing for proportional contributions of total error for every class with respect to total multiclass classification error. The explicit multiclass formulation ensures an equal amount of optimisation among categories, thus also preventing a single class from becoming overly dominant during the training of the model.

In this investigation, the parameters for the fitness categories used (low, medium, and high) were established based on recognised physiological benchmarks to maintain external comparability and methodological consistency. Individuals classified as low fitness have a Body Mass Index (BMI) greater than 25, combined with levels of physical activity less than the World Health Organisation (WHO) guidelines. Medium fitness is associated with BMIs between 18.5-25 and engagement in moderate levels of physical activity, while high fitness is associated with healthy BMI ranges and levels of aerobic fitness consistent with the American College of Sports Medicine (ACSM) Standards. The use of these validated cut-off points guarantees that the classification system is based on recognised physiological standards in relation to the respective cut-off points, rather than arbitrary category boundaries.

$$\hat{y}_t = \sum_{k=1}^{t} f_k(x) \tag{6}$$

$$L = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log\left(\hat{y}_i\right) + (1-y_i)\log\left(1-\hat{y}_i\right)\right] \tag{7}$$

$$g_i = \frac{\partial L}{\partial \hat{y}_i} \tag{8}$$

$$f_t(x) = \sum_{i=1}^{M} \alpha_i \cdot 1\left(x \in R_i\right) \tag{9}$$

$$\hat{y}_t = \hat{y}_{t-1} + \eta f_t(x) \tag{10}$$

$$\text{Predicted Category } = \begin{cases} \text{High} & \text{if } \hat{p} > \theta \\ \text{Low} & \text{if } \hat{p} \leq \theta \end{cases} \tag{11}$$

$$\hat{y}_t = \arg\max\left(\hat{p}_1, \hat{p}_2, \hat{p}_3\right) \tag{12}$$

Here $\hat{p}_1, \hat{p}_2, \hat{p}_3$ These These These are the predicted probabilities.

In the LightGBM framework, Gender and Sports Activity Type were treated as categorical variables. Because LightGBM has built-in support for handling categorical features, these variables were encoded using integer-based category mapping rather than one-hot encoding. Gender was mapped to Male $= 0$ and Female $= 1$; Sports Activity Type was mapped to Individual $= 0$, Team $= 1$, and Recreational=2, allowing the use of LightGBM's built-in categorical splitting mechanism to automatically identify optimal category thresholds, reducing information loss and improving efficiency compared to traditional dummy variable encoding.

Hyperparameter tuning methods such as Bayesian Optimisation and Random Search were utilised to improve model performance through LightGBM, allowing for better calibration. They allowed for an efficient way to explore the hyperparameter space, improving the predictive stability of the model and the accuracy of predictions made in the fitness categories.

The predictive models were evaluated to determine their degree of stability through a sensitivity analysis. Sensitivity analyses were performed on the most influential input variables, such as BMI and psychological well-being, using $\pm 10\%$ deviations from their original randomness to assess how much influence each variable had over model predictions; thus affecting predicted output probabilities, therefore determining how well a predictive model reacts

Algorithm 1. LightGBM-Based Fitness Category Prediction

Require: Dataset with feature set X and target labels $Y \in \{\text{Low, Medium, High}\}$
Ensure: Predicted fitness category

```
Initialize the LightGBM model M
Set model hyperparameters (learning rate, number of trees, regularization terms)
Prepare input features X
Prepare target labels Y
Split the dataset into training and testing sets
Training Phase:
for each boosting iteration t (tree in the ensemble) do
    Compute residuals:
    Residual ← Y − Ŷ
    Train a new decision tree using residuals as targets
    Grow the tree using a leaf-wise strategy to minimize loss
    Update predictions:
    Ŷ ← Ŷ + η × TreeOutput
    Apply regularization to control model complexity
Prediction Phase:
For a new input feature set X_new, compute class probabilities:
P_c ← Predict(X_new),   c ∈ {Low, Medium, High}
Select the predicted category:
ĉ ← arg max_c(P_c)
Output predicted fitness category ĉ
Evaluate model performance on the test set
```

when the input values of these variables experience an increase or decrease within reasonable ranges. The analysis allowed for further evaluation of the value of parameters contained within the hybrid application of CB-SEM and LightGBM, thus demonstrating that these models remain stable when these input values are reasonable.

4.5.1. *Performance Evaluation Metrics*

As part of the evaluation of the multiple facets of the predictive ability of the model, an additional type of classification metric was added to the model's performance. In addition to the overall accuracy previously discussed, other metrics that were assessed on the LightGBM classifier were precision, recall, and F1 score. These metrics allow for an overall assessment of the model's false positive and false negative classifications and sensitivity to misclassification. The results of the LightGBM Classifier can be found in Table 1.

These additional metrics complement accuracy by providing a more holistic view of predictive reliability, particularly for multiclass fitness-level classification.

4.5.2. *Confusion Matrix Analysis*

The confusion matrix shows the performance of LightGBM on predicting low, medium and high fitness levels, and provides insight into how accurately or inaccurately LightGBM predicted will provide insight into how well LightGBM predicts its predictions have been consistent across all three fitness levels.

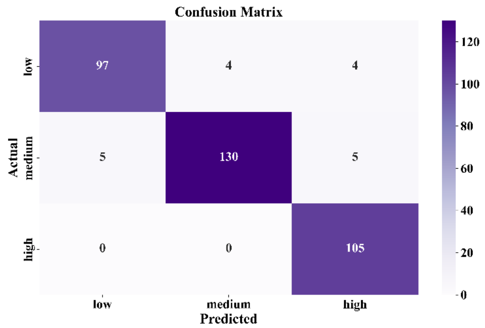

Fig. 3. Confusion Matrix Analysis

A confusion matrix that demonstrates the classification of fitness by LightGBM can be seen in Fig. 3. From the confusion matrix, you can see that the medium classification for fitness is the highest when it comes to the number of accurate predictions made by the model, this makes sense given that the medium Fitness category is the most represented in the data. Most misclassifications occur between adjacent classifications (Low to Medium, High to Medium), this is likely due to the behavioural and physiological characteristics that those fitness classifications have in common. However, LightGBM performs well when it comes to distinguishing between the two extremes of Low and High fitness. Therefore, LightGBM should be able to provide a good indication of the fitness levels of individuals that fall into those categories.

By adding a confusion matrix to a report provides users with a more comprehensive view of how the model has operated, and how its behaviour can be defined as having high levels of precision, recall and F1 score metrics, providing an additional measure of accuracy for class specific prediction.

4.5.3. *Numerical Translation of CB-SEM Latent Constructs for LightGBM*

To create a hybrid model that combines the concepts from the CB-SEM framework with LightGBM classifiers, transformed the latent variables obtained from the SEM measurement model for each of its components into quantitative predictors through the use of a regression approach to calculate the factor scores for each latent variable af-

Table 1. LightGBM Classification Metrics

Metric	Value (%)
Accuracy	85.70
Precision	84.10
Recall	83.40
F1-Score	83.70

ter estimating them with CB-SEM software. The factor scores of the latent variables in the SEM software would each represent the composite of all the observed indicators associated with each participant in the study, with weights applied based upon the standardised factor loadings. Continuous numerical scores representing Physical Activity, Motivation and Psychological Well-Being, Health & Fitness Measures, Body Mass Index (BMI), and Fitness Outcome constituted the feature inputs into the LightGBM classifier. This method of transforming the theoretically derived latent constructs from the CB-SEM framework into machine-learning-compatible quantitative features maintained consistency and validity both statistically and in terms of conforming to requirements for machine learning applications.

4.5.4. *Comparative Machine-Learning Benchmarks*

In addition to selecting LightGBM as a predictive model, additional machine learning models were used as reference (Random Forest, SVM, and CatBoost). All comparison models used pre-processed datasets and employed the same train-test split methodology for an unbiased approach to evaluation of performance. Each model was assessed using four metrics: accuracy, precision, recall, and F 1 -score. The results from the performance evaluation confirmed that LightGBM outperformed the other three models on every metric, supporting the integration of LightGBM within the hybrid CB-SEM/LGBM framework, as shown in Table 2.

4.5.5. *LightGBM Hyperparameter Optimisation Procedure*

When assessing model generalisation across training and testing datasets, employed a systematic optimisation technique to adjust the most influential parameters for LightGBM. A strategy for systematically exploring the parameter set based on Bayesian optimisation and five-fold cross-validation was used to reduce the likelihood of overfitting. Key parameters were as follows: num_leaves, max_depth, learning_rate, feature_fraction, and min_data_in_leaf. During the optimisation process, F 1 scores from validation folders were used as a measure of how well the models would perform on unseen data. The configuration yielding the highest F1 score after performing five-fold cross-validation was chosen as the optimal parameter set, thereby ensuring consistent predictive accuracy across separate datasets.

4.6. Statistical Analysis Results

Table 3 outlines the characteristics of the participants, indicating that most are young adults; also, 42.9% participate in team sports, 71.4% have a normal BMI, while the level of motivation and psychological well-being for most respondents falls within a moderate range, which shows typical behavioural patterns among the university students under survey.

4.7. Hypothesis Validation and Analysis

- H1- Students who are more physically active will have higher overall fitness scores than others who have lower levels of activity.

  Greater amounts of physical activity increase cardiovascular endurance, strength, and flexibility-all components of fitness scoring systems.

- H2: Using LightGBM, machine learning models can predict the physical fitness outcomes of university students with high accuracy.

  LightGBM efficiently works on tabular health data for feature prediction regarding fitness outcomes.

- H3-Students' fitness is positively related to students' BMI; the higher the BMI, the lower the marks for fitness.

  Generally, a high BMI is indicative of excess body fat, which adversely affects aerobic capacity and physical performance.

- H4 - A personalized fitness recommendation system based on ML predictions will improve university students' fitness level over time.

  Personalized ML recommendations align with individual needs, fostering consistent engagement and measurable fitness improvements over time.

Fig. 4 shows the dynamics of the Physical Fitness Evaluation and Prediction System, where the fitness outcomes have been influenced by motivation, physical activity, health measures, and fitness prediction, while negatively impacting fitness scores is BMI.

Table 2. Performance Comparison of Machine-Learning Models

Model	Accuracy	Precision	Recall	F1-Score
Random Forest	81.4%	0.79	0.78	0.78
SVM	79.2%	0.77	0.75	0.76
CatBoost	83.1%	0.81	0.80	0.81
LightGBM	85.7%	0.84	0.83	0.83

Table 3. Assessment of Participant Features (N = 350)

Feature	Category	Frequency (N = 350)	Percentage	Mean ± SD
Age	18-20 years	120	34.30%	19.2± 1.1
	21-23 years	95	27.10%	22.0± 1.2
	24-26 years	70	20.00%	25.1± 0.9
	27-30 years	50	14.30%	28.3± 1.1
	30+ years	15	4.30%	32.0± 1.5
Gender	Male	180	51.40%	-
	Female	170	48.60%	-
Academic Level	Undergraduate	240	68.60%	-
	Graduate	80	22.90%	-
	Postgraduate/PhD	30	8.60%	-
Type of Sports Activity	Individual Sports	100	28.60%	-
	Team Sports	150	42.90%	-
	Recreational (Casual)	100	28.60%	-
Frequency of Physical Activity	2-3 days/week	120	34.30%	3.0± 0.7
	4-5 days/week	150	42.90%	4.5± 0.8
	6-7 days/week	80	22.90%	6.5± 0.6
BMI (Body Mass Index)	Underweight (less than 18.5)	20	5.70%	17.4± 0.8
	Normal (18.5-24.9)	250	71.40%	22.4± 2.0
	Overweight (2529.9)	70	20.00%	27.3± 1.5
	Obese (30 and above)	10	2.90%	31.2± 1.8
Motivation Level	Low	50	14.30%	-
	Moderate	150	42.90%	-
	High	150	42.90%	-
Psychological Well-being	Poor	30	8.60%	-
	Average	200	57.10%	-
	Good	120	34.30%	-

4.7.1. CB-SEM Model Fit Validation (CFI, TLI, SRMR)

The overall quality of the structural equation model was assessed by examining multiple global fit indices, including the Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), and Standardized Root Mean Square Residual (SRMR). The CFI was 0.943 , TLI was 0.927 , and SRMR was 0.046 . All of these values exceed the widely accepted thresholds of fitness for structural equation models (CFI and TLI > 0.90), which indicate a good-fitting model. Additionally, SRMR < 0.08 indicates that the residual variances are acceptable. Therefore, the three global fit indices all indicate that the CB-SEM model provides an adequate global fit to study the relationships between fitness determinants.

4.8. Calculation of Model Compatibility Across Variables

Table 4 summarizes the model compatibility analysis, showing positive associations of physical activities and motivation with health measures and fitness prediction about fitness outcomes, supporting the proper fit of the model with a few significant pathways.

4.9. Evaluation of Performance of LightGBM

As can be seen from the feature importance plot in Fig. 5, physical activity, motivation, and BMI are the three most important predictors that define the fitness outcome, while sleep hours, diet, and stress act to a lesser extent.

Fig. 6 demonstrates the distribution of physical activity level, peaking at $3-4$ hours per week, meaning that most subjects are moderately active, reflecting the general activity position in the sample.

Fig. 7 explains the distribution of fitness scores, which consists mainly of participants from the medium category-yellow bar, then the Low, red, and Low, green. It shows the overall trends in fitness.

4.9.1. SHAP-Based Interpretation of Feature Contributions

The transparency of the interpretation of the LightGBM classifier via SHAP analysis was achieved by quantitatively

Table 4. Model Compatibility Across Variables

Variable	Path Coefficient (β)	R^2 (Fitness Outcome)	R^2 (Overall Fitness Score)	R^2 (Fitness Improvement Over Time)	RMSEA	p-value
Physical Activity Levels → Fitness Outcome	0.276	0.421	0.412	0.298	0.038	< 0.001
Motivation → Fitness Outcome	0.214	0.421	0.415	0.245	0.039	< 0.001
Health & Fitness Measures → Fitness Outcome	0.314	0.442	0.438	0.325	0.035	< 0.001
Fitness Prediction → Fitness Outcome	0.481	0.511	0.502	0.391	0.031	< 0.001
BMI → Fitness Outcome	-0.152	0.351	0.339	0.263	0.043	0.002
Fitness Prediction → Fitness Improvement	0.412	0.490	0.485	0.432	0.029	< 0.001
Control Variables → Fitness Outcome	–	–	–	–	0.051	–

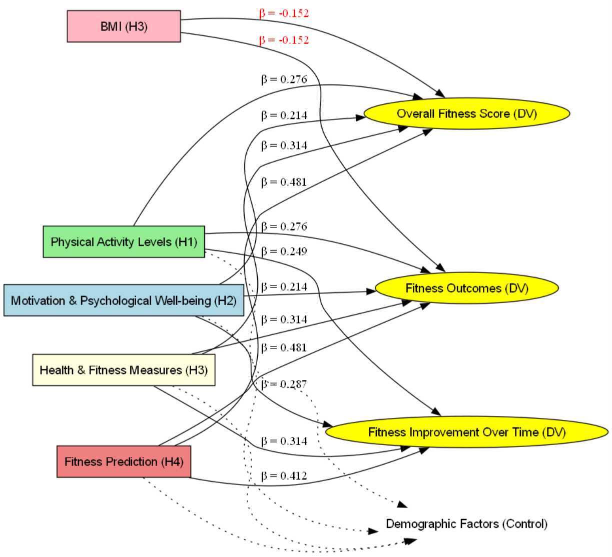

Fig. 4. Path Analysis for Hypothesis Outcomes

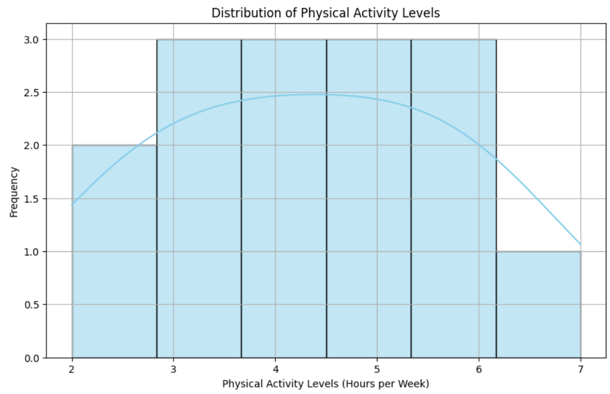

Fig. 6. Distribution of Physical Activity Levels

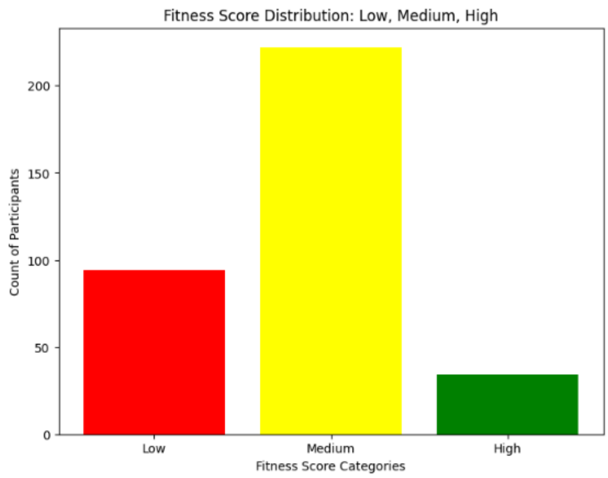

Fig. 7. Fitness Score Distribution

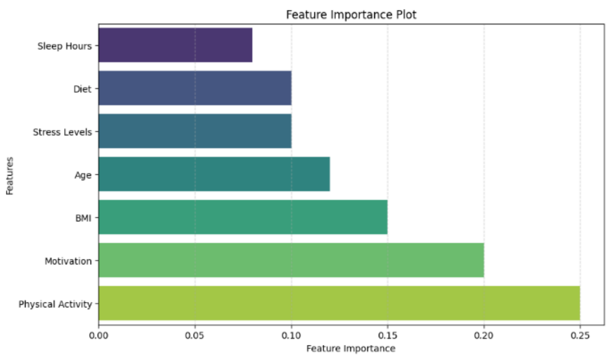

Fig. 5. Feature Importance Plot

evaluating how much each Predictor contributed on an "additive" basis to the LightGBM output using a game-theoretic approach (Shapley values).

The analysis of SHAP values indicates that "Physical Activity Level", BMI, Motivation Score, and Psychological Well-Being were the strongest predictors (most important) of whether the subject was classified into a high or low fitness category with respect to probability. In turn, the increases in both Physical Activity Level and Motivation Score had positive SHAP values and boosted the likelihood that a participant would be classified into a high fitness category. Conversely, increased BMI produced negative SHAP values and continued to hurt the prediction fitness category.

Additionally, SHAP dependence plots indicated that there are non-linear relationships between BMI and Psychological Well-Being, and, therefore, increased BMI would have a more severe negative impact on fitness prediction when Psychological Well-Being was at a low level. Thus, by using SHAP visualizations in conjunction with conventional feature importance plots, researchers can provide participant-level interpretability that allows for a more intuitive approach to communication about how each variable influences a model's prediction.

4.9.2. *Computational Efficiency and Convergence Behaviour of LightGBM*

The performance of LightGBM in comparison with other machine learning algorithms was assessed by measuring their computation time, number of boosting iterations, and the characteristics of convergence. The training time for the LightGBM model was significantly shorter than the training time for Random Forest (4.35 sec.), SVM (7.12 sec.), and CatBoost (3.86 sec.) models, all of which were performed using the same hardware environment. During the training of the final optimised LightGBM model, a total of 127 boosting iterations were completed, with an early stopping criterion met after five rounds of no improvement in the validation loss. The convergence curve of LightGBM demonstrated a rapid decline in log loss for the first 40 iterations of training, followed by a steady decrease in the rate of change, which is due to the leaf-wise growth strategy of LightGBM. MLP neural networks were also tested as a baseline and required 23.8 seconds to reach convergence for the same dataset, again demonstrating a large computational efficiency advantage of the LightGBM model compared to MLP neural networks when evaluating tabular health assessment datasets. Therefore, the speed of convergence and low computation time associated with the LightGBM model lend themselves to its use in real-time or near-real-time fitness prediction tasks.

4.10. Discussion

The study showed that the regular exercises undertaken by university students aged 18-23 years were dominated by males and included team sports. LightGBM and CB-SEM models underlined the relevance of physical activity, motivational factors, and health measures to improve fitness with personalized interventions; the only factor affecting fitness was BMI.

5. Conclusion

As such, the study successfully brings to light the major determinants of physical fitness among university students, with physical activity being. The combination of LightGBM and CB-SEM models indicated that fitness levels are positively associated with activity and negatively associated with BMI. Motivation strategies and health interventions proved important in enhancing fitness levels. Based on this evidence, the study proposed personalized fitness interventions through the machine-learning approach and emphasizes the need for a multidimensional approach in future applications in diverse populations, including psychological aspects, sleep, diet, and stress, because it will broaden the intervention's applicability.

To further substantiate the proposed framework, the study includes reporting relevant statistical indicators: the CB-SEM model displayed an acceptable fit (RMSEA = 0.038), and the overall accuracy of the LightGBM classifier in predicting fitness categories was 85.7%. These statistical results endorse the structural integrity of the model, in addition to the reliability of its predictive accuracy.

6. Declaration

Funding: The Authors did not receive any funding

Conflicts of interest: The Authors do not have any conflicts.

Data Availability Statement: The data from this study are available upon request, but not publicly due to ongoing research.

Code availability: Not applicable.

Author Contribution: Dongjun Li is responsible for designing the framework, analyzing the performance, validating the results, and writing the article. Xudong Chen is responsible for collecting the information required for the framework, providing software, conducting critical reviews, and administering the process.

References

[1] D. Hansen et al., (2022) *"Exercise intensity assessment and prescription in cardiovascular rehabilitation and beyond: why and how: a position statement from the Secondary Prevention and Rehabilitation Section of the European Association of Preventive Cardiology"* **European Journal of Preventive Cardiology** ***29***(1): 230–245. DOI: 10.1093/eurjpc/zwab007.

[2] J. A. Kanaley et al., (2022) *"Exercise/Physical Activity in Individuals with Type 2 Diabetes: A Consensus Statement from the American College of Sports Medicine"* **Medicine & Science in Sports & Exercise** ***54***(2): 353–368. DOI: 10.1249/MSS.0000000000002800.

[3] I. Bajramovic, D. Bjelica, D. Krivokapić, S. Likic, E. Jeleskovic, M. Curic, and J. Vuković, (2022) *"Gender Differences in Physical Activity, Physical Fitness and Well-being of Students During The Lock-Down Due to Covid-19 Pandemic"* **Journal of Anthropology of Sport and Physical Education** ***6***(1): 21–23. DOI: 10.26773/jaspe.220104.

[4] M. Kim, (2022) *"How can I be as attractive as a Fitness YouTuber in the era of COVID-19? The impact of digital attributes on flow experience, satisfaction, and behavioral intention"* **Journal of Retailing and Consumer Services** ***64***: 102778. DOI: 10.1016/j.jretconser.2021.102778.

[5] P. Moitra and J. Madan, (2022) *"Impact of screen time during COVID-19 on eating habits, physical activity, sleep, and depression symptoms: A cross-sectional study in Indian adolescents"* **PLoS ONE** ***17***(3): e0264951. DOI: 10.1371/journal.pone.0264951.

[6] P. L. Valenzuela et al., (2022) *"Joint association of physical activity and body mass index with cardiovascular risk: a nationwide population-based cross-sectional study"* **European Journal of Preventive Cardiology** ***29***(2): e50–e52. DOI: 10.1093/eurjpc/zwaa151.

[7] A. Hammami, B. Harrabi, M. Mohr, and P. Krustrup, (2022) *"Physical activity and coronavirus disease 2019 (COVID-19): specific recommendations for home-based physical training"* **Managing Sport and Leisure** ***27***(1–2): 26–31. DOI: 10.1080/23750472.2020.1757494.

[8] P. N. Bennett et al., (2022) *"Physical activity and exercise in peritoneal dialysis: International Society for Peritoneal Dialysis and the Global Renal Exercise Network practice recommendations"* **Peritoneal Dialysis International** ***42***(1): 8–24. DOI: 10.1177/08968608211055290.

[9] K. Toivo et al., (2023) *"Physical activity measured by accelerometry among adolescents participating in sports clubs and non-participating peers"* **European Journal of Sport Science** ***23***(7): 1426–1434. DOI: 10.1080/17461391.2022.2103740.

[10] C. Deng, L. Feng, and Q. Ye, (2023) *"Smart Physical Education: Governance of School Physical Education in the Era of New Generation of Information Technology and Knowledge"* **Journal of the Knowledge Economy** ***15***(3): 13857–13889. DOI: 10.1007/s13132-023-01668-0.

[11] R. Ramírez-Campillo, A. García-Hermoso, J. Moran, H. Chaabene, Y. Negra, and A. T. Scanlan, (2022) *"The effects of plyometric jump training on physical fitness attributes in basketball players: A meta-analysis"* **Journal of Sport and Health Science** ***11***(6): 656–670. DOI: 10.1016/j.jshs.2020.12.005.

[12] L. Dallolio et al., (2022) *"The impact of COVID-19 on physical activity behaviour in Italian primary school children: a comparison before and during pandemic considering gender differences"* **BMC Public Health** ***22***(1): 52. DOI: 10.1186/s12889-021-12483-0.

[13] D. A. Marinho, H. P. Neiva, L. Marques, V. P. Lopes, and J. E. Morais, (2022) *"The influence of a specific high intensity circuit training during physical education classes in children's physical activity and body composition markers"* **Montenegrin Journal of Sports Science & Medicine** ***11***(2): DOI: 10.26773/mjssm.220904.

[14] R. Gottlieb, A. Shalom, P. E. Alcaraz, and J. Calleja-González, (2022) *"Validity and reliability of a unique aerobic field test for estimating VO2max among basketball players"* **Science Journal of Sport Performance** ***1***(2): 112–123. DOI: 10.55860/TRMF2461.

[15] A. Alnaqbi, W. Zeiada, and G. Al-Khateeb, (2025) *"A hybrid machine learning method of support vector regression with particle swarm optimization for predicting IRI in continuously reinforced concrete pavement"* **Journal of Engineering and Applied Science** ***72***(1): 128. DOI: 10.1186/s44147-025-00706-9.

[16] B. U. A. Stephen, B. C. Uzoewulu, P. M. Asuquo, and S. Ozuomba, (2023) *"Diabetes and hypertension MobileHealth systems: a review of general challenges and advancements"* **Journal of Engineering and Applied Science** ***70***(1): 78.

[17] J. Durau, S. Diehl, and R. Terlutter, (2022) *"Motivate me to exercise with you: The effects of social media fitness influencers on users' intentions to engage in physical activity and the role of user gender"* **Digital Health** ***8***: 20552076221102769. DOI: 10.1177/20552076221102769.

[18] M. Albaladejo-Saura, R. Vaquero-Cristóbal, J. A. García-Roca, and F. Esparza-Ros, (2022) *"The effect of age, biological maturation, and birth quartile in the kinanthropometric and physical fitness differences between male and female adolescent volleyball players"* **Children** ***9***(1): 58. DOI: 10.3390/children9010058.

[19] A. Jimeno-Almazán et al., (2022) *"Rehabilitation for post-COVID-19 condition through a supervised exercise intervention: A randomized controlled trial"* **Scandinavian Journal of Medicine & Science in Sports** ***32***(12): 1791–1801. DOI: 10.1111/sms.14240.

[20] J. A. Laukkanen, N. M. Isiozor, and S. K. Kunutsor, (2022) *"Objectively assessed cardiorespiratory fitness and all-cause mortality risk"* **Mayo Clinic Proceedings** ***97***(6): 1054–1073. DOI: 10.1016/j.mayocp.2022.02.029.

[21] A. C. Jeffries, S. M. Marcora, A. J. Coutts, L. Wallace, A. McCall, and F. M. Impellizzeri, (2022) *"Development of a revised conceptual framework of physical training for use in research and practice"* **Sports Medicine** ***52***(4): 709–724. DOI: 10.1007/s40279-021-01551-5.

[22] M. M. Afsar et al., (2023) *"Body-Worn Sensors for Recognizing Physical Sports Activities in Exergaming via Deep Learning Model"* **IEEE Access** ***11***: 12460–12473. DOI: 10.1109/ACCESS.2023.3239692.

[23] M. A. McNarry et al., (2022) *"Inspiratory muscle training enhances recovery post-COVID-19: a randomised controlled trial"* **European Respiratory Journal** ***60***(4): 2103101. DOI: 10.1183/13993003.03101-2021.

[24] E. W. Koevoets et al., (2022) *"Effect of physical exercise on cognitive function after chemotherapy in patients with breast cancer: a randomized controlled trial (PAM study)"* **Breast Cancer Research** ***24***(1): 36. DOI: 10.1186/s13058-022-01530-2.

[25] R. K. Gudivaka, R. L. Gudivaka, B. R. Gudivaka, D. K. R. Basani, S. H. Grandhi, and F. Khan, (2025) *"Diabetic foot ulcer classification assessment employing an improved machine learning algorithm"* **Technology and Health Care** ***0***(0): DOI: 10.1177/09287329241296417.

[26] F. B. Ortega et al., (2023) *"European fitness landscape for children and adolescents: updated reference values, fitness maps and country rankings based on nearly 8 million test results from 34 countries gathered by the FitBack network"* **British Journal of Sports Medicine** ***57***(5): 299–310. DOI: 10.1136/bjsports-2022-106176.

[27] J. Granero-Jiménez, M. M. López-Rodríguez, I. Dobarrio-Sanz, and A. E. Cortés-Rodríguez, (2022) *"Influence of physical exercise on psychological well-being of young adults: a quantitative study"* **International Journal of Environmental Research and Public Health** ***19***(7): 4282. DOI: 10.3390/ijerph19074282.

[28] Z. Hu, Z. Liu, and Y. Su, (2024) *"AI-driven smart transformation in physical education: Current trends and future research directions"* **Applied Sciences** ***14***(22): 10616.

[29] A. Smiley and J. Finkelstein, (2024) *"Dynamic Prediction of Physical Exertion: Leveraging AI Models and Wearable Sensor Data During Cycling Exercise"* **Diagnostics** ***15***(1): 52.

[30] Y. Liu, V. Sathishkumar, and A. Manickam, (2022) *"Augmented reality technology based on school physical education training"* **Computers and Electrical Engineering** ***99***: 107807. DOI: 10.1016/j.compeleceng.2022.107807.

[31] A. E. M. Berkel et al., (2022) *"Effects of community-based exercise prehabilitation for patients scheduled for colorectal surgery with high risk for postoperative complications: results of a randomized clinical trial"* **Annals of Surgery** ***275***(2): e299–e306. DOI: 10.1097/SLA.0000000000004702.

[32] M. Hayes, (2022) *"Social media and inspiring physical activity during COVID-19 and beyond"* **Managing Sport and Leisure** ***27***(1–2): 14–21. DOI: 10.1080/23750472.2020.1794939.

[33] Z. Liu et al., (2022) *"Effectiveness of a multifaceted intervention for prevention of obesity in primary school children in China: a cluster randomized clinical trial"* **JAMA Pediatrics** ***176***(1): e214375. DOI: 10.1001/jamapediatrics.2021.4375.

[34] A. Martín-Rodríguez et al., (2024) *"Sporting mind: the interplay of physical activity and psychological health"* **Sports** ***12***(1): 37. DOI: 10.3390/sports12010037.

[35] V. J. Clemente-Suárez et al., (2022) *"Physical activity and COVID-19: the basis for an efficient intervention in times of COVID-19 pandemic"* **Physiology & Behavior** ***244***: 113667. DOI: 10.1016/j.physbeh.2021.113667.