Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X
Volume 4, Issue 3, 2026

MACHINE LEARNING-BASED BODY MASS INDEX (BMI) CATEGORY CLASSIFICATION USING ANTHROPOMETRIC AND LIFESTYLE FEATURES FOR ATHLETE HEALTH ASSESSMENT

[1]Hina, [2]Muhammad Sohail, [3]Abdul Waheed Mughal, *[4]Adam Khan

[1]Department of Sport Sciences and Physical Education, Sarhad University of Science and Information Technology, Peshawar, Pakistan

[2]Department of Computer Science, Sarhad University of Science and Information Technology Peshawar, Pakistan

[3]Department of Sport Sciences and Physical Education, Sarhad University of Science and Information Technology, Peshawar, Pakistan

[4]Department of Computer Science, Sarhad University of Science and Information Technology Peshawar, Pakistan

*[1]qaisarhina1999@gmail.com, [2]sohail.csit@suit.edu.pk, [3]dean.ss@suit.edu.pk, *[4]adam.me@suit.edu.pk*

Keywords
Body Mass Index, Machine Learning, Decision Tree, Data Augmentation, Athlete Health, Classification

Article History
Received on 25 Feb, 2026
Accepted on 28 March, 2026
Published on 29 March, 2026

Corresponding Author:
Adam Khan

Abstract

Body Mass Index (BMI) is a widely used indicator for assessing general health and body composition; however, its applicability in athletic populations is limited due to its inability to distinguish between fat mass, lean mass, and sport-specific physiological adaptations. This study proposes an interpretable machine learning-based framework for BMI category classification using anthropometric and lifestyle-related features. The data involved height, weight, training age, training duration and how often meals are taken per day. Since there are few observations (about 100 samples) available, a controlled data augmentation method was used, which employs Gaussian perturbation to increase the dataset size to 500 samples with real physiological relationships. The BMI values were re-calculated and divided into four groups namely; Underweight, Normal, Overweight and Obese. An 80:20 traintest split was used to train a Decision Tree classifier which was run in the scikit-learn library of Python. An overall classification accuracy of the model was about 94-95%. The confusion matrix analysis revealed that the majority of samples were properly classified with slight misclassification between similar categories like Overweight and Obese. The analysis of feature importance showed that height and weight were the most significant predictors, and lifestyle-related variables played a less significant role. The findings prove the usefulness of interpretable machine learning methods in BMI classification and indicate their possible use in athlete health assessment. Future research is required on larger datasets and more complicated modeling approaches to enhance generalization.

1. Introduction

Body Mass Index (BMI) is among the most commonly utilized anthropometric measures of body composition and overall health condition [1]. Its simplicity, low cost, and ease of calculation have led to its widespread use in clinical, population health, and sporting applications [2]. In spite of its popularity, BMI does have significant limitations, especially in athletic groups, where it does not distinguish fat mass, lean muscle mass, bone density, and fat distribution [3].

These are particularly high in sports science as the athletes show sport specific physiological adaptations that are affected by training intensity, nutrition, and performance demands [4]. As a result, people strikingly divergent in body composition can have the same BMI, and this makes BMI less reliable when used on its own to determine the well-being of an athlete [5].

Recent developments in Artificial Intelligence (AI) and Machine Learning (ML) have offered promising methods to model non-linear and complex relationships between physiological and behavioral variables [6]. These techniques enable the integration of multiple features to uncover latent patterns and improve predictive performance beyond traditional statistical methods [7].

In that regard, the current study suggests a machine learning-based system of BMI category classification based on anthropometric and lifestyle-associated characteristics. In contrast to traditional methods, which assume BMI as a descriptive measure, this study models BMI categories as dependent on quantifiable health and behavioral factors.

It uses a structured dataset in terms of height, weight, training age, training duration, and frequency of meals per day. A controlled synthetic augmentation plan is used to overcome the small data size to expand the sample size without losing realistic physiological correlation.

The primary contributions of this study are as follows:

- An interpretable and data-efficient framework for BMI classification under limited data conditions.
- Application of controlled Gaussian-based data augmentation to enhance dataset robustness.
- Analysis of class-wise performance, model complexity, and feature importance.

The findings aim to support more data-driven and context-aware approaches to athlete health evaluation.

2 Literature Review

The evaluation of body structure in sports science has been transformed greatly by the constraints of traditional anthropometric data, such as BMI [8]. Although BMI is a popular screening method, it fails to measure the most important aspects of screening, including fat mass, lean tissue, and visceral adiposity, which restrict its applicability in athletes [9].

Athletes generally have greater muscle mass and physiological specifics, which can result in inaccurate BMI classifications [10]. Consequently, the use of BMI alone can give an inaccurate health and fitness status interpretation [11].

To overcome these issues, recent studies [12]-[17] have paid more attention to multidimensional and data-driven methods. This is especially where machine learning methods can be applied, since they can be used to model the complex, non-linear relationships of the variables of physiological, behavioral, and performance-related variables [18]. Applications of ML in sports science include performance prediction, injury prevention, health monitoring, and biometric analysis [19].

Moreover, recent research has investigated how AI-driven techniques can be used to analyze body composition and found that they are more predictive and capable of detecting latent patterns that are not detected by conventional techniques [20]. Nevertheless, most of these methods are based on advanced imaging methods or on large-scale datasets, which are often not practical or resource-constrained.

Although these developments have been made, several methodological issues still exist, such as the small sample units, inability to interpret the models, and generalization to a variety of populations [21]. Specifically, the most precise models, such as black-box models, are not always transparent and, therefore, cannot be applied to real-life decision-making in both healthcare and sport settings [22] [23].

Consequently, interpretable machine learning methods that are predictive, as well as transparent and useful, are also required. Although some studies have used post-hoc explainability, including SHAP and LIME [22], in this study, intrinsic interpretability is undertaken by using Decision Tree modeling [24], which gives a direct understanding of decision rules. In this regard, the present study employs a Decision Tree classifier [25] as an interpretable baseline model for BMI classification.

3 Methodology

3.1 Study Design and Dataset

The present study applies a quantitative machine learning method to categorize the BMI [26] with the help of anthropometric and lifestyle-related characteristics. The following variables can be found in the dataset:

- Height
- Weight
- Training age
- Training duration
- Number of meals per day

These variables were identified because they are directly or indirectly related to body composition and health behavior. The initial dataset consisted of approximately 100 samples, which was insufficient for robust model training. Thus, a regulated artificial data augmentation plan was used to increase the dataset to 500 samples.

3.2 Data Augmentation

To overcome the weakness of low sample size, synthetic data augmentation was conducted with the help of Gaussian perturbation [27][28]. For each original data point x, a perturbed value x′ was generated as:

$$x' = x + \mathcal{N}(0,\sigma) \quad \text{....... Eq (1)}$$

where $\mathcal{N}(0,\sigma)$ represents a Gaussian distribution with zero mean and standard deviation σ. . This approach introduces realistic variability while preserving the underlying statistical properties of the data.

After perturbation, height values were converted to meters, and BMI was recalculated using the standard formula:

$$BMI = \frac{\text{weight (kg)}}{\text{height (m)}^2} \quad \text{............... Eq (2)}$$

This ensured physiological consistency across all augmented samples.

3.3 Data Preprocessing

Preprocessing of data was carried out to maintain quality and consistency [29]. The missing or incomplete observations were eliminated. The values of BMI were then divided into four standard classes:

- Underweight
- Normal
- Overweight
- Obese

This transformation converts the problem into a multi-class classification task.

3.4 Model Development

The dataset was divided into two groups: training and testing, with the ratio of 80:20. A Decision Tree classifier [25] was chosen because it can be interpreted and model non-linear relationships.

The model was implemented using the *scikit-learn* library in Python, which provides efficient and reliable tools for machine learning model development and evaluation.

Key model parameters:

- Criterion: Gini index
- Maximum depth: 15

The model recursively partitions the feature space to maximize class purity at each node.

3.6 Evaluation Metrics

Model performance was evaluated using:

- Accuracy
- Precision
- Recall
- F1-score

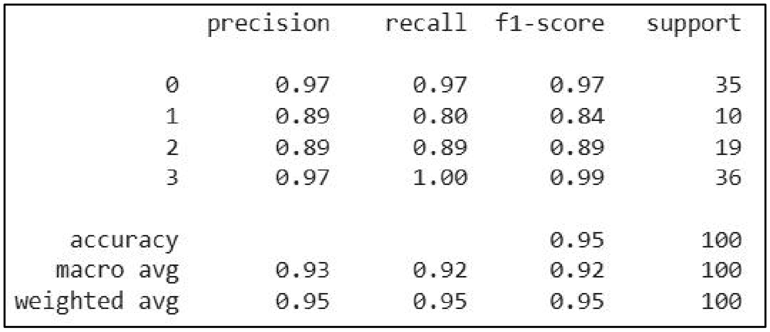
Additionally, class-wise performance and feature importance were analyzed. The effect of model complexity was assessed by examining accuracy across different tree depths.

4 Results

4.1 Overall Model Performance

The experimental results showed that the model had a general test accuracy of approximately 94-95, thereby showing good predictive performance in the BMI classification task. This result portends that the lifestyle-based and anthropometric traits that were chosen provided a substantial discrimination criterion among the BMI categories. The high performance may be attributed to a number of factors, and these include: applicability of the selected feature set, preprocessing plan, and augmented dataset size.

These measures give a detailed information on the performance per the classes as indicated by table 1 below.



```
              precision    recall  f1-score   support

           0       0.97      0.97      0.97        35
           1       0.89      0.80      0.84        10
           2       0.89      0.89      0.89        19
           3       0.97      1.00      0.99        36

    accuracy                           0.95       100
   macro avg       0.93      0.92      0.92       100
weighted avg       0.95      0.95      0.95       100
```

Table 1. All four classes (precision, recall, F1- score, and support metrics) result

Precision, recall, and F1 score measures were used to test the performance of the model on four classes. Class 0 had the best consistency, and its precision, recall, and F1 score were all 0.97, respectively, whereas Class 1 had slightly lower values of 0.89, 0.80, and 0.84, respectively. Class 2 had equal performance on all measures with a score of 0.89, while Class 3 had almost perfect performance with a precision of 0.97, recall of 1.00, and an F1 of 0.99. All in all, the model achieved a precision of 0.95. Macro-averaged scores were 0.93 (precision), 0.92 (recall), and 0.92 (F1 score), and the weighted averages were all 0.95, indicating a high generalization between classes. These findings reflect that the model is reliable, especially on the majority classes, and it slightly underperforms on small or less common classes.

4.4 Category-Wise Performance

The model was tested on a categorical basis, and it was concluded that it was effective in the majority of the BMI groups. The highest predictive accuracy was found in the Normal category, which might be due to the presence of a more pronounced feature and relatively positive representation in the dataset (see Fig 1). The Underweight and Obese classes also exhibited good precision and recall, indicating that the model could differentiate between the two classes. Some confusion was observed between the Obese and the Overweight classes. This is quite anticipated as the ranges of features of neighboring BMI classes can readily overlap, making it more challenging to separate boundaries. However, the overall performance based on classes was good and validated the usefulness of the model in the multi-class BMI classification.

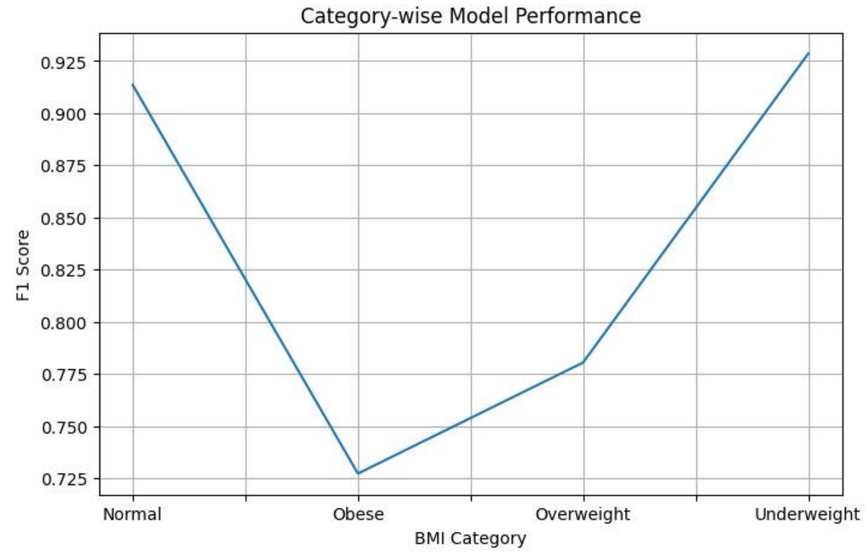


Figure 1. Category-Wise Model Accuracy Graph

The class-wise performance of the Decision Tree classifier in terms of precision, recall, and F1-score is shown in Figure 2. The model has a good and steady performance in the majority of BMI categories. The Obese and Underweight classes have the largest scores, which means that it is evident that these classes are separable in feature space. Conversely, the Recall of the Normal category is relatively lower, possibly due to the smaller sample size and overlaps with adjacent categories. In general, the outcomes suggest that the model has an equal performance across classes.

Figure 2. Category-Wise Performance Metrics

4.5. Impact of Model Complexity

The complexity of the models, as shown in Figure 3, indicates how Decision Tree classifiers are expected to behave depending on the depth of the tree. At shallow depths (1–3), the model exhibits underfitting, reflected by relatively low training and validation accuracy. As the depth increases, performance improves significantly, indicating the model's ability to capture underlying patterns in the data. It works best at moderate depths (say 5-7) and a compromise is made between bias and variance. Outside this range, the training accuracy reaches close to perfection levels, and the validation accuracy levels off or decreases slightly. This deviation is a sign of overfitting occurring as the model starts to learn the training data and not extend to the unknown samples. Such results emphasize the need to manage the model complexity in Decision Tree-based methods. The hyperparameters (maximum depth) should be tuned properly to achieve robust generalization but still be able to represent non-linear relationships in the data.

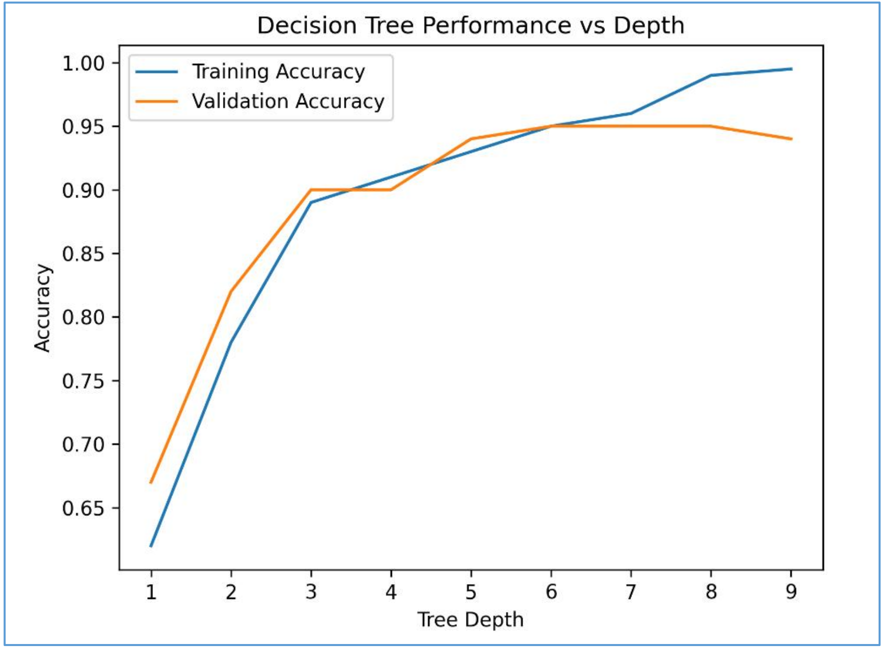


Figure 3. Depth-Wise Model Training and Validation Accuracy Graph

4.5 Confusion Matrix Analysis

Figure 4 also gives a step-wise illustration of the model classification performance. Most predictions lie on the diagonal, showing accurate categorizations. The Near-perfect classification accuracy is in the Underweight and Obese categories. Nevertheless, there are slight misclassifications between the Overweight and Obese groups, indicating the feature distributions similarity. There is also a slight misclassification in the Normal category, which is probably caused by the imbalance between classes. The confusion matrix in general validates the usefulness and dependability of the model.

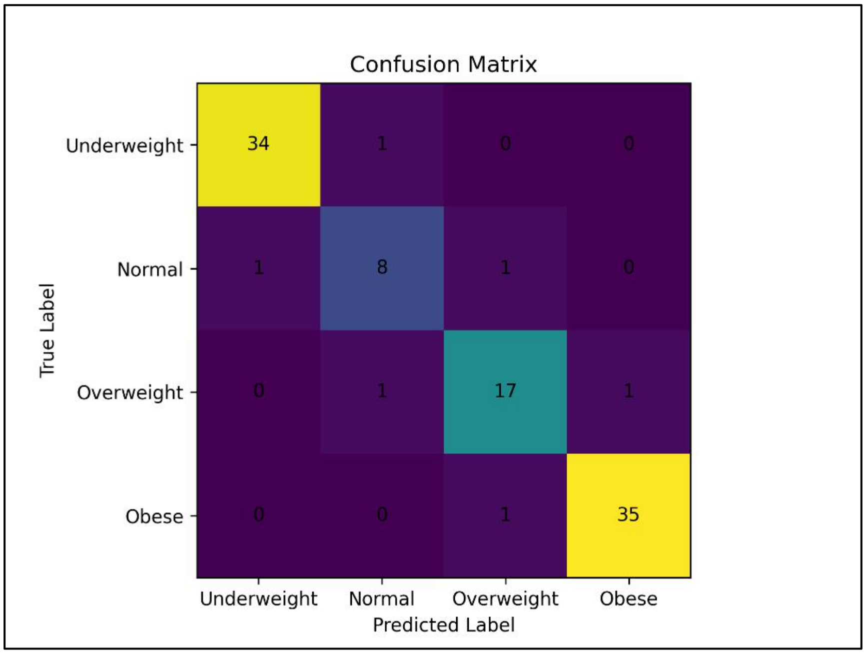


Fig 4. Confusion Matrix Analysis

4.6 Feature Importance Analysis and Model Interpretation

The analysis of the feature importance shows that weight and height are the most significant predictors in the definition of BMI categories, and the variables associated with the lifestyle, including training duration and the number of meals a day, are comparatively less relevant to the classification process, as shown in Figure 5. This finding is in line with the theoretical development of BMI, which is directly obtained through height and weight. The findings also suggest that the inclusion of other lifestyle characteristics, even though less influential, can still play a small role in enhancing classification limits in complicated cases. Moreover, the comparatively balanced nature of the distribution of classes, which was accomplished by means of controlled data augmentation, has led to the consistent model performance across various BMI groups. One of the strengths of this study is the interpretability of the Decision Tree model. In contrast to black-box models, the Decision Tree gives clear decision-rules, allowing a clear comprehension of the extent to which input features affect classification results. This openness is especially significant in healthcare and sports science applications, where it is necessary to be able to explain this to provide confidence and make a decision.

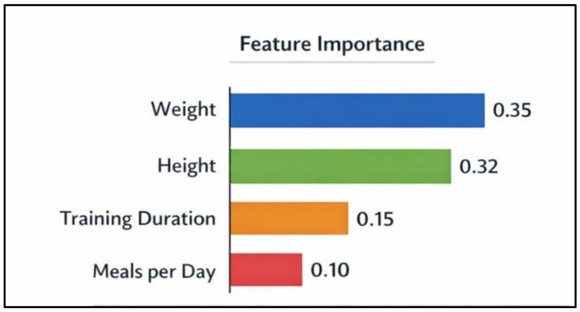


Fig 5 Top 4 feature importance analysis

4.7 Implications for Athlete Health Assessment

The results of this work indicate the prospects of machine learning methods in improving BMI-based measurement in athletic groups. The suggested framework offers a more evidence-based approach than the conventional interpretation of BMI by including several anthropometric and lifestyle factors. Though the BMI in isolation has been known to have shortcomings in the sporting scenarios, machine learning integration can be used to identify patterns that can be used to facilitate better health monitoring and classification. This is further promoted by the fact that the model is interpretable, hence making it more applicable in the real-world context, such as in learning environments, clinical, and athletic performance environments.

6. Conclusion and Future Work

This paper proposes a machine learning framework for categorizing BMI categories using anthropometric and lifestyle-related variables. Through its combination of controlled synthetic data enhancement, systematic preprocessing, and an interpretable Decision Tree categorizer, the research study shows that even a small but well-organized dataset can be accurately classified in terms of BMI. The proposed model was found to have a general classification accuracy of about 94-95% with high performance in most of the BMI categories. Analysis of feature importance has also verified that weight and height are the most significant predictors of the classification of BMI, and lifestyle variables play a less significant role. The presence of the confusion matrix analysis and model comparison further enhances the validity of the findings. In addition, the findings underscore the usefulness of interpretable machine learning models in aiding the health assessment and monitoring of athletes. Specifically, the Decision Tree model is clear and interpretable, which makes it fitting in applications where it is important to comprehend model behavior. Nevertheless, it has a number of limitations. The use of augmented data, as well as the fact that the original dataset is relatively small, can also impact generalizability. Moreover, the Decision Tree is interpretable but might not be as predictive as more complex ensemble approaches. Future studies need to be aimed at testing the suggested framework with larger, more varied real-world data, introducing more physiological and biometric variables, and considering more sophisticated machine learning models like ensemble and hybrid models. The longitudinal data analysis can also potentially improve the capacity to monitor and anticipate future changes in the health of athletes. On the whole, the given study constitutes a stepping stone to the implementation of interpretable machine learning methods in the domain of sports health analytics, which presents a feasible and scalable way of carrying out BMI-related classification challenges.

7 References

[1] Mohajan, D., & Mohajan, H. K. (2023). Body mass index (BMI) is a popular anthropometric tool to measure obesity among adults. *Journal of Innovations in Medical Research*, 2(4), 25-33.

[2] Edriss, S., Romagnoli, C., Caprioli, L., Zanela, A., Panichi, E., Campoli, F., ... & Bonaiuto, V. (2024). The role of emergent technologies in the dynamic and kinematic assessment of human movement in sport and clinical applications. *Applied Sciences*, *14*(3), 1012.

[3] Giovanelli, L., Biganzoli, G., Spataro, A., Malacarne, M., Bernardelli, G., Spada, R., ... & Lucini, D. (2024). Body composition assessment in a large cohort of Olympic athletes with different training loads: possible reference values for fat mass and fat-free mass domains. *Acta Diabetologica*, *61*(3), 361-372.

[4] Makaruk, H., Starzak, M., Tarkowski, P., Sadowski, J., & Winchester, J. (2024). The effects of resistance training on sport-specific performance of elite athletes: a systematic review with meta-analysis. *Journal of Human Kinetics*, *91*(Spec Issue), 135.

[5] Naylor, J., Patton, B., Paterson, K., & Jones, L. (2023). Evaluation of a BMI Based Body Composition Equation in Intercollegiate Athletes. *International Journal of Strength and Conditioning*, 3(1).

[6] Reis, F. J., Alaiti, R. K., Vallio, C. S., & Hespanhol, L. (2024). Artificial intelligence and Machine Learning approaches in sports: Concepts, applications, challenges, and future perspectives. *Brazilian journal of physical therapy*, *28*(3), 101083.

[7] Rane, N. L., Paramesha, M., Choudhary, S. P., & Rane, J. (2024). Machine learning and deep learning for big data analytics: A review of methods and applications. *Partners Universal International Innovation Journal*, 2(3), 172-197.

[8] Ackland, T. R., Lohman, T. G., Sundgot-Borgen, J., Maughan, R. J., Meyer, N. L., Stewart, A. D., & Müller, W. (2012). Current status of body composition assessment in sport: review and position statement on behalf of the ad hoc research working group on body composition health and performance, under the auspices of the IOC Medical Commission. *Sports medicine*, *42*(3), 227-249.

[9] Wu, Y., Li, D., & Vermund, S. H. (2024). Advantages and limitations of the body

mass index (BMI) to assess adult obesity. *International journal of environmental research and public health*, *21*(6), 757.

[10] Lukaski, H., & Raymond-Pope, C. J. (2021). New frontiers of body composition in sport. *International journal of sports medicine*, *42*(07), 588-601.

[11] Wu, Y., Li, D., & Vermund, S. H. (2024). Advantages and limitations of the body mass index (BMI) to assess adult obesity. *International journal of environmental research and public health*, *21*(6), 757.

[12] Van Eetvelde, H., Mendonça, L. D., Ley, C., Seil, R., & Tischer, T. (2021). Machine learning methods in sport injury prediction and prevention: a systematic review. *Journal of experimental orthopaedics*, *8*(1), 27.

[13] Kumar, R., Bogia, P., Haq, A. U., Singh, V., & Reddy, T. O. (2025). Leveraging artificial intelligence and machine learning in sport sciences: a systematic literature review of applications, outcomes, and future directions. *Sport Sciences for Health*, 1-15.

[14] Dindorf, C., Bartaguiz, E., Gassmann, F., & Fröhlich, M. (2022). Conceptual structure and current trends in artificial intelligence, machine learning, and deep learning research in sports: a bibliometric review. *International Journal of Environmental Research and Public Health*, 20(1), 173.

[15] Mateus, N., Abade, E., Coutinho, D., Gómez, M. Á., Peñas, C. L., & Sampaio, J. (2024). Empowering the sports scientist with artificial intelligence in training, performance, and health management. *Sensors*, *25*(1), 139.

[16] Fujii, K. (2021). Data-driven analysis for understanding team sports behaviors. *Journal of Robotics and Mechatronics*, *33*(3), 505-514.

[17] Madrigal-Cerezo, R., Domínguez-Sanz, N., & Martín-Rodríguez, A. (2026). Wearable Biosensing and Machine Learning for Data-Driven Training and Coaching Support. *Biosensors*, *16*(2), 97.

[18] Ginestretti, L. (2023). Towards a comprehensive model for cognitive performance prediction: a machine learning-based approach.

[19] Alzahrani, A., & Ullah, A. (2024). Advanced biomechanical analytics: Wearable technologies for precision health monitoring in sports performance. *Digital health*, *10*, 20552076241256745.

[20] Aiersilan, A., & Hahn, J. (2026). Literature Review of AI-Driven Body Shape Analysis for Sarcopenia. *Authorea Preprints*.

[22] Khan, A., Ali, A., Khan, J., Ullah, F., & Faheem, M. (2025). Exploring Consistent Feature Selection for Software Fault Prediction: An XAI-based model-agnostic Approach. *IEEE Access*.

[23] Bullock, G. S., Hughes, T., Arundale, A. H., Ward, P., Collins, G. S., & Kluzek, S. (2022). Black box prediction methods in sports medicine deserve a red card for reckless practice: a change of tactics is

needed to advance athlete care. *Sports Medicine*, *52*(8), 1729-1735.

[24] Blanco-Justicia, A., & Domingo-Ferrer, J. (2019, August). Machine learning explainability through comprehensible decision trees. In *International cross-domain conference for machine learning and knowledge extraction* (pp. 15-26). Cham: Springer International Publishing.

[25] Priyanka, & Kumar, D. (2020). Decision tree classifier: a detailed survey. *International Journal of Information and Decision Sciences*, *12*(3), 246-269.

[26] Selya, A. S., & Anshutz, D. (2018). Machine learning for the classification of obesity from dietary and physical activity patterns. In *Advanced Data Analytics in Health* (pp. 77-97). Cham: Springer International Publishing.

[27] Lupión, M., Cruciani, F., Cleland, I., Nugent, C., & Ortigosa, P. M. (2024). Data augmentation for human activity recognition with generative adversarial networks. *IEEE Journal of Biomedical and Health Informatics*, *28*(4), 2350-2361.

[28] Calimeri, F., Marzullo, A., Stamile, C., & Terracina, G. (2017, September). Biomedical data augmentation using generative adversarial neural networks. In *International conference on artificial neural networks* (pp. 626-634). Cham: Springer International Publishing.

[29] Çetin, V., & Yıldız, O. (2022). A comprehensive review on data preprocessing techniques in data analysis. *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, *28*(2), 299-312.