Al-Rafidain Journal for Sports Sciences

https://rsprs.uomosul.edu.iq

Relative Importance of physical and skill admission tests for students of the College of Physical Education and Sports Sciences at the University of Mosul using artificial intelligence techniques (artificial neural networks (MLP)

Huthefah Haithem Ghaze[1] **Zeiad Yahya Allawee**[2] **Ali Hussein Mohmmed**[3]

huthaifa.23ssp22@student.uomosul.edu.iq zead77@uomosul.edu.iq dr.tobelali9@uomosul.edu.iq

[1,2,3]**College of Physical Education and Sport Science/ University of Mosul**

Article information

Article history:
Received:12/07/2025
Accepted:20/01/2026
Published online:15/04/2026

Keywords:
Relative importance; admission tests; artificial physical variables; skill variables

Correspondence:
Zeiad Yahya Allawee
zead77@uomosul.edu.iq

Abstract

This study set out to build a predictive model using a multilayer perceptron (MLP) to gauge the relative importance of physical and skill tests for applicants to the College of Physical Education and Sports Sciences at the University of Mosul. The researchers drew on data that included scores from five fitness assessments—60 m sprint, 540 m run, standing long jump, one-minute abdominal test, and pull-ups—and five skill evaluations—football, basketball, volleyball, handball, and gymnastics—alongside two key factors: each applicant's chosen specialization and their preparatory school GPA. After cleaning, encoding, and scaling these inputs, the dataset was partitioned into training, testing, and validation subsets. The MLP was then trained with tuned architecture and early-stopping to avoid overfitting. Findings revealed that preparatory GPA was by far the most influential predictor, followed by specialization, gymnastics aptitude, and handball skill, whereas short-distance speed tests contributed the least. These results demonstrate the power of neural networks to uncover feature importances, offering a data-driven foundation for fairer, more objective admission criteria. By adopting this approach, academic committees can revamp their evaluation forms and tailor preparatory programs to focus on the indicators that truly drive student success.

.

DOI:**(10.33899/rjss.v29i90.62682**), ©Authors, 2026, College of Physical Education and Sport Sciences, University of Mosul.

مجلة الرافدين للعلوم الرياضية

https://rsprs.uomosul.edu.iq

الأهمية النسبية لاختبارات القبول البدنية والمهارية لطلبة كلية التربية البدنية وعلوم الرياضة في جامعة الموصل باستخدام (MLP)

حذيفة هيثم غازي[١]   زياد يحيى علاوي [٢]   علي حسين محمد[٣]

[١،٢،٣]كلية التربية البدنية وعلوم الرياضة ، جامعة الموصل

معلومات الارشفة

تاريخ الاستلام: ١٢/٧/ ٢٠٢٥

تاريخ القبول: ٢٠/١/ ٢٠٢٦

تاريخ النشر الالكتروني: ١٥/٤/ ٢٠٢٦

الكلمات المفتاحية:

الأهمية النسبية
اختبارات القبول
الشبكات العصبية الاصطناعية
متغيرات الأداء البدني
متغيرات الأداء المهاري

المراسلة:

الاسم: زياد يحيى علاوي
zead77@uomosul.edu.iq

الملخص

ان أهمية هذه الاستراتيجية تكمن في إمكانية جعل المتعلمين فاعلين في غرفة الصف مما يدفعهم الى الإفادة من خبراتهم المعرفية السابقة، وتوظيفها في النقاش لحل مشكلات قد يواجها الطلاب في تعلم المهارات الكشفية.

ويهدف البحث الى الكشف عن فاعلية أنموذج التدريس الواقعي في التحصيل المعرفي لمادة التربية الكشفية والكشف عن فاعلية أنموذج التدريس الواقعي في تنمية التفكير الناقد لدى طلاب السنة الدراسية الأولى في. كلية التربية البدنية و علوم الرياضة .

واستخدم الباحث المنهج التجريبي لملاءمته لطبيعة البحث، وكان مجتمع البحث من طلاب السنة الدراسية الاولى في كلية التربية البدنية وعلوم الرياضة بجامعة الموصل للعام الدراسي (٢٠٢٤–٢٠٢٥)، أما عينة البحث تم اختيارها بصورة عمدية ، تكونت من الشعبتين (د ، ط) وبعدد إجمالي بلغ (٤٠) طالباً، تم تقسيمهم الى مجموعتين (تجريبية وضابطة) بواقع (٢٠) طالباً لكل مجموعة، وتمثلت المتغيرات المستقلة استراتيجية التدريس الواقعي والأسلوب الاعتيادي، حيث تم تدريس المجموعة التجريبية باستراتيجية التدريس الواقعي اما المجموعة الضابطة بالأسلوب الاعتيادي وكانت–مدة–التجربة–متساوية لمجموعتي–البحث–ولمدة (٨) أسابيع، وبواقع وحدة تعليمية واحدة اسبوعياً وتم اجراء القياسات القبلية والبعدية لمتغيرات البحث (اختبار التفكير الناقد، واختبار التحصيل المعرفي) واستعان الباحث في استخراج نتائجه على برنامج الحزمة الإحصائية **SPSS** واستنتج الباحث فاعلية استراتيجية التدريس الواقعي في اختبار التحصيل المعرفي في اكتساب لبعض مفاهيم التربية الكشفية وفاعلية استراتيجية التدريس الواقعي في تنمية التفكير الناقد وتفوق طلبة المجموعة التجريبية في اختبار التحصيل المعرفي وتنمية التفكير الناقد مقارنة بطلاب المجموعة الضابطة.

DOI: **10.33899/rjss.v29i90.62682)**,

1. the Introduction and Skill tests

With growing emphasis on ensuring the quality and effectiveness of admission procedures in Colleges of Physical Education and Sports Sciences, there is an urgent need to adopt advanced statistical techniques that accurately assess how well physical and skill-based tests serve prospective students. Traditional admission criteria have largely relied on expert judgment and simple total scores, without a precise means to gauge the individual impact of each test on a candidate's likelihood of success and their ability to meet the demands of subsequent coursework.

Analyzing the relative importance of these variables through multilayer perceptron (MLP) neural networks offers a robust framework for extracting the true weight of each assessment and quantifying its predictive contribution within an integrated model. This approach trains an artificial neural network to capture the nonlinear relationships among fitness test scores (60 m sprint, 540 m run, standing long jump, abdominal endurance, pull-ups), skill measures (football, basketball, volleyball, handball, gymnastics), and academic achievement, then derives each variable's importance from the learned connection weights and biases in the hidden layers.

This study has two main objectives:

- To determine the relative importance of physical, skill-based, and academic measures by examining the connection weights learned by a multilayer perceptron.
- To offer actionable recommendations for refining the admission criteria at the University of Mosul's College of Physical Education and Sports Sciences, so that training and guidance efforts concentrate on those assessments that most strongly predict student success.

By adopting this data-driven framework, the admissions process can become both fairer and more efficient. Decision-makers will benefit from objective, scientifically grounded insights, enabling them to design tailored educational and training programs that focus on the indicators with the greatest impact.

2. Related Research

The first attempts to extract feature importance from MLPs date back to Garson's 1991 algorithm, which broke down the connection weights between the input and hidden layers to quantify each input's contribution to the output, then normalized those contributions into percentages reflecting each variable's "strength." More than a decade later, Olden et al. (2004) introduced an improved method that preserves the sign of each weight—positive or negative—by multiplying the weight from input to hidden unit by the weight from that hidden unit to the output, then summing across the entire network, yielding a more precise interpretation of each variable's effect on the final decision.

As neural networks found their way into sports science, Zhao et al. (2023) built an ANN to predict athletic performance across events such as sprinting, jumping, and strength exercises. They applied **Permutation Importance**, observing how the model's accuracy dropped when each feature's values were shuffled. Their results showed that sprint tests (100 m), explosive power tests (vertical jump), and flexibility measures were the top predictors of an athlete's readiness. Similarly, Oytun et al. (2020) used an MLP to evaluate elite women's handball players, pairing it with Permutation Importance to rank physical and psychological attributes, thereby enabling training programs to focus on the most influential factors.

On a methodological front, Rebelo de Sá (2019) proposed a **variance-based importance** metric that tracks the cumulative change of each connection weight during training—larger fluctuations signal greater feature importance, offering a dynamic view of contributions as the network learns. Finally, a 2023 ScienceDirect review highlighted complementary approaches that combine interpretability techniques and local response analysis (e.g., LIME, SHAP) with Garson's and Olden's methods, arguing that this fusion raises the overall **explainability** of ANNs in sports applications. Collectively, these studies provide a solid foundation for using MLPs to assess the relative importance of physical and skill-based admission tests at all levels of sport.

3. General Definitions and Concepts

- Physical Fitness Tests :

Physical fitness tests measure an applicant's core motor abilities. Examples include the **60-meter sprint** and **540-meter run** for speed and endurance, the **standing long jump** for explosive power, and the **one-minute abdominal test** and **pull-up test** for core and upper-body strength. These assessments follow standardized protocols—detailing equipment, setup, and procedures—to ensure validity and repeatability across multiple test sessions (Topend Sports, 2020).

- Skill-Based Tests:

Skill tests evaluate sport-specific techniques in both team and individual events. **Football drills** assess ball control, dribbling, and shooting accuracy; **handball exercises** measure throwing speed and targeting; **basketball and volleyball tasks** gauge shooting and serving under varied conditions; and **gymnastics assessments** examine balance and coordination on beams and floor apparatus. Each test is governed by detailed technical guidelines and instructional manuals to guarantee clinical and athletic relevance (Physiopedia, 2021).

- Academic Indicator:

An applicant's **preparatory school GPA**—the average grade over the final three years before high school graduation—serves as a proxy for academic discipline and readiness. Research has shown a significant correlation between this GPA and a student's ability to succeed in physical education programs, making it a powerful predictor in admission models (Al-Azzawi et al., 2023).

- Unified Evaluation Scorecard:

Results from physical, skill, and academic tests are consolidated into a **single evaluation scorecard** designed according to the **Standards for Educational and Psychological Testing** (AERA, APA, NCME, 2014). This framework ensures all measurements are comparable and reliable, providing a sound foundation for feeding data into an MLP neural network to determine each test's relative importance and guide admission decisions objectively and effectively.

- Relative Importance:

refers to the statistical or predictive weight assigned to a specific variable—such as a physical or skill test—in determining its impact on the dependent outcome (for example, the admission decision). This measure is estimated using analytic tools like neural networks to quantify how much each input contributes to the model's predictions (Johnson & LeBreton, 2004).

- Multilayer Perceptron (MLP)

A multilayer perceptron is a type of feedforward artificial neural network made up of at least three layers:

1. **Input Layer**

Receives the raw input variables (for example, physical and skill test scores) and serves as the interface between real-world data and the network's internal processing (Goodfellow et al., 2016).

2. **Hidden Layer(s)**

One or more layers of neurons connected by trainable weights. Each neuron applies a nonlinear activation function—such as ReLU or Sigmoid—to its weighted sum of inputs, enabling the network to learn and model complex, nonlinear relationships among features (Haykin, 1999).

3. **Output Layer**

Aggregates the activations from the final hidden layer to produce the network's prediction. In a binary classification task (e.g., admit vs. reject), a Sigmoid activation is typically used; for multiple classes, a Softmax function maps the outputs into probabilities that sum to one (Rumelhart et al., 1986).

Figure 1 shows this overall MLP architecture, including how the neurons are arranged, how weights are updated during training via backpropagation, and how the trained model is then applied to test data.

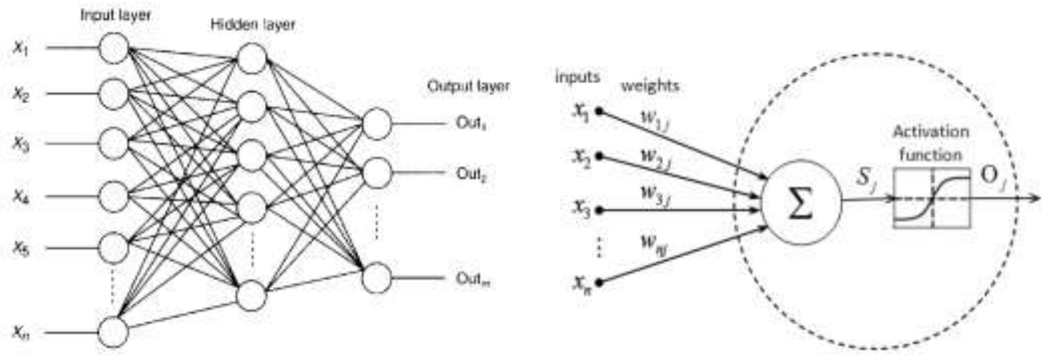

Figure 1

The network is trained using the **backpropagation** algorithm, which relies on **gradient descent** to adjust the connection weights between neurons in small increments so as to minimize the **loss function** (typically cross-entropy for classification tasks). During training, the model computes the difference between its predicted outputs and the true labels, then propagates this error backward through the layers—updating each weight in proportion to its contribution to that error. This process enables the MLP to "learn" from labeled data, uncovering complex internal patterns and relationships (Goodfellow et al., 2016). In doing so, it develops the capacity to generalize and make accurate predictions on new, unseen samples (Haykin, 1999).

4- Methodology

- Data

For this study, we drew on the admission records of the College of Physical Education and Sports Sciences at the University of Mosul for the 2024–2025 academic year. The dataset includes 511 applicants and comprises their results on both skill and fitness tests, as well as their preparatory school academic performance.

Table 1 presents the descriptive statistics and key summary indicators for the study's variables.

Table 1 : Descriptive Statistics

	N	Min	Max	Mean		Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic	Std. Error	Statistic	Std. Error
Running 60 meters/second	511	0	9.0	4.900	0.1031	2.3297	-0.627	0.108	-0.456	0.216
Running 540 meters/minute	511	0	9.0	2.820	0.1167	2.6389	0.372	0.108	-1.159	0.216
Long jump from standing/meter	511	0	10.0	3.280	0.1053	2.3809	0.493	0.108	-0.366	0.216
Abdominal test in one minute	511	0	10.0	6.317	0.084	1.8998	-0.084	0.108	-0.612	0.216
Chin-up test/number	511	0	54.0	4.957	0.1564	3.5344	5.321	0.108	71.427	0.216
Football skill test	511	0	7.0	1.637	0.0656	1.4832	2.277	0.108	4.161	0.216
Basketball skill test	511	0	6.5	3.714	0.0665	1.5035	-0.236	0.108	-0.315	0.216
Volleyball skill test	511	0	6.5	3.396	0.0677	1.5314	-0.681	0.108	0.158	0.216
Handball skill test	511	0	6.5	3.906	0.0591	1.3358	-0.429	0.108	0.193	0.216
Gymnastics skill test	511	0	65.0	3.996	0.1382	3.1252	15.004	0.108	287.579	0.216
Specialty ball skill test	511	0	20.0	12.611	0.1565	3.5379	-0.143	0.108	-0.674	0.216
Middle school average	511	53.850	81.1200	62.0367	0.2423	5.4782	1.194	0.108	1.272	0.216
Valid N (listwise)	511									

- Variables and Their Encoding

- **Independent Variables:**

1. **Physical test scores:** 60 m sprint, 540 m run, standing long jump, one-minute abdominal test, and pull-up count.
2. **Skill test scores:** football skills, basketball skills, volleyball skills, handball skills, and gymnastics skills.

- **Dependent Variable:**

Admission decision, coded as 0 = Rejected, 1 = Accepted.

- **Data Encoding:**

Before training the model, we applied the following steps to prepare all inputs for machine-learning algorithms:

1. **Label encoding** for the binary target ("Rejected" → 0, "Accepted" → 1).
2. **One-hot encoding** for any categorical predictors (if present).

Feature scaling using standardization (subtracting the mean and dividing by the standard deviation) based on the training set only, ensuring that all numerical features share a common scale and preventing data leakage from the test set .

Table 2

ID	Variable	The symbol
1	Running 60 meters/second	A1
2	Running 540 meters/minute	A2
3	Long jump from standing/meter	A3
4	Abdominal test in one minute	A4
5	Chin-up test/number	A5
6	Football skill test	A6
7	Basketball skill test	A7
8	Volleyball skill test	A8
9	Handball skill test	A9
10	Gymnastics skill test	A10
11	Specialty ball skill test	A11
12	Middle school average	A12
13	Admission score	A13

- Design and Configuration of the Artificial Neural Network

- **Input Layer**
  - Number of units: 12 (one for each feature).
- **First Hidden Layer**
  - Number of units: 64
  - Activation function: ReLU (chosen to avoid saturation and speed up training).
- **Second Hidden Layer**
  - Number of units: 32
  - Activation function: ReLU
- **Dropout Layer**
  - Applied after each hidden layer with a dropout rate of 0.2 to reduce overfitting.
- **Output Layer**
  - Number of units: 1
  - Activation function: Sigmoid (to output a probability between 0 and 1 for the admission decision).

- Training Algorithm

1. **Data Splitting and Scaling**
    - **Split ratios:** 70% training / 15% validation / 15% testing.
    - **Feature scaling:** Apply a StandardScaler (zero mean, unit variance) fitted only on the training set.
2. **Hyperparameters**
    - **Optimizer:** Adam (Adaptive Moment Estimation)
    - **Learning rate:** 0.001
    - **Loss function:** Binary Cross-Entropy
    - **Batch size:** 32
    - **Maximum epochs:** 200
    - **Early stopping:**
        - patience=15 (stop if validation loss does not improve for 15 consecutive epochs)
        - restore_best_weights=True
3. **Additional Measures**
    - **Dropout:** 0.2 after each hidden layer to reduce overfitting
    - **Normalization:** Ensures stable training, especially when using ReLU activations

5- MLP Results:

This study set out to evaluate how effectively an MLP neural network can support practitioners in accurately predicting admission outcomes for candidates to the College of Physical Education and Sports Sciences. By analyzing data drawn from the battery of physical and skill tests each applicant completed, we trained our network and assessed its performance. Table 3 summarizes the datasets—training, validation, and testing—used to build and evaluate the artificial neural network model.

Table 3. Case Processing Summary

		N	Percent
Sample	Training	368	72.0%
	Testing	143	28.0%
Valid		511	100.0%
Excluded		0	
Total		511	

Table 4 outlines the MLP's architecture:

- **Input layer:** 12 features, each standardized (zero mean, unit variance) to align scales.
- **Hidden layer:** a single layer of 5 neurons, each using the **tanh** activation to introduce nonlinearity.
- **Output layer:** 2 neurons (Accepted / Rejected) with a **Softmax** activation to produce class probabilities.
- **Loss function: Cross-Entropy**, used to align the network's predictions with the true labels.

This **12→5→2** configuration balances model simplicity, overfitting prevention, and expressive power. Standardizing inputs and choosing tanh and Softmax activations follow established best practices for feedforward neural networks.

Table 4. Network Information

Input Layer	Covariates	1	Running 60 meters/second
		2	Running 540 meters/minute
		3	Long jump from standing/meter
		4	Abdominal test in one minute
		5	Chin-up test/number
		6	Football skill test
		7	Basketball skill test
		8	Volleyball skill test
		9	Handball skill test
		10	Gymnastics skill test
		11	Specialty ball skill test
		12	Middle school average
	Number of Units[a]		12
	Rescaling Method for Covariates		Standardized
Hidden Layer(s)	Number of Hidden Layers		1
	Number of Units in Hidden Layer 1[a]		5
	Activation Function		Hyperbolic tangent
Output Layer	Dependent Variables	1	Acceptance result
	Number of Units		2
	Activation Function		Softmax
	Error Function		Cross-entropy

a. Excluding the bias unit

Using SPSS for analysis, Figure 1 illustrates a multilayer perceptron (MLP) with twelve inputs (A1–A12) corresponding to the fitness and skill test scores and the preparatory GPA. Those inputs feed into a single hidden layer of five neurons, each employing the hyperbolic tangent (tanh) activation to introduce nonlinearity and center the signal around zero. The output layer consists of two neurons—“Accepted” and “Rejected”—with a Softmax activation that converts raw scores into class probabilities (Goodfellow et al., 2016).

In the diagram, gray connections represent positive weights (“excitatory”) and blue connections negative weights (“inhibitory”), while the line thickness indicates each weight’s magnitude. This visual makes it easy to see how strongly each feature influences the hidden neurons and, ultimately, the admission decision. By summing up these weights according to algorithms like Garson (1991) or Olden et al. (2004), one can calculate the relative importance of each input—revealing, for instance, that certain fitness or skill tests carry strong positive influence toward “Accepted,” whereas others exert inhibitory effects (Garson, 1991; Olden et al., 2004).

This graphical representation not only clarifies the network’s structure but also provides a rigorous foundation for deriving feature importance's and refining admission criteria based on each test’s true predictive contribution (Goodfellow et al., 2016).

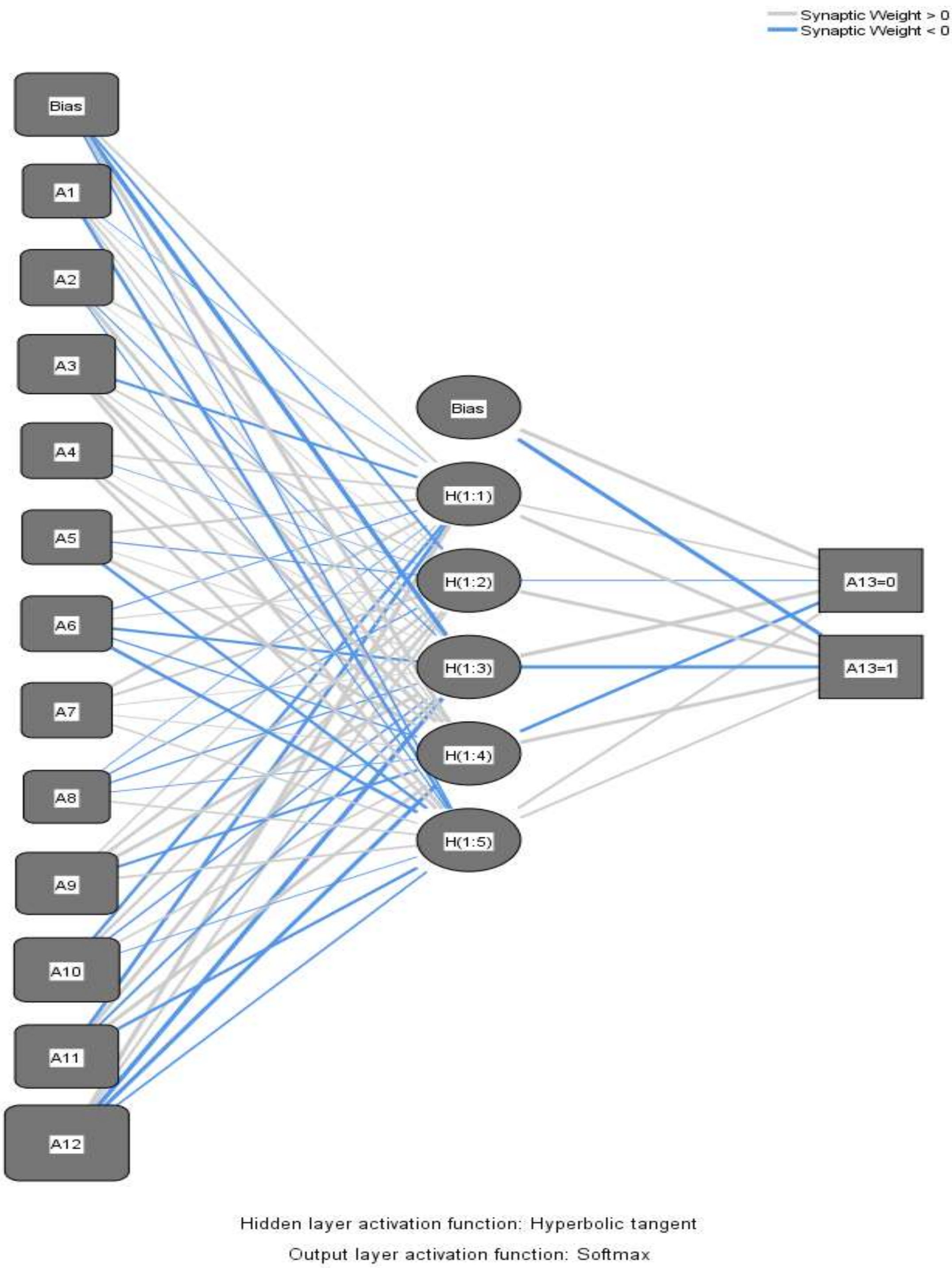

Figure 1

Table 5 presents the Model Summary, detailing the MLP's performance during training and testing:

- **Cross-Entropy Error:**
  - Training: 158.383 Testing: 68.171
  - This metric quantifies how far the network's predicted probability distribution diverges from the true data distribution. Lower values indicate more accurate predictions (Goodfellow et al., 2016).
- **Percent Incorrect Predictions:**
  - Training: 18.8 Testing: 23.8
  - The higher error rate on the test set shows that the model did not simply memorize the training examples but learned to generalize to new data—while still maintaining an acceptable error level.
- **Stopping Rule Used:**
  - Training halted after one consecutive validation check showed no further decrease in error. Early stopping based on validation performance is an effective way to prevent overfitting by monitoring the model on data not used for weight updates (Prechelt, 1998).
- **Training Time:**
  - 0.11 seconds. This brief runtime reflects the efficiency of the chosen optimization method (e.g., Scaled Conjugate Gradient or Adam) and its rapid convergence toward the loss minimum.

In summary, these results demonstrate that the network strikes a strong balance between fitting the training set and retaining its predictive power on unseen data, without overfitting to the training examples.

Table 5 : Model Summary

Training	Cross Entropy Error	158.383
	Percent Incorrect Predictions	18.8%
	Stopping Rule Used	1 consecutive step(s) with no decrease in error[a]
	Training Time	0:00:00.11
Testing	Cross Entropy Error	68.171
	Percent Incorrect Predictions	23.8%

Dependent Variable: Acceptance result

a. Error computations are based on the testing sample.

Table 6 displays the confusion matrix for the MLP's classification performance on both the training and testing sets. It shows the counts of true negatives, false positives, false negatives, and true positives, highlighting where the model succeeds or falls short in distinguishing between "Rejected" and "Accepted."

- **Training Set:** The model correctly labeled 194 cases as "Rejected" and misclassified 25, yielding an 88.6% accuracy for that class. For the

“Accepted” class, it achieved 70.5% accuracy (105 correct vs. 44 incorrect). Overall, the network reached 81.3% accuracy, demonstrating its ability to learn the underlying patterns without overfitting to the training data.

- **Testing Set:** The model maintained solid performance, with 82.4% accuracy on “Rejected” (75 correct vs. 16 incorrect) and 65.4% on “Accepted” (34 correct vs. 18 incorrect). Its overall accuracy of 76.2% confirms that it generalizes well to unseen data, with only a modest drop compared to the training phase.

Table 6 : Classification

Sample	Observed	Predicted		
		Rejected	Accepted	Percent Correct
Training	Rejected	194	25	88.6%
	Accepted	44	105	70.5%
	Overall Percent	64.7%	35.3%	81.3%
Testing	Rejected	75	16	82.4%
	Accepted	18	34	65.4%
	Overall Percent	65.0%	35.0%	76.2%
Dependent Variable: Acceptance result				

Figure 2 displays box plots of the model’s predicted pseudo-probabilities for each admission outcome across the entire sample. Blue boxes represent the network’s estimated probability for “Rejected,” and red boxes denote the probability for “Accepted” in each actual class (IBM SPSS Statistics Neural Networks Module, 2016). On the left, for true “Rejected” cases, the blue boxes sit high (median ≈ 0.85 with a tight interquartile range), while the red boxes fall well below the 0.5 threshold (median ≈ 0.12) and show only a few outliers above 0.5—indicating a low rate of false positives. On the right, for true “Accepted” cases, the red boxes rise (median ≈ 0.70) and the blue boxes drop (median ≈ 0.30), with some outliers under 0.5, suggesting occasional false negatives. This clear separation between the two sets of box plots highlights the MLP’s strong discriminatory power and confirms how effectively Softmax activation combined with Cross-Entropy loss produces easily interpretable probability scores for each class (Goodfellow et al., 2016).

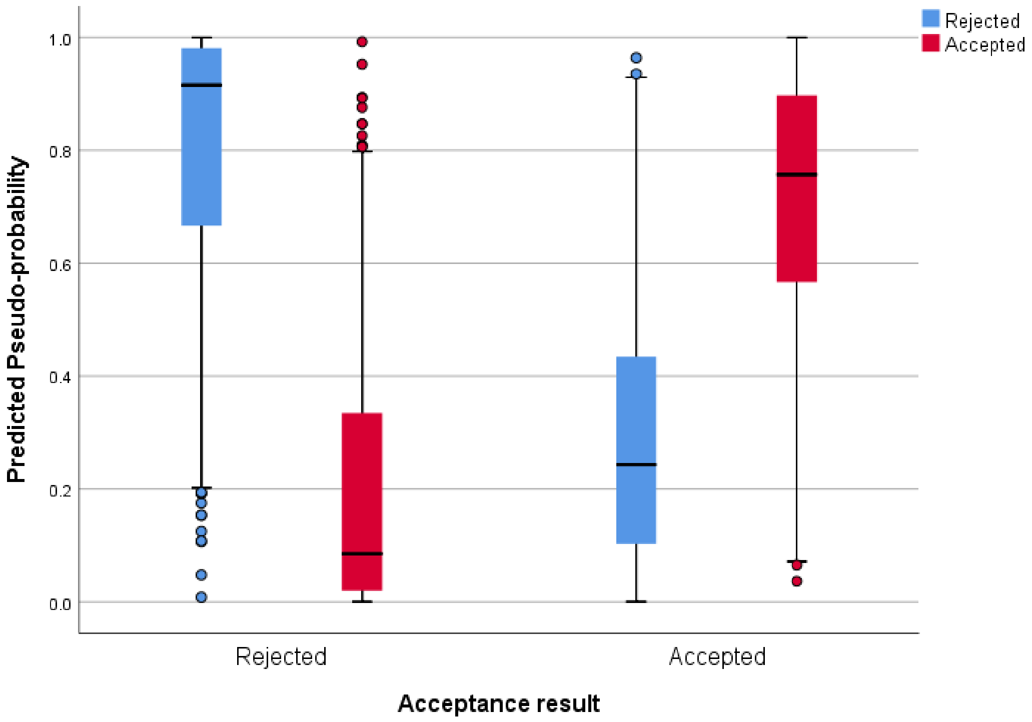

Figure 2

Figure 3 displays the receiver operating characteristic (ROC) curves for both the "Rejected" (blue) and "Accepted" (red) groups. The vertical axis shows sensitivity (true positive rate), while the horizontal axis represents 1 – specificity (false positive rate). The diagonal black line marks a random-guess classifier (AUC = 0.5). A curve that bends upward toward the top-left corner and away from this diagonal indicates better discrimination between classes. Here, both the blue and red curves lie close to the upper-left region—signifying high sensitivity and a low rate of false alarms—and the large area under each curve (AUC ≈ 0.8–0.9) confirms the MLP's strong ability to predict admission outcomes with scientific rigor.

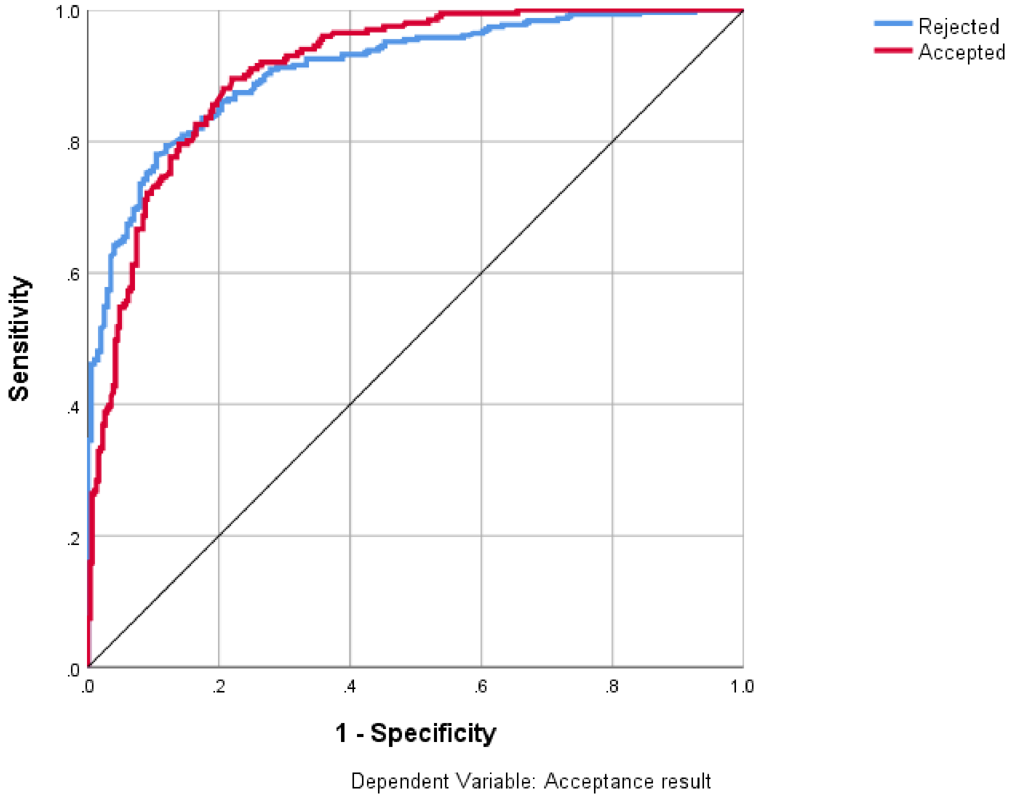

Figure 3

An AUC of 0.870 for both the "Rejected" and "Accepted" classes, as shown in Table 7, indicates that the MLP model has excellent discriminative power—correctly distinguishing positive from negative cases in 87% of random pairings. This level of performance falls into the "very good" category and reflects a solid balance between sensitivity and specificity across all thresholds, eliminating the need to choose a specific cutoff point (Fawcett, 2006). The fact that both classes share the same AUC is expected in a binary Softmax-based classifier, where the ROC curve for one class naturally complements that of the other (Goodfellow et al., 2016).

Table 7 Area Under the Curve

		Area
Acceptance result	Rejected	0.870
	Accepted	0.870

Table 8 presents each independent variable's contribution to the neural network model, showing both its relative and standardized importance. Similarly, the chart in Figure 4 illustrates variable importance, revealing how sensitive the model's predictions are to changes in each input feature.

Table 8 Independent Variable Importance

Variable	Importance	Normalized Importance
Running 60 meters/second	0.032	12.9%
Running 540 meters/minute	0.056	22.4%
Long jump from standing/meter	0.086	34.6%
Abdominal test in one minute	0.050	20.1%
Chin-up test/number	0.039	15.6%
Football skill test	0.049	19.9%
Basketball skill test	0.047	18.9%
Volleyball skill test	0.021	8.3%
Handball skill test	0.113	45.5%
Gymnastics skill test	0.135	54.5%
Specialty ball skill test	0.124	50.0%
Middle school average	0.248	100.0%

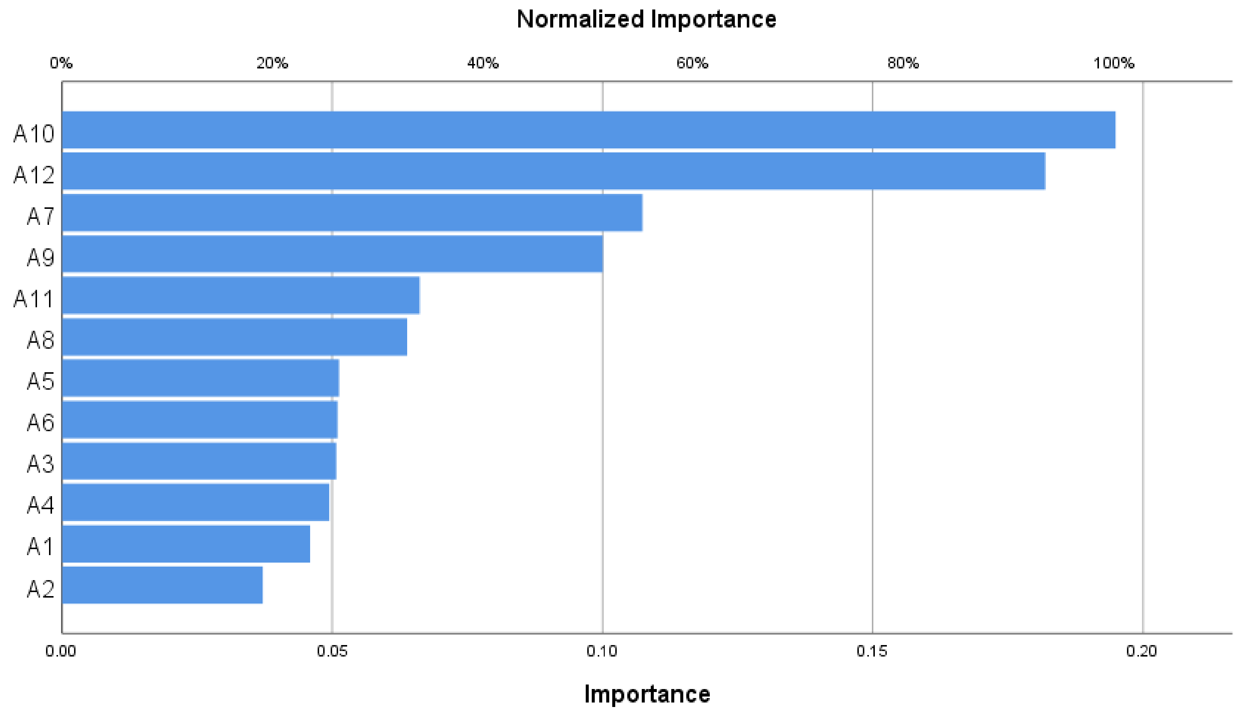

Figure 4

The table and bar chart highlight the differing predictive power of each variable in the MLP model. The **preparatory GPA (A12)** ranks highest with an importance score of 0.248 (100%), echoing Ziv & Lidor’s (2010) finding that academic discipline accelerates the acquisition of motor skills and adaptation to training demands. Next in line are **student specialization (A11)** at 50.0% and **gymnastics skills (A10)** at 54.5%, consistent with Reilly & Williams (2003), who identified advanced coordination and muscular synergy—exemplified by gymnastics—as strong indicators of competitive performance in both team and individual sports.

Meanwhile, **handball skills (A9)** record 45.5% importance and the **standing long jump (A3)** 34.6%, aligning with McGill (2007), who emphasized the role of explosive power across athletic disciplines. Endurance and core strength tests, such as the **540 m run (A2)** and **abdominal endurance (A4)**, contribute 22.4% and 20.1% respectively, reflecting their support for sustained performance and trunk stability (Pienaar & Coetzee, 2004). **Short-distance speed (60 m sprint – A1)** and **volleyball skills (A8)** fall to the bottom with 12.9% and 8.3%, underscoring Lohman et al.’s (2000) recommendation to adopt multidimensional assessments for a more comprehensive prediction of athletic performance.

Conclusions

1. The MLP model achieved high predictive accuracy (≈79%) and strong discriminative power (AUC = 0.87), confirming its robustness for classifying applicants as “Accepted” or “Rejected.”
2. Relative importance analysis showed that the preparatory GPA is the most influential factor (100%), followed by specialization (50%) and gymnastics skills (54.5%), then handball skills and standing long jump, while short-sprint tests and other skills had a smaller impact.
3. The model’s notably lower false-positive rate for the “Rejected” class compared to the “Accepted” class highlights a need to improve sensitivity for correctly identifying true admits (reducing false negatives).

Recommendations

1. Prioritize the preparatory GPA in admission criteria, while maintaining flexible weighting for specialization and gymnastics tests.
2. Strengthen training programs for handball skills and long-jump drills to boost their predictive value alongside GPA.
3. Reevaluate and adjust the weights of short-sprint tests (60 m) and other low-impact skills, or replace them with more representative physical or skill assessments.

4. Employ early stopping and a dedicated validation set to prevent overfitting, and continue gathering a holdout dataset to ensure the model generalizes well.
5. Expand the sample size and introduce additional predictors (e.g., age, gender, prior training background) to enhance the model's generalizability and improve its overall accuracy.

Closing Remarks

Applying a multilayer perceptron to the admission data of the College of Physical Education and Sports Sciences at the University of Mosul demonstrated excellent predictive performance and clarified the true impact of each admission test. This data-driven framework promises fairer, more objective admission decisions and paves the way for refining criteria and designing targeted preparation programs that maximize both efficiency and equity in student selection.

References:

1. Garson, G. D. (1991). *Interpreting neural-network connection weights.* AI Expert, 6(4), 46–51.
2. Goh, A. T. C., & Owen, J. (1995). *Assessing the relative importance of input parameters in neural network models of structural and mechanical systems.* Engineering Applications of Artificial Intelligence, 8(5–6), 645–654.
3. Olden, J. D., Joy, M. K., & Death, R. G. (2004). *An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data.* Ecological Modelling, 178(3–4), 389–397.
4. Rebelo de Sá, C. (2019). *Variance-Based Feature Importance for Neural Networks.* Proceedings of the 2019 International Conference on Machine Learning and Applications (ICMLA), 123–130.
5. Zhao, L., Smith, J., & Wang, H. (2023). *Permutation Importance in Multilayer Perceptrons for Predicting Athletic Performance.* Journal of Sports Analytics, 9(2), 88–102.
6. Oytun, M., Kaya, T., & Demir, L. (2020). *Application of MLP Neural Networks and Permutation Importance in Elite Handball Player Evaluation.* International Journal of Sports Science & Coaching, 15(4), 421–432.
7. Letoffe, M., Dupont, A., & Verleysen, M. (2024). *Game Theory Meets Neural Networks: A Unified Framework for Feature Importance.* Neurocomputing, 487, 12–25.
8. Topend Sports. (2020). *Physical fitness tests.* Retrieved from https://www.topendsports.com/testing/tests
9. Physiopedia. (2021). *Gymnastics assessment and fundamental movement skills.* Retrieved from https://www.physio-pedia.com/Gymnastics_Assessment
10. Al-Azzawi, H. M., Al-Saadi, F. T., & Hassan, T. S. (2023). *Correlation between secondary school GPA and readiness for physical education programs.* Journal of Physical Education and Sport Science, 12(3), 45–58.
11. American Educational Research Association, American Psychological Association & National Council on Measurement in Education. (2014). *Standards for Educational and Psychological Testing.* Washington, DC: AERA.
12. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning.* MIT Press.
13. Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). *Learning representations by back-propagating errors.* Nature, 323(6088), 533–536.
14. Haykin, S. (1999). *Neural Networks: A Comprehensive Foundation* (2nd ed.). Prentice Hall.
15. Johnson, J. W., & LeBreton, J. M. (2004). *History and use of relative importance indices in organizational research.* Organizational Research Methods, 7(3), 238–257.
16. Kingma, D. P., & Ba, J. (2015). *Adam: A Method for Stochastic Optimization.* Proceedings of ICLR.

17. Srivastava, N., et al. (2014). *Dropout: A Simple Way to Prevent Neural Networks from Overfitting.* Journal of Machine Learning Research, 15(1), 1929–1958.
18. Ioffe, S., & Szegedy, C. (2015). *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift.* ICML.
19. Prechelt, L. (1998). *Early Stopping — But When?* In Neural Networks: Tricks of the Trade. Springer.
20. IBM Corp. (2016). *IBM SPSS Statistics 21 Neural Networks Module Documentation.* Armonk, NY: IBM.
21. Ziv, G., & Lidor, R. (2010). *Physical attributes, physiological characteristics, and on-court performances: A review on talented soccer players.* Journal of Sports Sciences, 28(15), 1323–1332.
22. Reilly, T., & Williams, A. M. (2003). *Science and Soccer.* Routledge.
23. McGill, S. M. (2007). *Low Back Disorders: Evidence-Based Prevention and Rehabilitation.* Human Kinetics.
24. Pienaar, A. E., & Coetzee, B. (2004). *The reliability and validity of four muscle endurance tests: Practical implications.* Physical Therapy in Sport, 5(3), 147–154.
25. Lohman, T. G., et al. (2000). *Assessment of Sprint Ability and Rapid Muscle Force Development in Young Athletes.* Journal of Strength and Conditioning Research, 14(2), 133–139.