

Fernández, E.; Izquierdo, J.M.; Zarauz, A.; Redondo, J.C. (202x) Prediction of Sports Talent in Young Throwers Using Machine Learning. Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte vol. (\*) pp. \*. [Http://cdeporte.rediris.es/revista/\\_\\_\\_\\*](http://cdeporte.rediris.es/revista/)

## ORIGINAL

# PREDICCIÓN DEL TALENTO DEPORTIVO EN JÓVENES LANZADORES UTILIZANDO MACHINE LEARNING

## PREDICTION OF SPORTS TALENT IN YOUNG THROWERS USING MACHINE LEARNING

Fernández, E.<sup>1</sup>; Izquierdo, J.M.<sup>2</sup>; Zarauz, A.<sup>3</sup> y Redondo, J.C.<sup>4</sup>

<sup>1</sup> Licenciado en Ciencias de la Actividad Física y Deporte. Universidad de León (España) etcefm00@estudiantes.unileon.es

<sup>2</sup> Doctor en Ciencias de la Actividad Física y del Deporte. Departamento de Educación Física y Deportiva. Universidad de León (España) jizqv@unileon.es

<sup>3</sup> Graduado en Matemáticas. Universidad de Almería (España) azm630@inlumine.ual.es

<sup>4</sup> Doctor en Ciencias de la Actividad Física y del Deporte. Departamento de Educación Física y Deportiva. Universidad de León (España) jcredc@unileon.es

**Código UNESCO / UNESCO code:** 5899 Otras Especialidades Pedagógicas (Educación Física y Deporte)

**Consejo de Europa / Council of Europe classification:** 4. Educación Física y deporte comparado / Physical Education and sport compared

**Recibido** 18 de septiembre de 2021 **Received** September 18, 2021

**Aceptado** 6 de marzo de 2022 **Accepted** March 6, 2022

### RESUMEN

El estudio aborda la detección de factores de rendimiento en los lanzamientos atléticos utilizando técnicas de Machine Learning, en 662 lanzadores (15,67 ± 1,01 años) del Programa Nacional de Tecnificación Deportiva de la Real Federación Española de Atletismo entre 1997 y 2015, mediante diferentes algoritmos (Logistic Regression, Random Forest y XGBoost). Al medir la importancia de las variables en función del rendimiento, el triple salto (0,20) destaca sobre el resto de variables: lanzamiento dorsal (0,14), envergadura (0,11), salto vertical (0,10), masa corporal, estatura (0,07) y flexibilidad (0,03). En cada disciplina, el triple salto encabeza la importancia en los lanzamientos de peso (0,20), disco (0,21) y martillo (0,29), mientras que el lanzamiento dorsal lo hace en la jabalina (0,20). El resto de las variables se reordena de forma particular modificando su importancia para cada disciplina. Estos hallazgos permiten mejorar la detección inicial de posibles talentos, así como su posterior especialización deportiva.

**PALABRAS CLAVE:** atletismo; rendimiento; fuerza; detección de talento.

## ABSTRACT

The study addresses the detection of performance factors in athletic throws using Machine Learning techniques, in 662 throwers ( $15.67 \pm 1.01$  years) of the National Program for Sports Technification of the Royal Spanish Athletics Federation between 1997 and 2015, using different algorithms (Logistic Regression, Random Forest and XGBoost). Measuring the variables as a function of performance, the triple jump (0.20) stands out over the other variables: dorsal throw (0.14), wingspan (0.11), vertical jump (0.10), body mass, height (0.07) and flexibility (0.03). In each discipline, the triple jump leads the importance in the shot put (0.20), discus (0.21) and hammer (0.29), while the dorsal throw does it in the javelin (0.20). The rest of the variables are rearranged in a particular way, modifying their importance for each discipline. These findings make it possible to improve the initial detection of potential talents, as well as their subsequent sports specialization.

**KEY WORDS:** athletics; performance; strength; talent detection.

## 1. INTRODUCCIÓN

En el ámbito atlético de los lanzamientos, como en otras disciplinas deportivas, la selección de talentos es un proceso complejo que persigue detectar en los deportistas, las características potencialmente determinantes para alcanzar el éxito deportivo. Puesto que los recursos son limitados, es necesario combinar a la perfección cualidades físicas y habilidades motrices tales como la fuerza, la potencia, el equilibrio o el ritmo, dependiendo en gran medida de un alto nivel de excelencia técnica (Judge et al., 2010), llegar al alto nivel requiere una gran cantidad de trabajo muy exigente durante años. Por eso, los procesos de selección temprana han ido ganando importancia hasta convertirse en una parte esencial del éxito (Ortigosa-Márquez et al., 2018; Maszczyk et al., 2011).

Tradicionalmente, los entrenadores han monitorizado, de forma periódica, los progresos alcanzados durante las sesiones de entrenamiento; usando estos datos para predecir el rendimiento, basado en muchas ocasiones en la propia experiencia fruto de la práctica. En el ámbito de los lanzamientos, los test de potencia y fuerza explosiva han sido una fuente habitual de información, aun cuando todo sugiere que no existe relación entre el rendimiento alcanzado por lanzadores noveles y experimentados (Maszczyk et al., 2014) debido a que el rendimiento forma parte de un esquema complejo (Musa et al., 2019; García-Martín et al., 2016).

El estudio de los factores de rendimiento en el deporte se ha realizado mediante la descripción de variables antropométricas, fisiológicas, psicológicas, biomecánicas o de condición física, utilizando la estadística descriptiva y modelos lineales, a pesar de que el análisis estadístico no ha sido capaz de describir con precisión las relaciones entre variables dependientes e independientes (Maszczyk et al., 2011). Actualmente contamos con modelos que

han generado eficientes sistemas computerizados de decisión, utilizados por ejemplo para disponer a jugadores de un equipo en las posiciones del campo más adecuadas (Woods et al., 2018) o determinando los factores de rendimiento para adecuar las cargas de entrenamiento en corredores de 400 metros vallas (Przednowek et al., 2017) obteniendo una mayor rentabilidad y disminuyendo el riesgo de lesionarse.

Técnicas como el Machine Learning permiten establecer nuevas relaciones entre los datos, debido a un uso eficiente de la información al discriminar entre relevantes e irrelevantes, teniendo en consideración únicamente los primeros (Musa et al., 2019). Así, recientemente encontramos modelos de predicción para el rendimiento en lanzadores de peso (Tuo & Li, 2020) o en corredores de 400 metros vallas (Przednowek et al., 2016). Estas técnicas también son de aplicación en la optimización del entrenamiento adecuando las cargas (Przednowek et al., 2017), en la selección de ejercicios más rentables (Flores & Redondo, 2020), en la disminución del tiempo de entrenamiento (Bunker & Thabtah, 2019) con programas individualizados o en el análisis del desempeño deportivo (Soto-Valero, 2018).

Hasta el momento, no hemos constatado la existencia de estudios que, apartados de una metodología tradicional, aborden los factores de rendimiento, orientados a la detección y selección de talentos deportivos en atletas de categorías menores en los lanzamientos. Por ello, motivado por el avance de las nuevas tecnologías y las posibilidades de realizar el tratamiento de los datos de una manera forma que ha demostrado ser más precisa, el objetivo de este estudio es identificar algunos de los factores más determinantes en el rendimiento deportivo de lanzadores jóvenes para mejorar los procesos de detección y selección temprana de talentos deportivos en el ámbito de los lanzamientos en general y en cada una de las pruebas en particular.

## **2. METODOLOGÍA**

### **2.1. Aproximación experimental al problema**

Este trabajo forma parte de una tesis doctoral centrada en la cuantificación y control de los factores de rendimiento de la condición física en lanzadores jóvenes de alto rendimiento. Después de haber estudiado el efecto de la edad relativa (Redondo et al., 2019) nos planteamos este estudio para analizar las variables que potencialmente pueden permitirnos determinar el rendimiento específico en lanzadores de categorías sub-18. Los sujetos del estudio participaron en alguna de las 106 concentraciones realizadas entre 1997 y 2015, organizadas por la Real Federación Española de Atletismo (RFEA) desarrollando el Programa Nacional de Tecnificación Deportiva (PNTD), convocado anualmente desde 1989 por el Consejo Superior de Deportes (CSD). Los participantes eran convocados, a dichas concentraciones, que se realizaban tres veces a lo largo del año, coincidiendo con los periodos de vacaciones escolares. En ellas, siempre realizaban los mismos test de forma estandarizada, precedida por un calentamiento y bajo las instrucciones de un licenciado en ciencias de la actividad física y del deporte, facilitando la familiarización de los lanzadores con el protocolo específico para cada prueba. Los test se realizaban siempre en el

mismo orden: pruebas antropométricas, capacidad de aceleración, salto vertical, triple salto y flexibilidad. Esta investigación se realizó respetando los principios establecidos por la Declaración de Helsinki, adoptada por la 18ª Asamblea Médica Mundial (Helsinki, Finlandia, junio de 1964) y enmendada por la 64ª Asamblea General (Fortaleza, Brasil, octubre de 2013); siendo aprobada, el 16 de abril de 2021, por el Comité de Ética (ETICA-ULE-022-2021) para la Investigación Científica de la Universidad de León, Comunidad de Castilla y León, en España. Los tutores legales dieron su consentimiento informado por escrito, previamente a la participación de los atletas en las diferentes actividades, y todos los participantes recibieron una explicación de los protocolos de prueba antes de la recopilación de datos.

## 2.2. Participantes

Como se observa en la Tabla 1, en el estudio participaron 662 lanzadores (393 hombres y 269 mujeres) con una edad decimal media de  $15,67 \pm 1,01$  años.

**Tabla 1.** Descripción de los deportistas participantes en el estudio (edad decimal media  $\pm$  DE)

Disciplina	Género	N	Edad (años)
Lanzamiento de Peso	Femenino	67	15,65 $\pm$ 0,95
	Masculino	102	15,59 $\pm$ 1,04
Lanzamiento de Disco	Femenino	55	15,53 $\pm$ 1,06
	Masculino	112	15,81 $\pm$ 1,01
Lanzamiento de Martillo	Femenino	79	15,59 $\pm$ 0,97
	Masculino	106	15,65 $\pm$ 1
Lanzamiento de Jabalina	Femenino	68	15,62 $\pm$ 1,03
	Masculino	73	15,83 $\pm$ 1,05

## 2.3. Variables del rendimiento en lanzadores

### 2.3.1. Medidas antropométricas

Las variables antropométricas fueron obtenidas siguiendo las directrices técnicas de medición y protocolo de la Sociedad Internacional para el avance de la Cineantropometría (Esparza-Ros et al., 2019). La medición y registro de estas variables formaron parte del reconocimiento médico realizado individualmente a cada atleta por médicos colegiados los servicios médicos designados para cada concentración por la RFEA.

La talla de los padres fue extraída de los informes aportados y recogidos al inicio de la actividad junto con todas las autorizaciones pertinentes.

### 2.3.2. Capacidad de aceleración (30m)

En la pista sintética de atletismo, los sujetos realizaban 2 carreras de velocidad de 30 metros lanzados con un tiempo de recuperación de 3 minutos entre cada intento. Los tiempos se tomaron utilizando un sistema de cronometraje

electrónico (Prosport TMR ESC 2100, Tumer Engineering, Ankara, Turquía). Para el análisis únicamente se registró el menor de los tiempos.

### **2.3.3. Salto vertical**

El rendimiento del salto vertical se evaluó utilizando una plataforma de fuerza portátil (Newtest, Finland). Los atletas ejecutaron un salto vertical con contramovimiento (CMJ) de acuerdo con Bosco (Bosco et al., 1983). Cada sujeto realizó 3 saltos con un tiempo de recuperación de 3 minutos entre cada intento. Los atletas fueron animados a saltar lo más alto posible, siendo el registro más elevado el registrado en centímetros para realizar el análisis.

### **2.3.4. Triple salto**

Desde una posición estática, se ejecutaron tres saltos horizontales consecutivos con los pies paralelos, siendo el último en un foso de arena. La distancia obtenida se midió con una cinta métrica, extendida a lo largo del suelo, desde el punto de inicio hasta la huella más próxima a este sobre la arena. Después de ejecutar tres intentos con una recuperación de 3 minutos, se registró la distancia más larga obtenida en metros y centímetros, truncando al centímetro más próximo siguiendo las normas de la Asociación Internacional de Federaciones de Atletismo (IAAF, 2019).

### **2.3.5. Lanzamiento dorsal (Dorsal)**

Los atletas realizaron 6 lanzamientos dorsales (Ekstrand et al., 2013) con la bola de peso, de acuerdo con las normas IAAF de competición en función de la categoría y el género (IAAF, 2019), con una recuperación entre cada intento de aproximadamente 3 minutos. El lanzamiento se iniciaba de espaldas a la zona de caída, tras separar los pies aproximadamente la anchura de los hombros, flexionando las rodillas y caderas hasta colocar el peso aproximadamente entre las rodillas, para lanzarlo sobre la cabeza con una brusca extensión de piernas, tronco y brazos. Se registró la distancia más larga obtenida, medida en metros y centímetros (truncando al más próximo) con una cinta métrica extendida a lo largo del suelo, desde el punto de lanzamiento hasta la huella de caída marcada en el suelo de hierba o arena más próxima a dicho punto de lanzamiento.

### **2.3.6. Flexibilidad**

Se midió la flexibilidad de la articulación coxo-femoral utilizando el ejercicio denominado sit-and-reach (López-Miñarro et al., 2009). Se midió la distancia alcanzada. Tras dos intentos, se registró el mejor resultado.

### **2.3.7. Marca personal (PB IAAF)**

Es el registro de la mejor marca personal obtenida en competición oficial y reconocida por la RFEA durante la temporada en la que participó en las actividades de tecnificación. Los datos fueron extraídos de las memorias que anualmente publica la RFEA en su memoria, entre los años 1997 y 2004, y de la

publicación de dichos rankings en la página web oficial de la Federación desde 2005 hasta 2015. Los registros de las marcas fueron normalizados, convertidos a su equivalente en puntos de acuerdo con las tablas oficiales de puntuación de la IAAF, herramienta oficial para comparar el rendimiento de los atletas y entre diferentes especialidades (Spiriev & Spiriev, 2017).

#### 2.4. Análisis estadístico

Se emplearon métodos de árboles de decisión para medir la importancia de las características, que han demostrado ser modelos robustos (Biau et al., 2008). La robustez es una condición necesaria, dado que nuestro conjunto de datos contiene una gran cantidad de variables (11) en comparación con la cantidad de muestras (662).

En cuanto a la visualización de datos, se realizó un análisis de componentes principales (PCA) (O'Donoghue, 2008). Este análisis se basa en la hipótesis de que nuestra variable se puede describir como una combinación lineal de nuestras características, basada en la covarianza o en las matrices de correlación. PCA se ha convertido recientemente en un tema muy popular entre la investigación de deportistas de élite (Gløersen et al., 2018). Para comprender las dependencias lineales mencionadas anteriormente, también se introdujo una matriz de correlación.

Para medir la importancia de la variable en un conjunto de datos, se ajustaron tres algoritmos basados en árboles de decisión (Breiman, 2001) a los datos. Durante el proceso de ajuste, se realizó un seguimiento del error de salida de la bolsa para cada punto de datos y se promedia en el bosque.

Los algoritmos anteriores no se basan en si el problema es de clasificación o de regresión; es decir, no se basan en la naturaleza de nuestra variable. En este caso, se proporcionó una variable continua y, por ello, las estimaciones de error fueron dadas por la función de pérdida de error cuadrático medio.

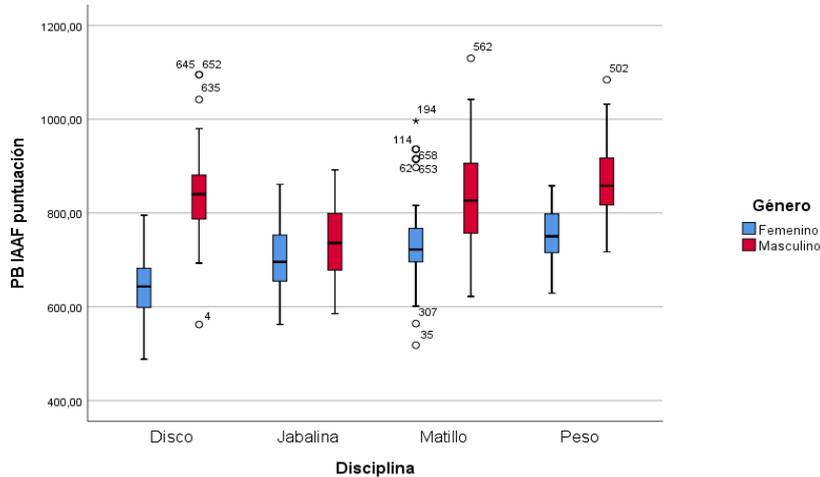
Para ser más precisos, los modelos de regresión logística, bosque aleatorio y aumento de gradiente extremo están entrenados para proporcionar estimaciones de la importancia de las características, y los resultados se ensamblan con los mismos criterios.

El problema de magnitud escalar de cada uno de los componentes se resolvió estandarizando las variables cuantitativas escalando al intervalo unitario, y las categóricas se codificaron numéricamente, ya que la información categórica no requirió un análisis extenso de procesamiento del lenguaje natural para poder encontrar un coeficiente de similitud entre variables (Singh & Singh, 2020).

El rendimiento de los modelos se evaluó mediante el lenguaje de programación Python con varias bibliotecas (pandas: 1.4.0; scikit-learn: 1.0.2; numpy: 1.22; matplotlib: 3.5.1; seaborn: 0.11.2).

### 3. RESULTADOS

La figura 1, muestra la relación existente en las distintas disciplinas en función de la categoría con el rendimiento obtenido según las tablas de puntuación de la IAAF. Se observa que el conjunto de datos no contiene muchos valores atípicos, lo que hace que los resultados sean consistentes; destaca que la mayoría de ellos pertenecen a la disciplina de lanzamiento de martillo.



**Figura 1.** Distribución de atletas ordenada por género, categoría y rendimiento.

En la figura 2, se representa la matriz que muestra la dependencia lineal entre las distintas variables. Las tres variables con una mayor dependencia con el rendimiento, expresado en puntos IAAF, son lanzamiento dorsal (0,54), triple salto (0,53) y salto vertical (0,49).

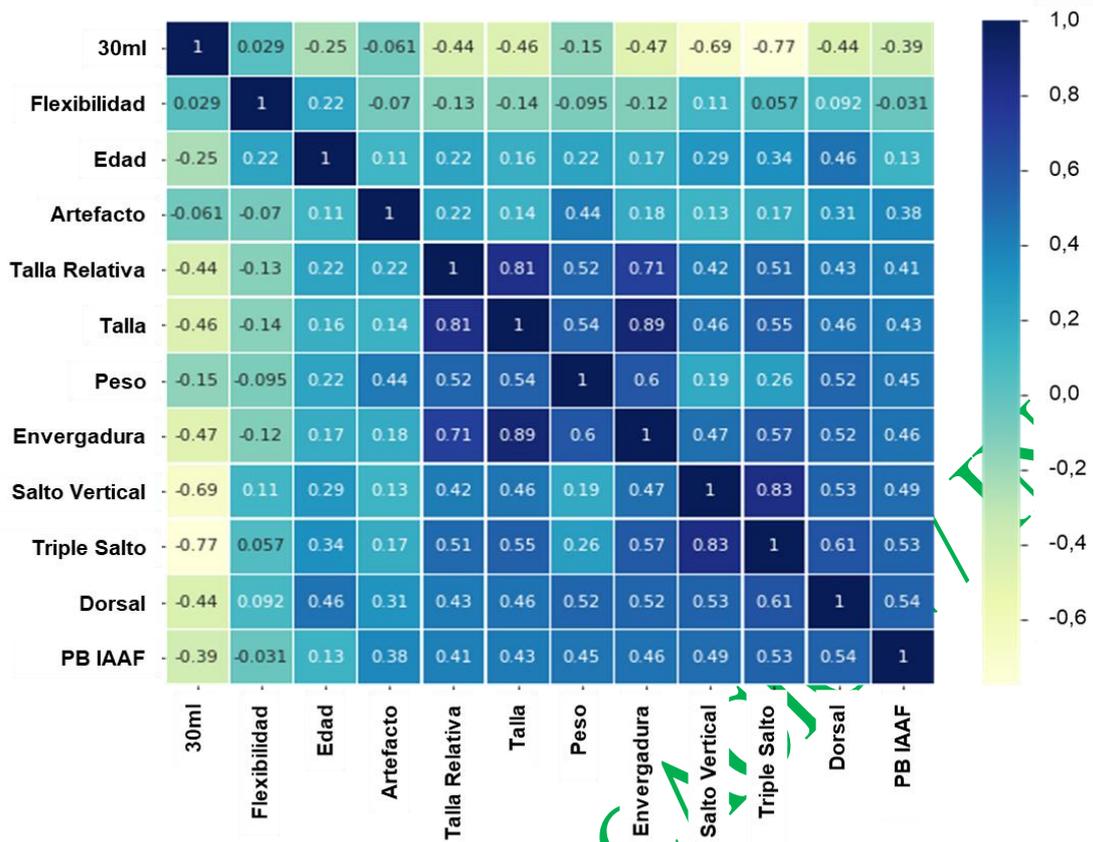


Figura 2. Mapa de calor de la dependencia lineal entre las variables.

Se observa una fuerte correlación entre el triple salto y el salto vertical (0,83), lo que podría llevarnos a eliminar una de esas variables según sus respectivos coeficientes lineales con respecto al rendimiento expresado en puntos IAAF. Sin embargo, al observar la distribución conjunta de ambas variables (Figura 3) se muestra un cierto grado de asimetría que nos lleva a conservar ambas.

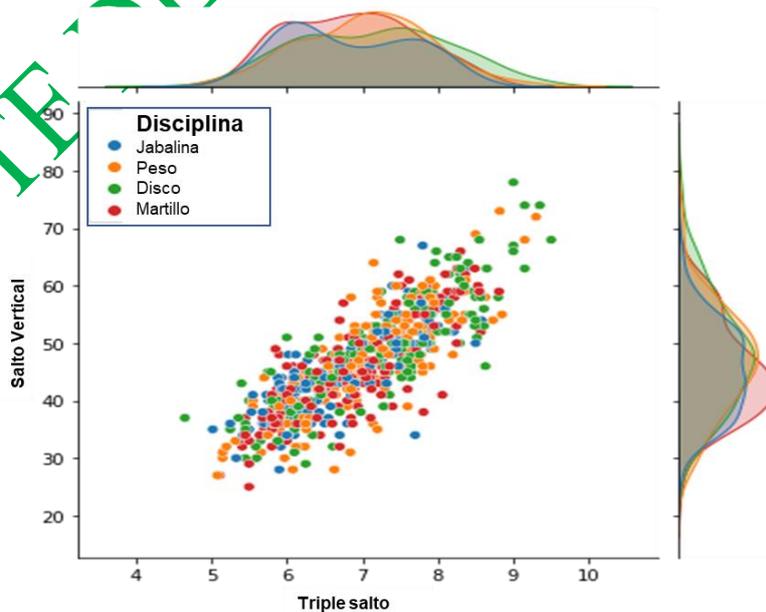


Figura 3. Distribución conjunta de salto vertical y triple salto, agrupada por disciplina.

Del mismo modo, como se observa en la figura 4, existe una fuerte correlación negativa entre el triple salto y la velocidad en 30 metros (-0,77). Observando más de cerca la distribución conjunta de ambas variables, según la disciplina respecto al rendimiento obtenido en puntos IAAF, se advierte cierto grado de asimetría, por lo que se decide nuevamente preservar ambas variables.

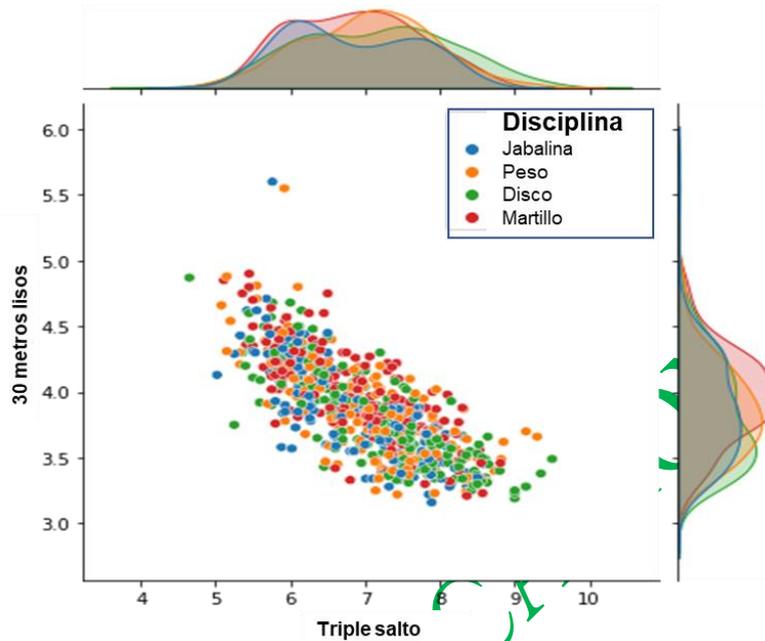


Figura 4. Distribución conjunta de 30 metros lisos y triple salto, agrupada por disciplina.

### 3.1. Modelado de la importancia de las variables

En esta sección, se muestra una herramienta cuantitativa para medir la importancia de las características para las siguientes diez variables: edad, artefacto, altura relativa, altura, peso, brazos extendidos, 30 m, flexibilidad, salto dorsal, vertical y triple salto. Como se mencionó anteriormente, se prueban tres modelos para proporcionar una medida efectiva de la importancia de la característica, y los resultados se agregan por modelo y característica (Figura 5):

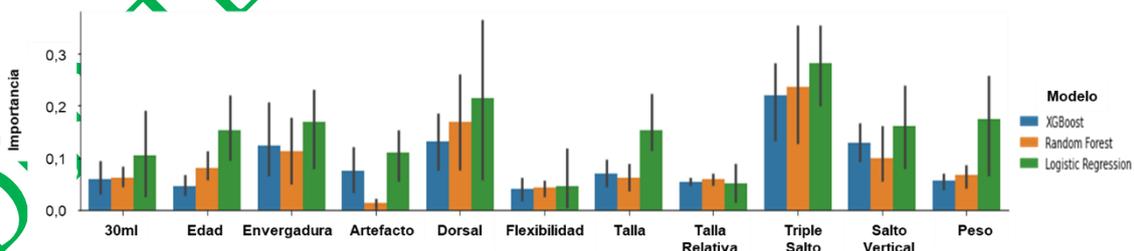


Figura 5. Distribución de la importancia de las variables por modelo.

Como se puede apreciar, el modelo logístico es un poco más "optimista" que otros estimadores en la mayoría de las variables, ya que Random Forest y XGBoost tienen en común la mayoría de sus puntuaciones. Esto es el resultado de la naturaleza de un modelo diferente: mientras que Logistic es un algoritmo único, Random Forest y XGBoost utilizan árboles de decisión con dos técnicas de regularización diferentes, que son el ensacado y el impulso.

Considerando el desempeño de los tres métodos, se ensamblan por la media ponderada, obteniendo una única representación numérica de importancia característica (figura 10). Es particularmente importante señalar el significado abrumadoramente fuerte de la variable de triple salto:

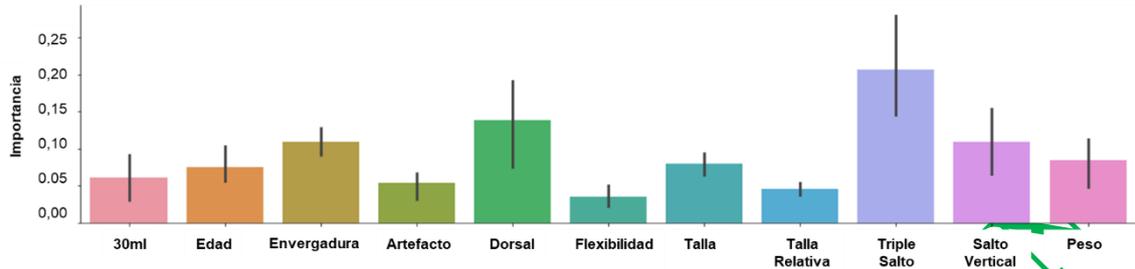


Figura 6. Distribución de la importancia agrupada por variable.

La Figura 6, muestra claramente la importancia de cada variable para todas las disciplinas. Observamos que de entre el conjunto de variables, el triple salto es la más destacada (0,20), una prueba específica en la que se mide la explosividad y la fuerza. En segundo lugar, se encuentran el lanzamiento dorsal (0,14), la envergadura (0,11) y el salto vertical (0,10), todos ellos de la misma naturaleza excepto la envergadura, que a su vez se correlaciona estrictamente con el género y categoría de los deportistas. Por otro lado, la flexibilidad (0,03) resulta ser la menos relevante según este análisis.

Sin embargo, este análisis no es capaz de detectar la significación de cada variable en cada disciplina y, a pesar de que todas pertenecen a la misma área, tienen características propias. Para lograr esto, la importancia de las características ahora se agrupa por modelo, característica y disciplina (Figura 7):

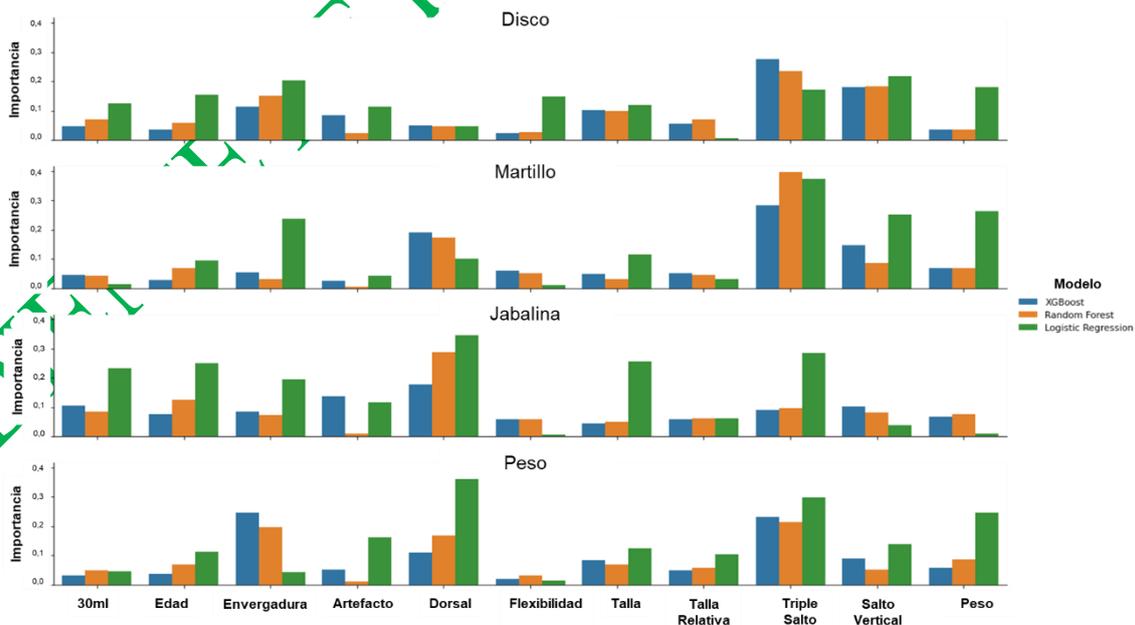


Figura 7. Distribución de la importancia de las variables por disciplina.

Como se ha reflejado en la Figura 7, el modelo logístico es más sensible a posibles valores atípicos, un factor técnico que hace que Random Forest y XGBoost sean más confiables. También es importante recordar que los modelos muestran una distribución muy diferente según la disciplina en variables como envergadura y peso.

Nuevamente, al ensamblar estimadores utilizando la media ponderada del desempeño de cada modelo, se muestra una perspectiva simplificada de los resultados (Figura 8):

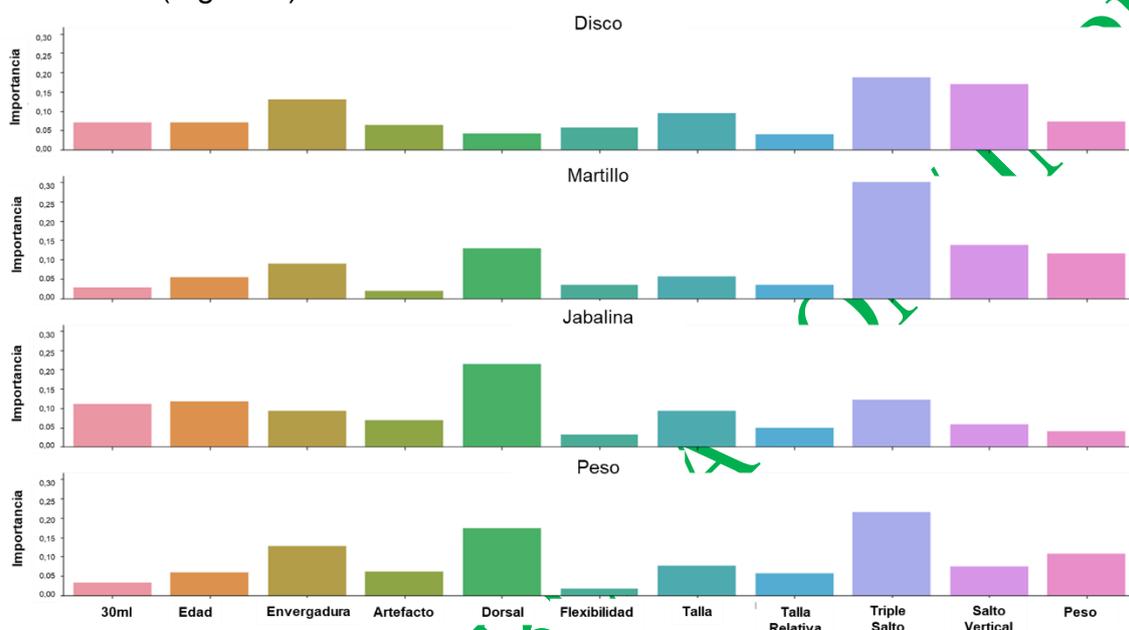


Figura 8. Distribución de la importancia agregada por variables diferenciada por disciplina.

#### 4. DISCUSIÓN

Este estudio presenta una nueva perspectiva para la selección temprana de talentos deportivos a través de la evaluación de los factores de rendimiento en jóvenes lanzadores de alto nivel. Los resultados confirman la hipótesis de los autores sobre la posibilidad de detectar el talento deportivo utilizando las nuevas tecnologías, propiciadas por los avances informáticos sin la inconsistencia de las regresiones lineales debido a las desviaciones. Así, nuestro estudio confirma las explicaciones de los autores de las tablas de puntuación de la IAAF, quienes afirman que no tratan de ser una herramienta de comparación del rendimiento entre hombre y mujeres, ya que el propio modelo es capaz de diferenciar entre el desempeño masculino y femenino debido a obvias diferencias biológicas (Spiriev & Spiriev, 2017).

El análisis de los componentes principales ha permitido descubrir la existencia de dos variables que, por sí solas, explican el 81% de la variabilidad de los datos; son la edad y la disciplina. Acorde con estos resultados, podemos afirmar que los jóvenes lanzadores consiguen mejores resultados por el hecho de hacerse mayores, a pesar de que evidentemente esta no es la causa directa. Evaluando nuestros resultados, podría especularse que el rendimiento estaría directamente asociado con el crecimiento en edad cronológica, debido al desarrollo de otra

serie de factores asociados como son los antropométricos y la evolución de cualidades físicas como la fuerza y la velocidad, que se muestran ligadas a la edad (Cobley et al., 2009). Tampoco podemos olvidarnos del mayor desarrollo que a las habilidades motoras proporciona el aprendizaje, fruto de la práctica (Davids et al., 2000) y que pudiera estar condicionado por la participación en un mayor número de concentraciones, en las que se realizan actividades técnicas dirigidas por entrenadores especializados; a lo que habría que añadir la posterior transferencia de estos aprendizajes a la competición y el incremento de la experiencia por la simple participación en un mayor número de competiciones (Folgar et al., 2014).

El presente estudio ha revelado que el triple salto es la variable más importante al determinar el rendimiento en los lanzamientos, acorde con los resultados obtenidos por Kyriazis et al. (2009), quienes descubrieron una fuerte correlación entre la capacidad de salto y el rendimiento en lanzadores, lo que contrasta con los descubrimientos de Aoki et al. (2015) quienes no descubrieron esta correlación entre la capacidad de salto y el rendimiento de forma específica y genérica en los lanzamientos. Por otra parte, el lanzamiento dorsal se mostró como la siguiente variable en orden de importancia, lo que podría refrendar la sugerencia de Zaras y col. (2019), quienes sugieren que un incremento de un 2% en el lanzamiento dorsal junto con un incremento de un 5% en el grosor de la porción vasto lateral del músculo cuádriceps pudiera incrementar el rendimiento un 6.9%.

La envergadura, según el modelo estudiado, aparece como la tercera variable en importancia, siendo la más destacada de entre las variables antropométricas, lo que concuerda con Dayal (2007), sobre las condiciones morfológicas de los lanzadores; afirmando que deben ser altos y pesados, con unas extremidades superiores más largas en relación con las inferiores. En una línea similar, las investigaciones realizadas por Till et al. (2015) dieron como resultado que los atletas seleccionados tienen en común unas características antropométricas particulares, relacionadas con el rendimiento, al mostrar una estatura elevada para incrementar la altura de liberación del artefacto. Del mismo modo, Zaras (2021) da mucha importancia al incremento en la altura de liberación del artefacto durante los lanzamientos y restando importancia a la masa corporal.

Diferenciando por disciplinas, se ha observado que en el lanzamiento de peso el factor de rendimiento más importante es el triple salto, lo que concuerda con Tuo & Li (2020) quienes, creando un modelo de red neuronal, confirmaron que el triple salto está muy relacionado con el rendimiento en esta disciplina. En este sentido, cabe destacar que, según Reis & Ferreira (2003), a pesar de que los ejercicios de saltos, por sí solos, no parecen ser indicadores precisos del rendimiento en esta disciplina, sugieren de su entrenamiento para mejorar la potencia explosiva en las piernas de los lanzadores de peso. Sus resultados mostraron una combinación de variables con ejercicios de fuerza, lanzamientos y saltos alcanzando una mayor precisión al predecir el rendimiento.

Abordando de forma específica los resultados obtenidos en la disciplina de lanzamiento de disco, las variables más importantes fueron el triple salto y el salto vertical, seguidos de la envergadura, la talla y el peso, ocupando

llamativamente el lanzamiento dorsal la penúltima posición en importancia. Siendo estos resultados similares a lo reflejado por Maeda et al. (2018) que dan una gran importancia a la envergadura y el peso.

Los resultados evidencian que, para la disciplina de lanzamiento de martillo, las variables más importantes son triple salto, salto vertical, lanzamiento dorsal y la masa corporal del lanzador. En el caso del salto vertical, nuestros resultados contrastan con los obtenidos por Sung & Lee (2011). Destacando que el rendimiento en el lanzamiento dorsal está ligado de forma significativa al rendimiento en el lanzamiento de martillo junto con la importancia del peso corporal justificado por el incremento de la masa muscular, que contribuye significativamente al rendimiento en esta disciplina (Terzis et al., 2010).

Rompiendo la tendencia de las disciplinas anteriores, es el lanzamiento dorsal la variable más importante para obtener una mayor distancia en la disciplina de lanzamiento de jabalina. El triple salto le sigue en importancia, lo que parece estar en la línea con los resultados de Ihalainen (2018) al exponer la existencia de una gran correlación entre el rendimiento y los saltos horizontales, siendo el triple salto el que tiene una correlación más estrecha con las marcas personales; concluyendo no ser necesario la inclusión de velocidad lanzada de 20 m en la batería de test de campo para determinar el rendimiento, algo que también se evidencia en nuestros resultados en el caso de los 30 m.

## 5. CONCLUSIONES

El presente estudio ha identificado la importancia de variables directamente relacionadas con el rendimiento para el desarrollo y la mejora de la detección de jóvenes lanzadores. Triple salto, lanzamiento dorsal, la envergadura del atleta y el salto vertical se han manifestado como marcadores eficientes del potencial rendimiento en lanzadores. La realización de estos test de campo es sencilla por la ausencia de materiales específicos, que puedan ser caros o complejos, y que están al alcance de atletas y entrenadores; ni la ejecución ni la medición conllevan ningún tipo de dificultad técnica.

Las técnicas de Machine Learning utilizadas en este estudio, han mostrado la capacidad de determinar los factores de rendimiento en jóvenes lanzadores. Los entrenadores deberían tener en cuenta estos ejercicios y, junto con las características morfológicas, seleccionar atletas con el fin de formar lanzadores para posteriormente, atendiendo a las características particulares de cada especialidad, orientarlos hacia ellas para alcanzar el mayor nivel posible. Como consecuencia de esta investigación podemos seleccionar lanzadores para cada una de las disciplinas, lanzamiento de peso, disco, martillo y jabalina.

### 5.1. Propuestas de mejora

Puesto que las tablas de puntos de la IAAF no tienen correlación entre sí en función del rendimiento y del género, podría hacer pensar que las variables estudiadas tuvieran un comportamiento diferente en función del género. Además, aunque los atletas estudiados son de alto nivel, existen diferencias en cuanto a

su rendimiento y podría ser productivo abordar el estudio contemplando diferentes rangos de rendimiento buscando una mayor proyección deportiva, lo que implicaría estudiar el comportamiento de las variables en función del nivel de rendimiento, definiendo diferentes rangos de éxito deportivo. Y por último, debido a la dependencia directa de las variables más importantes de la fuerza y la potencia muscular, sería conveniente tener en cuenta nuevas variables de esta índole para elaborar un modelo que nos permita alcanzar una mayor precisión de predicción

## 6. REFERENCIAS BIBLIOGRÁFICAS

- Aoki, K., Kohmura, Y., Sakuma, K., Koshikawa, K., & Naito, H. (2015). Relationships between field tests of power and athletic performance in track and field athletes specializing in power events. *International Journal of Sports Science & Coaching*, 10(1), 133-144. <https://doi.org/10.1260%2F1747-9541.10.1.133>
- Biau, G., Devroye, L., y Lugosi, G. (2008). Consistency of random forests and other averaging classifiers. *Journal of Machine Learning Research*, 9(9).
- Bosco, C., Luhtanen, P., y Komi, P.V. (1983). A simple method for measurement of mechanical power in jumping. *European journal of applied physiology and occupational physiology*, 50(2), 273-282.
- Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.
- Bunker, R. P., y Thabtah, F. (2019). A machine learning framework for sport result prediction. *Applied computing and informatics*, 15(1), 27-33. <https://doi.org/10.1016/j.aci.2017.09.005>
- Cobley, S., Baker, J., Wattie, N., y McKenna, J. (2009). Annual age-grouping and athlete development. *Sports medicine*, 39(3), 235-256.
- Davids, K., Lees, A., y Burwitz, L. (2000). Understanding and measuring coordination and control in kicking skills in soccer: Implications for talent identification and skill acquisition. *Journal of sports sciences*, 18(9), 703-714.
- Dayal, D. (2007). *Comparative Study of Anthropometrical Characteristics of Indian Elite Male Throwers of Different Throwing Events*, Aligarh Muslim University].
- Ekstrand, L.G., Battaglini, C.L., McMurray, R.G., y Shields, E.W. (2013). Assessing explosive power production using the backward overhead shot throw and the effects of morning resistance exercise on afternoon performance. *Journal of Strength and Conditioning Research*, 27(1), 101-106. <https://doi.org/10.1519/JSC.0b013e3182510886>
- Esparza Ros, F.; Vaquero Cristobal, R.; Marfell Jones, M. (2019). *Protocolo Internacional Para la Valoración Antropométrica; UCAM Universidad Católica de Murcia: Murcia, España.*
- Flores, F.J., y Redondo, J.C. (2020). Proposal for selecting weightlifting exercises on the basis of a cybernetic model. *International Journal of Advanced Research* 8, 906-914. <http://dx.doi.org/10.21474/IJAR01/10857>
- Folgar, M.I., Cristóbal, R.V., González, J.C.V., y Boubeta, A.R. (2014). El efecto de la edad relativa en la participación en el Plan nacional de tecnificación deportiva y la consecución de grandes éxitos deportivos en piragüismo. *Apunts. Educación física y deportes*, 1(115), 46-53.

- García-Martín, A.; Antúnez, A. e Ibáñez, S.J. (2016). Análisis del proceso formativo en jugadores expertos: validación de instrumento/ Analysis of Expert Players' Training Process: Validation of Tools. *Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte*, 16(61), 157-182. <https://doi.org/10.15366/rimcafd2016.61.012>
- Gløersen, Ø., Myklebust, H., Hallén, J., y Federolf, P. (2018). Technique analysis in elite athletes using principal component analysis. *Journal of sports sciences*, 36(2), 229-237. <https://doi.org/https://doi.org/10.1080/02640414.2017.1298826>
- IAAF. (2019). Reglamento de competición. Reglas técnicas (Vol. Libro C-C2.1).
- Ihalainen, S. (2018). Javelin throwing performance - Field test results and Biomechanics. World Javelin Conference, Olympic Training Center. Kuortane, Finland. [https://www.researchgate.net/publication/329040484\\_Javelin\\_throwing\\_performance\\_-\\_Field\\_test\\_results\\_and\\_Biomechanics](https://www.researchgate.net/publication/329040484_Javelin_throwing_performance_-_Field_test_results_and_Biomechanics)
- Judge, W.L., Bellar, D., McAtee, G., y Judge, M. (2010). Predictors of personal best performance in the hammer throw for US collegiate throwers. *International Journal of Performance Analysis in Sport*, 10(1), 54-65. <https://doi.org/10.1080/24748668.2010.11868501>
- Kyriazis, T.A., Terzis, G., Boudolos, K., y Georgiadis, G. (2009). Muscular power, neuromuscular activation, and performance in shot put athletes at preseason and at competition period. *Journal of Strength and Conditioning Research*, 23(6), 1773-1779. <https://doi.org/10.1519/jsc.0b013e3181b3f91e>
- López-Miñarro, P.A., de Baranda Andújar, P.S., y Rodríguez-García, P. L. (2009). A comparison of the sit-and-reach test and the back-saver sit-and-reach test in university students. *Journal of sports science and medicine*, 8(1), 116.
- Maeda, K., Byun, K.O., Hirose, K., y Ogata, M. (2018). Relationships between throwing distance and parameters about morphology and physical strength in male discus throw: Estimating physical strength requirement corresponding to throwing distance. *The Japan Journal of Coaching Studies*, 31(2), 175-184. [https://doi.org/10.24776/jcoaching.31.2\\_175](https://doi.org/10.24776/jcoaching.31.2_175)
- Maszczyk, A., Gołaś, A., Pietraszewski, P., Roczniok, R., Zając, A., y Stanula, A. (2014). Application of neural and regression models in sports results prediction. *Procedia-Social Behavior Science*, 117, 482-487.
- Maszczyk, A., Zając, A., y Ryguła, I. (2011). A neural network model approach to athlete selection. *Sports Engineering*, 13(2), 83-93.
- Musa, R. M., Majeed, A.P.A., Taha, Z., Chang, S.W., Nasir, A.F.A., y Abdullah, M.R. (2019). A machine learning approach of predicting high potential archers by means of physical fitness indicators. *Plos one*, 14(1). <https://doi.org/10.1371/journal.pone.0209638>
- Ortigosa-Márquez, J.M.; Reigal, R.E.; Serpa, S. y Hernández-Mendo, A. (2018). Efectos de la edad relativa en el proceso de selección nacional de triatletas / Relative Age Effect on National Selection Process in Triathlon. *Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte*, 18(70), 199-211. <http://dx.doi.org/10.15366/rimcafd2018.70.001>
- O'Donoghue, P. (2008). Principal components analysis in the selection of key performance indicators in sport. *International Journal of Performance*

- Analysis in Sport*, 8(3), 145-155.  
<https://doi.org/https://doi.org/10.1080/24748668.2008.11868456>
- Przednowek, K., Iskra, J., Maszczyk, A., y Nawrocka, M. (2016). Regression shrinkage and neural models in predicting the results of 400-metres hurdles races. *Biology of Sport*, 33(4), 415.  
<https://doi.org/10.5604/20831862.1224463>
- Przednowek, K., Iskra, J., Wiktorowicz, K., Krzeszowski, T., y Maszczyk, A. (2017). Planning training loads for the 400 m hurdles in three-month mesocycles using artificial neural networks. *Journal of human kinetics*, 60(1), 175-189. <https://doi.org/10.6018/cpd.37831>
- Redondo, J.C., Fernández-Martínez, E. y Izquierdo, J.M. (2019). Efecto de la edad relativa en las disciplinas de lanzamientos de los participantes españoles en el plan nacional de tecnificación de atletismo. *Cuadernos de Psicología del Deporte*, 19(3), 156-167.
- Soto-Valero, C. (2018). Aplicación de métodos de aprendizaje automático en el análisis y la predicción de resultados deportivos. *Retos: nuevas tendencias en educación física, deporte y recreación*, (34), 377-382.
- Reis, V.M., y Ferreira, A.J. (2003). The validity of general and specific strength tests to predict the Shot Put performance—a pilot study. *International Journal of Performance Analysis in Sport*, 3(2), 112-120.
- Singh, D., y Singh, B. (2020). Investigating the impact of data normalization on classification performance. *Applied Soft Computing*, 97, 105524. <https://doi.org/https://doi.org/10.1016/j.asoc.2019.105524>
- Spiriev, B., y Spiriev, A. (2017). IAAF scoring tables of athletics: 2017 revised edition. IAAF.
- Sung, B., y Lee, Y. (2011). An Examination of Basic and Specific Fitness of Elite National Throwing Athletes (Javelin, Shot put, Discus, Hammer Throwers). *Koran Journal of Sport Science*, 22(3), 2220-2236.
- Terzis, G., Spengos, K., Kavouras, S., Manta, P., y Georgiadis, G. (2010). Muscle fibre type composition and body composition in hammer throwers. *Journal of sports science & medicine*, 9(1), 104.
- Till, K., Cogley, S., O'Hara, J., Morley, D., Chapman, C., y Cooke, C. (2015). Retrospective analysis of anthropometric and fitness characteristics associated with long-term career progression in rugby league. *Journal of science and medicine in sport*, 18(3), 310-314. <https://doi.org/10.1016/j.jsams.2014.05.003>
- Tuo, X., y Li, T. (2020). Construction of a neural network model for performance prediction in shot put athletes. (Ed.),^(Eds.). *Journal of Physics: Conference Series*.
- Woods, C.T., Veale, J., Fransen, J., Robertson, S., y Collier, N.F. (2018). Classification of playing position in elite junior Australian football using technical skill indicators. *Journal of Sports Sciences*, 36(1), 97-103.
- Zaras, N., Stasinaki, A.-N., Methenitis, S., Karampatsos, G., Fatouros, I., Hadjicharalambous, M., y Terzis, G. (2019). Track and field throwing performance prediction: training intervention, muscle architecture adaptations and field tests explosiveness ability. *Journal of Physical Education and Sport*, 19, 436-443.

Zaras, N., Stasinaki, A.-N., y Terzis, G. (2021). Biological Determinants of Track and Field Throwing Performance. *Journal of Functional Morphology and Kinesiology*, 6(2), 40.

**Número de citas totales / Total references:** 39 (100%)

**Número de citas propias de la revista / Journal's Own References:** 2 (5%)

PENDIENTE DE PUBLICACIÓN / IN PRESS