






Machine learning en medicina deportiva. Evaluación del riesgo de tendinopatía rotuliana en atletas de alto nivel

Facundo M. Segura,¹ Florencio P. Segura,¹ María P. Lucero Zudaire,¹
Agustina Trevisson,² Florencio V. Segura¹

1. Segura Centro Privado de Ortopedia y Traumatología. Argentina

2. Sanatorio Dupuytren. Argentina

RESUMEN

Introducción: la tendinopatía rotuliana afecta comúnmente a atletas de alto rendimiento, especialmente en deportes con saltos repetitivos como el básquet. La identificación temprana de atletas en riesgo sigue siendo un desafío clínico, donde el aprendizaje automático podría ofrecer soluciones efectivas.

Objetivos: desarrollar un modelo de *machine learning* para predecir el riesgo de tendinopatía rotuliana en jugadores de baloncesto, utilizando variables clínicas, biomecánicas, antropométricas y de carga de entrenamiento.

Materiales y métodos: se realizó un estudio observacional y transversal en ochenta y un jugadores profesionales y semiprofesionales. Se evaluaron variables clínicas (VISA-P, EVA), ecográficas (diámetro del tendón en AP5 y AP10), antropométricas y funcionales (SLS). Los datos se analizaron en Python utilizando el algoritmo *Random Forest*, con validación cruzada de 10-folds.

Resultados: el modelo alcanzó un 93.94 % de precisión en el conjunto de prueba. La variable más influyente fue el score VISA-P (55.3 %), seguida por AP10 y EVA. La matriz de confusión mostró alta sensibilidad y especificidad, con mínima tasa de falsos negativos. La inclusión de un árbol de decisión facilitó la interpretación clínica.

Conclusión: el modelo *Random Forest* alcanzó una precisión superior al 93 % y permitió identificar de forma temprana el riesgo de tendinopatía rotuliana en jugadores de baloncesto; el score VISA-P y el AP10 se destacaron como los principales predictores. Su precisión y capacidad de interpretación lo convierten en una herramienta útil para la prevención personalizada en el deporte de alto rendimiento.

Palabras clave: Tendinopatía Rotuliana, Machine Learning, Score, Ecografía, Alto Rendimiento, Básquet, Prevención, Inteligencia Artificial

Nivel de evidencia: IV. Estudio de Cohorte Retrospectiva

Autor para correspondencia: Facundo M. Segura, facusegura@gmail.com

Recibido: 04/04/2025 Aceptado: 03/05/2025

DOI: <https://doi.org/10.63403/re.v32i2.409>

Cómo citar: Segura F, Segura P, Zudaire MP, Trevisson A, Segura F. Machine learning en medicina deportiva. Evaluación del riesgo de tendinopatía rotuliana en atletas de alto nivel. Relart 2025;32(2): 119-127.

Machine Learning in Sports Medicine. Risk Assessment of Patellar Tendinopathy in Elite Athletes

ABSTRACT

Introduction: patellar tendinopathy is prevalent among elite athletes, particularly in sports involving frequent jumping like basketball. Early identification of at-risk players remains a clinical challenge, where machine learning could offer innovative solutions.

Objectives: to develop a machine learning model to predict the risk of patellar tendinopathy in basketball players using clinical, biomechanical, anthropometric, and training load variables.

Materials and methods: a cross-sectional observational study was conducted on eighty-one professional and semi-professional basketball players. Clinical (VISA-P, VAS), ultrasound (AP5 and AP10), anthropometric, and functional (SLS) data were collected. The Random Forest algorithm was used with 10-fold cross-validation in Python.

Results: the model achieved a 93.94% accuracy in the test set. The most influential variable was VISA-P (55.3%), followed by AP10 and VAS. The confusion matrix showed high sensitivity and specificity with minimal false negatives. A simplified decision tree enhanced clinical applicability.

Conclusion: the Random Forest model achieved accuracy above 93% and enabled early identification of patellar tendinopathy risk in basketball players, highlighting the VISA-P score and AP10 as the main predictors. Its accuracy and interpretability make it a valuable tool for personalized prevention in elite sports settings.

Keywords: Patellar Tendinopathy, Basketball, Prevention, Machine Learning, Elite Athletes, Random Forest Model, Artificial Intelligence

Level of Evidence: IV. Retrospective Cohort Study

INTRODUCCIÓN

La tendinopatía rotuliana, comúnmente conocida como “rodilla del saltador”, es una lesión prevalente en atletas que realizan actividades con saltos repetitivos y cargas elevadas sobre el tendón rotuliano, como el básquet y el voleibol. Esta condición se caracteriza por dolor localizado en el polo inferior de la rótula y puede llevar a una disminución significativa del rendimiento deportivo e, incluso, a la incapacidad para competir.^{1,2}

En los últimos años, el aprendizaje automático ha emergido como una herramienta prometedora en el ámbito de la ortopedia y la traumatología, ofreciendo nuevas perspectivas para el diagnóstico, tratamiento y pronóstico de diversas patologías tendinosas,³ por ejemplo, se ha desarrollado un sistema de aprendizaje profundo para la medición automática de la altura rotuliana, demostrando una precisión comparable a la de radiólogos experimentados.¹ Además, se ha implementado esta herramienta para analizar imágenes ecográficas en escala de grises del tendón rotuliano, clasificándolas como normales o con tendinopatía, con una precisión cercana al 83 %, similar a la de clínicos expertos.⁴

Mendonça *et al.* identificaron que factores de la cadera y el pie/tobillo, como la rotación interna de cadera y la alineación pierna-antepié, se asocian con tendinopatía rotuliana en atletas. Mediante árboles de decisión, se logró clasificar con buena precisión la presencia de la lesión. El estudio destaca la utilidad de enfoques no lineales para detectar perfiles de riesgo más allá de los factores locales de rodilla.⁵ Adicionalmente, redes neuronales convolucionales han sido utilizadas para descubrir signos tempranos de

tendinopatía en imágenes de ultrasonido del manguito rotador, lo que ha demostrado que estas técnicas pueden aplicarse de manera similar en el estudio del tendón rotuliano.⁶

Uno de los mayores desafíos en la prevención de la tendinopatía rotuliana en atletas de alto rendimiento es la identificación temprana de factores de riesgo. Recientes avances en el análisis biomecánico y de imágenes médicas han permitido la detección de patrones que podrían predisponer a los deportistas a esta lesión. Se ha probado que los modelos de aprendizaje profundo pueden analizar dinámicas de salto y cambios en la biomecánica de la rodilla para predecir posibles lesiones.^{7,8}

A pesar de los avances en la evaluación clínica y en el uso de modelos de aprendizaje automático para el análisis de lesiones deportivas, la literatura actual presenta una limitada aplicación de estos métodos en poblaciones específicas, como los jugadores de básquet. En particular, no se han desarrollado modelos predictivos que integren variables clínicas, funcionales y ecográficas para estimar el riesgo de tendinopatía rotuliana en esta población. Este trabajo busca cubrir ese vacío, proponiendo un enfoque basado en aprendizaje automático que permite identificar atletas en riesgo mediante herramientas accesibles y de aplicación clínica directa.

Objetivos

Desarrollar un modelo de *machine learning* para predecir el riesgo de tendinopatía rotuliana en jugadores de baloncesto, utilizando variables clínicas, biomecánicas, antropométricas y de carga de entrenamiento.

MATERIALES Y MÉTODOS

Se realizó un estudio observacional, descriptivo y transversal con el objetivo de desarrollar un modelo de *machine learning* (ML) capaz de predecir la probabilidad de tendinopatía rotuliana en jugadores de baloncesto de alto rendimiento. La recolección de datos se llevó a cabo entre septiembre y octubre de 2024, en coincidencia con el período preparatorio de la temporada. La muestra fue seleccionada por conveniencia, en función de la disponibilidad de los equipos durante el período de recolección de datos.

La muestra estuvo conformada por ochenta y un jugadores (ciento sesenta y dos rodillas) de seis equipos de básquet profesionales y semiprofesionales de la Liga Nacional A, el Torneo Nacional de Ascenso y la Liga Federal. Se incluyeron jugadores de dieciocho años o más, con participación en entrenamientos y competencias durante el período de estudio, quienes firmaron el consentimiento informado. Se excluyeron aquellos con lesiones previas de rodilla que impedirían la práctica deportiva normal, cirugías previas en la rodilla evaluada, uso de corticosteroides en los tres meses previos al estudio o presencia de patologías sistémicas o metabólicas que afectarían la salud musculoesquelética.

Las variables analizadas incluyeron medidas antropométricas, clínicas, deportivas y ecográficas. Se registraron la altura, el peso y el índice de masa corporal (IMC), así como la posición en el campo, las horas de entrenamiento semanales y los años de experiencia profesional. La variable "posición en el campo" fue codificada numéricamente en cinco categorías: base (1), ayuda base (2), alero (3), ala-pívot (4) y pívot (5). A nivel clínico, se evaluó la función y el dolor mediante la puntuación VISA-P (Victorian Institute of Sport Assessment for Patellar Tendinopathy) y la escala de dolor EVA (Escala Visual Analógica). Para el análisis estructural del tendón rotuliano, se realizaron mediciones ecográficas del diámetro anteroposterior en dos puntos: a 5 mm (AP5) y 10 mm (AP10) del polo inferior de la rótula, obtenidas mediante un transductor lineal de alta frecuencia (7-15 MHz). Se calculó, además, el índice AP Ratio (AP5/AP10) como un marcador de engrosamiento tendinoso.

Los procedimientos incluyeron la administración del cuestionario VISA-P y la realización del Single Leg Squat Test (SLS) para evaluar estabilidad y funcionalidad. Posteriormente, se efectuaron ecografías de ambos tendones rotulianos con la rodilla en flexión de 90° mediante protocolos estandarizados de ganancia y profundidad, y se determinó el diámetro anteroposterior del tendón rotuliano a 5 y 10 mm del polo distal de la rótula.

Para el desarrollo del modelo de *machine learning*, los datos fueron procesados en Google Colab utilizando Python y las librerías *Pandas*, *NumPy* y *Scikit-Learn*. Se eliminaron columnas irrelevantes y las variables categóricas se transformaron en numéricas. El 6.8 % de los datos presentaban valores nulos en alguna de las variables consideradas, los cuales fueron imputados

con la media correspondiente. La variable dependiente se definió a partir del puntaje VISA-P, para el que se establecieron tres categorías: asintomáticos (VISA-P ≥ 90), en riesgo (VISA-P 80-89) y sintomáticos (VISA-P < 80).

Para la clasificación de la variable "Diagnóstico" se utilizó el algoritmo de agrupamiento K-Means. Inicialmente, se aplicó sobre la variable "Score" del conjunto de datos, definiendo tres *clusters*. Los grupos resultantes fueron denominados "Asintomático", "En Riesgo" y "Sintomático", en función de la distancia de cada punto de datos al centroide correspondiente. Posteriormente, se realizó una segunda clasificación empleando K-Means sobre las variables "Score" y "EVA", manteniendo la agrupación en tres *clusters*. Para evaluar la distribución de los datos, se generaron representaciones gráficas en un plano cartesiano, en las cuales la posición de cada centroide sirvió como criterio de clasificación (Fig. 1).

Los datos fueron divididos en un conjunto de entrenamiento (80 %) y otro de prueba (20 %) mediante la función *train_test_split* de *Scikit-Learn*. Para asegurar una representación equilibrada de cada clase en el conjunto de entrenamiento y prueba, se aplicó estratificación según la variable "Diagnóstico". La evaluación del modelo se hizo por medio de métricas estándar de clasificación, incluidas la matriz de confusión, el F1-score y la precisión global (*accuracy*).

Con el objetivo de predecir el riesgo de tendinopatía rotuliana en jugadores de básquet de alto rendimiento, se desarrolló un modelo de *machine learning* basado en el algoritmo *Random Forest* (RF). Este fue seleccionado por su robustez frente al sobreajuste, su capacidad para manejar variables no lineales y por su interpretabilidad, especialmente útil en contextos clínicos.

El conjunto de datos incluyó variables clínicas, funcionales y ecográficas que permitieron clasificar a los atletas en tres categorías: asintomáticos, en riesgo y sintomáticos, utilizando como referencia clínica el score VISA-P. Dado el desbalance de clases, se incorporó el parámetro *class_weight="balanced"* para compensar esta distribución.

El protocolo del siguiente estudio fue aprobado por el Comité de Ética de nuestra Institución y todos los pacientes firmaron un consentimiento informado aceptando su participación.

RESULTADOS

El modelo *Random Forest* fue entrenado inicialmente con las variables seleccionadas por su mayor relevancia clínica y estadística: score VISA-P, AP10, AP5, edad y posición en el campo. En el conjunto de prueba ($n = 33$), el modelo alcanzó un nivel de precisión del 93.94 %, clasificando correctamente a la mayoría de los jugadores en cada clase. El rendimiento detallado por clase se resume en la Tabla 1, se pueden observar los valores de precisión, *recall* y F1-score para cada una de las clases (Asintomático, En Riesgo y Sintomático), junto con el soporte o número de casos por clase. El

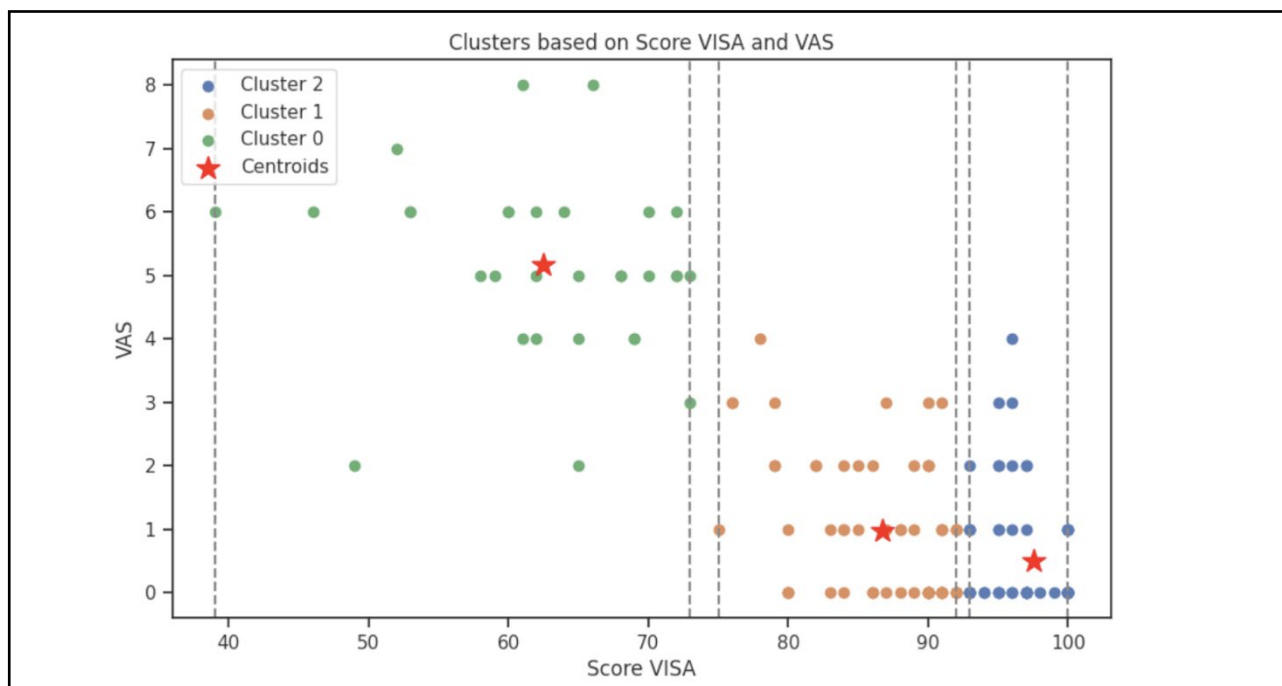


Figura 1. Clasificación de los datos mediante el algoritmo K-Means. Inicialmente, este se aplicó sobre la variable *Score*, generando tres *clusters*: “Asintomático”, “En Riesgo” y “Sintomático”, según la proximidad de los datos a sus centroides. Posteriormente, se realizó una clasificación considerando *Score* y EVA (VAS en el gráfico, por su nombre en inglés), manteniendo la división en tres grupos. La distribución de los datos y la posición de los centroides se presentan en un plano cartesiano, lo que evidencia la segmentación de los patrones identificados.

modelo demostró un rendimiento alto y equilibrado, con clasificación perfecta en las clases Asintomático y Sintomático, y un leve descenso en el *recall* de la clase En Riesgo.

Tabla 1. Métricas de desempeño del modelo *random forest* en el conjunto de prueba

Clase	Precisión	Recall	F1-score	Soporte
Asintomático (0)	0.90	1.00	0.95	19
En Riesgo (1)	1.00	0.78	0.88	9
Sintomático (2)	1.00	1.00	1.00	5

En términos globales, el modelo mostró una precisión promedio ponderada (*weighted avg*) de 0.95, un *recall* de 0.94 y un F1-score de 0.94. La matriz de confusión evidenció que todos los casos sintomáticos y asintomáticos fueron correctamente clasificados y solo un caso en riesgo fue clasificado como asintomático (falso negativo), lo cual representa un desafío clínico importante (Fig. 2).

Además del valor global de precisión, se calcularon métricas específicas por clase para evaluar el rendimiento discriminativo del modelo. La clase “Asintomáticos” alcanzó una sensibilidad del 100 %, una precisión del 90 % y una especificidad del 86 %, con un F1-score de 0.95. Para los atletas “En Riesgo”, el modelo mostró una sensibilidad del 78 %, precisión del

100 % y especificidad del 100 %, con un F1-score de 0.88. Finalmente, la clase “Sintomáticos” obtuvo valores perfectos en todas las métricas evaluadas (sensibilidad, precisión, especificidad y F1-score: 100 %). Estos resultados reflejan una capacidad robusta del modelo para distinguir entre los distintos perfiles clínicos, entre los que se destaca, particularmente, su eficacia para detectar correctamente los casos sintomáticos sin generar falsos negativos.

Optimización y validación cruzada del modelo

Para asegurar la estabilidad del modelo y verificar su capacidad de generalización, se realizó una validación cruzada de 10-*folds*. El nivel de precisión promedio obtenido fue de 93.93 %, con una desviación estándar de apenas 0.0001, lo que sugiere un modelo altamente confiable, con bajo riesgo de sobreajuste. En la Tabla 2 se puede observar cómo cada *fold* representa una partición distinta de los datos, y en todos los casos el modelo mantuvo un nivel de precisión constante del 94 %. Esta uniformidad en el rendimiento indica una alta estabilidad del modelo y baja variabilidad entre los subconjuntos, lo que refuerza su capacidad de generalización. En el contexto clínico, estos resultados sugieren que el modelo es confiable para predecir el riesgo de tendinopatía rotuliana en diferentes grupos de atletas, sin pérdida de precisión.

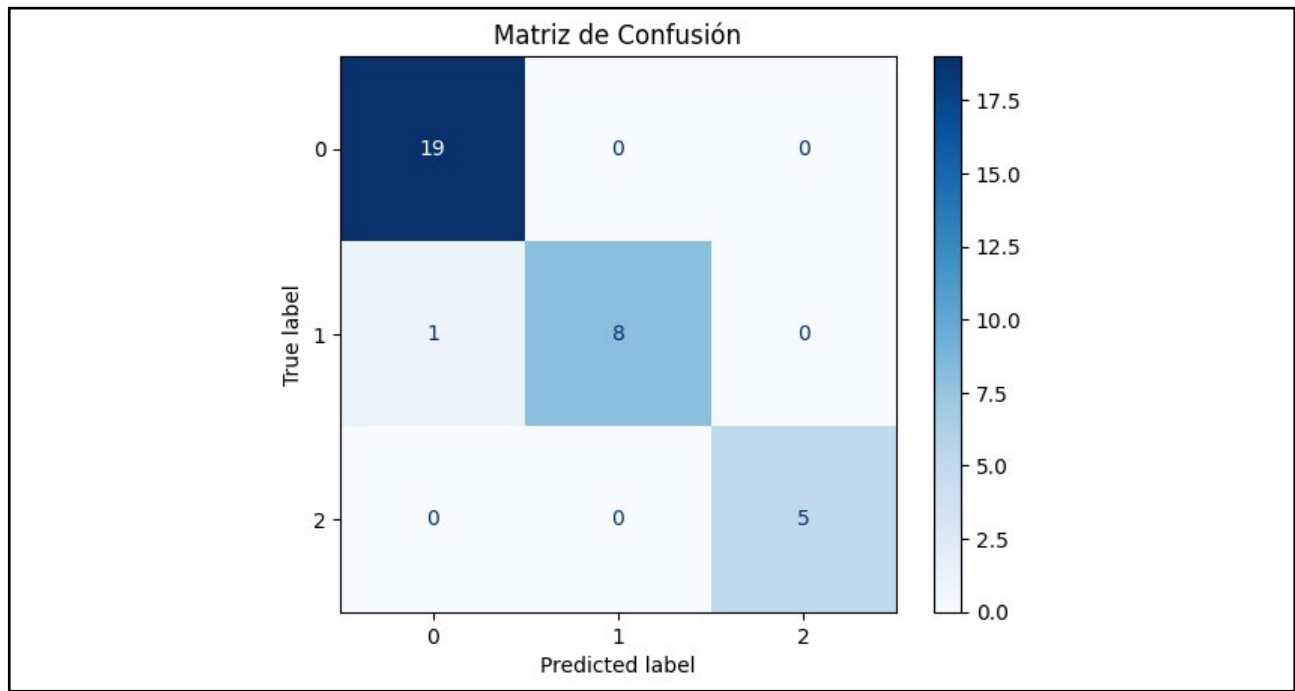


Figura 2. Matriz de confusión del modelo *Random Forest* aplicada al conjunto de prueba. Se observa una clasificación perfecta en las clases Asintomático (0) y Sintomático (2), y un único error de clasificación en la clase En Riesgo (1), donde un caso fue clasificado como asintomático. La matriz ilustra el alto nivel de precisión del modelo, especialmente en la identificación de sujetos sintomáticos, lo cual es clínicamente relevante para estrategias de intervención temprana.

Tabla 2. Resultados individuales de la validación cruzada de 10-folds aplicada al modelo *random forest*

Fold	Nivel de precisión
1	0.94
2	0.94
3	0.94
4	0.94
5	0.94
6	0.94
7	0.94
8	0.94
9	0.94
10	0.94
Promedio	0.94

Este comportamiento uniforme en todas las particiones del conjunto de entrenamiento valida el uso del modelo *Random Forest* como una herramienta predictiva sólida para aplicaciones clínicas.

Importancia de las variables predictoras

El análisis de importancia de características (*feature importance*) reveló que la variable *score VISA-P* fue la más determinante en el modelo, aportando un 55.3 % de la capacidad predictiva. Le siguieron, en orden

descendente, el diámetro del tendón rotuliano a 10 mm (AP10), EVA, edad, y posición en el campo (Fig. 3).

Aplicación clínica: árbol de decisión e interpretación de reglas

Se generó un árbol de decisión simplificado derivado del modelo *Random Forest* para facilitar su aplicación práctica en entornos clínicos. El flujo de decisiones se inicia con el valor de AP10, seguido por el *score VISA-P*, las horas de entrenamiento semanales, altura y la prueba funcional SLS. Esta estructura permite clasificar al atleta en una de las tres clases con un razonamiento clínico fácilmente interpretable (Fig. 4).

Ejemplo clínico ilustrativo sobre un caso real

En la Tabla 3 se pueden observar los valores de la evaluación de un caso real: un base (1) de equipo de Liga Nacional A.

Tabla 3. Ejemplo clínico ilustrativo sobre un caso real

Variable	Valor
AP10	5.2 mm
Score VISA-P	86
Horas entrenamiento	14
Altura	190 cm
SLS	0.7

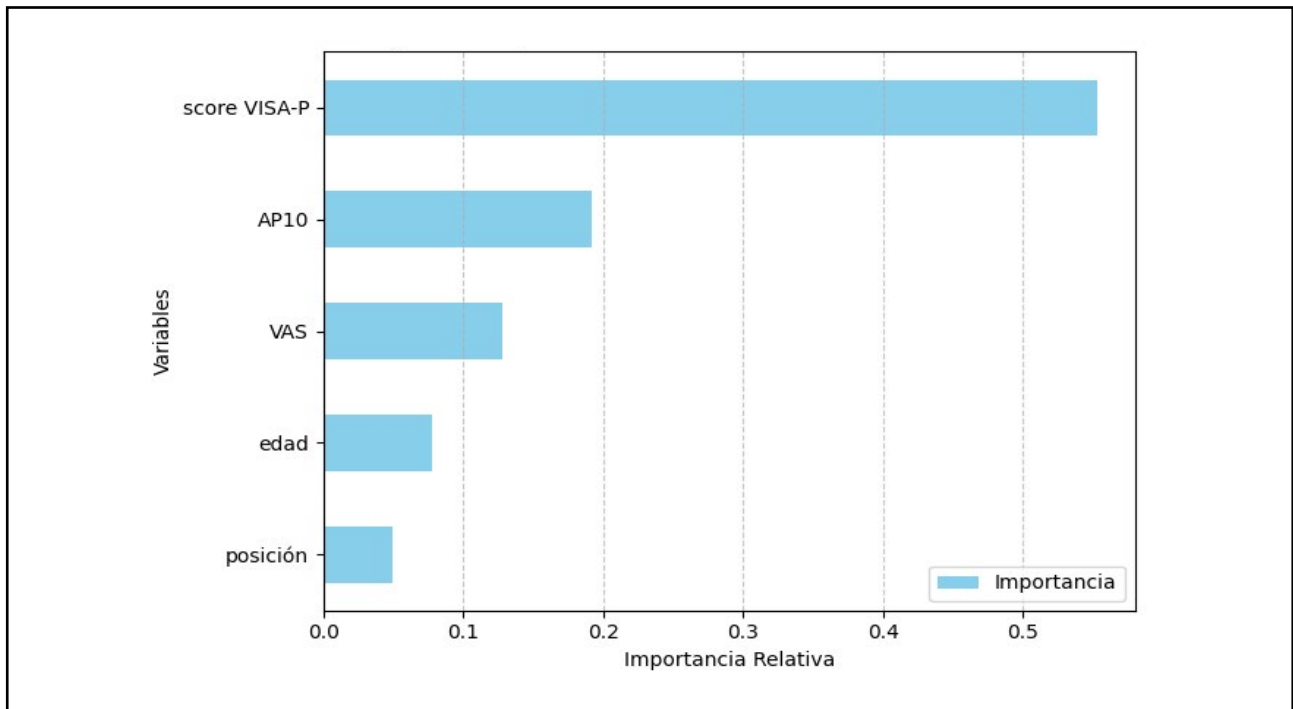


Figura 3. Importancia relativa de las variables utilizadas en el modelo *Random Forest* para la predicción de tendinopatía rotuliana. El score VISA-P fue la variable más influyente, seguida por el diámetro anteroposterior del tendón rotuliano a 10 mm (AP10), la escala de dolor EVA, la edad y la posición en el campo. Estos resultados confirman que las variables funcionales y estructurales tienen un peso predictivo mayor que los factores antropométricos. Se destaca la utilidad del score VISA-P como herramienta clave en la evaluación de riesgo clínico en atletas.

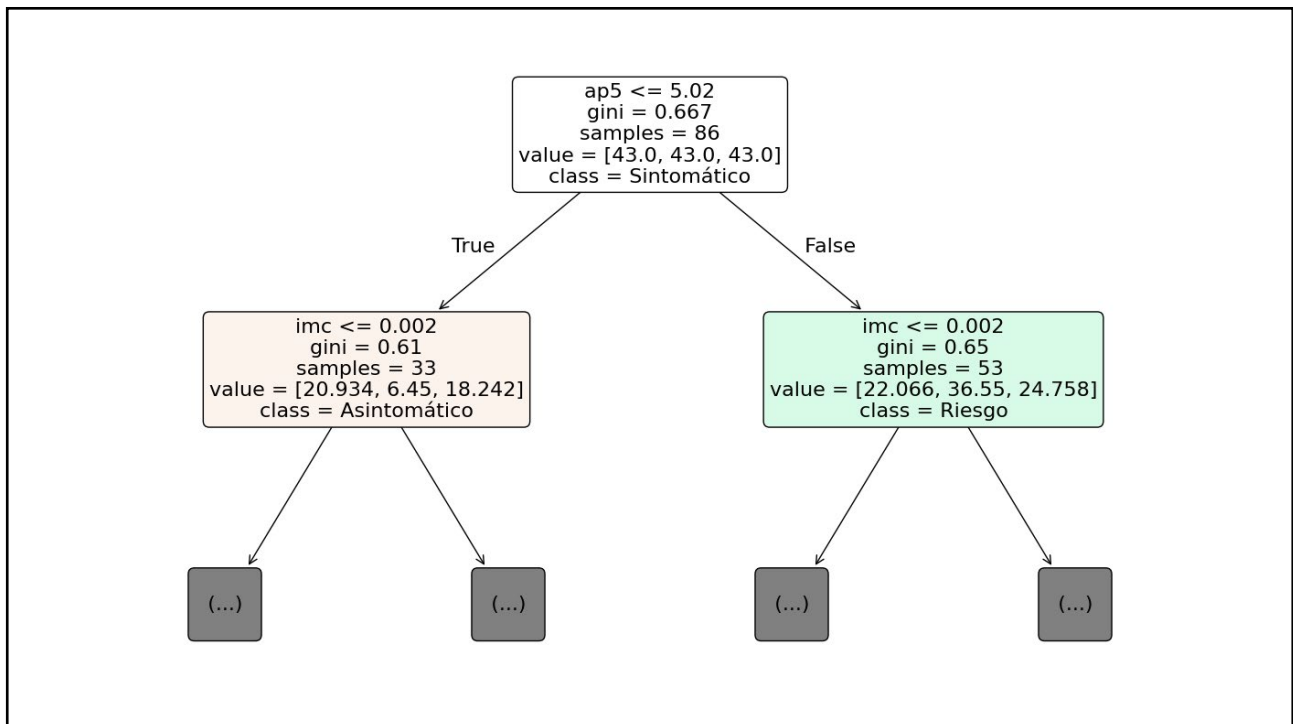


Figura 4. Visualización ilustrativa de un árbol de decisión individual perteneciente al modelo *Random Forest*. En esta figura se representan únicamente los primeros niveles del árbol (hasta profundidad 1) con el objetivo de facilitar la interpretación visual del proceso de decisión inicial. *Esta imagen no muestra el árbol completo y se incluye solo con fines ilustrativos para la presente publicación.

Ruta seguida en el árbol de decisión

1. **AP10 \leq 6.11** \rightarrow \checkmark (el diámetro del tendón está levemente engrosado. Continúa la evaluación).
2. **AP10 \leq 5.1** \rightarrow \times (el valor es 5.2, apenas por encima del umbral. Esto desvía al siguiente nodo).
3. **Horas de entrenamiento $>$ 12.5** \rightarrow \checkmark (el jugador entrena intensamente, lo que suma carga sobre el tendón. Se continúa).
4. **Score VISA-P \leq 89** \rightarrow \checkmark (un score de 86 indica síntomas funcionales moderados. Ya no es un jugador asintomático).
5. **Altura \leq 197.5 cm** \rightarrow \checkmark (altura dentro de parámetros normales. No modifica la trayectoria, pero se continúa).
6. **SLS \leq 0.5** \rightarrow \checkmark (valor de 0.7 indica buena función en la prueba de control neuromuscular).

El árbol de decisión ubica al jugador en la categoría de "Sintomático", a pesar de no presentar alteraciones funcionales severas en el SLS, esto se debe a:

- Un score VISA-P ya por debajo del rango saludable (\leq 89).
- Carga elevada de entrenamiento semanal ($>$ 12.5 h).
- Engrosamiento estructural leve (AP10 = 5.2 mm).

DISCUSIÓN

En el presente estudio se desarrolló un modelo de *machine learning* basado en el algoritmo *Random Forest* para predecir el riesgo de tendinopatía rotuliana en jugadores de básquet de alto rendimiento. El modelo, entrenado con variables clínicas, funcionales y ecográficas, alcanzó un desempeño sobresaliente con una exactitud del 93.94 % en el conjunto de prueba y una validación cruzada 10-fold con un nivel de precisión (*accuracy*) promedio de 94 %, sin evidencias de sobreajuste. Estos resultados destacan el potencial del modelo como herramienta predictiva en entornos clínicos y deportivos.

La aplicación de modelos de aprendizaje automático (ML) para la predicción de tendinopatías ha sido abordada en investigaciones recientes, principalmente mediante técnicas de *deep learning* e imagenología. Por ejemplo, una revisión sistemática y metaanálisis evaluó el rendimiento de modelos de redes neuronales en el diagnóstico de patologías tendinosas encontrando una sensibilidad del 91 % y especificidad del 95 %, lo que indica un alto grado de precisión en la detección de alteraciones tendinosas.⁹ Aunque nuestro modelo no se basa en imágenes, sino en datos clínicos y funcionales, el nivel de precisión obtenido se encuentra en línea con estos estudios, lo cual reafirma la utilidad de los algoritmos clásicos de ML, como *Random Forest*, en contextos clínicos estructurados.

Otro estudio reciente mostró que los modelos de ML pueden predecir la tendinopatía con alta precisión en poblaciones, tanto de atletas de élite como no élite, aunque resaltan la necesidad de identificar las características clínicas más relevantes para maximizar la capacidad predictiva.¹⁰ Este hallazgo es consistente con nuestro enfoque, donde el análisis de

importancia de variables permitió identificar factores clave que contribuyeron significativamente al rendimiento del modelo.

En nuestro caso, las variables más influyentes para el modelo fueron:

- **Score VISA-P:** este cuestionario específico para tendinopatía rotuliana resultó ser la variable más importante, representando más del 50 % de la importancia total del modelo. Su validez y confiabilidad han sido confirmadas previamente como instrumento de evaluación funcional en esta condición.¹¹
- **Escala Visual Analógica (EVA):** esta herramienta subjetiva de medición del dolor es comúnmente utilizada en la evaluación de patologías musculoesqueléticas, incluida la tendinopatía rotuliana.¹² En el presente modelo, su contribución también fue significativa.
- **Prueba funcional Single Leg Squat (SLS):** esta prueba evaluativa de control neuromuscular y estabilidad del miembro inferior ha sido previamente relacionada con alteraciones funcionales en atletas con tendinopatía.¹³
- **AP10 (diámetro del tendón rotuliano a 10 mm):** la medición ecográfica estructural del tendón mostró una alta correlación con la presencia de síntomas, reforzando su valor como marcador diagnóstico.¹⁴

Estos resultados coinciden con la literatura, que identifica factores como el aumento de la carga de entrenamiento, alteraciones biomecánicas durante el salto y el aterrizaje y deficiencias en la flexibilidad como contribuyentes a la aparición de la tendinopatía rotuliana.¹⁵ Además, investigaciones recientes han demostrado que la combinación de herramientas clínicas como el umbral de dolor a la presión (UDP), evaluaciones ecográficas y escalas como VISA-P o EVA mejoran significativamente la precisión diagnóstica en poblaciones deportivas.¹⁶

La elección del algoritmo *Random Forest* fue acertada dado su buen equilibrio entre desempeño, robustez frente al desbalance de clases y capacidad de interpretación. A diferencia de modelos de caja negra, como redes neuronales profundas, *Random Forest* permite visualizar las reglas de decisión dominantes, extraer árboles simplificados y comprender el papel de cada variable en el resultado final. Esto es especialmente valioso en el entorno clínico, donde la trazabilidad de la decisión diagnóstica es tan importante como el resultado en sí.

Si bien el modelo demostró un rendimiento excelente en esta cohorte, es importante considerar que el tamaño muestral relativamente reducido (ochenta y un atletas, ciento sesenta y dos rodillas) podría limitar la generalización de los resultados. Modelos de aprendizaje automático pueden sobreajustarse en muestras pequeñas, especialmente en escenarios con desbalance de clases, lo que podría comprometer su desempeño en poblaciones externas. Diversos autores han destacado la necesidad de

contar con grandes volúmenes de datos para garantizar la estabilidad y reproducibilidad de modelos predictivos clínicos. En este sentido, la validación externa en cohortes más amplias y diversas será fundamental para confirmar la utilidad práctica de esta herramienta.¹⁶

La implementación de este modelo podría contribuir significativamente a la detección temprana de jugadores en riesgo, optimizando la prevención de lesiones y mejorando los tiempos de intervención médica. Además, su estructura permite una fácil adaptación a nuevas variables clínicas, biomecánicas o imagenológicas en estudios futuros.

No obstante, es importante señalar que, si bien los resultados fueron prometedores, el modelo se construyó sobre una muestra de tamaño moderado y centrada exclusivamente en atletas de básquet. Por lo tanto, como ya mencionáramos, futuras investigaciones deberían validar este modelo en poblaciones más amplias y diversas, incluyendo atletas de otras disciplinas y diferentes niveles competitivos. La incorporación de nuevas variables, como marcadores biomecánicos en tiempo real, sensores portátiles o resonancia magnética funcional podría enriquecer la capacidad predictiva del modelo.

CONCLUSIÓN

El modelo *Random Forest* alcanzó una precisión superior al 93 % y permitió identificar de forma temprana el riesgo de tendinopatía rotuliana en jugadores de baloncesto; el score VISA-P y el AP10 se destacaron como los principales predictores. Su precisión y capacidad de interpretación lo convierten en una herramienta útil para la prevención personalizada en el deporte de alto rendimiento.

Contribuciones de autoría: Conceptualización, Metodología: FMS, FPS, FVS. Investigación: FMS, AT, MPLZ. Análisis formal: FMS, MPLZ. Redacción - borrador original: FMS. Revisión y edición: FMS, FVS. Supervisión: FVS. Visualización: FMS, AT.

Conflictos de intereses: los autores declaran no tener conflictos de interés relacionados con este estudio.

Financiamiento: los autores declaran que no hubo financiamiento externo para la realización de este estudio.

REFERENCIAS

1. Liu Z, Wu J, Gao X, Qin Z, Tian R, Wang C. Deep learning-based automatic measurement system for patellar height: a multicenter retrospective study. *J Orthop Surg Res.* 2024 May;19(1):324. doi: <https://www.doi.org/10.1186/s13018-024-04809-6>
2. Cheung CA. Machine learning approach for the objective sonographic evaluation of patellar tendinopathy [tesis]. Virginia: Virginia Polytechnic Institute and State University; 2021. 76 p. Disponible en: <https://vtechworks.lib.vt.edu/handle/10919/102728>

3. Mendonça LD, Ocarino JM, Bittencourt NFN, Macedo LG, Fonseca ST. Association of Hip and Foot Factors With Patellar Tendinopathy (Jumper's Knee) in Athletes. *J Orthop Sports Phys Ther.* 2018 Sep;48(9):676-684. doi: <https://www.doi.org/10.2519/jospt.2018.7426>
4. Bayramoglu N, Nieminen MT, Saarakkala S. Automated detection of patellofemoral osteoarthritis from knee lateral view radiographs using deep learning: data from the Multicenter Osteoarthritis Study (MOST). *Osteoarthritis Cartilage.* 2021 Oct;29(10):1432-1447. doi: <https://www.doi.org/10.1016/j.joca.2021.06.011>
5. Llombart R, Mariscal G, Barrios C, Llombart-Ais R. The best current research on patellar tendinopathy: a review of published meta-analyses. *Sports (Basel).* 2024 Feb;12(2):46. doi: <https://www.doi.org/10.3390/sports12020046>
6. Jahanifar M, Tajeddin NZ, Hasani M, Shekarchi B, Azema K. Automatic recognition of the supraspinatus tendinopathy from ultrasound images using convolutional neural networks. *arXiv preprint arXiv:2011.11777.* 2020. doi: <https://doi.org/10.48550/arXiv.2011.11777>
7. Verma A, Bhattacharya P, Zuhair M, Tanwar S, Kumar N. VaCoChain: Blockchain-Based 5G-Assisted UAV Vaccine Distribution Scheme for Future Pandemics. *IEEE J Biomed Health Inform.* 2022 May;26(5):1997-2007. doi: <https://www.doi.org/10.1109/JBHI.2021.3103404>
8. Siouras A, Moustakidis S, Giannakidis A, Chalatsis G, Liampas I, Vlychou M, et al. Knee Injury Detection Using Deep Learning on MRI Studies: A Systematic Review. *Diagnostics (Basel).* 2022 Feb;12(2):537. <https://www.doi.org/10.3390/diagnostics12020537>
9. Droppelmann G, Rodríguez C, Smague D, Jorquera C, Feijoo F. Deep learning models for tendinopathy detection: a systematic review and meta-analysis of diagnostic tests. *EFORT Open Rev.* 2024 Oct;9(10):941-952. <https://doi.org/10.1530/EOR-24-0016>
10. Muir D, Elgebaly A, Kim WJ, Althaher A, Narvani A, Imam MA. Predictive utility of the machine learning algorithms in predicting tendinopathy: a meta-analysis of diagnostic test studies. *Eur J Orthop Surg Traumatol.* 2025 Feb;35(1):73. doi: [10.1007/s00590-025-04197-5](https://doi.org/10.1007/s00590-025-04197-5)
11. Visentini PJ, Khan KM, Cook JL, Kiss ZS, Harcourt PR, Wark JD. The VISA score: an index of severity of symptoms in patients with jumper's knee (patellar tendinosis). *Victorian Institute of Sport Tendon Study Group. J Sci Med Sport.* 1998 Jan;1(1):22-8. [https://www.doi.org/10.1016/s1440-2440\(98\)80005-4](https://www.doi.org/10.1016/s1440-2440(98)80005-4)
12. Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant

Osteoarthritis Pain (ICOAP). *Arthritis Care Res (Hoboken)*. 2011 Nov;63 Suppl 11:S240-52. doi: <https://www.doi.org/10.1002/acr.20543>

13. Crossley KM, Zhang WJ, Schache AG, Bryant A, Cowan SM. Performance on the single-leg squat task indicates hip abductor muscle function. *Am J Sports Med*. 2011 Apr;39(4):866-873. doi: <https://www.doi.org/10.1177/0363546510395456>.

14. Cook JL, Purdam CR. Is tendon pathology a continuum? A pathology model to explain the clinical presentation of load-induced tendinopathy. *Br J Sports Med*. 2009 Jun;43(6):409-16. doi: <https://www.doi.org/10.1136/bjism.2008.051193>.

15. Tiemessen IJ, Kuijer PP, Hulshof CT, Frings-Dresen MH. Risk factors for developing jumper's knee in sport and occupation: a review. *BMC Res Notes*. 2009 Jul ;2:127. <https://www.doi.org/10.1186/1756-0500-2-127>

16. Karhade AV, Thio QCBS, Ogink PT, Shah AA, Bono CM, Oh KS, Saylor PJ, Schoenfeld AJ, Shin JH, Harris MB, Schwab JH. Development of Machine Learning Algorithms for Prediction of 30-Day Mortality After Surgery for Spinal Metastasis. *Neurosurgery*. 2019 Jul 1;85(1):E83-E91. doi: <https://www.doi.org/10.1093/neuros/nyy469>.