Módulo de Machine Learning – Comparación de Modelos

# 🎯 Objetivo

Predecir el potencial de los jugadores (Alto, Medio, Bajo) en función de métricas de rendimiento, y analizar qué variables son más influyentes según la posición del jugador.

# 📌 Modelos utilizados

1. Random Forest Classifier:

- Ensamble de árboles de decisión.

- Captura relaciones no lineales y maneja bien datos ruidosos.

- Permite obtener la importancia de cada variable directamente.

- Robusto y con buen rendimiento general.

2. Regresión Logística:

- Modelo estadístico lineal.

- Buena interpretabilidad: los coeficientes indican el peso de cada variable en la predicción.

- Eficiente computacionalmente.

- Útil como modelo base o comparativo.

# 🧪 Métricas calculadas

- Precision: Exactitud de las predicciones positivas.

- Recall: Capacidad para identificar todos los casos reales.

- F1-Score: Media armónica entre precision y recall.

Estas métricas se calculan por clase (Alto, Medio, Bajo).

# 📊 Visualización de Resultados

- Gráfico de barras horizontal con la importancia de variables:

- En Random Forest: basado en ganancia de información.

- En Regresión Logística: valor absoluto de los coeficientes.

- Tabla con métricas de clasificación por clase.

# 🧠 Conclusiones esperadas

Comparar modelos permite evaluar desempeño y variables relevantes.

El usuario puede seleccionar la posición y el modelo desde la interfaz del dashboard, lo cual permite un análisis dinámico y exploratorio.

Esto aporta valor en la toma de decisiones deportivas y procesos de scouting.

# ✅ Valor añadido al TFM

Integra análisis estadístico, machine learning y visualización interactiva.

Permite exploración de hipótesis y validación en tiempo real.

Aporta claridad, profundidad técnica y valor práctico al proyecto final de máster.