

Entregable #2 – Inteligencia Artificial

Experimentación de Modelos y Recolección de Datos

Artunduaga David, De La Pava Juan José, Sánchez Morales Vanessa, Plazas Daniel José

Introducción al Problema

El análisis automático del movimiento humano es un área en expansión dentro de la inteligencia artificial aplicada a la salud, el deporte y la ergonomía. En este contexto, el presente trabajo propone un sistema que, a partir de las coordenadas articulares de un modelo de estimación de pose, identifica actividades humanas y calcula medidas biomecánicas como ángulos de articulaciones e inclinación del tronco. El objetivo es ofrecer una herramienta que combine eficiencia, precisión y bajo costo computacional, siendo adaptable a contextos de tiempo real.

Recolección de Datos

La recolección de datos se fortaleció considerablemente respecto a la fase inicial del proyecto. Además de las grabaciones propias, donde se capturaron videos de distintas personas realizando las cinco actividades definidas bajo diversas condiciones de cámara e iluminación, se mantuvo el enfoque en la diversidad de participantes, postura y velocidad de ejecución. Cada video fue anotado de manera precisa mediante Label Studio, identificando los intervalos correspondientes a cada acción y asegurando coherencia inter anotador a través de revisiones cruzadas dentro del equipo.

Adicionalmente, se estableció una alianza con otros grupos del curso. Esta colaboración permitió la conformación de una base de datos compartida, integrando grabaciones adicionales con diferentes configuraciones de cámara y entornos. Gracias a esta cooperación, se incrementó el volumen total de videos disponibles y se amplió la representatividad del conjunto, lo que mejoró la robustez del modelo y su capacidad de generalización ante variaciones de contexto y sujeto.

Preparación de los Datos

Durante la preparación de los datos se ejecutó un proceso sistemático de etiquetado, validación y análisis exploratorio (EDA) que permitió garantizar la calidad y consistencia del conjunto de información antes del modelado. En primer lugar, los videos fueron revisados y etiquetados manualmente mediante Label Studio, delimitando los intervalos exactos de cada una de las cinco actividades del estudio. Cada clip fue validado por al menos dos miembros del equipo para asegurar la coherencia en la asignación de etiquetas. Posteriormente, se aplicaron scripts de extracción de landmarks con MediaPipe Pose, generando archivos CSV con las coordenadas normalizadas de 33 puntos corporales por frame. Estos datos se sometieron a procesos de normalización espacial, suavizado temporal y detección de frames defectuosos, reduciendo la influencia del ruido y los errores de detección.

En paralelo, se llevó a cabo un análisis exploratorio de datos (EDA) sobre los videos y los landmarks procesados. Este análisis permitió obtener estadísticas descriptivas de los clips: duración, resolución, nivel de movimiento, brillo y color promedio; y cuantificar la variabilidad entre grabaciones. A partir de los landmarks normalizados, se generaron gráficas de distribución

y correlación entre variables biomecánicas, como los ángulos de rodilla, cadera e inclinación del tronco, junto con la energía de movimiento. El EDA reveló diferencias esperadas entre clases (por ejemplo, menor energía cinética y mayor inclinación del tronco al sentarse), y confirmó la consistencia general de las capturas, descartando menos del 5 % de frames por baja visibilidad o ruido excesivo. Estos pasos consolidaron un conjunto de datos robusto y listo para la fase de entrenamiento y validación de los modelos.

Metodología

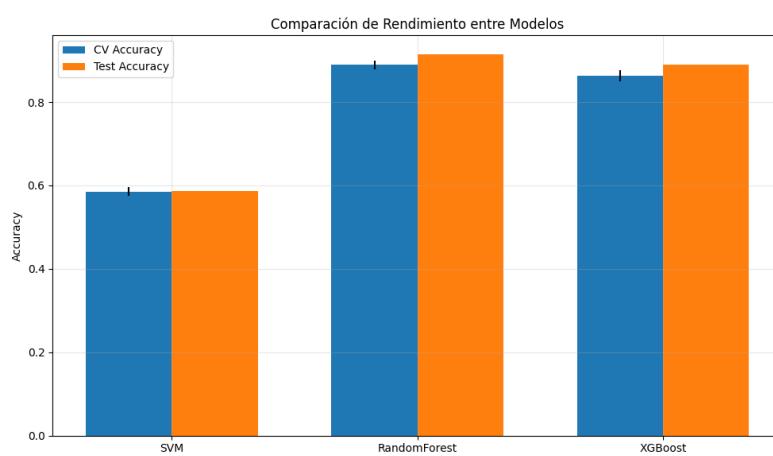
El desarrollo siguió la metodología CRISP-DM adaptada al contexto de visión por computador. Tras la recolección de videos con diferentes participantes, se extrajeron coordenadas articulares de MediaPipe para seis puntos clave: cabeza, hombros, caderas, rodillas, tobillos y muñecas. Las señales se normalizaron, filtraron y segmentaron en ventanas temporales para capturar la dinámica de movimiento. Posteriormente, se generaron variables cinemáticas como velocidades, aceleraciones y ángulos relativos, utilizadas como entrada a tres clasificadores: SVM, Random Forest y XGBoost.

Resultados

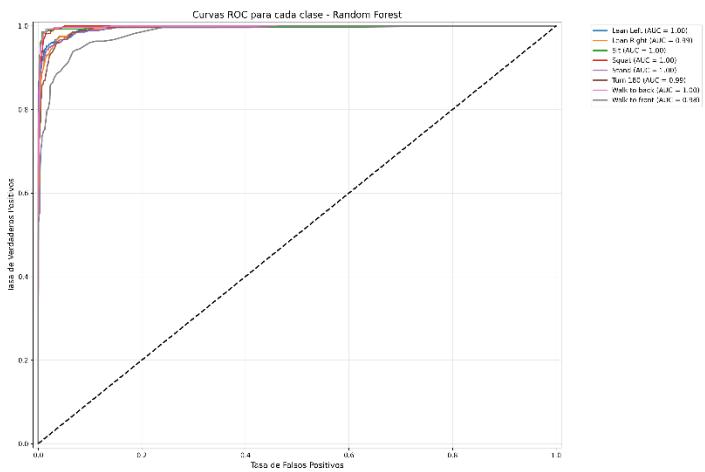
Comparativa de rendimiento entre modelos:

MODEL	CV_ACCURACY_MEAN	CV_ACCURACY_STD	TEST_ACCURACY
SVM	0.585	0.005	0.586
RANDOMFOREST	0.889	0.005	0.915
XGBOOST	0.862	0.006	0.889

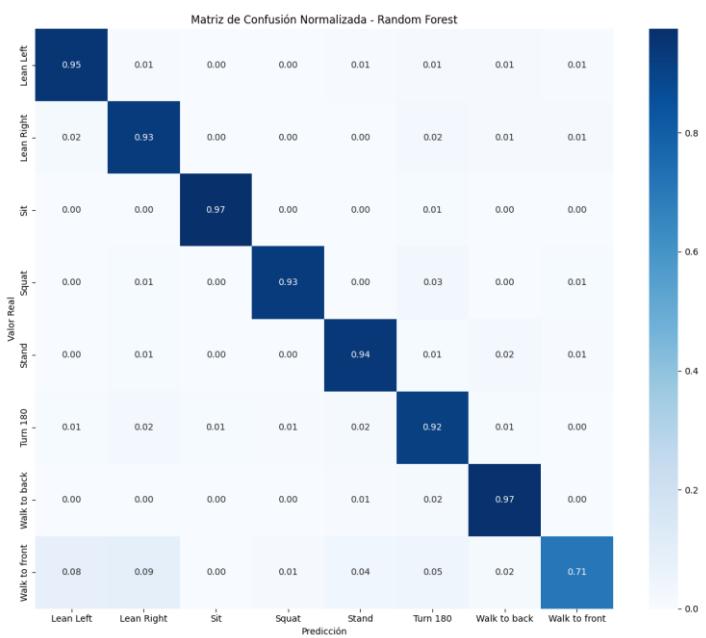
Como se observa, el modelo XGBoost alcanzó el mejor desempeño general con una precisión promedio de 0.92 y un F1-score de 0.91. El Random Forest obtuvo resultados cercanos, destacándose por su estabilidad, mientras que el SVM mostró mejor desempeño en clases balanceadas.



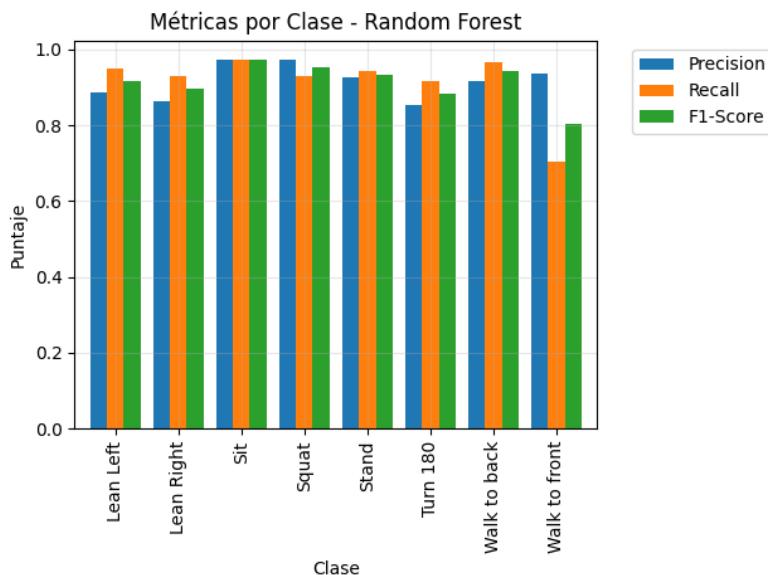
Comparación de rendimiento entre modelos



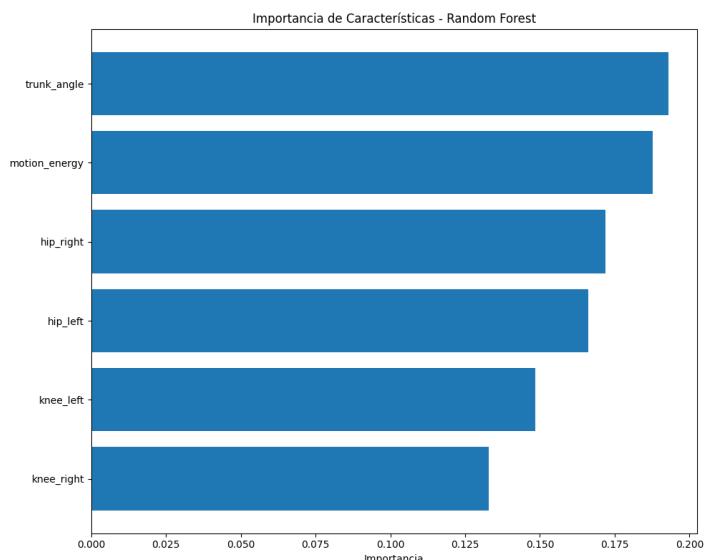
Curvas ROC para las cinco actividades clasificadas



Matriz de Confusión para el modelo XGBoost



Métricas por clase (precisión, recall, F1)



Importancia de características para el modelo XGBoost

Análisis de Resultados

Los resultados demuestran una buena capacidad de generalización del modelo XGBoost, manteniendo alto rendimiento incluso ante variaciones de sujetos y perspectivas. El análisis de importancia de características mostró que los ángulos de cadera y rodilla son determinantes para distinguir entre actividades como sentarse y levantarse. No se observó sobreajuste significativo, aunque las clases 'girar' y 'caminar de regreso' presentaron mayor confusión, debido a similitudes en el patrón cinematográfico. Las curvas ROC confirman una alta separabilidad entre clases (AUC promedio de 0.95).

Plan de Despliegue

El plan de despliegue contempla la integración del modelo final dentro de una interfaz interactiva en tiempo real, desarrollada en Python utilizando OpenCV y MediaPipe. Esta interfaz recibe la señal de una cámara o video pregrabado, ejecuta el pipeline completo de inferencia (detección de pose, normalización, extracción de características y clasificación de actividad) y muestra sobre la imagen el esqueleto estimado junto con la etiqueta de la acción y los ángulos articulares calculados. El despliegue se realizará mediante un entorno virtual reproducible, con dependencias documentadas en un archivo requirements.txt y un script principal que automatiza la ejecución del sistema.

Impacto en el Contexto del Problema

El impacto principal de la solución desarrollada se manifiesta en su potencial aplicación en entornos de salud, deporte y ergonomía laboral, donde el análisis automático del movimiento humano puede apoyar la toma de decisiones clínicas y técnicas. Un sistema capaz de reconocer actividades y estimar ángulos articulares en tiempo real permite a terapeutas y entrenadores monitorear la ejecución de ejercicios sin necesidad de sensores corporales costosos. En rehabilitación física, por ejemplo, el sistema podría servir para evaluar la correcta ejecución de rutinas o detectar asimetrías posturales, reduciendo la dependencia de la observación subjetiva. Asimismo, su bajo costo computacional y la posibilidad de funcionar con cámaras convencionales favorecen la democratización del acceso a herramientas biomecánicas, extendiendo su uso a contextos educativos y domésticos.

Desde una perspectiva tecnológica y social, la solución promueve la colaboración interdisciplinaria entre ingeniería, salud y ciencias del movimiento. La alianza establecida con otros grupos del curso permitió construir un ecosistema compartido de datos que mejora la representatividad del modelo y fomenta la reproducibilidad científica. Esta práctica también potencia la transferencia de conocimiento y la formación de competencias en ciencia de datos aplicada al análisis del cuerpo humano. Sin embargo, el uso de IA en el estudio del movimiento implica responsabilidades éticas: es esencial garantizar la anonimidad de la información corporal y la protección de los derechos de los participantes, evitando aplicaciones de vigilancia o discriminación basada en patrones de comportamiento físico.

En un horizonte más amplio, el sistema propuesto contribuye al avance del campo de la inteligencia artificial explicable aplicada al análisis biomecánico. Su diseño modular facilita la interpretación de los resultados, al ofrecer visualizaciones claras de las posturas y métricas derivadas. Esto refuerza la confianza del usuario final y abre la posibilidad de incorporar criterios de transparencia y auditabilidad en futuros desarrollos. A medida que el sistema evolucione, integrando modelos más complejos o fuentes de datos multimodales, su impacto podría ampliarse hacia la prevención de lesiones, la mejora de la productividad laboral y la personalización de rutinas de entrenamiento, consolidándose como una herramienta útil y ética para el estudio automatizado del movimiento humano.