

# Sistema de Reconocimiento de Movimiento y Seguimiento de Articulaciones

## **Integrantes:**

Pablo Fernando Pineda Patiño

Juan David Colonia Aldana

Miguel Ángel González Arango

**Docente:** Uram Aníbal Sosa Aguirre

**Asignatura:** Inteligencia Artificial I

Universidad ICESI

Facultad de Ingeniería, Diseño y Ciencias Aplicadas

Departamento de Computación y Sistemas Inteligentes

Ingeniería de Sistemas

2025

# Índice

1. Preguntas de Interés	3
2. Tipo de Problema	3
3. Metodología	3
4. Métricas para Medir el Progreso	5
5. Datos Recolectados	5
6. Estrategias para Conseguir Más Datos	5
7. Análisis Exploratorio de los Datos	6
8. Próximos Pasos del Proyecto	6
9. Aspectos Éticos por Considerar	7

# 1 Preguntas de Interés

Este proyecto busca desarrollar una solución de software capaz de identificar y analizar en tiempo real distintas actividades humanas, como caminar hacia o alejándose de la cámara, girar, sentarse y levantarse. El enfoque está en el monitoreo preciso de posturas y movimientos articulares, con especial atención en parámetros como la inclinación del tronco y la movilidad de articulaciones clave (muñecas, rodillas, caderas).

A partir de este objetivo, se plantean las siguientes preguntas de interés que guiarán el desarrollo del sistema:

- ¿Por qué es relevante o interesante abordar este problema en el contexto actual?
- ¿Cómo se adaptará la metodología CRISP-DM al proyecto? ¿Qué etapas se priorizarán?
- ¿Qué métricas se utilizarán para evaluar el rendimiento del modelo (precisión, recall, F1-Score, etc.)?
- ¿Cómo se medirá la precisión en la clasificación de actividades y el seguimiento de articulaciones?
- ¿Qué criterios se usarán para determinar el éxito del proyecto?
- ¿Qué datos se necesitan y cómo se recolectarán (videos, perspectivas, variaciones de velocidad, etc.)?
- ¿Existen posibles sesgos en los datos o el modelo que deban ser mitigados?
- ¿Qué soluciones similares existen en la literatura o en la industria? ¿Cómo se comparará el proyecto con ellas?

# 2 Tipo de Problema

El problema abordado en este proyecto corresponde a una tarea de **clasificación supervisada** en el ámbito de la visión por computadora. El objetivo principal es desarrollar un modelo que pueda reconocer actividades humanas específicas —como caminar, girar o sentarse— a partir de la información obtenida del seguimiento de puntos articulares (landmarks) extraídos de secuencias de video.

Dado que las actividades a clasificar ocurren a lo largo del tiempo, el problema se enmarca además en el análisis de **series temporales**, donde cada muestra consiste en una secuencia de coordenadas corporales asociadas a distintas articulaciones. Este tipo de enfoque es fundamental en aplicaciones que requieren interpretación automatizada del comportamiento humano, como el análisis biomecánico en deportes, la monitorización de pacientes en entornos clínicos o la detección de eventos críticos en sistemas de videovigilancia.

# 3 Metodología

El desarrollo seguirá la metodología CRISP-DM, adaptada al contexto del proyecto:

### 3.1. Comprensión del Negocio

**Objetivo:** Definir claramente el alcance del sistema: clasificar actividades humanas y analizar posturas en tiempo real mediante el uso de landmarks articulares.

### 3.2. Comprensión de los Datos

**Objetivo:** Explorar los datos disponibles y sus características:

- Grabación de videos de personas realizando actividades definidas.
- Etiquetado manual o semi-automático con herramientas como LabelStudio o CVAT.
- Análisis exploratorio de las posiciones y ángulos articulares durante distintas acciones.
- Evaluación de la calidad del seguimiento con MediaPipe.

### 3.3. Preparación de los Datos

**Objetivo:** Transformar los datos crudos en un formato utilizable:

- Extracción de landmarks con MediaPipe (hombros, caderas, rodillas, tobillos, muñecas).
- Normalización de coordenadas para independencia de distancia a la cámara.
- Aplicación de filtros suaves para reducir el ruido en las señales temporales.
- Segmentación de las secuencias en ventanas temporales para entrenamiento.

### 3.4. Modelado

**Objetivo:** Entrenar y optimizar modelos predictivos:

- Entrenamiento de modelos supervisados como SVM, Random Forest y XGBoost.
- Selección de características relevantes (ángulos entre articulaciones, velocidades, aceleraciones).
- Ajuste de hiperparámetros mediante validación cruzada.

### 3.5. Evaluación

**Objetivo:** Validar los resultados del modelo:

- Comparación de predicciones contra etiquetas reales.
- Cálculo de precisión, recall, F1-score y matrices de confusión.
- Evaluación del error en estimaciones angulares (ej. RMSE).
- Pruebas de tiempo real para medir latencia y estabilidad.

### 3.6. Despliegue

**Objetivo:** Implementar el sistema final:

- Visualización en tiempo real con información de actividad y ángulos.
- Documentación detallada del funcionamiento, limitaciones y requisitos técnicos.
- Posibilidad de exportar resultados para análisis posteriores.

## 4 Métricas para Medir el Progreso

Para evaluar el desempeño del sistema y monitorear el avance del proyecto, se utilizarán diversas métricas enfocadas tanto en la calidad de la clasificación como en la utilidad práctica del sistema. Entre ellas se destacan:

- **Precisión general (accuracy):** Indicará qué proporción de las actividades fueron correctamente identificadas por el modelo, reflejando su desempeño global.
- **Precisión positiva (precision):** Ayudará a conocer cuántas de las actividades que el modelo predijo como positivas realmente lo eran, evitando falsas alarmas.
- **Recall (sensibilidad):** Medirá la capacidad del modelo para detectar todas las actividades que efectivamente ocurrieron, especialmente importante para no omitir acciones relevantes.
- **F1-Score:** Permitirá encontrar un balance entre precisión y sensibilidad, útil cuando existe desequilibrio entre clases.
- **Matriz de confusión:** Se utilizará para visualizar qué actividades tienden a ser confundidas entre sí, lo cual orientará ajustes en el modelo o los datos.

## 5 Datos Recolectados

La base de datos del proyecto está compuesta por grabaciones de video donde se observan personas realizando una serie de actividades específicas, como caminar hacia la cámara, alejarse, girar, sentarse y levantarse. Estos videos se han capturado desde distintos ángulos y con variaciones de velocidad y estilo para enriquecer la diversidad del conjunto de datos. Además, se ha comenzado el proceso de etiquetado de las actividades presentes en cada segmento, utilizando herramientas especializadas para anotar los fotogramas y facilitar la futura extracción de puntos articulares.

## 6 Estrategias para Conseguir Más Datos

Dado que la calidad y diversidad del conjunto de datos son esenciales para el éxito del proyecto, se consideran las siguientes estrategias para su expansión:

- **Grabación adicional:** Se realizarán nuevas sesiones de captura con distintos participantes, escenarios y condiciones ambientales para aumentar la robustez del sistema.
- **Aumento de datos (data augmentation):** Se aplicarán transformaciones sintéticas sobre los videos existentes, como rotación, inversión horizontal o alteraciones en la velocidad de reproducción.
- **Uso de datasets públicos:** Se explorará la posibilidad de integrar bases de datos ya existentes, como el NTU RGB+D o MPII Pose Dataset, siempre que sean compatibles con el enfoque y el formato del proyecto.

- **Simulación de escenas:** Se evaluará el uso de entornos virtuales para generar datos sintéticos realistas que complementen los casos reales.

## 7 Análisis Exploratorio de los Datos

El análisis exploratorio tiene como finalidad comprender mejor la estructura y distribución del conjunto de datos antes del modelado. En esta etapa se han considerado los siguientes aspectos:

- **Distribución de clases:** Se examina qué tan equilibrada está la representación de cada actividad en los datos recolectados, identificando posibles desbalances.
- **Calidad de los landmarks:** Se inspecciona el ruido o posibles errores en los puntos articulares detectados automáticamente, especialmente en condiciones de iluminación o movimiento complejas.
- **Patrones articulares:** Se analizan las trayectorias de articulaciones clave (como caderas, rodillas y muñecas) durante las distintas actividades para identificar rasgos discriminativos.
- **Relación entre ángulos y actividades:** Se estudian los ángulos entre segmentos corporales para validar su utilidad como características relevantes para la clasificación.

## 8 Próximos Pasos del Proyecto

Para asegurar un desarrollo estructurado del proyecto y alcanzar los objetivos planteados, se han definido las siguientes etapas con sus respectivas tareas:

- **Anotación de datos:** Se continuará con la recopilación de nuevas grabaciones, asegurando que las perspectivas de cámara y condiciones de iluminación sean adecuadas para una correcta detección de movimientos. Posteriormente, se realizará el etiquetado de las actividades clave utilizando herramientas como *LabelStudio* o *CVAT*, lo que permitirá generar un conjunto de datos organizado para entrenar los modelos.
- **Preprocesamiento de datos:** Una vez recolectados y anotados los videos, se procederá a normalizar y filtrar la información de seguimiento articular obtenida con herramientas como *MediaPipe*. Esto implica aplicar técnicas de suavizado para reducir el ruido en los puntos articulares y estandarizar las coordenadas, minimizando la influencia de la altura del sujeto o la distancia a la cámara. El objetivo es construir características robustas que contemplen la velocidad articular, ángulos relativos e inclinación del tronco.
- **Selección y entrenamiento de modelos:** Se evaluarán diferentes algoritmos de clasificación supervisada, tales como SVM, Random Forest y XGBoost. Para este proceso, los datos serán divididos en conjuntos de entrenamiento y prueba, manteniendo una estructura de aproximadamente 80 % para entrenamiento y 20 % para validación.

Inicialmente se busca alcanzar una precisión mínima del 60 % en la clasificación de actividades como punto de partida.

- **Ajuste de hiperparámetros:** Con base en los resultados obtenidos en las primeras pruebas, se procederá a optimizar los modelos mediante técnicas de búsqueda como *Grid Search* o *Random Search*. El propósito de esta fase es refinar los parámetros clave de los modelos para mejorar el desempeño, priorizando métricas como F1-Score, con la meta de superar un valor de 0.85.
- **Desarrollo de la interfaz gráfica:** Paralelamente al entrenamiento, se diseñará una interfaz sencilla que permita visualizar las actividades reconocidas junto con los ángulos articulares en tiempo real. Esta primera versión funcional mostrará las predicciones generadas por el modelo y reflejará, de forma gráfica, el seguimiento de las articulaciones clave durante la ejecución de las actividades.
- **Validación y pruebas del sistema:** Una vez el modelo y la interfaz estén integrados, se realizarán pruebas con múltiples participantes para evaluar el desempeño del sistema en diversas condiciones (variación de posiciones, distancias y sujetos). Se planea llevar a cabo al menos 10 pruebas controladas comparando las predicciones del sistema con las etiquetas reales para validar la robustez y generalización del modelo.
- **Documentación e informe final:** Finalmente, se elaborará un informe detallado que incluirá los resultados obtenidos, la arquitectura y diseño implementado, así como la comparación con soluciones similares existentes. Además, se documentarán las lecciones aprendidas, mejoras potenciales y se preparará la entrega final, asegurando que tanto el código como la documentación estén actualizados en el repositorio correspondiente.

## 9 Aspectos Éticos por Considerar

Dado que el proyecto involucra el uso de imágenes y datos derivados del cuerpo humano, es importante tener en cuenta algunos principios éticos clave para asegurar un desarrollo responsable:

- **Privacidad y consentimiento:** Antes de recopilar cualquier video, se solicitará el consentimiento de los participantes de forma clara y accesible.
- **Diversidad y equidad:** Se intentará incluir una variedad de perfiles en el conjunto de datos (por ejemplo, en términos de género, edad o contextura física), con el fin de que el sistema sea más justo y no favorezca ciertos grupos sobre otros.
- **Seguridad de la información:** Los datos se almacenarán de forma segura, limitando el acceso solo a personas autorizadas y eliminando aquellos elementos sensibles una vez que dejen de ser necesarios.
- **Uso consciente:** Se fomentará que la aplicación del sistema tenga propósitos positivos y no se utilice en contextos que comprometan la privacidad o la dignidad de las personas. El enfoque está en el apoyo a tareas como la actividad física o la salud, no en la vigilancia intrusiva.