

# **Sistema de Anotación de Video y Seguimiento de Articulaciones**

**Facultad de Ingeniería, Diseño y Ciencias Aplicadas**  
**Departamento de Computación y Sistemas Inteligentes**  
**Ingeniería de Sistemas**

**Docentes:** Uram Sosa Aguirre, Milton Sarria Paja

**Asignatura:** Inteligencia Artificial I

**Institución:** Universidad Icesi

**Integrantes:**

Samuel Alvarez, Gerson de Jesus Hurtado Borja, Rafaela Sofia Ruiz Pizarro

**Día de Entrega:**

4/05/2025

<b>Caso de Estudio Propuesto</b>	<b>3</b>
Pregunta(s) de Interés	3
Tipo de Problema	4
<b>Metodologías</b>	<b>4</b>
Preparación de los Datos:	5
Modelado:	5
Evaluación:	5
Despliegue:	5
<b>Cronograma de Trabajo</b>	<b>5</b>
<b>Métricas para Medir el Progreso</b>	<b>6</b>
• Precisión, Recall, F1-Score:	6
• Error medio absoluto (MAE) en estimación de ángulos:	6
• Tasa de falsos positivos y falsos negativos:	6
• Tiempo de inferencia (latencia):	6
• Evaluación subjetiva por usuarios:	6
<b>Siguientes Pasos</b>	<b>7</b>
<b>Estrategias para Ampliar el Conjunto de Datos</b>	<b>8</b>
<b>Análisis de Aspectos Éticos</b>	<b>9</b>
1. Consentimiento y privacidad	9
2. Seguridad de los datos	9
3. Evitar sesgos algorítmicos	9
4. Transparencia y explicabilidad	9
<b>Justificación de la Viabilidad del Proyecto</b>	<b>10</b>

## Caso de Estudio Propuesto

Este proyecto busca desarrollar un sistema inteligente capaz de reconocer y clasificar actividades humanas específicas (caminar hacia la cámara, caminar de regreso, girar, sentarse y ponerse de pie) mediante el análisis de video en tiempo real. El sistema realizará un seguimiento de puntos articulares clave (muñecas, rodillas, caderas, tobillos, hombros, cabeza) para evaluar tanto la actividad como aspectos posturales como inclinación del tronco.

**Entradas:** Video en tiempo real capturado por cámara.

**Salidas:** Clasificación de la actividad realizada y análisis de inclinaciones laterales y movimientos articulares clave (muñecas, rodillas, caderas, etc.).

## Pregunta(s) de Interés

- **¿Puede un sistema de IA identificar en tiempo real actividades humanas específicas a partir del análisis de video?**

Esta pregunta busca validar la viabilidad técnica del proyecto, considerando los retos computacionales y algorítmicos que implica el análisis en tiempo real. Identificar actividades humanas en video requiere modelos robustos, capaces de adaptarse a diferentes condiciones de iluminación, perspectiva y velocidad de movimiento.

- **¿Qué tan efectivos son los modelos de clasificación supervisada para detectar movimientos articulares y posturales a partir de datos articulares?**

El uso de algoritmos como Support Vector Machines (SVM), Random Forest y XGBoost ha demostrado ser eficaz en tareas de clasificación. Sin embargo, en este caso el desafío está en trabajar con datos temporales (secuencias de coordenadas articulares) que requieren técnicas de ingeniería de características específicas para capturar dinámicas del movimiento.

- **¿Cuál es la mejor combinación de características biomecánicas para clasificar con precisión dichas actividades?**

Esta pregunta se enfoca en la fase de ingeniería de características. Los ángulos relativos entre articulaciones, las velocidades angulares y la inclinación del tronco son hipótesis iniciales, pero es necesario validar empíricamente cuáles de estas variables aportan mayor poder discriminativo.

para la clasificación.

- **¿Cómo afecta la variabilidad interindividual (tamaño corporal, estilo de caminar, edad) al rendimiento del modelo?**

La robustez del sistema ante diferentes tipos de usuarios es clave para su aplicabilidad general. Por eso es importante analizar el rendimiento del modelo con múltiples participantes y condiciones.

## Tipo de Problema

Este proyecto se sitúa en el campo de la Inteligencia Artificial aplicada a la visión por computador, concretamente en el subcampo de Reconocimiento de Actividades Humanas (HAR). Desde un punto de vista técnico, se trata de un problema de clasificación multiclase supervisada, cuya entrada es una secuencia temporal de posiciones articulares y cuya salida es una etiqueta que representa la actividad realizada.

El sistema combina técnicas de seguimiento de poses humanas (pose estimation), procesamiento de señales temporales (filtrado, normalización), extracción de características biomecánicas (ángulos, inclinaciones, velocidades) y clasificación. La dificultad adicional radica en que el sistema debe operar en tiempo real, lo cual impone restricciones de eficiencia computacional y latencia de inferencia.

En cuanto a la naturaleza de los datos, se trata de datos espaciales y temporales (secuencias de puntos en el espacio a lo largo del tiempo), lo cual requiere una adecuada gestión del ruido, regularización y posiblemente estrategias de segmentación para delimitar claramente el inicio y fin de cada actividad.

## Metodologías

Se emplea metodologías tipo **CRISP-DM**, adaptada a las necesidades del proyecto:

Comprensión del Negocio y del Problema: Análisis del contexto de aplicación y definición de los objetivos del sistema.

**Comprensión de los datos:** Grabación y anotación de videos con diferentes actividades y participantes.

**Preparación de los Datos:**

- Seguimiento de articulaciones usando MediaPipe o OpenPose.
- Normalización de coordenadas.
- Aplicación de filtros suaves.
- Extracción de características relevantes (velocidades, ángulos relativos, inclinación del tronco).

**Modelado:**

- Pruebas con modelos supervisados como SVM, Random Forest o XGBoost.
- Ajuste de hiper parámetros.

**Evaluación:**

- Validación cruzada y cálculo de métricas como precisión, recall y F1-Score.

**Despliegue:**

- Implementación de una interfaz gráfica que visualice en tiempo real las actividades detectadas y las medidas posturales.

**Cronograma de Trabajo**

Fase	Fecha estimada	Actividad Principal
Recolección de datos	5 - 10 mayo 2025	Grabar videos con participantes, organización de datos
Preprocesamiento de datos	11 - 15 mayo 2025	Seguimiento de articulaciones, normalización, filtrado
Ingeniería de características	16 - 20 mayo 2025	Cálculo de ángulos, inclinaciones y velocidades
Entrenamiento de modelos	21 - 25 mayo 2025	Entrenamiento y validación de SVM, Random Forest y XGBoost

Fase	Fecha estimada	Actividad Principal
Evaluación del sistema	26 - 28 mayo 2025	Métricas, matriz de confusión, análisis de errores, test con nuevos usuarios
Despliegue	29 mayo - 1 junio 2025	Interfaz gráfica + visualización en tiempo real
Documentación final y entrega	2 - 3 junio 2025	Revisión del informe, formato, anexos

## Métricas para Medir el Progreso

- **Precisión, Recall, F1-Score:**

Métricas estándar en clasificación multiclase. F1-score permite un equilibrio entre precisión y sensibilidad.

- **Error medio absoluto (MAE) en estimación de ángulos:**

Evalúa qué tan cerca están las medidas articulares calculadas del valor esperado (en casos donde hay etiquetas de referencia).

- **Tasa de falsos positivos y falsos negativos:**

Clave en entornos sensibles como salud, donde una mala clasificación puede tener implicaciones prácticas importantes.

- **Tiempo de inferencia (latencia):**

La velocidad del sistema debe ser menor a 100ms por frame para considerarse en tiempo real.

- **Evaluación subjetiva por usuarios:**

Encuestas o entrevistas a usuarios que interactúen con la interfaz para validar la usabilidad y claridad del sistema.

## **Siguientes Pasos**

### **Ampliación del dataset**

- Continuar la recolección de videos con al menos 10–15 participantes, incluyendo diversidad de género, edad y complejidad física.
- Grabar en condiciones variadas: distintos ángulos de cámara, iluminación, y fondos.

### **Extracción y análisis de características**

- Calcular ángulos entre segmentos corporales (por ejemplo, muslo-pierna, brazo-tronco).
- Medir inclinación del tronco y su variación temporal.
- Incluir velocidades angulares y aceleraciones.

### **Entrenamiento de modelos**

- Probar clasificadores como SVM, Random Forest y XGBoost con validación cruzada.
- Evaluar métricas por clase: accuracy, F1-score, matriz de confusión.

### **Desarrollo de prototipo funcional**

- Implementar una interfaz gráfica simple que muestre video en vivo, pose detectada y actividad reconocida en tiempo real.
- Medir tiempo de inferencia y optimizar la eficiencia.

## Estrategias para Ampliar el Conjunto de Datos

Dado que el tamaño del dataset impacta directamente la capacidad del modelo para generalizar, se proponen las siguientes estrategias para incrementar la cantidad y diversidad de los datos:

- **Recolección manual de más videos**

- Invitar voluntarios con distintas características físicas (edad, género, complexión).

Grabar actividades en interiores y exteriores.

- **Uso de datasets públicos complementarios**

Incorporar secuencias de datasets como:

- **NTU RGB+D Dataset** (acciones humanas con datos 3D).
- **Kinetics Pose** (acciones extraídas de YouTube).
- **Human3.6M** (movimientos humanos con anotaciones de articulaciones).

Estos pueden usarse para pre-entrenamiento o como conjunto de validación.

- **Síntesis de datos mediante simulación**

Usar software como Blender para generar animaciones realistas con poses humanas controladas.

Generar secuencias de movimiento artificiales con condiciones variadas.

- **Aumento de datos**



Realizar transformaciones a los datos ya existentes (rotaciones leves, escalados, inversión horizontal) para simular nuevos movimientos.

## **Análisis de Aspectos Éticos**

### **1. Consentimiento y privacidad**

- Toda persona grabada debe otorgar su consentimiento informado de manera explícita.
- Se debe garantizar la anonimización de datos visuales, evitando mostrar rostros si no es necesario.
- No se debe utilizar el sistema para vigilancia o monitoreo sin autorización expresa.

### **2. Seguridad de los datos**

- Los videos y datos procesados deben almacenarse con protocolos de seguridad adecuados (encriptación, control de acceso).
- Cumplimiento de la legislación sobre protección de datos personales, como la Ley Habeas Data o GDPR, si aplica.

### **3. Evitar sesgos algorítmicos**

- Si los datos provienen de un grupo homogéneo, el modelo puede aprender sesgos (por ejemplo, mal rendimiento con personas mayores o de distintas etnias).
- Es necesario garantizar diversidad demográfica en el conjunto de entrenamiento.

### **4. Transparencia y explicabilidad**

- El sistema debe explicar, al menos de forma visual o textual, cómo se llegó a una clasificación.
- Se deben evitar los “modelos caja negra” cuando se utilicen para toma de decisiones con impacto real.

## **Justificación de la Viabilidad del Proyecto**

Este proyecto es viable dentro del marco del curso debido al uso de herramientas y recursos accesibles, tanto tecnológicos como humanos. Se emplearán bibliotecas de código abierto como MediaPipe y OpenPose para el seguimiento de articulaciones, así como scikit-learn y XGBoost para el modelado. Además, se complementará la recolección de datos propia con el uso de datasets públicos como NTU RGB+D y Kinetics Pose, lo cual reduce el tiempo y esfuerzo requerido en la etapa de adquisición de datos. El equipo cuenta con conocimientos previos en programación, aprendizaje automático y procesamiento de datos, y se ha distribuido el trabajo de forma colaborativa para garantizar un avance constante. El cronograma propuesto permite alcanzar un prototipo funcional antes del cierre del semestre, incluyendo recolección de datos, modelado y evaluación, por lo que se considera que el proyecto es técnica y logísticamente factible.