

# Prueba de progreso 2: Búsqueda de patrones en el movimiento humano. Clustering y procesamiento de señales.

2025

Usando un conjunto de datos con grabaciones de las posiciones de las principales articulaciones de un cuerpo humano, encuentra una agrupación de diferentes tipos de movimientos que pueden hacerse en 4 segundos. Puedes utilizar cualquier técnica de clustering, pero deberás justificar su elección. Este trabajo requiere una considerable labor de limpieza y preparación de datos.

## Datos

Debe usarse [este conjunto de datos sobre marcha humana](#), que contiene múltiples grabaciones del movimiento para cada una de un total de 26 personas andando siguiendo diferentes rutas. Hay una carpeta para cada persona y un fichero para cada ruta que esa persona ha seguido. Los ficheros, en formato CSV, almacenan la posición de 23 de las articulaciones de cada persona y la velocidad instantánea en cada fotograma. Para cada una de estas articulaciones (y también para la velocidad), se tienen 3 componentes: la X, la Y y la Z. Estas componentes representan cada una de las coordenadas espaciales orientadas tal y como muestra la Figura 1. Dichas grabaciones están realizadas a una frecuencia de 60 hercios.

## Normas y estructura

- Esta tarea se hará **individualmente**.
- El código deberá estar alojado en un **repositorio git**, ya sea GitHub, GitLab, Codeberg o cualquier otro.
  - Antes de empezar la tarea cada grupo deberá compartir acceso al repositorio donde esté alojado el código al profesor, para que este pueda hacer seguimiento de los cambios y avances en el trabajo.

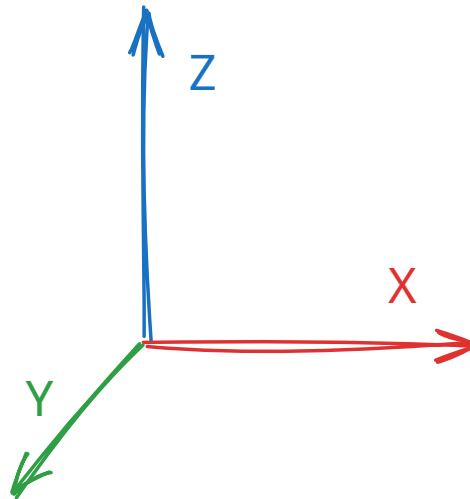


Figura 1: Dirección de los ejes que representan las coordenadas 3D del conjunto de datos.

- El **envío** consistirá en uno (preferible) o varios cuadernos de **Jupyter Notebook bien documentados**, valorando particularmente el uso combinado de código Python y Markdown.

### Procedimiento

En esta sección se explicitan los pasos a dar necesarios para llevar a cabo la prueba satisfactoriamente. Esto no significa que no puedan darse otros pasos similares en sustitución o complementándolos si crees que son adecuados. El orden de los mismos tampoco tiene por qué ser estrictos.

- **Carga de datos**
  - En este caso, los datos están distribuidos entre varios ficheros, así que será necesario combinarlos de manera apropiada.
- **Eliminar la correlación entre articulaciones**
  - Como todos los puntos se mueven en la dirección de la persona, la posición general de estos afecta mucho más que el movimiento relativo de cada uno con respecto al resto.
  - Para solucionar esto, será necesario restar la posición de la pelvis (que es buena representante de la posición general, al encontrarse en una posición central) al resto de articulaciones para que lo que quede sea la variabilidad propia de cada una.
- **Ventaneo**

- Será necesario segmentar el conjunto total de datos en ventanas de 4 segundos, tal y como pide el enunciado.

- **Normalización**

- Por un lado será importante hacer una normalización de los puntos de cada persona para que la altura y otras dimensiones de cada persona no afecten en las agrupaciones en función del movimiento.
- Por otro lado, será importante normalizar las rutas en cada ventana para asegurarse de que todas empiezan desde la posición  $x=0$ ,  $y=0$  y  $z=0$ . Hay que tener en cuenta que la coordenada de la pelvis es ahora la única que tenemos de referencia para la posición, ya que todas las demás deberían haber sido extraídas en un paso anterior.

- **Extracción de características**

- Por cada articulación y cada ventana será necesario extraer características que se vean afectadas por el movimiento. A continuación os propongo algunas ideas, pero os invito a buscar alguna otra.
  - \* Media
  - \* Desviación típica
  - \* Mínimo y máximo
  - \* Energía de la señal
  - \* Parámetros de Hjorth

- **Elección de algoritmo**

- Selecciona un algoritmo de clustering para llevar a cabo la agrupación.

La elección del algoritmo debe ser justificada.

- **Interpretación de resultados**

- Una vez generados los grupos, es importante analizar qué representa cada uno. Algunas técnicas para poder hacerlo son:
  - \* Resúmenes (medias, desviaciones, etc.)
  - \* Visualización del movimiento con gráficas o incluso animaciones.
- Esta es la etapa más fundamental de una tarea de clustering, así que se valorará especialmente.

## **Evaluación**

La valoración final de la prueba está formada por:

- Un **15%** perteneciente a la **segunda prueba de progreso**.
- Un **8%** perteneciente a las **prácticas de laboratorio**.

En total hace un **23%** del total de la asignatura que podrá ser recuperado si hace falta.

## **Envío**

Deberá entregarse en el campus virtual un enlace a la carpeta del repositorio donde se encuentre el **cuaderno de Jupyter** (fichero .ipynb).

La fecha límite para el envío será el **4 de mayo de 2025**.