

# Desmitificando las 189 Características (Features) del Modelo HAR

•

Este documento explica cómo el modelo de Inteligencia Artificial "ve" el movimiento humano. A diferencia de nosotros, que vemos un video continuo, el modelo recibe **189 números** cada vez que intenta adivinar qué estás haciendo.

Estos números provienen de tomar **2.56 segundos** de datos de los sensores (una "ventana") y resumirlos matemáticamente.

## Mapa del Tesoro: ¿Qué es la posición #180?

Antes de entender los ejemplos, debemos saber qué significa cada posición en el array de 189 números que recibe el modelo.

El sistema tiene **27 canales de sensores** (Acelerómetros, Giroscopios, Magnetómetros en pecho, tobillo y brazo, más magnitudes calculadas).

Para cada uno de estos 27 canales, calculamos 7 estadísticas. El orden es secuencial:

Rango de Índices	Estadística	¿Qué representa?
0 - 26	Promedio (Mean)	La orientación promedio del cuerpo.
27 - 53	Desviación Estándar (Std)	Qué tanto varía el movimiento (intensidad).
54 - 80	Mínimos	El pico más bajo de la señal.
81 - 107	Máximos	El pico más alto de la señal.
108 - 134	Mediana	El valor central (ignora picos raros).
135 - 161	Skewness	La asimetría de la señal.
162 - 188	Energía (FFT)	La fuerza o periodicidad del movimiento.

### Localizando el Ejemplo: Posición #180

Si miramos el último bloque (Energía, empieza en 162), la posición **#180** corresponde al índice 18 de ese bloque. En nuestra lista de sensores, el canal 18 es el **Giroscopio del Brazo (Eje Y)**.

Feature #180 = Energía del Giro en el Brazo (Y)

## ¿Por qué usamos estas matemáticas? (Con Ejemplos)

El modelo usa estas estadísticas para descartar opciones hasta quedarse con la correcta. Veamos cómo piensa:

## 1. El Promedio (Mean): "La Brújula de la Gravedad"

El promedio nos dice hacia dónde apunta el sensor la mayor parte del tiempo, aprovechando que la gravedad de la Tierra siempre tira hacia abajo.

- **Ejemplo (Acelerómetro Pecho X):**

- **De Pie (Standing):** La gravedad afecta al eje Y o Z, pero el X (lateral) debería ser cercano a **0**.
- **Acostado de lado (Lying):** Ahora el eje X recibe toda la gravedad. El promedio salta a **~9.8**.
- *El modelo piensa:* "Si el promedio en X es alto, seguro está acostado".

## 2. Desviación Estándar (Std): "El Detector de Agitación"

Mide qué tanto se alejan los datos del promedio. Es decir, ¿estás quieto o te mueves como loco?

- **Ejemplo (Acelerómetro Tobillo):**

- **Sentado (Sitting):** Tu pie está quieto. La desviación es **casi 0**.
- **Corriendo (Running):** Tu pie golpea el suelo y sube violentamente. La desviación es **gigante (> 15)**.
- *El modelo piensa:* "Std bajita = Sedentario. Std alta = Ejercicio".

## 3. Máximos y Mínimos: "Los Límites del Golpe"

Detectan impactos fuertes o movimientos bruscos en una dirección.

- **Ejemplo (Saltar / Jump front & back):**

- Al saltar y caer, hay un impacto seco. El **Máximo** del acelerómetro vertical se disparará momentáneamente a valores extremos (ej. **20 m/s<sup>2</sup>**), aunque el promedio sea normal.
- *El modelo piensa:* "Promedio normal pero Máximo explosivo = Saltos".

## 4. Energía (FFT): "El Ritmo del Movimiento"

Aquí está nuestro ejemplo estrella. La energía se calcula usando la Transformada de Fourier, que mide qué tan "repetitiva" y fuerte es una señal.

### 💡 El Caso de la Posición #180 (Energía Giroscopio Brazo Y)

Imagina dos actividades: **Ciclismo** vs **Trotar**.

#### 1. Ciclismo:

- Tus manos están fijas en el manubrio. Aunque tus piernas se mueven rápido, tus brazos están relativamente estables.
- **Dato #180:** Será un valor **BAJO** (ej. 50 - 200).

#### 2. Trotar (Jogging):

- Tus brazos se balancean rítmicamente adelante y atrás para impulsarte.
- Ese balanceo constante genera una señal sinusoidal perfecta en el giroscopio.
- **Dato #180:** Será un valor **ALTO** (ej. 2000 - 5000).

### La Deducción del Modelo:

"Veo que la Energía en las piernas es alta en ambos casos (ambos cansan). PERO, miro el **Dato #180**.

- ¿Es **bajo**? -> Los brazos están quietos -> Debe ser **Ciclismo**.
- ¿Es **alto**? -> Los brazos se balancean -> Debe ser **Trotar**."

## Resumen de Rangos Típicos

Para darte una idea de los valores que ve el modelo:

Actividad	Promedio (Ejes Verticales)	Desviación Estándar	Energía (Pos #180 - Brazo)
Dormir / Quieto	~9.8 (Gravedad constante)	< 0.5 (Casi nula)	<b>0 - 10</b> (Inexistente)
Caminar	~9.8 (Con ruido)	2.0 - 5.0 (Moderada)	<b>500 - 1500</b> (Balanceo suave)
Correr	Variable	> 10.0 (Caótica)	<b>&gt; 3000</b> (Balanceo intenso)
Ciclismo	Estable	Alta en piernas, Baja en brazos	<b>Baja</b> (Manos en manubrio)

El modelo no "adivina". Resuelve un sistema de 189 desigualdades matemáticas simultáneamente para encontrar la única actividad que encaja con todas las pistas.

## ¿Qué son las "Ventanas" y los 2.56 segundos?

Para entender el movimiento, no podemos mirar un solo instante (como una foto); necesitamos ver una secuencia (como un clip de video).

- Muestreo (50 Hz):**
  - Los sensores toman una medición **50 veces por segundo**.
- La Ventana (Window):**
  - El modelo agrupa **128 de estas mediciones consecutivas** en un bloque para analizarlo.
  - A este bloque le llamamos **"Ventana"**.
- El Cálculo de 2.56s:**
  - Si tomamos 128 muestras a una velocidad de 50 por segundo:
  - **128 muestras / 50 muestras/segundo = 2.56 segundos**.
  - Cada vez que el modelo hace una predicción, está juzgando lo que hiciste en esos exactos 2.56 segundos.

## Sobre el Solapamiento (Overlap)

Para no perder detalles entre una ventana y otra, usamos un solapamiento del 50% (64 muestras). Esto significa que la siguiente ventana comienza a la mitad de la anterior.


- **Ventana 1:** Segundos 0.00 a 2.56
- **Ventana 2:** Segundos 1.28 a 3.84

- Esto nos permite detectar actividades que ocurren justo en el corte de los bloques.

## Lista Completa de Índices y Bloques

El vector de **189 características (features)** se construye procesando **27 señales base**. Para cada señal, calculamos 7 estadísticas en el siguiente orden de bloques:

1. **Mean (Promedios):** Índices 0-26
2. **Std (Desviación):** Índices 27-53
3. **Min (Mínimos):** Índices 54-80
4. **Max (Máximos):** Índices 81-107
5. **Median (Mediana):** Índices 108-134
6. **Skewness (Asimetría):** Índices 135-161
7. **Energy (Energía):** Índices 162-188

 Las 27 Señales Base (y su orden dentro de cada bloque)

Dentro de cada uno de los 7 bloques anteriores, los datos siguen este estricto orden de sensores:

Índice Relativo	Sensor / Señal	Ubicación
+0	Aceleración X	Pecho
+1	Aceleración Y	Pecho
+2	Aceleración Z	Pecho
+3	ECG 1	Pecho
+4	ECG 2	Pecho
+5	Aceleración X	Tobillo
+6	Aceleración Y	Tobillo
+7	Aceleración Z	Tobillo
+8	Giroscopio X	Tobillo
+9	Giroscopio Y	Tobillo
+10	Giroscopio Z	Tobillo
+11	Magnetómetro X	Tobillo
+12	Magnetómetro Y	Tobillo
+13	Magnetómetro Z	Tobillo
+14	Aceleración X	Brazo
+15	Aceleración Y	Brazo
+16	Aceleración Z	Brazo
+17	Giroscopio X	Brazo

Índice Relativo	Sensor / Señal	Ubicación
+18	Giroscopio Y	Brazo
+19	Giroscopio Z	Brazo
+20	Magnetómetro X	Brazo
+21	Magnetómetro Y	Brazo
+22	Magnetómetro Z	Brazo
+23	Magnitud Acc.	(Calculado Tobillo)
+24	Magnitud Giro.	(Calculado Tobillo)
+25	Magnitud Acc.	(Calculado Brazo)
+26	Magnitud Giro.	(Calculado Brazo)

### Cómo calcular un índice específico

Si quieres encontrar, por ejemplo, el **Máximo del Magnetómetro X del Brazo**:

1. Busca el bloque **Max**: Empieza en el índice **81**.
2. Busca la señal **Magnetómetro X (Brazo)**: Es el relativo **+20**.
3. Suma:  $81 + 20 = 101$ .
4. Esa feature está en la posición **#101**.

## ¿Por qué funciona esta suma? (La Lógica de la Posición)

La razón es simple: **Concatenación (Pegar listas una tras otra)**.

El modelo no entiende de estructuras complejas, solo ve una **única lista larga** de números. Para construirla, el código toma los bloques de estadísticas y los pone en fila india.

Imagina un tren con **7 vagones**:

- Cada vagón es una estadística (Promedio, Desviación, etc.).
  - Cada vagón tiene exactamente **27 asientos** (uno para cada sensor).
1. **Vagón 1 (Promedios)**: Ocupa los asientos del **0 al 26**.
  2. **Vagón 2 (Desviación)**: Se conecta justo detrás. Su primer asiento no es el 0, es el **27** (porque ya pasaste los 27 asientos del primer vagón).
  3. **Vagón 3 (Mínimos)**: Se conecta detrás. Empieza en el **54** (27 del primero + 27 del segundo).
    - ... Y así sucesivamente.

**La Fórmula:**

$$\text{Posición Final} = (\text{Inicio del Vagón}) + (\text{Asiento del Sensor})$$

- **"Inicio del Vagón"**: Es el *Índice Base* del bloque (por ejemplo, el bloque **Max** empieza en el 81 porque hay  $27 * 3 = 81$  números antes de él).

- **"Asiento del Sensor":** Es el *Índice Relativo*. El sensor "Magnetómetro X Brazo" siempre se sienta en el asiento número **20** de *cualquier* vagón.

Por eso, la suma determina la posición única: **81 (Saltar los vagones anteriores) + 20 (Caminar hasta el asiento del sensor) = Posición #101**. Posición #101\*\*. El orden es estricto y nunca cambia.

? ¿Por qué se repiten los números (+0 al +26)?

Te estarás preguntando por qué tenemos el mismo índice **+0** o **+5** siete veces diferentes. La respuesta es que **son los mismos sensores**, pero vistos con lentes diferentes.

Imagina un examen médico a **27 pacientes** (los sensores):

1. **Ronda 1 (Promedio):** Les tomas la temperatura a los 27 pacientes (en orden del 0 al 26).
2. **Ronda 2 (Máximo):** Les tomas la presión a los **mismos 27 pacientes** (en el mismo orden del 0 al 26).
3. **Ronda 3 (Energía):** Les mides la altura a los **mismos 27 pacientes**.

Los pacientes (sensores) son los mismos y están formados en el mismo orden. Lo único que cambia es **qué les estás midiendo** (la estadística). Por eso el índice relativo (+0 a +26) se repite en cada bloque.



## Detalles Técnicos "Invisibles" (Pero Vitales)

Para cerrar, hay dos cosas que ocurren "bajo el capó" y son cruciales para que esto funcione:

### 1. La Normalización (Scaler)

Aunque tú veas un **9.8** en el acelerómetro y un **5000** en la energía, el modelo recibiría un shock si ve números tan diferentes. Antes de entrar al modelo, **todos los números se transforman (Estandarización Z-Score)**.

- Al 9.8 le restamos el promedio histórico y lo dividimos por la desviación.
- Quizás el 9.8 se convierte en un **0.1** y el 5000 se convierte en un **1.2**.
- **¿Por qué?** Para que la Energía (que es enorme) no eclipse a la Aceleración (que es pequeña) a la hora de votar. Es como convertir peras y manzanas a "puntos" para poder sumarlos.

### 2. La Clase 0 (El "Limbo")

En los datos originales, verás muchas filas con la etiqueta **label: 0**.

- Estas filas representan momentos donde el sujeto no estaba haciendo ninguna de las 12 actividades definidas (ej. esperando instrucciones, bebiendo agua, ajustándose el sensor).
- **Importante:** El sistema **elimina** automáticamente todo lo que sea 0. No entrenamos con "basura" ni intentamos predecirla. Si subes un archivo lleno de ceros, el sistema te dirá que no encontró ventanas válidas.

### 3. Las Unidades Reales

Para que los ejemplos de "Máximos" o "Promedios" tengan sentido, recuerda las unidades físicas:

- **Acelerómetros:** Metros por segundo cuadrado ( $m/s^2$ ). (Gravedad  $\approx 9.8$ ).

- **Giroscopios:** Radianes por segundo ( $\text{rad/s}$ ) o Grados por segundo ( $\text{deg/s}$ ) según configuración. Miden velocidad de giro.
- **Magnetómetros:** Micro-Teslas ( $\mu\text{T}$ ). Miden el campo magnético (como una brújula 3D).