

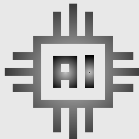
Modelo de IA

RETO - 1

Andrés Alejandro Guzmán González -----	A01633819
Ernesto Reynoso Lizárraga -----	A01639915
Joel Isaías Solano Ocampo -----	A01639289
Luis Rodolfo Bojorquez Pineda -----	A01250513
Tania Sayuri Guizado Hernandez -----	A01640092

Introducción

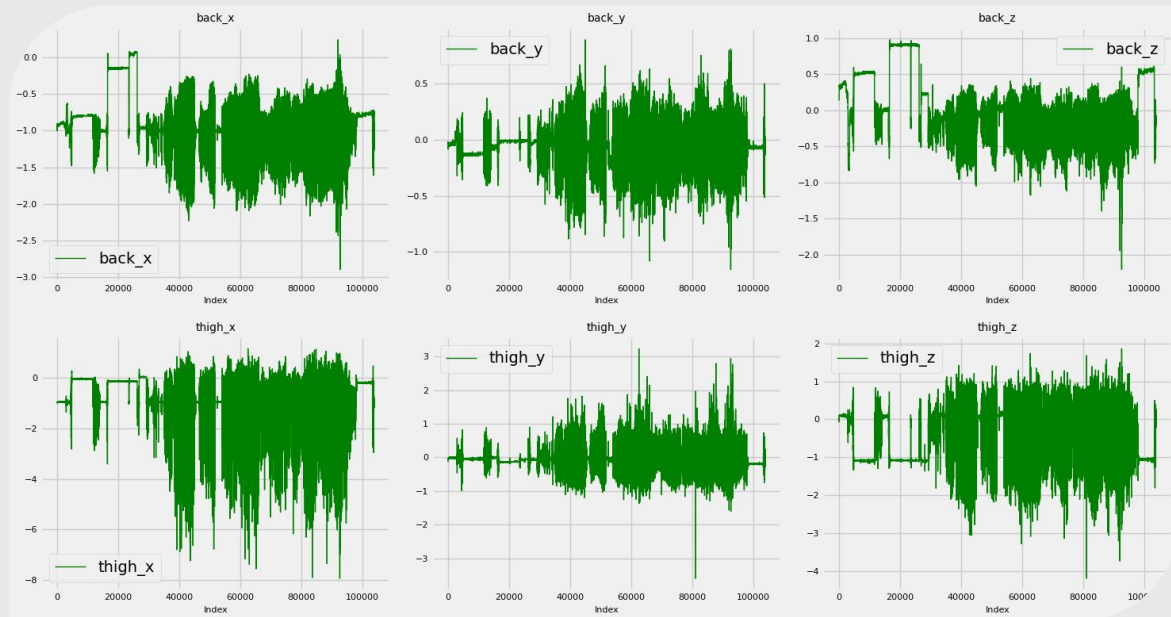
Data Set: Activity Recognition in Senior Citizens

Variables	Categorías	Objetivo
<p>timestamp: date and time of sample</p> <p>back_x, back_y, back_z: acceleration in x,y or z-direction</p> <p>Thigh_x, thigh_y, thigh_z: acceleration in x,y or z-direction</p> <p>label: annotated activity code</p>	<p>1: walking</p> <p>3: shuffling</p> <p>4: stairs (ascending)</p> <p>5: stairs (descending)</p> <p>6: standing</p> <p>7: sitting</p> <p>8: lying</p>	<p>Generar un modelo</p>  <p>que prediga las acciones de acuerdo con un conjunto de datos dado.</p>

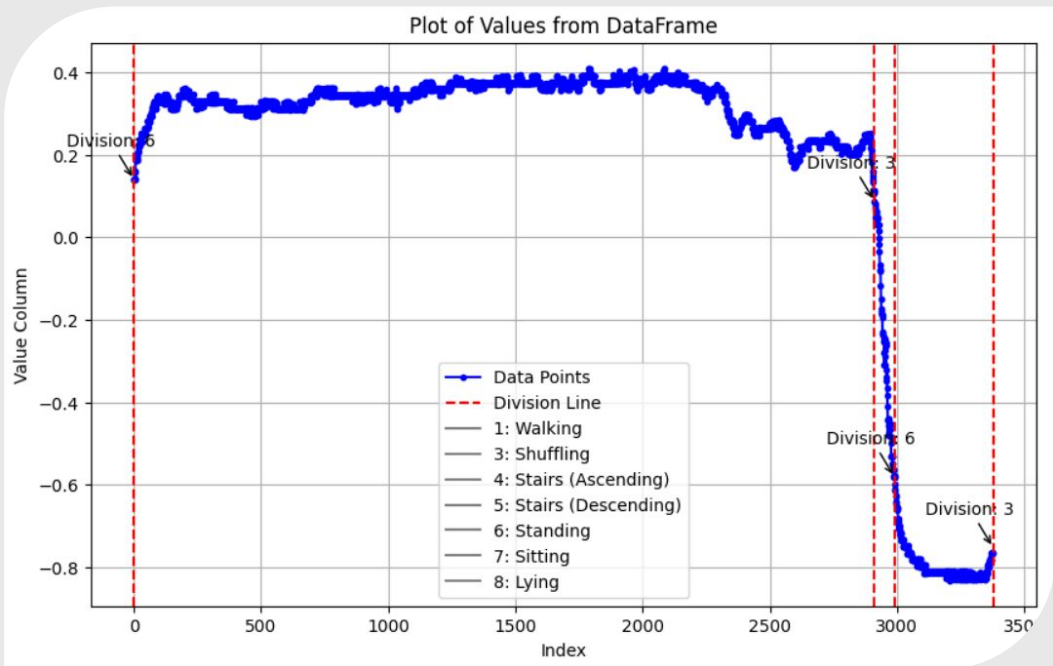
Análisis Exploratorio

1. Identificar cómo se comportan los datos.
2. Definir qué estrategia a seguir para la construcción del modelo.

¿Por dónde empezar?,
¿Qué nos decían los datos?,
¿Cómo los podríamos interpretar?



Razones de cambio.



Pusimos fronteras en los datos de acuerdo a sus labels

¿Qué identificamos?

- **Modelo categórico.**
- **Valores nulos para ciertas categorías en la primer muestra.**
- **Rango de valores de aceleración por acción.**

¿Cómo preparamos los datos?

DATA SET

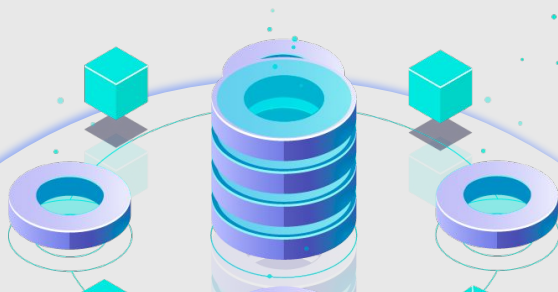
Fusionamos todos los datos (15 archivos) en un solo data frame.

CLASES

Aseguramos que todas las clases tengan muestras.

MODELOS

Revisamos los tipos de modelos para probar la muestra de datos.



Pruebas de modelos



01

RANDOM FOREST

Accuracy: 0.94 F1 score



03

LOGISTIC REGRESSION

Accuracy: 0.76 F1 score



02

SVC LINEAR

Accuracy: 0.00 F1 score



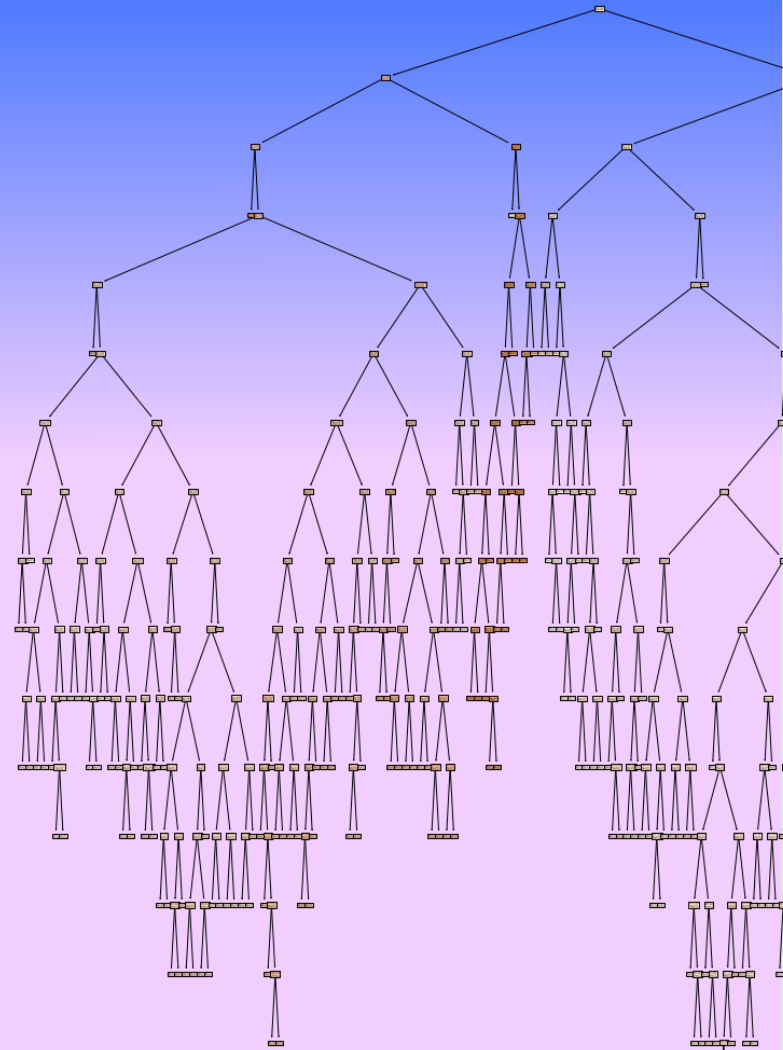
04

LINEAR REGRESSION

Accuracy: 0.00 F1 score

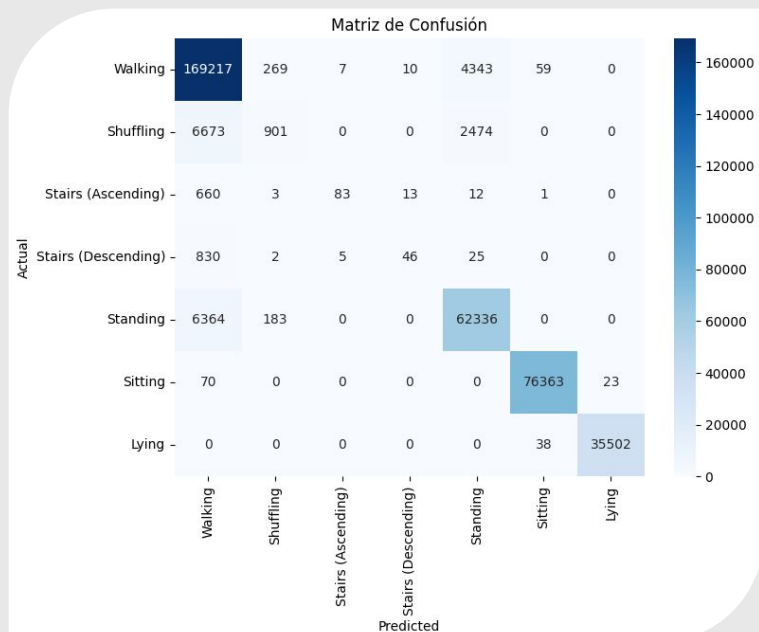
Random Forest

- Capacidad de manejar múltiples características.
- Manejo de relaciones no lineales.
- Es menos propenso a sobre ajuste .
- Manejo de datos desequilibrados.
- Combina múltiples árboles



Modelos

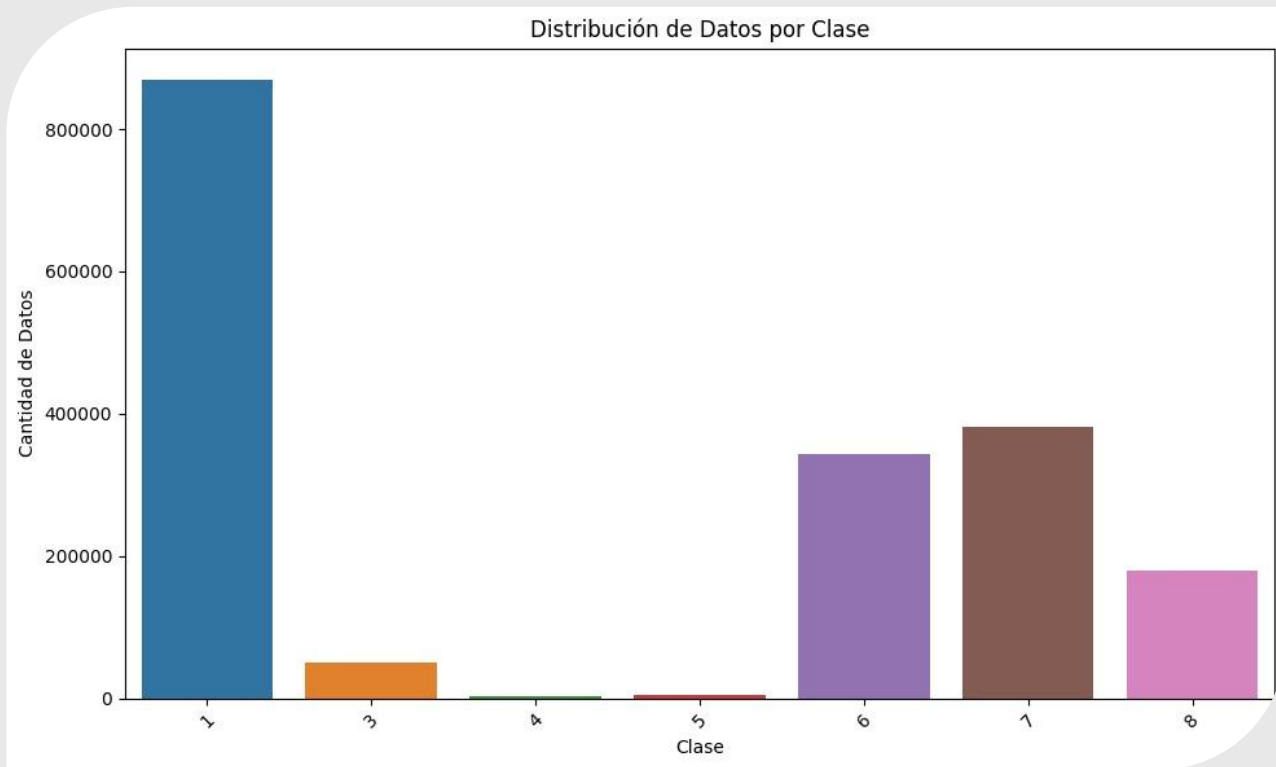
Datos desbalanceados **94%**



Datos Balanceados **92%**



¿Por qué estos resultados?



¿Cómo mejoramos el modelo?

Evaluamos los rangos de aceleración de las clases y con ello:

- Unificamos las clases:
 - **1: walking y 3: shuffling**
 - **4: stairs (ascending) y 5: stairs (descending)**

Accuracy: 98%



The background is a light gray with several large, organic, blue and white shapes that resemble liquid or smoke. A large, central, blue speech bubble-like shape contains the text. The text is in a bold, black, sans-serif font.

**¿Qué se puede
mejorar?**

Gracias

Andrés Alejandro Guzmán González ----- A01633819
Ernesto Reynoso Lizárraga ----- A01639915
Joel Isaias Solano Ocampo ----- A01639289
Luis Rodolfo Bojorquez Pineda ----- A01250513
Tania Sayuri Guizado Hernandez ----- A01640092

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik**



Referencias

Logacjov,Aleksej and Ustad,Astrid. (2023). HAR70+. UCI Machine Learning Repository.
<https://doi.org/10.24432/C5CW3D>.

Logacjov,Aleksej and Ustad,Astrid. (2023). Activity Recognition in Senior Citizens. Kaggle:
<https://www.kaggle.com/datasets/anshtanwar/adult-subjects-70-95-years-activity-recognition>

Geron, A. (2019). Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). An Introduction to Statistical Learning: With Applications in R.

