



NOMBRE y DNI:

**PROBLEMA 1:**

Se propone utilizar un algoritmo de agrupamiento Fuzzy-C-means para un conjunto de datos compuestos de 1000 casos que poseen 6 variables de entrada. El objetivo es agrupar pacientes según su riesgo cardíaco, alto y bajo. Las variables, sus unidades y sus rangos son: Colesterol (mg/dl) [0:300], Edad [0:100], Fumador [0:1], Días que realiza actividad física por semana [0:7], Glucemia (mg/dl) [0:200] y Nivel de estrés [0:100%].

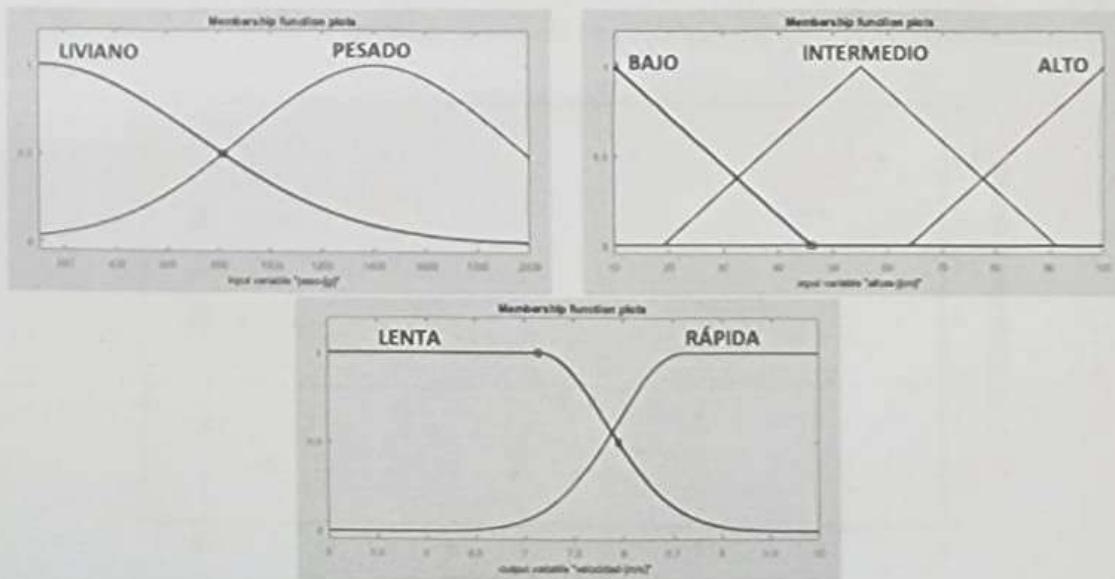
- ¿Cómo influye la normalización de los datos para asegurar el adecuado funcionamiento del FCM?
- Proponga valores de los hiperparámetros necesarios en la implementación del algoritmo.
- ¿Cuál es la expresión de la función objetivo a minimizar en este problema? Explíquela.
- Indique la dimensión de los centros de cluster obtenidos una vez cumplido el criterio de detención.
- Indique la dimensión de la Matriz de Pertenencia.
- ¿Cómo utiliza el agrupamiento obtenido si quiere asignar un grupo a un dato nuevo, no incluido en los datos de entrada al FCM?
- ¿Cómo implementaría un Algoritmo Genético (AG) para optimizar la cantidad de grupos a encontrar? Proponga hiperparámetros necesarios (función de fitness, cromosoma, criterios de selección, mutación y detención).

**PROBLEMA 2:**

Dado un objeto que tiene una altura de 70 cm y un peso de 1200 g, mostrar la forma del conjunto difuso de la salida del siguiente sistema de inferencia difusa con una sola regla, antes de la defuzzificación. Utilizar los operadores [producto] para la conjunción, [suma probabilística] para la disyunción y el operador [mínimo] para la implicación.

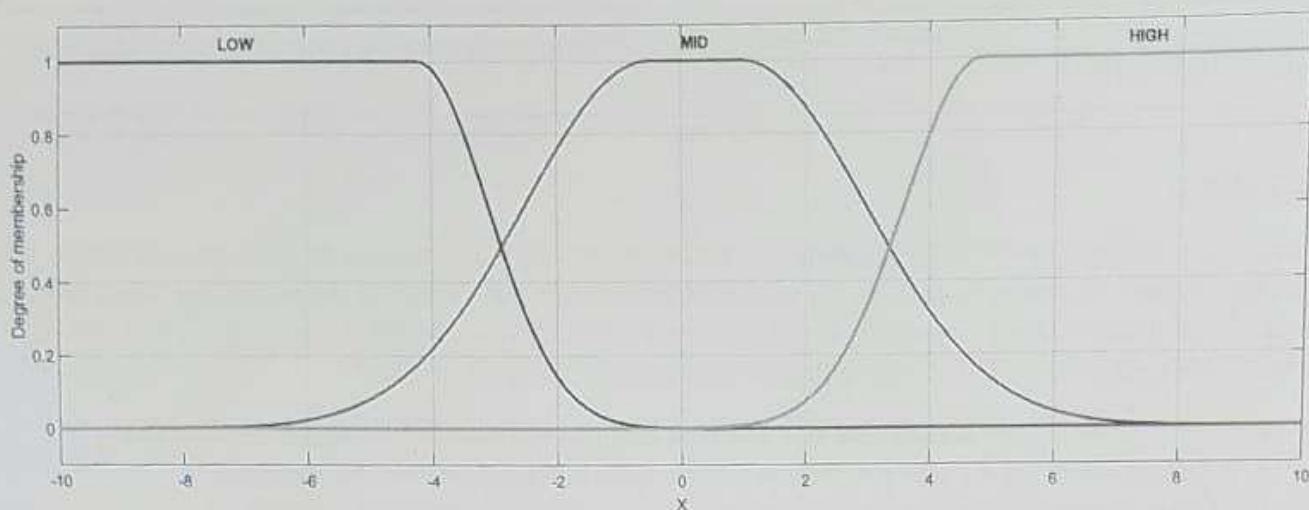
*SI el objeto es bajo y es pesado ENTONCES la velocidad es lenta.*

*SI el objeto es alto y es liviano ENTONCES la velocidad es rápida.*



### PROBLEMA 3:

Se tienen las siguientes funciones de pertenencia y las reglas de un sistema de inferencia difusa de Sugeno de una entrada  $X$  y una salida  $Y$ .



IF  $X$  IS LOW THEN  $Y = X + 2$

IF  $X$  IS MID THEN  $Y = 3X$

IF  $X$  IS HIGH THEN  $Y = 4X + 1$

- ¿Cuál regla se activa más?
- Calcule la salida para  $X = 3$ .

---

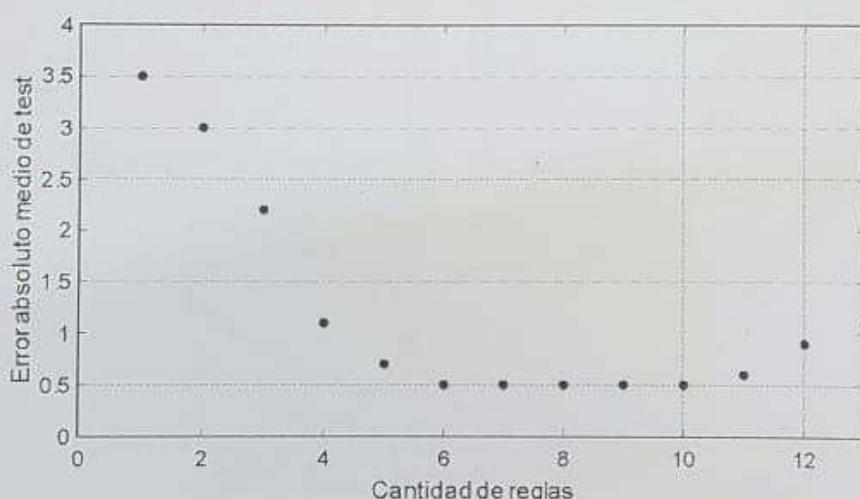
### PROBLEMA 4:

Se intenta diseñar un modelo de Sugeno para estimar un valor de alarma (RIESGO) según dos variables: la cantidad de usuarios activos (ACT) y el monto promedio de las operaciones efectuadas (MONTO).

Para generar el modelo se dispone de 1600 casos en los que se conoce ACT, MONTO y su respectivo valor RIESGO.

Se ha decidido utilizar el 10% para datos de **test** y el resto para datos de "entrenamiento" (generación del modelo). Se utiliza el error absoluto medio para evaluar el comportamiento del modelo. Se utiliza el método de hold-out para la evaluación con 20 iteraciones.

- Indique las dimensiones de las matrices que contienen los datos para entrenamiento y los datos para test.
- Se varió la cantidad de reglas de diversos modelos y se observa lo siguiente:



Seleccione la cantidad de reglas que estima adecuada para elegir uno de los modelos. Explique cuál puede ser la causa del aumento del error en la última parte.

c) Calcule el error para los siguientes 3 datos, según lo que estimó el modelo elegido:

| ACT  | MONTO | RIESGO estimado | RIESGO real |
|------|-------|-----------------|-------------|
| 110  | 746   | 1.2             | 0.5         |
| 12   | 512   | 9.7             | 9.1         |
| 5056 | 1031  | 2.2             | 3.0         |

**PROBLEMA 4:**

Los gráficos a y b representan la recta de separación obtenida mediante un perceptrón simple.

a) Calcule cuales son los pesos y bias iniciales (gráfico a)

b) Calcule cuales son los pesos y bias finales (gráfico b)

c) Que ocurre si se agrega un nuevo punto de la clase 2 en la posición (0.25, 0.5)?

¿Es necesario modificar los pesos y el bias para lograr una correcta clasificación?

¿Cuántos pasos de iteración se necesitan aproximadamente para lograr la clasificación?

