

ENUNCIADO:

Como ingeniero de una empresa recibes unos datos (datos.csv) para crear un modelo de sistema experto que ayude a predecir tendencias:

Se recomienda al alumno explorar los datos, ya que las columnas no están etiquetadas, para ver posibles evoluciones, correlaciones, etc.

EJ01.- Para predecir tendencias de evolución, nos fijamos en las dos primeras columnas de datos:

- a) La primera tarea será predecir la evolución del modelo cuando se aumenten los números de la primera columna, estudiando la correlación con los datos de la segunda columna → Modelo $y=f(x)$. Para ello, escogiendo los primeros 80% de los datos para crear el modelo, y los últimos (el 20% restante del final) para validarlo. *[tal y como vienen en el archivo sin reordenar las filas]*
- b) En una segunda prueba generamos un segundo modelo, esta vez utilizamos el 80% escogiéndolos aleatoriamente, para validarlo con el 20% restante.
- c) Relaciona los resultados obtenidos en a) y b). Explica la similitud o diferencia de los errores obtenidos en la validación, por qué crees que ocurre y con qué modelo te quedarías.

EJ02.- Para crear modelos de clasificación correspondiente a la cuarta columna que indica con un 1 o un 0 la validez del producto obtenido. Se trata de tener un modelo que clasifique los productos en Válido o No_Válido, según predecir los datos de las 2 primeras columnas y también según los colores asignados por el código RGB de las columnas 3, 4 y 5.

- d) ¿Usarías alguna técnica de clustering? ¿Para qué y con qué columnas?
- e) Usa todas las técnicas de clasificación que conozcas y obtén un modelo que clasifique entre Válido y No_válido para cada una de ellas.
- f) Compara los resultados obtenidos en el apartado e) y decide con qué modelo te quedas y fundamenta bien tu decisión.

EJ03.- Midiendo de nuevo obtenemos un nuevo dato con el que no contábamos anteriormente. Una nueva columna con una nueva medición.

- g) Examina la nueva información y decide si utilizarías otra técnica diferente a las empleadas en el EJ02 para hallar el modelo e impleméntala.