

Posibles pasos a seguir para el preprocesado y modelado de datos:

=====

- 1.- Comprobación de datos perdidos y eliminación de aquellos atributos que tengan demasiados.
Recordad que hay que tener las mismas variables de entrada en entrenamiento y test si queremos comprobar el rendimiento de un modelo.
- 2.- De aquellos atributos que se hayan dejado y tengan pocos datos perdidos, usar un método de Reemplazamiento de esos datos perdidos.
A la parte de test habría que hacerle también el Reemplazamiento (para hacerlo más justo debería ser por el valor de Reemplazamiento de train).
Lo mismo ocurre con una Normalización, habría que aplicarla también a test (para hacerlo más justo los valores min y max deberían ser los de train).
- 3.- Unificación de medidas similares, por ejemplo adoptar la misma medida para presión en pascales y en hectopascales.
- 4.- Visualizar los datos para seguir con el preprocesamiento. En función de ello realizar preprocesado:
 - Datos extremos.
 - Correlaciones/selección de características.
 - Normalización.
- 5.- Probar con algunos algoritmos de aprendizaje.