

Aprendizaje Automático.

Ejercicios para modelos de clasificación y clustering.

1. Implementación de una regresión stepwise con eliminación hacia atrás.

Utilizando como ejemplo la implementación del algoritmo de regresión stepwise del archivo “02_02_Stepwise.ipynb” realizar una implementación del algoritmo con eliminación hacia atrás. En esta implementación el modelo ha de comenzar con todas la variables e ir eliminando en cada paso la menos significativa hasta que la eliminación de una nueva variable aumenta el error de manera significativa.

2. Estimación de la supervivencia de los pasajeros del Titanic.

Partiendo de los datos del archivo “titanic.csv”, que contiene la información del pasaje del Titanic, crear un modelo que estime si un pasajero sobrevivir al hundimiento del barco en el momento de embarque seleccionando las variables más con mayor capacidad de predicción. Las variables “boat” ni “body” no pueden ser utilizadas a la hora de crear el modelo, ya que estas contienen información a posteriori al momento del embarque.

En el conjunto de datos existes valores NaN en las columnas “age” y “fare”, para eliminar las filas con estos valores se puede utilizar el siguiente comando:

```
mask = titanic[['age', 'fare']].applymap(lambda x: math.isnan(x))
titanic = titanic[-mask.any(axis = 1)]
```

3. Segmentación de los vinos tintos y blancos

Partiendo de los datos contenidos en los archivos “winequality-red.csv” y “winequality-white.csv” realizar una segmentación de los distintos tipos de vinos que tintos y blancos. Para seleccionar la cantidad de segmentos en los que se han de dividir los vinos tintos y blancos utilizar el método de la distorsión seleccionando el valor en el que esta empieza a reducir más lentamente. Una vez obtenido el número de segmentos para cada tipo de vino obtener la posición de los centroides que identifican a cada segmento.