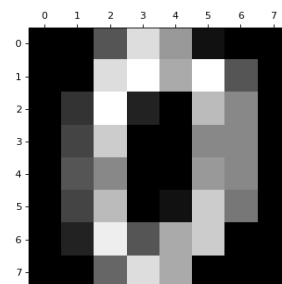


## 1 ML- REGRESIÓN LOGÍSTICA MULTICLASE

La clasificación multiclase implica que el target u objetivo no es un valor binario, sino que puede tener múltiples valores.

Para ver este tipo de regresión vamos a usar un dataset que viene en la librería sklearn que contiene dígitos del 0 al 9 escritos a mano. Estos dígitos están almacenados en una matriz 7x7 de valores en escala de grises y por tanto tendremos que identificar los 10 valores posibles {0,1,2,3,4,5,6,7,8,9}



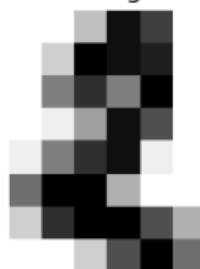
Training: 0



Training: 1



Training: 2



Training: 3



¿Cómo se tratan los datos de este dataset?

Primero lo que hacemos es cargar el DataSet:

```
from sklearn.datasets import load_digits
digits = load_digits()
```

Podemos ver los atributos y datos del objeto digits de la siguiente manera:

```
dir(digits)

['DESCR', 'data', 'feature_names', 'frame', 'images', 'target', 'target_names']
```

Podemos mostrar cada uno a ver qué contiene:



GENERALITAT  
VALENCIANA



UNIÓN EUROPEA  
Fondo Social Europeo  
El FSE invierte en tu futuro

```
digits.data
array([[ 0.,  0.,  5., ...,  0.,  0.,  0.],
       [ 0.,  0.,  0., ..., 10.,  0.,  0.],
       [ 0.,  0.,  0., ..., 16.,  9.,  0.],
       ...,
       [ 0.,  0.,  1., ...,  6.,  0.,  0.],
       [ 0.,  0.,  2., ..., 12.,  0.,  0.],
       [ 0.,  0., 10., ..., 12.,  1.,  0.]])
```

Aquí tenemos todos los datos de las muestras

digits.feature\_names → Los nombres de las columnas

digits.target\_names → Los nombres de los objetivos (del 0 al 9)

digits.target → la columna target para cada dato

## 1.1 MODELO DE IA. REGRESIÓN LOGÍSTICA

```
from sklearn.linear_model import LogisticRegression
model = LogisticRegression()
```

Separamos los datos en test y train tanto los datos como el target

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test =
train_test_split(digits.data, digits.target, test_size=0.2)
```

20% Test

Entrenamos el modelo:

```
model.fit(X_train, y_train)
```

Medimos el resultado con el conjunto de datos de test:

```
model.score(X_test, y_test)
```

Se puede predecir pasándole un conjunto de datos:

```
model.predict(digits.data[0:5])
```



GENERALITAT  
VALENCIANA



UNIÓN EUROPEA  
Fondo Social Europeo  
El FSE invierte en tu futuro

03008915 C/ Ferrocarril, 22, 03570 La Vila Joiosa Tel 966870140 Fax 966870141 <http://portal.edu.gva.es/iesmarcoszaragoza>

Para comprobar la eficacia del modelo podemos verificar con una matriz de confusión con los valores que se han predicho.

```
y_predicted = model.predict(X_test)

from sklearn.metrics import confusion_matrix
cm = confusion_matrix(y_test, y_predicted)
cm

import seaborn as sn
plt.figure(figsize = (10,7))
sn.heatmap(cm, annot=True)
plt.xlabel('Predicted')
plt.ylabel('Truth')
```