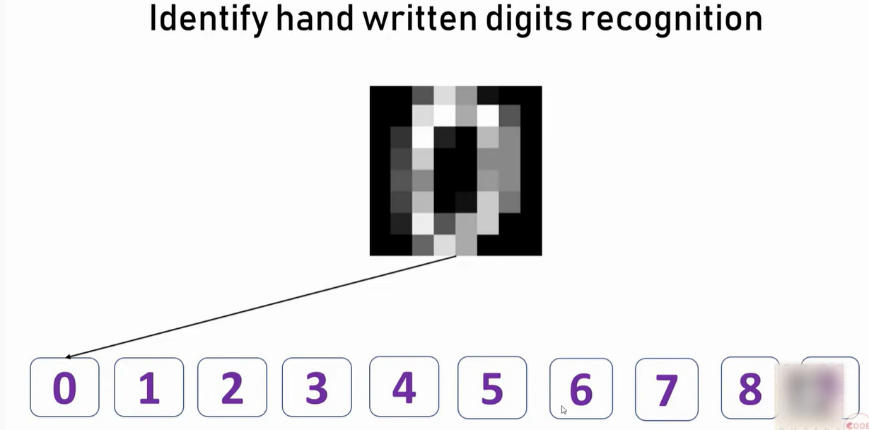


Se crean varios random simples, de cada uno se crea un decisión tree

En vida real, se llama a varios amigos, cada uno da su opinión y se vota por la mayoría



**import** pandas **as** pd

**from** sklearn.datasets **import** load\_digits

digits **=** load\_digits()

dir(digits)

Out[4]:

['DESCR', 'data', 'images', 'target', 'target\_names']

**%matplotlib** inline

**import** matplotlib.pyplot **as** plt

plt**.**gray()

**for** i **in** range(4):

plt**.**matshow(digits**.**images[i])

IMÁGENES 8X8

df **=** pd**.**DataFrame(digits**.**data)

df['target'] **=** digits**.**target

df[0:12]

**Train and the model and prediction**

X **=** df**.**drop('target',axis**=**'columns')

y **=** df**.**target

**from** sklearn.model\_selection **import** train\_test\_split

X\_train, X\_test, y\_train, y\_test **=** train\_test\_split(X,y,test\_size**=**0.2)

**from** sklearn.ensemble **import** RandomForestClassifier

model **=** RandomForestClassifier(n\_estimators**=**20)

model**.**fit(X\_train, y\_train)

model**.**score(X\_test, y\_test)

Out[15]:

0.96666666666666667

y\_predicted **=** model**.**predict(X\_test)

**Confusion Matrix**

**from** sklearn.metrics **import** confusion\_matrix

cm **=** confusion\_matrix(y\_test, y\_predicted)

cm

array([[34, 0, 0, 0, 0, 0, 0, 0, 0, 0],

[ 0, 30, 0, 0, 0, 0, 0, 0, 0, 0],

[ 0, 1, 38, 0, 0, 0, 0, 0, 0, 0],

[ 0, 0, 0, 30, 0, 1, 0, 0, 0, 0],

[ 0, 0, 0, 0, 39, 0, 0, 1, 0, 0],

[ 0, 0, 0, 0, 0, 25, 0, 0, 0, 0],

[ 1, 0, 0, 0, 0, 1, 45, 0, 0, 0],

[ 0, 0, 0, 0, 0, 0, 0, 38, 0, 0],

[ 0, 1, 1, 0, 0, 0, 0, 0, 33, 0],

[ 0, 1, 0, 1, 0, 1, 0, 2, 0, 36]])

**%matplotlib** inline

**import** matplotlib.pyplot **as** plt

**import** seaborn **as** sn

plt**.**figure(figsize**=**(10,7))

sn**.**heatmap(cm, annot**=True**)

plt**.**xlabel('Predicted')

plt**.**ylabel('Truth')

|  |  |
| --- | --- |
|  | This means that some of the images were not able to predict correctly  1 vez era 6 y predijo 0  2 veces era 9 y predijo 7 |