

TO START

```
# IMPORT DATA LIBRARIES
import pandas as pd
import numpy as np

# IMPORT VIS LIBRARIES
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline

# IMPORT MODELLING LIBRARIES
from sklearn.model_selection import train_test_split

# libraries for decision trees
from sklearn.tree import DecisionTreeClassifier
from sklearn.metrics import classification_report, confusion_matrix

# libraries for random forest
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import classification_report, confusion_matrix
```

PRELIMINARY OPERATIONS

df = pd.read_csv('data.csv')	import data
sns.pairplot(df,hue='col')	pairplot
df.info()	check info df
df.describe()	check stats df
df.head()	check head df

TRAIN MODEL - DECISION TREES

□ SPLIT DATASET

X = df[['col1','col2',etc.]]	create df features
y = df['col']	create df var to predict
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3)	split df in train and test df

... FIT THE MODEL

tree = DecisionTreeClassifier()	instantiate model
tree.fit(X_train, y_train)	train/fit the model

◎ MAKE PREDICTIONS

pred = tree.predict(X_test)	make predictions
-----------------------------	------------------

TRAIN MODEL - DECISION TREES (cont)

✓ EVALUATE MODEL

```
print(classification_report(y_test,pred))
print(confusion_matrix(y_test,pred))
```

TRAIN MODEL - RANDOM FOREST

□ SPLIT DATASET

X = df[['col1','col2',etc.]]	create df features
y = df['col']	create df var to predict
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3)	split df in train and test df

... FIT THE MODEL

rfc = RandomForestClassifier(n_estimators=200)*	instantiate model
rfc.fit(X_train, y_train)	train/fit the model

◎ MAKE PREDICTIONS

rfc_pred = rfc.predict(X_test)	make predictions
--------------------------------	------------------

✓ EVALUATE MODEL

```
print(confusion_matrix(y_test,rfc_pred))
print(classification_report(y_test,rfc_pred))
```

n_estimators: number of trees to be used in the forest.



By **DarioPittera** (aggialavura)

Not published yet.
Last updated 1st July, 2019.
Page 1 of 1.

Sponsored by **CrosswordCheats.com**
Learn to solve cryptic crosswords!
<http://crosswordcheats.com>