```
import numpy as np
import pandas as pd
import matplotlib as mpl
import matplotlib.pyplot as plt
%matplotlib inline
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import *
from sklearn.metrics import confusion_matrix, classification_report
from sklearn.preprocessing import StandardScaler
```

```
!wget -q -O '/kaggle/input/bank-full.csv' https://raw.githubusercontent.com/jbrownlee/Datasets/master/bank.csv?token=AMVCCZ3ZCCCCCCCCCCCCCCCCCCCCC
```

```
bank = pd.read_csv('/kaggle/input/bank-full.csv')
bank.head()
```

	age	job	marital	education	default	balance	housing	loan	contact	day	month	duration	campaign	pdays	previous	poutcome	y
0	58	management	married	tertiary	no	2143	yes	no	unknown	5	may	261	1	-1	0	unknown	no
1	44	technician	single	secondary	no	29	yes	no	unknown	5	may	151	1	-1	0	unknown	no
2	33	entrepreneur	married	secondary	no	2	yes	yes	unknown	5	may	76	1	-1	0	unknown	no
3	47	blue-collar	married	NaN	no	1506	yes	no	unknown	5	may	92	1	-1	0	unknown	no
4	33	NaN	single	NaN	no	1	no	no	unknown	5	may	198	1	-1	0	unknown	no

```
bank.shape
```

```
(45211, 17)
```

```
bank = pd.read_csv('/kaggle/input/bank-full.csv', sep=';')
```

```
bank
```

	age	job	marital	education	default	balance	housing	loan	contact	day	month	duration	campaign	pdays	previous	poutcome	y
0	58	management	married	tertiary	no	2143	yes	no	unknown	5	may	261	1	-1	0	unknown	no
1	44	technician	single	secondary	no	29	yes	no	unknown	5	may	151	1	-1	0	unknown	no
2	33	entrepreneur	married	secondary	no	2	yes	yes	unknown	5	may	76	1	-1	0	unknown	no
3	47	blue-collar	married	NaN	no	1506	yes	no	unknown	5	may	92	1	-1	0	unknown	no
4	33	NaN	single	NaN	no	1	no	no	unknown	5	may	198	1	-1	0	unknown	no
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
45206	51	technician	married	secondary	no	825	no	no	cellular	17	nov	977	3	-1	0	unknown	yes
45207	71	retired	divorced	primary	no	1228	no	no	cellular	17	nov	456	2	-1	0	unknown	yes
45208	72	retired	married	secondary	no	8223	no	no	cellular	17	nov	1127	5	184	3	success	yes
45209	57	blue-collar	married	secondary	no	1778	no	no	telephone	17	nov	508	4	-1	0	unknown	no
45210	37	entrepreneur	married	secondary	no	2787	no	no	cellular	17	nov	361	2	188	11	other	no

45211 rows × 17 columns

```
bank.shape
```

```
(45211, 17)
```

```
bank = pd.read_csv('/kaggle/input/bank-full.csv', sep=';')
```

```
bank.head()
```

	age	job	marital	education	default	balance	housing	loan	contact	day	month	duration	campaign	pdays	previous	poutcome	y
0	58	management	married	tertiary	no	2143	yes	no	unknown	5	may	261	1	-1	0	unknown	no
1	44	technician	single	secondary	no	29	yes	no	unknown	5	may	151	1	-1	0	unknown	no
2	33	entrepreneur	married	secondary	no	2	yes	yes	unknown	5	may	76	1	-1	0	unknown	no
3	47	blue-collar	married	NaN	no	1506	yes	no	unknown	5	may	92	1	-1	0	unknown	no
4	33	NaN	single	NaN	no	1	no	no	unknown	5	may	198	1	-1	0	unknown	no

45211 rows × 17 columns

```
bank.shape
```

```
(45211, 17)
```

```
bank.describe()
```

	age	balance	day	duration	campaign	pdays	previous
count	45211.000000	45211.000000	45211.000000	45211.000000	45211.000000	45211.000000	45211.000000
mean	40.936210	1362.272058	15.806419	258.163080	2.763841	40.197828	0.580323
std	10.618762	3044.765829	8.322476	257.527812	3.098021	100.128746	2.303441
min	18.000000	-8019.000000	1.000000	0.000000	1.000000	-1.000000	0.000000
25%	33.000000	72.000000	8.000000	103.000000	1.000000	-1.000000	0.000000
50%	39.000000	448.000000	16.000000	180.000000	2.000000	-1.000000	0.000000
75%	48.000000	1428.000000	21.000000	319.000000	3.000000	-1.000000	0.000000
max	95.000000	102127.000000	31.000000	4918.000000	63.000000	871.000000	275.000000

```
!pip install kaggle
```

```
!mkdir ~/.kaggle
```

```
!cp kaggle.json ~/.kaggle/
```

```
!chmod 600 ~/.kaggle/kaggle.json
```

```
!kaggle datasets download -d jeffreyevans/satellite-land-cover-classification
```

```
Dataset URL: https://kaggle.com/datasets/jeffreyevans/satellite-land-cover-classification
License(s): unknown
Traceback (most recent call last):
  File "/usr/local/bin/kaggle", line 5, in from kaggle.cli import cli, main
  File "/usr/local/lib/python3.10/dist-packages/kaggle/cli.py", line 23, in from kaggle.api.kaggle_api_extended import KaggleApi
  File "/usr/local/lib/python3.10/dist-packages/kaggle/api/kaggle_api_extended.py", line 32, in from urllib3.exceptions import MaxRetryError
ImportError: cannot import name 'urlencode' from 'urllib3.exceptions' (/usr/local/lib/python3.10/dist-packages/urllib3/exceptions.py)
```

```
!unzip -q -o -d /kaggle/input/satellite-land-cover-classification -p /kaggle/input/satellite-land-cover-classification.zip > /dev/null
```

```
!ls -lh /kaggle/input/satellite-land-cover-classification
```

```
ls: cannot access '/kaggle/input/satellite-land-cover-classification': No such file or directory
```

```
!kaggle datasets download -d jeffreyevans/satellite-land-cover-classification
```

```
Dataset URL: https://kaggle.com/datasets/jeffreyevans/satellite-land-cover-classification
License(s): unknown
Traceback (most recent call last):
  File "/usr/local/bin/kaggle", line 5, in from kaggle.cli import cli, main
  File "/usr/local/lib/python3.10/dist-packages/kaggle/cli.py", line 23, in from kaggle.api.kaggle_api_extended import KaggleApi
  File "/usr/local/lib/python3.10/dist-packages/kaggle/api/kaggle_api_extended.py", line 32, in from urllib3.exceptions import MaxRetryError
ImportError: cannot import name 'urlencode' from 'urllib3.exceptions' (/usr/local/lib/python3.10/dist-packages/urllib3/exceptions.py)
```

```
!unzip -q -o -d /kaggle/input/satellite-land-cover-classification -p /kaggle/input/satellite-land-cover-classification.zip > /dev/null
```

```
!ls -lh /kaggle/input/satellite-land-cover-classification
```

```
ls: cannot access '/kaggle/input/satellite-land-cover-classification': No such file or directory
```

```
bank.shape
```

```
(45211, 17)
```

```
bank.describe()
```

	age	balance	day	duration	campaign	pdays	previous
count	45211.000000	45211.000000	45211.000000	45211.000000	45211.000000	45211.000000	45211.000000
mean	40.936210	1362.272058	15.806419	258.163080	2.763841	40.197828	0.580323
std	10.618762	3044.765829	8.322476	257.527812	3.098021	100.128746	2.303441
min	18.000000	-8019.000000	1.000000	0.000000	1.000000	-1.000000	0.000000
25%	33.000000	72.000000	8.000000	103.000000	1.000000	-1.000000	0.000000
50%	39.000000	448.000000	16.000000	180.000000	2.000000	-1.000000	0.000000
75%	48.000000	1428.000000	21.000000	319.000000	3.000000	-1.000000	0.000000
max	95.000000	102127.000000	31.000000	4918.000000	63.000000	871.000000	275.000000

```
bank.info()
```

```
<class'pandas.core.frame.DataFrame'>
RangeIndex: 45211 entries, 0 to 45210
Data columns (total 17 columns):
 0   age        45211 non-null  int64
 1   job        45211 non-null  object
 2   marital    45211 non-null  object
 3   education  45211 non-null  object
 4   default    45211 non-null  object
 5   balance    45211 non-null  int64
 6   housing    45211 non-null  object
 7   loan       45211 non-null  object
 8   contact    45211 non-null  object
 9   day        45211 non-null  int64
 10  month      45211 non-null  object
 11  duration   45211 non-null  int64
 12  campaign   45211 non-null  int64
 13  pdays      45211 non-null  int64
 14  previous   45211 non-null  int64
 15  poutcome   45211 non-null  object
 16  y          45211 non-null  object
dtype: int64, object: 10
memory usage: 5.9+ MB
```

```
bank.isnull()
```

	age	job	marital	education	default	balance	housing	loan	contact	day	month	duration	campaign	pdays	previous	poutcome	y
0	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
45206	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False
45207	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False
45208	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False
45209	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False
45210	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False

45211 rows × 17 columns

```
!pip install matplotlib-inline
```

```
!pip install -q matplotlib-inline
```

```
%matplotlib inline
```

```
The rich_traceback module is not available.
Traceback (most recent call last):
  File "<bin_ipykernel_workdir>/bin/ipkernel", line 5, in from ipykernel.kernelapp import main
  File "/usr/local/lib/python3.10/dist-packages/ipykernel/kernelapp.py", line 3, in import asyncio
  File "/usr/local/lib/python3.10/dist-packages/asyncio/__init__.pyc", line 1, in ImportEnvironmentVariable: Environ[b'PYTHONSTARTUP'] = b'/usr/local/lib/python3.10/dist-packages/jedi/inject.py'
```

```
%matplotlib inline
```

```
The rich_traceback module is not available.
Traceback (most recent call last):
  File "<bin_ipykernel_workdir>/bin/ipkernel", line 5, in from ipykernel.kernelapp import main
  File "/usr/local/lib/python3.10/dist-packages/ipykernel/kernelapp.py", line 3, in import asyncio
  File "/usr/local/lib/python3.10/dist-packages/asyncio/__init__.pyc", line 1, in ImportEnvironmentVariable: Environ[b'PYTHONSTARTUP'] = b'/usr/local/lib/python3.10/dist-packages/jedi/inject.py'
```

```
!pip install -q matplotlib-inline
```

```
!pip install -q seaborn
```

```
!pip install -q scikit-learn
```

```
!pip install -q imbalanced-learn
```

```
!pip install -q xgboost
```

```
!pip install -q lightgbm
```

```
!pip install -q catboost
```

```
!pip install -q shap
```

```
!pip install -q optuna
```

```
!pip install -q joblib
```

```
import matplotlib, joblib, sklearn, xgboost, lightgbm, catboost, shap, optuna
print('matplotlib', matplotlib.__version__)
print('joblib', joblib.__version__)
print('sklearn', sklearn.__version__)
print('xgboost', xgboost.__version__)
print('lightgbm', lightgbm.__version__)
print('catboost', catboost.__version__)
print('shap', shap.__version__)
print('optuna', optuna.__version__)
```

```
matplotlib 3.7.2
joblib 1.4.2
sklearn 1.6.1
xgboost 2.1.4
lightgbm 4.5.0
catboost 1.2.7
shap 0.46.0
optuna 4.1.0
```