```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
import sklearn
from sklearn.utils import resample
from sklearn.svm import SVC
from sklearn.model_selection import train_test_split, cross_val_score
from sklearn.metrics import confusion_matrix, classification_report, accuracy_score
from sklearn import preprocessing
from sklearn.preprocessing import LabelEncoder


from google.colab import drive
drive.mount('/content/gdrive')
```

```
Mounted at /content/gdrive
```

```
datas = pd.read_csv("/content/gdrive/MyDrive/bank-full.csv")
```

```
datas.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 45211 entries, 0 to 45210
Data columns (total 17 columns):
 #   Column     Non-Null Count  Dtype 
---  ------     --------------  ----- 
 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
dtypes: int64(7), object(10)
memory usage: 5.9+ MB
```

```
datas.head()
```

	age	job	marital	education	default	balance	housing	loan	contact	da
0	58	management	married	tertiary	no	2143	yes	no	unknown	
1	44	technician	single	secondary	no	29	yes	no	unknown	
2	33	entrepreneur	married	secondary	no	2	yes	yes	unknown	
3	47	blue-collar	married	unknown	no	1506	yes	no	unknown	
4	33	unknown	single	unknown	no	1	no	no	unknown	

```
datas = datas.drop_duplicates()
```

```
datas = datas[datas.job != 'unknown']
```

```
datas = datas[datas.education != 'unknown']
```

```
datas.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 43193 entries, 0 to 45210
Data columns (total 17 columns):
 #   Column     Non-Null Count  Dtype 
---  ------     --------------  ----- 
 0   age        43193 non-null  int64 
 1   job        43193 non-null  object
 2   marital    43193 non-null  object
 3   education  43193 non-null  object
 4   default    43193 non-null  object
 5   balance    43193 non-null  int64 
```

```
 6   housing    43193 non-null  object
 7   loan       43193 non-null  object
 8   contact    43193 non-null  object
 9   day        43193 non-null  int64
 10  month      43193 non-null  object
 11  duration   43193 non-null  int64
 12  campaign   43193 non-null  int64
 13  pdays      43193 non-null  int64
 14  previous   43193 non-null  int64
 15  poutcome   43193 non-null  object
 16  y          43193 non-null  object
dtypes: int64(7), object(10)
memory usage: 5.9+ MB
```

```
encoder = preprocessing.LabelEncoder()
datas['job'] = encoder.fit_transform(datas['job'])
datas['marital'] = encoder.fit_transform(datas['marital'])
datas['education'] = encoder.fit_transform(datas['education'])
datas['default'] = encoder.fit_transform(datas['default'])
datas['housing'] = encoder.fit_transform(datas['housing'])
datas['loan'] = encoder.fit_transform(datas['loan'])
datas['poutcome'] = encoder.fit_transform(datas['poutcome'])
datas['y'] = encoder.fit_transform(datas['y'])
datas['contact'] = encoder.fit_transform(datas['contact'])
datas['month'] = encoder.fit_transform(datas['month'])
```

```
datas.head()
```

	age	job	marital	education	default	balance	housing	loan	contact	day	month
0	58	4	1	2	0	2143	1	0	2	5	8
1	44	9	2	1	0	29	1	0	2	5	8
2	33	2	1	1	0	2	1	1	2	5	8
5	35	4	1	2	0	231	1	0	2	5	8
6	28	4	2	2	0	447	1	1	2	5	8

```
del datas['default']
```

```
del datas['contact']
```

```
datas.head()
```

	age	job	marital	education	balance	housing	loan	day	month	duration	campai
0	58	4	1	2	2143	1	0	5	8	261	
1	44	9	2	1	29	1	0	5	8	151	
2	33	2	1	1	2	1	1	5	8	76	
5	35	4	1	2	231	1	0	5	8	139	
6	28	4	2	2	447	1	1	5	8	217	

```
X= datas.iloc[:,0:14]
X[:]
```

```
            age  job  marital  education  balance  housing  loan  day  month  duration  ca
```

```
y = datas.iloc[:,14]
y[:]
```

```
     0        0
     1        0
     2        0
     5        0
     6        0
             ..
     45206    1
     45207    1
     45208    1
     45209    0
     45210    0
     Name: y, Length: 43193, dtype: int64
```

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=22)
```

```
X_train.shape, y_train.shape
```

```
     ((30235, 14), (30235,))
```

```
X_test.shape, y_test.shape
```

```
     ((12958, 14), (12958,))
```

```
svc_model = SVC()
```

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

```
     ▾ SVC
     SVC()
```

```
predictions = svc_model.predict(X_test)
```

```
print(confusion_matrix(y_test,predictions))
```

```
     [[11498     7]
      [ 1437    16]]
```

```
accuracy_score(y_test, predictions)
```

```
     0.8885630498533724
```

```
accuracy_score(y_train, svc_model.predict(X_train))
```

```
     0.8820902927071275
```

```
print(classification_report(y_test,predictions))
```

```
                   precision    recall  f1-score   support

                0       0.89      1.00      0.94     11505
                1       0.70      0.01      0.02      1453

         accuracy                           0.89     12958
        macro avg       0.79      0.51      0.48     12958
     weighted avg       0.87      0.89      0.84     12958
```

✓ 0s completed at 22:22