```python
# Load necessary libraries
import pandas as pd
from sklearn.model_selection import train_test_split

# Load the dataset
file_path = 'Downloads/heart disease.csv'
df = pd.read_csv(file_path, encoding='UTF-8-SIG')

# Display the head of the dataframe to understand its structure
print(df.head())
print(df.info())
# Split the data into training and test sets
X = df.drop('target', axis=1)
y = df['target']

# 80% training and 20% test
X_train, X_test, y_train, y_test = train_test_split(X, y,
test_size=0.2, random_state=42)

print('Training set size:', X_train.shape)
print('Test set size:', X_test.shape)
```

```
     age  sex  cp  trestbps  chol  fbs  restecg  thalach  exang  oldpeak
slope  \
0    63    1   3       145   233    1        0      150      0      2.3
0
1    37    1   2       130   250    0        1      187      0      3.5
0
2    41    0   1       130   204    0        0      172      0      1.4
2
3    56    1   1       120   236    0        1      178      0      0.8
2
4    57    0   0       120   354    0        1      163      1      0.6
2

     ca  thal  target
0    0     1       1
1    0     2       1
2    0     2       1
3    0     2       1
4    0     2       1
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 303 entries, 0 to 302
Data columns (total 14 columns):
 #   Column   Non-Null Count  Dtype
---  ------   --------------  -----
 0   age      303 non-null    int64
 1   sex      303 non-null    int64
 2   cp       303 non-null    int64
```

```
 3    trestbps  303 non-null    int64
 4    chol      303 non-null    int64
 5    fbs       303 non-null    int64
 6    restecg   303 non-null    int64
 7    thalach   303 non-null    int64
 8    exang     303 non-null    int64
 9    oldpeak   303 non-null    float64
 10   slope     303 non-null    int64
 11   ca        303 non-null    int64
 12   thal      303 non-null    int64
 13   target    303 non-null    int64
dtypes: float64(1), int64(13)
memory usage: 33.3 KB
None
Training set size: (242, 13)
Test set size: (61, 13)
```