

```
import numpy as np
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import accuracy_score
```

```
from google.colab import files
uploaded = files.upload()
```

[Choose files](#) Heart Disease Dataset.csv

- **Heart Disease Dataset.csv**(text/csv) - 11024 bytes, last modified: 05/01/2020 - 100% done  
Saving Heart Disease Dataset.csv to Heart Disease Dataset.csv

```
hd = pd.read_csv('Heart Disease Dataset.csv')
```

```
hd.head()
```

```
↗
```

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpe
0	63	1	3	145	233	1	0	150	0	2
1	37	1	2	130	250	0	1	187	0	3
2	41	0	1	130	204	0	0	172	0	1
3	56	1	1	120	236	0	1	178	0	0
4	57	0	0	120	354	0	1	163	1	0

```
hd.tail()
```

Automatic saving failed. This file was updated remotely or in another tab.

[diff](#)

[Show](#)

299	45	1	3	110	264	0	1	132	0	
300	68	1	0	144	193	1	1	141	0	
301	57	1	0	130	131	0	1	115	1	
302	57	0	1	130	236	0	0	174	0	

```
hd.shape
```

```
(303, 14)
```

```
hd.info()
```

```
<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
```

```
hd.isnull().sum()
```

```
age         0
sex         0
cp          0
trestbps    0
chol        0
fbs         0
```

Automatic saving failed. This file was updated remotely or in another tab.

[Show](#)

[diff](#)

```
oldpeak     0
slope       0
ca          0
thal        0
target      0
dtype: int64
```

```
hd.describe()
```

	age	sex	cp	trestbps	chol	f
<b>count</b>	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000
<b>mean</b>	54.366337	0.683168	0.966997	131.623762	246.264026	0.1485
<b>std</b>	9.082101	0.466011	1.032052	17.538143	51.830751	0.3561
<b>min</b>	29.000000	0.000000	0.000000	94.000000	126.000000	0.0000
<b>25%</b>	47.500000	0.000000	0.000000	120.000000	211.000000	0.0000
<b>50%</b>	55.000000	1.000000	1.000000	130.000000	240.000000	0.0000
<b>75%</b>	61.000000	1.000000	2.000000	140.000000	274.500000	0.0000
<b>max</b>	77.000000	1.000000	3.000000	200.000000	564.000000	1.0000

```
hd['target'].value_counts()
```

```
1    165
0    138
Name: target, dtype: int64
```

```
x = hd.drop(columns='target', axis=1)
y = hd['target']
```

```
print(x)
```

	age	sex	cp	trestbps	chol	...	exang	oldpeak	slope	ca
0	63	1	3	145	233	...	0	2.3	0	0
1	37	1	2	130	250	...	0	3.5	0	0
2	41	0	1	130	204	...	0	1.4	2	0
3	56	1	1	120	236	...	0	0.8	2	0
4	57	0	0	120	354	...	1	0.6	2	0

Automatic saving failed. This file was updated remotely or in another tab. [Show diff](#)

300	68	1	0	144	193	...	0	3.4	1	2
301	57	1	0	130	131	...	1	1.2	1	1
302	57	0	1	130	236	...	0	0.0	1	1

```
[303 rows x 13 columns]
```

```
print(y)
```

```

0      1
1      1
2      1
3      1
4      1
      ..
298    0
299    0
300    0
301    0
302    0
Name: target, Length: 303, dtype: int64

```

```

x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.2,

print(x.shape, x_train.shape, x_test.shape)

(303, 13) (242, 13) (61, 13)

```

## Model Training

```
model = LogisticRegression()
```

```
model.fit(x_train, y_train)
```

```

/usr/local/lib/python3.7/dist-packages/sklearn/linear_model/_logistic.py:1181:
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

```

```

Increase the number of iterations (max_iter) or scale the data as shown at:
https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
https://scikit-learn.org/stable/modules/linear\_model.html#logistic-regression
extra warning msg= LOGISTIC SOLVER CONVERGENCE MSG)

```

Automatic saving failed. This file was updated remotely or in another tab. [Show diff](#)

```

model = LogisticRegression(max_iter=1000, solver='lbfgs', tol=0.0001,
random_state=None, solver='lbfgs', tol=0.0001,
warm_start=False)

```

```

x_train_prediction = model.predict(x_train)
training_data_accuracy = accuracy_score(x_train_prediction, y_train)

```

```
print('Accuracy on Training data : ', training_data_accuracy)
```

```
Accuracy on Training data : 0.8512396694214877
```

```
x_test_prediction = model.predict(x_test)
test_data_accuracy = accuracy_score(x_test_prediction, y_test)
```

```
print('Accuracy on Test data : ', test_data_accuracy)
```

```
Accuracy on Test data : 0.819672131147541
```

```
input_data = (62,0,0,140,268,0,0,160,0,3.6,0,2,2)
```

```
input_data_as_numpy_array= np.asarray(input_data)
```

```
input_data_reshaped = input_data_as_numpy_array.reshape(1,-1)
```

```
prediction = model.predict(input_data_reshaped)
print(prediction)
```

```
if (prediction[0]== 0):
    print('The Person does not have a Heart Disease')
else:
    print('The Person has Heart Disease')
```

```
[0]
The Person does not have a Heart Disease
```

Automatic saving failed. This file was updated remotely or in another tab. [Show diff](#)

✓ 0s completed at 22:45



Automatic saving failed. This file was updated remotely or in another tab. [Show diff](#)