

In [3]:

```
import pandas as pd  
import numpy as np
```

In [4]:

```
from sklearn.datasets import make_blobs
```

In [5]:

```
data_assignment4 = make_blobs(n_samples=100,cluster_std=2.0,centers=2,n_features=2)
```

In [6]:

```
data_assignment4
```

Out[6]:

```
(array([[ 3.64865597,  0.23183381],
 [ 5.87274282,  1.87666592],
 [ 5.14792647, -5.24363253],
 [ 0.95719737, -0.67613436],
 [ 5.08052857, -5.66533091],
 [ 4.40764894, -5.70132454],
 [ 4.41673836, -0.26833365],
 [ 2.29877396,  2.68260054],
 [ 2.59256884, -1.18635874],
 [ 2.60287305, -6.65648008],
 [ 5.46559247, -0.75859349],
 [ 2.18734012,  0.69609011],
 [ 1.90758184, -6.96954345],
 [ 4.03173529, -9.39503102],
 [ 3.27406635, -1.09850522],
 [ 0.91033435, -8.48774068],
 [ 2.64256061, -0.70200959],
 [ 3.94955178, -4.30206704],
 [ 5.66992811, -0.82437355],
 [ 6.62590717, -6.55671679],
 [ 1.71137777,  2.54807819],
 [ 2.42504965, -3.25865298],
 [-1.21510533, -1.966984  ],
 [ 0.9854124 , -2.91690159],
 [ 4.59394857,  1.49803418],
 [-0.79936428, -0.32996557],
 [ 1.85001202, -8.81042034],
 [ 0.67930544, -2.57309197],
 [ 4.78746508, -7.08364931],
 [ 1.29559677, -7.43567848],
 [ 1.19514295, -7.38641443],
 [ 3.37738575, -7.33594713],
 [ 3.39678981,  4.17262186],
 [ 2.26040383, -7.2984445 ],
 [ 5.14728616, -10.67968462],
 [ 2.73372806, -6.52547286],
 [-1.21854904, -7.54310107],
 [ 5.75388608, -0.30823208],
 [ 3.67766444,  0.48572563],
 [ 1.05713846, -7.07547837],
 [ 4.571904  , -1.02128412],
 [ 1.69460884, -5.97299976],
 [ 3.46751115, -9.06275673],
 [ 0.4534438 , -7.41394058],
 [ 1.47777208, -9.36967592],
 [ 1.84705357, -11.0941951 ],
 [-0.4704296 , -7.14105864],
 [ 2.19960099, -4.52381309],
 [ 5.611646  , -0.01481545],
 [ 0.5433219 , -8.25898349],
 [ 2.04801162,  0.64422234],
 [ 4.63400892, -8.70376212],
 [ 1.44054327,  0.78931356],
 [ 2.09108017, -9.19077823],
 [ 2.2938964 , -7.56125104],
```

```

[ -0.1426864 , -2.61842188],
[ 3.1338796 , 0.50699116],
[ 5.49306905, -7.88614144],
[ 7.73088619, -2.8274941 ],
[ 2.22746091, -4.19827812],
[ 3.33931503, 1.23207231],
[ 3.85067795, -10.27560308],
[ 2.71363292, -10.10783752],
[ 2.26912246, -7.05661692],
[ 2.33190026, -11.48580964],
[ 4.13033476, -5.6124879 ],
[ 1.81720904, -0.50443449],
[ 3.11640869, 1.58201634],
[ 5.19189092, 0.23002045],
[ -0.18897332, -9.07715383],
[ 5.31781784, 1.52675699],
[ 2.12479067, -5.77391606],
[ 2.38917603, -6.61333172],
[ 3.78994273, -3.19953881],
[ 5.09166431, -0.76002405],
[ 1.63935834, -5.74568152],
[ 4.25752739, -5.94283939],
[ 2.04634474, -11.76015294],
[ 4.55312614, -2.53104791],
[ 6.89988439, 0.88294105],
[ 3.33884299, -7.05971676],
[ 2.67448951, -7.68317112],
[ 6.53227453, -0.09408995],
[ 5.20290611, -7.85607349],
[ 6.97116582, -3.40387039],
[ 2.03683305, -2.60605219],
[ 3.626855 , -7.73873825],
[ 1.9723558 , 0.12797218],
[ 5.77647922, -0.57389725],
[ 5.00767139, -2.69351047],
[ 0.51971374, -7.47836087],
[ 5.20264962, -0.99254295],
[ 2.30864264, -0.26830251],
[ 4.20791281, -0.91415758],
[ 4.35965119, -6.16946883],
[ -2.23608737, -4.82705349],
[ 1.47980524, -1.71863725],
[ 4.604727 , -1.60389555],
[ 7.85769951, -1.26617926],
[ 5.88642025, -10.43632335]],
array([1, 1, 0, 1, 0, 0, 1, 1, 1, 1, 1, 0, 0, 1, 0, 1, 0, 1, 1,
1,
1, 1, 1, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 1, 0, 1, 0, 0, 0, 0,
0,
0, 1, 1, 0, 1, 0, 1, 0, 0, 1, 1, 0, 1, 0, 1, 0, 0, 0, 0, 1, 1,
1,
0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 1, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0,
1,
1, 1, 0, 0, 1, 1, 1, 0]))

```

In [7]:

```

import matplotlib.pyplot as plt
%matplotlib inline

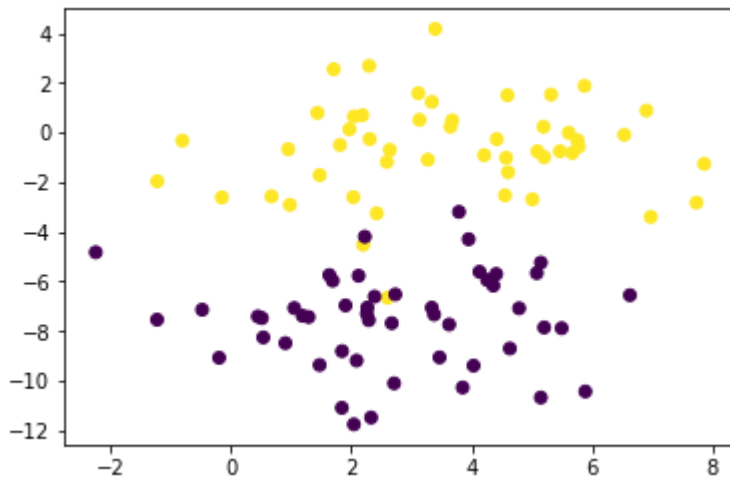
```

In [9]:

```
plt.scatter(data_assignment4[0][:,0],data_assignment4[0][:,1], c=data_assignment4[1], cmap=
```

Out[9]:

```
<matplotlib.collections.PathCollection at 0x27dc95708d0>
```



In [10]:

```
df = pd.DataFrame()
```

In [11]:

```
df["Feature1"] = data_assignment4[0][:,0]
df["Feature2"] = data_assignment4[0][:,1]
df["Class"] = data_assignment4[1]
```

In [12]:

```
df.head()
```

Out[12]:

	Feature1	Feature2	Class
0	3.648656	0.231834	1
1	5.872743	1.876666	1
2	5.147926	-5.243633	0
3	0.957197	-0.676134	1
4	5.080529	-5.665331	0

In [13]:

```
from sklearn.model_selection import train_test_split

X = df.drop("Class", axis=1)
y = df["Class"]

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.30)
```

In [14]:

```
from sklearn.naive_bayes import GaussianNB
```

In [15]:

```
GNB = GaussianNB()
```

In [16]:

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

Out[16]:

```
GaussianNB(priors=None)
```

In [17]:

```
predict = GNB.predict(X_test)
```

In [18]:

```
from sklearn.metrics import accuracy_score  
print('Accuracy: %.2f' %accuracy_score(y_test, predict))
```

```
Accuracy: 1.00
```

In []: