In [124]:

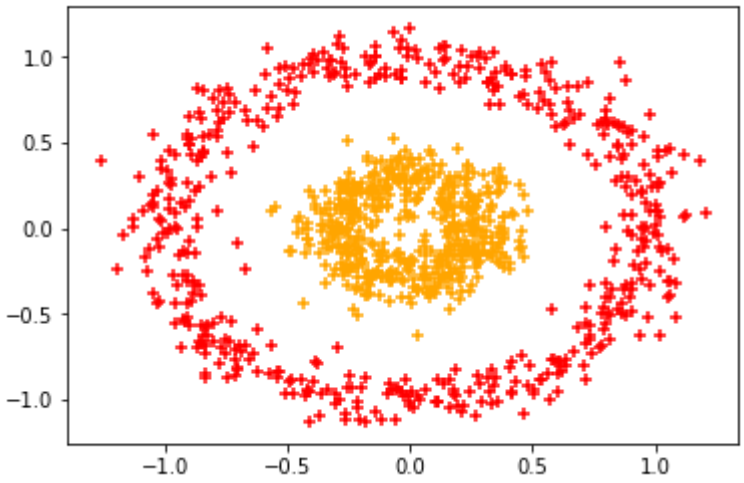
```
import sklearn
import pandas as pd
import matplotlib.pyplot as plt
from mpl_toolkits import mplot3d
%matplotlib inline
```

In [125]:

```
dataset1 = pd.read_csv('C:\Python37\datasets\circles0.3.csv')
```

In [138]:

```
df1=pd.DataFrame(dataset1)
colors=['red' if label==0 else 'orange' for label in df1.label]
plt.scatter(df1.x1,df1.x2,c=colors,marker='+')
plt.show()
```



In [139]:

```
df1.describe()
```

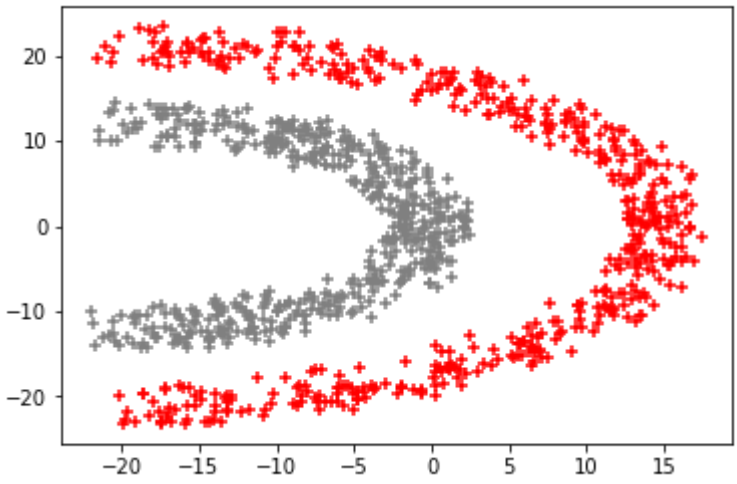
Out[139]:

	x1	x2	label
count	1000.000000	1000.000000	1000.00000
mean	-0.001126	-0.000579	0.50000
std	0.529914	0.527955	0.50025
min	-1.265528	-1.124877	0.00000
25%	-0.290869	-0.299358	0.00000
50%	-0.004964	-0.001341	0.50000
75%	0.310900	0.299989	1.00000
max	1.207588	1.162153	1.00000

The above dataset values plotted on the graph thereby shows shape of concentric circles separated by different colors 'orange' and 'red' in the above case, the figure shows the scatter describing the classification having some parametric values like mean of x1 = -0.001126 and that of x2 = -0.000579 and there by labelled as 0.5 because as per the scatter graph the values on x and y axis are present between -1 to +1 so all other answers has to be between then in the particular range, also the circles dataset comprises of standard deviation using the formula of approximate upto 0.529914 for x1 and 0.527955 for x2 thereby labelled as 0.50025 thus explaining that standard deviation is more inclined towards 1 also the max for x1 and x2 is shown as 1.207588 and 1.162153, similarly the minimum values are -1.265528 and -1.124877 for x1 and x2.

In [140]:

```
dataset2 = pd.read_csv('C:\Python37\datasets\halfkernel.csv')
df2=pd.DataFrame(dataset2)
colors=['grey' if label==0 else 'red' for label in df2.label]
plt.scatter(df2.x1,df2.x2,c=colors,marker='+')
plt.show()
```



In [141]:

```
df2.describe()
```

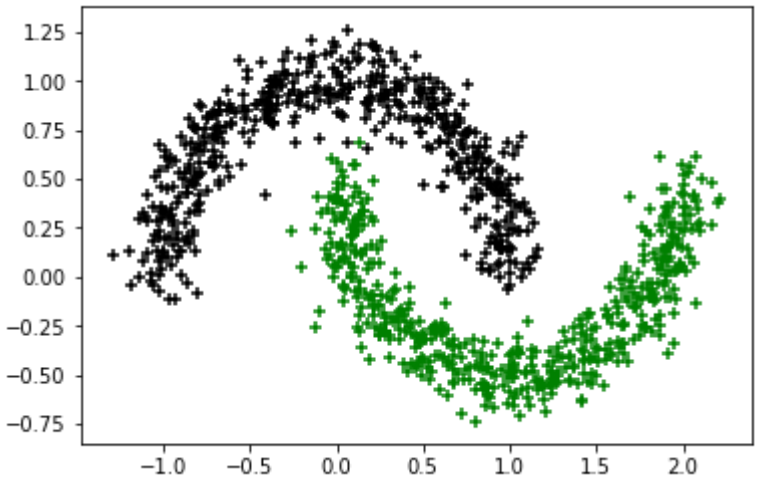
Out[141]:

	x1	x2	label
count	1000.000000	1000.000000	1000.00000
mean	-3.162222	-0.084438	0.50000
std	10.287057	12.533138	0.50025
min	-21.900000	-23.200000	0.00000
25%	-11.500000	-10.600000	0.00000
50%	-3.785000	0.359000	0.50000
75%	3.810000	10.500000	1.00000
max	17.400000	23.400000	1.00000

The above dataset values plotted on the graph separated by different colors 'grey' and 'red' in the above case, the figure shows the scatter describing the classification having some parametric values like mean of x1 = -3.162222 and that of x2 = -0.084438 and there by labelled as 0.5 because as per the scatter graph the values on x and y axis are present between -20 to +20 so all other answers has to be between then in the particular range, also the dataset comprises of standard deviation using the formula of approximate upto 10.287057 for x1 and 12.533138 for x2 thereby labelled as 0.50025 thus explaining that standard deviation is more inclined towards 1 also the max for x1 and x2 is shown as 17.400000 and 23.400000, similarly the minimum values are -21.900000 and -23.200000 for x1 and x2, these values exeeding as seen in the graph few points are scattered out of bound

In [128]:

```
dataset3 = pd.read_csv('C:\Python37\datasets\moons1.csv')
df3=pd.DataFrame(dataset3)
colors=['black' if label==0 else 'green' for label in df3.label]
plt.scatter(df3.x1,df3.x2,c=colors,marker='+')
plt.show()
```



In [142]:

```
df3.describe()
```

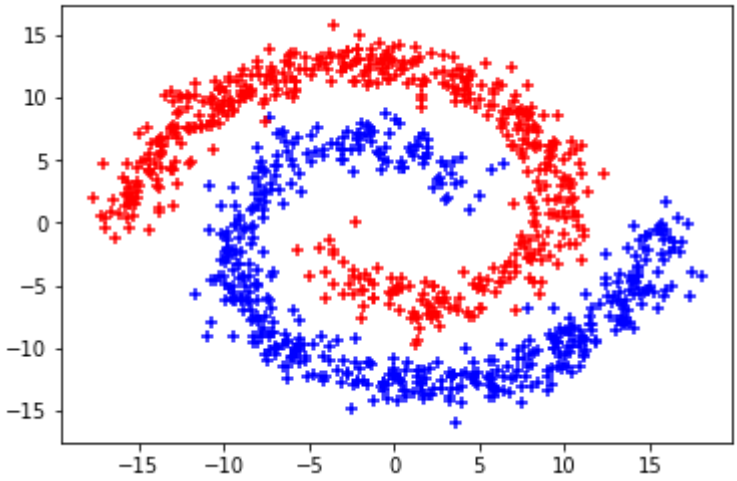
Out[142]:

	x1	x2	label
count	1000.000000	1000.000000	1000.00000
mean	0.502575	0.252113	0.50000
std	0.870595	0.507203	0.50025
min	-1.290328	-0.736386	0.00000
25%	-0.036783	-0.201826	0.00000
50%	0.503604	0.246921	0.50000
75%	1.055629	0.698716	1.00000
max	2.208051	1.265113	1.00000

The above dataset values plotted on the graph separated by different colors 'black' and 'green' in the above case, the figure shows the scatter describing the classification having some parametric values like mean of x1 = 0.502575 and that of x2 = 0.252113 and there by labelled as 0.5 because as per the scatter graph the values on x and y axis are present between -0.75 to 1.25 for y axis and -1.0 to 2.0 (approx) on x a xis so all other answers has to be between then in the particular range, also the dataset comprises of standard deviation using the formula of approximate upto for x1 0.870595 and 0.507203 for x2 thereby labelled as 0.50025 thus explaining that standard deviation is more inclined towards 1 also the max for x1 and x2 is shown as 2.208051 and 1.265113, similarly the minimum values are -1.290328 and -0.736386 for x1 and x2, these values exeeding as seen in the graph few points are scattered out of bound

In [129]:

```
dataset4 = pd.read_csv('C:\Python37\datasets\spiral1.csv')
df4=pd.DataFrame(dataset4)
colors=['blue' if label==0 else 'red' for label in df4.label]
plt.scatter(df4.x,df4.y,c=colors,marker='+')
plt.show()
```



In [143]:

```
df4.describe()
```

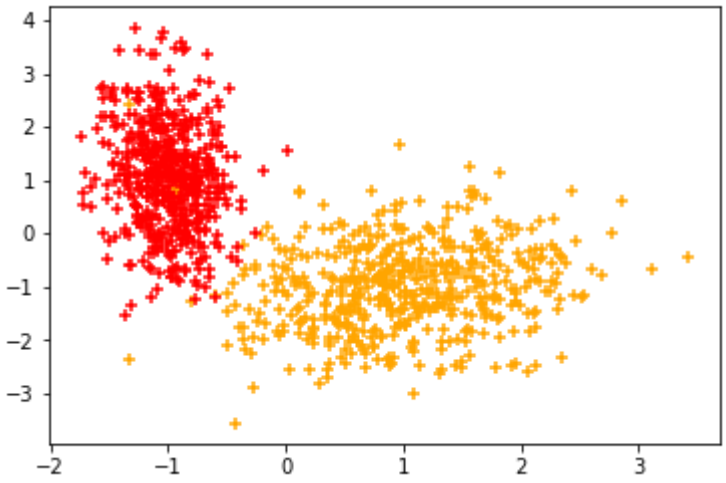
Out[143]:

	x	y	label
count	1000.000000	1000.000000	1000.00000
mean	0.031292	-0.067469	0.50000
std	8.577650	8.405857	0.50025
min	-17.734170	-15.939320	0.00000
25%	-7.374865	-7.330285	0.00000
50%	0.099885	-0.243010	0.50000
75%	7.449595	7.260847	1.00000
max	18.057730	15.734560	1.00000

The above dataset values plotted on the graph separated by different colors 'blue' and 'red' in the above case, the figure shows the scatter describing the classification having some parametric values like mean of x = 0.031292 and that of y = -0.067469 and there by labelled as 0.5 because as per the scatter graph the values on x and y axis are present between -15 to 15 (approx) for both axis so all other answers has to be between then in the particular range, also the dataset comprises of standard deviation using the formula of approximate upto for x 0.870595 and 0.507203 for y thereby labelled as 0.50025 thus explaining that standard deviation is more inclined towards 1 also the maximum for x and y is shown as 18.057730 and 15.734560, similarly the minimum values are -17.734170 and -0.15.939320 for x and y, these values exeeding as seen in the graph few points are scattered out of bound

In [130]:

```
dataset5 = pd.read_csv('C:\\Python37\\datasets\\twogaussians33.csv')
df5=pd.DataFrame(dataset5)
colors=['red' if label==0 else 'orange' for label in df5.label]
plt.scatter(df5.x1,df5.x2,c=colors,marker='+')
plt.show()
```



In [151]:

```
df5.describe()
```

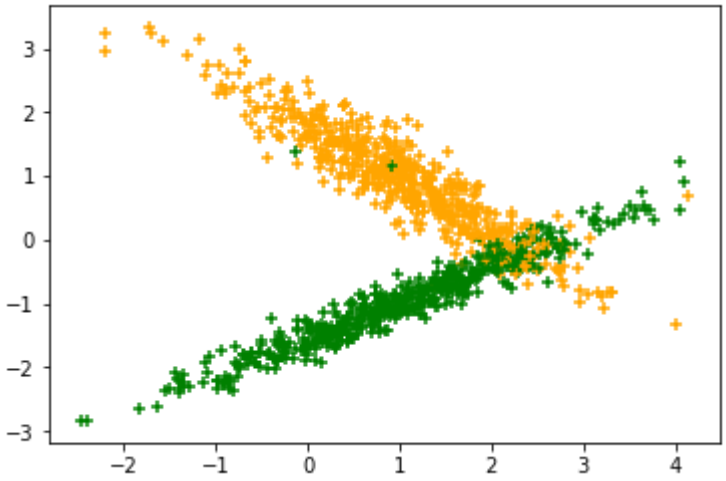
Out[151]:

	x1	x2	label
count	1000.000000	1000.000000	1000.000000
mean	0.001293	0.013081	0.502000
std	1.142285	1.364468	0.500246
min	-1.745742	-3.558012	0.000000
25%	-0.991735	-1.044560	0.000000
50%	-0.433449	-0.096095	1.000000
75%	0.969287	1.055700	1.000000
max	3.417897	3.862728	1.000000

The above dataset values plotted on the graph separated by different colors 'red' and 'yellow' in the above case, the figure shows the scatter describing the classification having some parametric values like mean of x1 = 0.001293 and that of x2 = 0.013081 and there by labelled as 0.5 because as per the scatter graph the values on x and y axis are present between -2 to 3 for y axis and -2 to 4 (approx) on x axis so all other answers has to be between then in the particular range, also the dataset comprises of standard deviation using the formula of approximate upto for x1 1.142285 and 1.364468 for x2 thereby labelled as 0.500246 thus explaining that standard deviation is more inclined towards 1 also the max for x1 and x2 is shown as 3.417897 and 3.862728, similarly the minimum values are -1.745742 and -3.558012 for x1 and x2, these values exeeding as seen in the graph few points are scattered out of bound

In [131]:

```
dataset6 = pd.read_csv('C:\\Python37\\datasets\\twogaussians42.csv')
df6=pd.DataFrame(dataset6)
colors=['green' if label==0 else 'orange' for label in df6.label]
plt.scatter(df6.x1,df6.x2,c=colors,marker='+')
plt.show()
```



In [152]:

```
df6.describe()
```

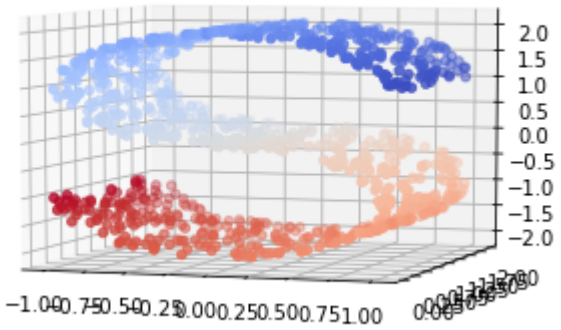
Out[152]:

	x1	x2	label
count	1000.000000	1000.000000	1000.000000
mean	1.025840	-0.012693	0.499000
std	1.071457	1.225378	0.500249
min	-2.472718	-2.850971	0.000000
25%	0.307209	-0.984268	0.000000
50%	1.023750	-0.102945	0.000000
75%	1.724713	0.973550	1.000000
max	4.138715	3.342864	1.000000

The above dataset values plotted on the graph separated by different colors 'green' and 'yellow' in the above case, the figure shows the scatter describing the classification having some parametric values like mean of x1 = 1.025840 and that of x2 = -0.012693 and there by labelled as 0.5(approx) because as per the scatter graph the values on x and y axis are present between -3 to 3 for y axis and -2 to 4 (approx) on x axis so all other answers has to be between then in the particular range, also the dataset comprises of standard deviation using the formula of approximate upto for x1 1.071457 and 1.225378 for x2 thereby labelled as 0.500249 thus explaining that standard deviation is more inclined towards 1 also the max for x1 and x2 is shown as 4.138715 and 3.342864, similarly the minimum values are -2.472718 and -2.850970 for x1 and x2, these values exeeding as seen in the graph few points are scattered out of bound

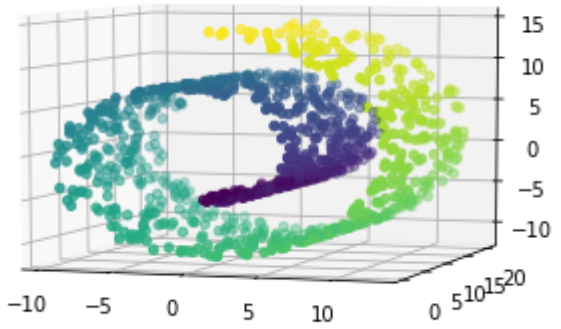
In [132]:

```
import sklearn.datasets
a,color=sklearn.datasets.make_s_curve(n_samples=1000, noise=0.0, random_state=0)
fig = plt.figure()
sroll = fig.add_subplot(111, projection='3d')
sroll.scatter(a[:, 0], a[:, 1], a[:, 2], c=color, cmap=plt.cm.coolwarm)
sroll.view_init(6,-72)
```



In [133]:

```
b,color=sklearn.datasets.make_swiss_roll(n_samples=1000, noise=0.0, random_state=None)
fig = plt.figure()
swissroll = fig.add_subplot(111, projection='3d')
swissroll.scatter(b[:, 0], b[:, 1], b[:, 2], c=color, cmap=plt.cm.viridis)
swissroll.view_init(6,-72)
```



In [ ]:

```
```