In [1]:
```
from sklearn.datasets import make_classification
```

In [2]:
```
#x independent features
#y dependant features
x,y=make_classification(n_samples=1000,n_features=2,n_redundant=0,n_clusters_per_class=1,weights=[0.90],random_state=
```

In [3]:
```
x.shape
```

Out[3]:
```
(1000, 2)
```

In [9]:
```
import pandas as pd
df1=pd.DataFrame(x,columns=['f1','f2'])
```

In [11]:
```
df2=pd.DataFrame(y,columns=['target'])
```

In [13]:
```
final_df=pd.concat([df1,df2],axis=1)
```

In [14]:
```
final_df.head()
```

Out[14]:

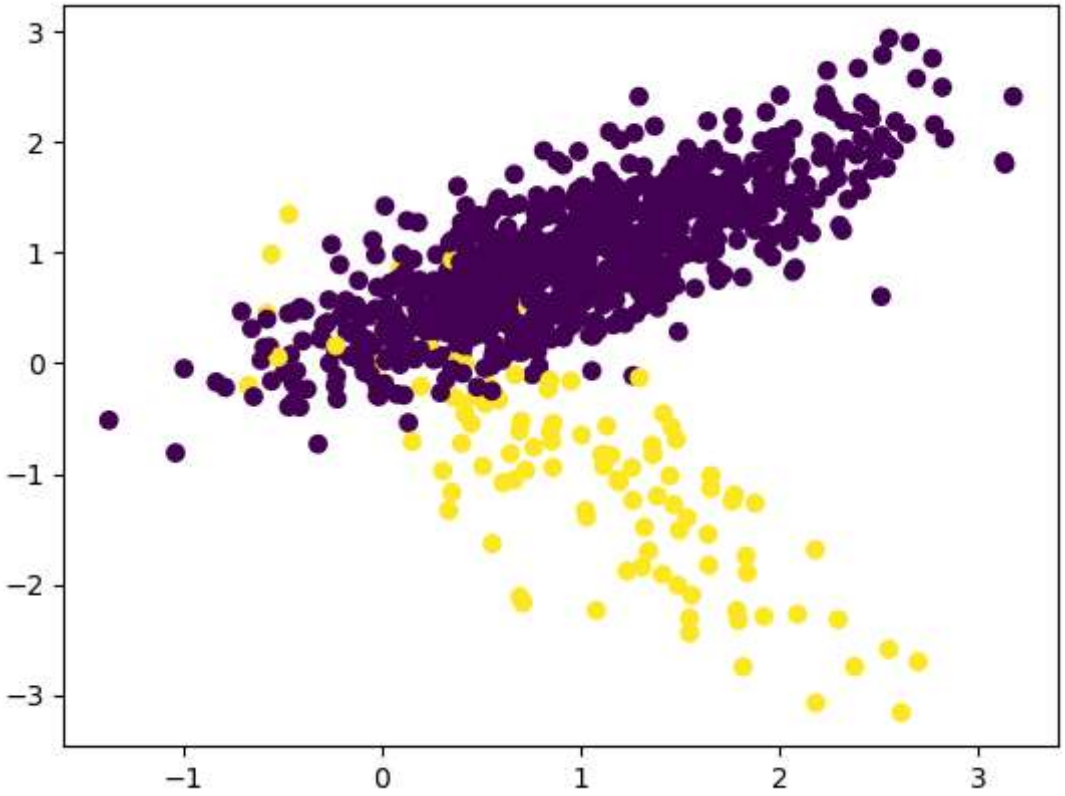
	f1	f2	target
0	1.536830	-1.398694	1
1	1.551108	1.810329	0
2	1.293619	1.010946	0
3	1.119889	1.632518	0
4	1.042356	1.121529	0

In [16]:
```
final_df['target'].value_counts()
```

Out[16]:
```
0    894
1    106
Name: target, dtype: int64
```

In [17]:
```
import matplotlib.pyplot as plt
plt.scatter(final_df['f1'],final_df['f2'],c=final_df['target'])
```

Out[17]:
```
<matplotlib.collections.PathCollection at 0x7f85e169a110>
```



In [18]:
```
!pip install imblearn
```

```
Collecting imblearn
  Downloading imblearn-0.0-py2.py3-none-any.whl (1.9 kB)
Collecting imbalanced-learn
  Downloading imbalanced_learn-0.10.1-py3-none-any.whl (226 kB)
     ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 226.0/226.0 kB 1.5 MB/s eta 0:00:00a 0:00:01
```

```
Requirement already satisfied: scikit-learn>=1.0.2 in /opt/conda/lib/python3.10/site-packages (from imbalanced-learn->imblearn) (1.2.0)
Requirement already satisfied: scipy>=1.3.2 in /opt/conda/lib/python3.10/site-packages (from imbalanced-learn->imblearn) (1.9.3)
Requirement already satisfied: threadpoolctl>=2.0.0 in /opt/conda/lib/python3.10/site-packages (from imbalanced-learn->imblearn) (3.1.0)
Requirement already satisfied: joblib>=1.1.1 in /opt/conda/lib/python3.10/site-packages (from imbalanced-learn->imblearn) (1.2.0)
Requirement already satisfied: numpy>=1.17.3 in /opt/conda/lib/python3.10/site-packages (from imbalanced-learn->imblearn) (1.23.5)
Installing collected packages: imbalanced-learn, imblearn
Successfully installed imbalanced-learn-0.10.1 imblearn-0.0
```

In [19]:
```
from imblearn.over_sampling import SMOTE
```

In [20]:
```
##transform dataset
oversample=SMOTE()
x,y=oversample.fit_resample(final_df[['f1','f2']],final_df['target'])
```

In [21]:
```
x.shape
```

Out[21]:
```
(1788, 2)
```

In [22]:
```
y.shape
```

Out[22]:
```
(1788,)
```

In [26]:
```
y[y==0]
```

Out[26]:
```
1      0
2      0
3      0
4      0
5      0
      ..
995    0
996    0
```

```
997    0
998    0
999    0
Name: target, Length: 894, dtype: int64
```

In [27]:
```
len(y[y==0])
```

Out[27]: 894

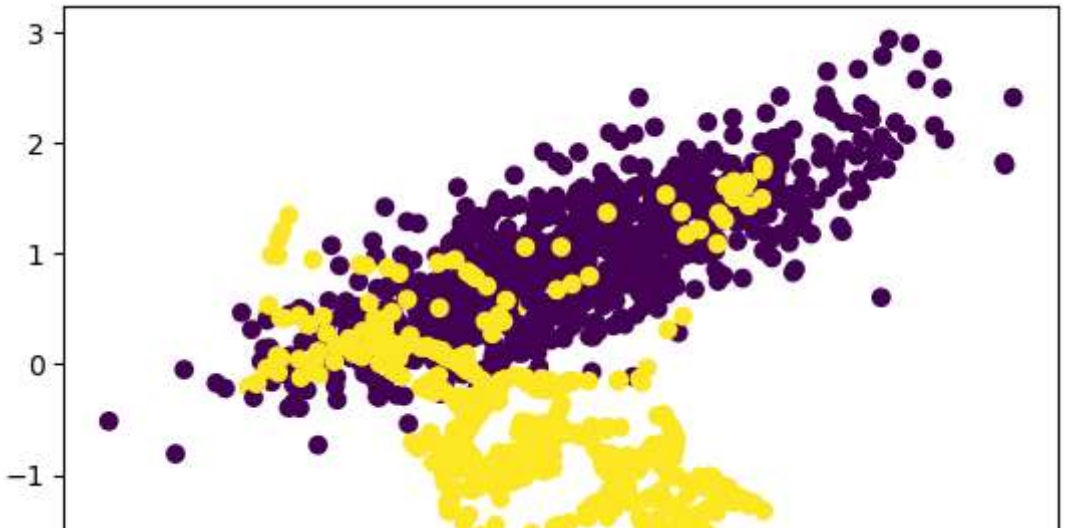
In [28]:
```
len(y[y==1])
```

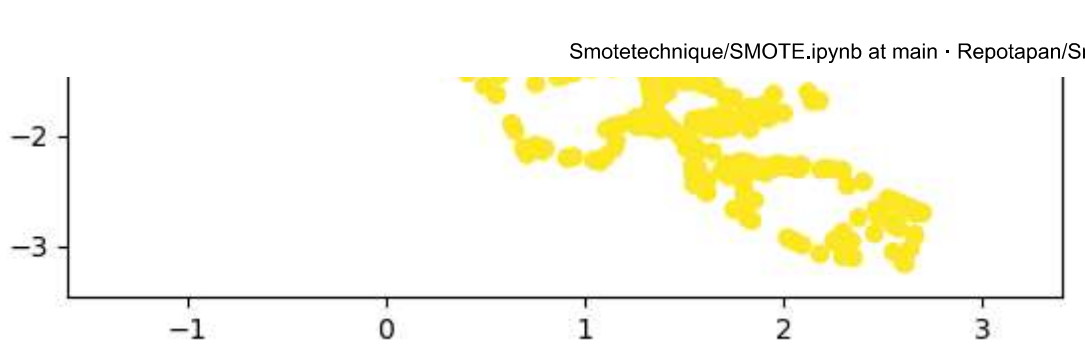
Out[28]: 894

In [29]:
```
import pandas as pd
df1=pd.DataFrame(x,columns=['f1','f2'])
df2=pd.DataFrame(y,columns=['target'])
oversampled_df=pd.concat([df1,df2],axis=1)
```

In [32]:
```
import matplotlib.pyplot as plt
plt.scatter(oversampled_df['f1'],oversampled_df['f2'],c=oversampled_df['target'])
```

Out[32]: <matplotlib.collections.PathCollection at 0x7f85f70d9d50>




−2
−3
−1
0
1
2
3