Experiment No: 05

```
[2] import pandas as pd
    import seaborn as sns
    dataset=pd.read_csv("admission.csv")
```

```
dataset
```

	admit	gre	gpa	rank
0	0.0	380.0	3.61	3.0
1	1.0	660.0	3.67	3.0
2	1.0	800.0	4.00	1.0
3	1.0	640.0	3.19	4.0
4	0.0	520.0	2.93	4.0
...	...	...	...	...
395	0.0	620.0	4.00	2.0
396	0.0	560.0	3.04	3.0
397	0.0	460.0	2.63	2.0
398	0.0	700.0	3.65	2.0
399	0.0	600.0	3.89	3.0

400 rows × 4 columns

```
[4] X=dataset.iloc[:,1:4]
```

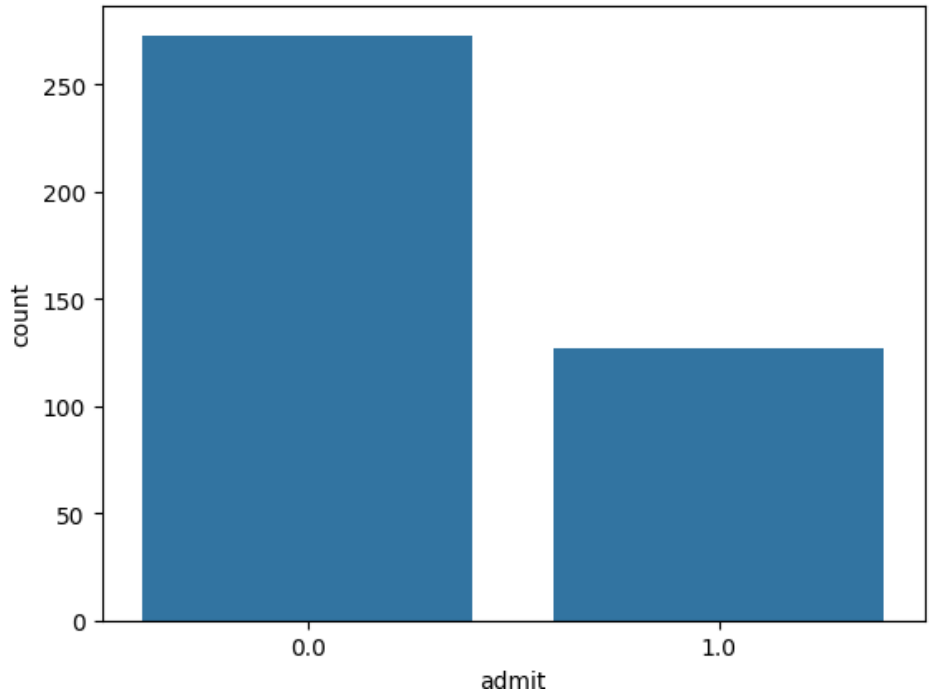
```
[5] Y=dataset.iloc[:,0:1]
```

```
[6] Y.value_counts()
```

```
admit
0.0      273
1.0      127
dtype: int64
```

```
sns.countplot(x="admit",data=dataset)
```

```
<Axes: xlabel='admit', ylabel='count'>
```



```
[8] from sklearn.model_selection import train_test_split
    X_train,X_test,Y_train,Y_test=train_test_split(X,Y,test_size=0.3,random_state=0)
```

```
    len(X_train)
    len(Y_test)
```

```
120
```

```
[10] from imblearn.over_sampling import RandomOverSampler
     ros=RandomOverSampler()
     X_ros,Y_ros=ros.fit_resample(X_train,Y_train)
```

```
[11] len(Y_ros)
```

```
382
```

```
[12] Y_ros.value_counts()
```

```
admit
0.0      191
1.0      191
dtype: int64
```

```
[13] from imblearn.under_sampling import RandomUnderSampler
     rus=RandomOverSampler()
     X_rus,Y_rus=rus.fit_resample(X_train,Y_train)
```

```
[14] len(Y_rus)
```

```
382
```

```
Y_rus.value_counts()
```

```
admit
0.0      191
1.0      191
dtype: int64
```

```
[16] from imblearn.over_sampling import SMOTE
     X_smote,Y_smote=SMOTE(k_neighbors=3).fit_resample(X_train,Y_train)
     Y_smote.value_counts()
```

```
admit
0.0      191
1.0      191
dtype: int64
```

Conclusion: Thus, in this experiment we have implemented SMOTE technique to generate synthetic data.(to solve the problem of class imbalance)