In [1]:

```
import pandas as pd
from sklearn.linear_model import LogisticRegression

from sklearn.preprocessing import StandardScaler
```

In [2]:

```
df=pd.read_csv("C2_test.gender_submission.csv")
df1=pd.read_csv("C2_train.gender_submission.csv")
```

In [3]:

```
df_=df.drop(["Cabin","Name","Embarked","Ticket","PassengerId","Sex"],axis=1)
df1_=df1.drop(["Survived","Cabin","Name","Embarked","Ticket","PassengerId"],axis=1)
print(df_)
print(df1_)
```

```
     Pclass   Age  SibSp  Parch      Fare
0         3  34.5      0      0    7.8292
1         3  47.0      1      0    7.0000
2         2  62.0      0      0    9.6875
3         3  27.0      0      0    8.6625
4         3  22.0      1      1   12.2875
..      ...   ...    ...    ...       ...
413       3   NaN      0      0    8.0500
414       1  39.0      0      0  108.9000
415       3  38.5      0      0    7.2500
416       3   NaN      0      0    8.0500
417       3   NaN      1      1   22.3583

[418 rows x 5 columns]
     Pclass     Sex   Age  SibSp  Parch     Fare
0         3    male  22.0      1      0   7.2500
1         1  female  38.0      1      0  71.2833
2         3  female  26.0      0      0   7.9250
3         1  female  35.0      1      0  53.1000
4         3    male  35.0      0      0   8.0500
..      ...     ...   ...    ...    ...      ...
886       2    male  27.0      0      0  13.0000
887       1  female  19.0      0      0  30.0000
888       3  female   NaN      1      2  23.4500
889       1    male  26.0      0      0  30.0000
890       3    male  32.0      0      0   7.7500

[891 rows x 6 columns]
```

In [4]:

```
df1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891 entries, 0 to 890
Data columns (total 12 columns):
 #   Column       Non-Null Count  Dtype  
---  ------       --------------  -----  
 0   PassengerId  891 non-null    int64  
 1   Survived     891 non-null    int64  
 2   Pclass       891 non-null    int64  
 3   Name         891 non-null    object 
 4   Sex          891 non-null    object 
 5   Age          714 non-null    float64
 6   SibSp        891 non-null    int64  
 7   Parch        891 non-null    int64  
 8   Ticket       891 non-null    object 
 9   Fare         891 non-null    float64
 10  Cabin        204 non-null    object 
 11  Embarked     889 non-null    object 
dtypes: float64(2), int64(5), object(5)
memory usage: 83.7+ KB
```

In [5]:

```
df_=df_.dropna()
df1_=df1_.dropna()
df1_.info()
df_.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 714 entries, 0 to 890
Data columns (total 6 columns):
 #   Column  Non-Null Count  Dtype  
---  ------  --------------  -----  
 0   Pclass  714 non-null    int64  
 1   Sex     714 non-null    object 
 2   Age     714 non-null    float64
 3   SibSp   714 non-null    int64  
 4   Parch   714 non-null    int64  
 5   Fare    714 non-null    float64
dtypes: float64(2), int64(3), object(1)
memory usage: 39.0+ KB
<class 'pandas.core.frame.DataFrame'>
Int64Index: 331 entries, 0 to 415
Data columns (total 5 columns):
 #   Column  Non-Null Count  Dtype  
---  ------  --------------  -----  
 0   Pclass  331 non-null    int64  
 1   Age     331 non-null    float64
 2   SibSp   331 non-null    int64  
 3   Parch   331 non-null    int64  
 4   Fare    331 non-null    float64
dtypes: float64(2), int64(3)
memory usage: 15.5 KB
```

In [6]:
```
y=df1_["Sex"]
x=df1_.drop(["Sex"],axis=1)
f=StandardScaler().fit_transform(x)
lo=LogisticRegression()
lo.fit(f,y)
```

Out[6]:
```
LogisticRegression()
```

In [7]:
```
lo.predict(df_)
```

Out[7]:
```
array(['male', 'male', 'male', 'male', 'male', 'male', 'male', 'male',
       'male', 'male', 'male', 'female', 'male', 'male', 'male', 'male',
       'male', 'male', 'male', 'male', 'male', 'female', 'female', 'male',
       'female', 'male', 'male', 'male', 'male', 'male', 'male', 'male',
       'male', 'male', 'male', 'male', 'male', 'male', 'male', 'male',
       'male', 'male', 'male', 'male', 'male', 'female', 'male', 'male',
       'male', 'female', 'male', 'male', 'male', 'male', 'female', 'male',
       'male', 'male', 'female', 'male', 'male', 'male', 'male', 'female',
       'female', 'male', 'male', 'male', 'male', 'female', 'male', 'male',
       'male', 'female', 'male', 'male', 'male', 'male', 'male', 'male',
       'male', 'male', 'male', 'male', 'male', 'male', 'male', 'male',
       'male', 'male', 'male', 'male', 'female', 'male', 'female', 'male',
       'male', 'male', 'male', 'male', 'male', 'male', 'male', 'male',
       'male', 'male', 'male', 'male', 'male', 'male', 'male', 'male',
       'female', 'female', 'female', 'male', 'male', 'male', 'male',
       'male', 'female', 'male', 'male', 'male', 'female', 'male', 'male',
       'male', 'male', 'male', 'male', 'male', 'male', 'male', 'male',
       'male', 'male', 'male', 'male', 'male', 'male', 'male', 'male',
       'male', 'male', 'female', 'female', 'male', 'male', 'male', 'male',
       'male', 'male', 'male', 'female', 'male', 'female', 'male', 'male',
       'female', 'female', 'male', 'male', 'male', 'male', 'male', 'male',
       'male', 'female', 'male', 'male', 'male', 'female', 'female',
       'male', 'male', 'male', 'male', 'male', 'male', 'male', 'male',
       'female', 'male', 'male', 'male', 'male', 'male', 'male', 'male',
       'male', 'male', 'male', 'female', 'male', 'male', 'male', 'male',
       'female', 'male', 'female', 'male', 'male', 'male', 'male', 'male',
       'male', 'male', 'male', 'male', 'male', 'male', 'male', 'female',
       'male', 'male', 'male', 'male', 'male', 'male', 'female', 'male',
       'female', 'male', 'female', 'male', 'male', 'male', 'male',
       'female', 'male', 'male', 'male', 'male', 'male', 'male', 'female',
       'male', 'male', 'male', 'male', 'male', 'male', 'male', 'male',
       'male', 'male', 'male', 'male', 'male', 'male', 'male', 'male',
       'female', 'male', 'female', 'male', 'male', 'male', 'male', 'male',
       'male', 'male', 'male', 'male', 'male', 'male', 'male', 'male',
       'female', 'male', 'male', 'male', 'male', 'male', 'male', 'male',
       'female', 'male', 'female', 'male', 'male', 'male', 'female',
       'male', 'male', 'male', 'male', 'male', 'male', 'male', 'male',
       'female', 'male', 'male', 'male', 'female', 'male', 'male', 'male',
       'female', 'male', 'male', 'female', 'male', 'male', 'male', 'male',
       'female', 'male', 'male', 'male', 'male', 'female', 'male', 'male',
       'male', 'male', 'female', 'male', 'female', 'male', 'male', 'male',
       'male', 'female', 'male', 'male', 'male', 'female', 'male'],
      dtype=object)
```

In [8]:

```
obs=[[1,23,1,1,3232]]
lo.predict(obs)
```

Out[8]:

```
array(['female'], dtype=object)
```

In [9]:

```
lo.predict_proba(obs)
```

Out[9]:

```
array([[1.00000000e+000, 3.62458485e-161]])
```

In [ ]: