In [98]:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import accuracy_score
from sklearn import svm
```

In [99]:

```
train_data=pd.read_csv("train.csv")
test_data=pd.read_csv("test.csv")
```

In [100…

```
train_data.head()
```

Out[100]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	E
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	NaN	
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN	

In [101…

```
train_data.describe()
```

Out[101]:

	PassengerId	Survived	Pclass	Age	SibSp	Parch	Fare
count	891.000000	891.000000	891.000000	714.000000	891.000000	891.000000	891.000000
mean	446.000000	0.383838	2.308642	29.699118	0.523008	0.381594	32.204208
std	257.353842	0.486592	0.836071	14.526497	1.102743	0.806057	49.693429
min	1.000000	0.000000	1.000000	0.420000	0.000000	0.000000	0.000000
25%	223.500000	0.000000	2.000000	20.125000	0.000000	0.000000	7.910400
50%	446.000000	0.000000	3.000000	28.000000	0.000000	0.000000	14.454200
75%	668.500000	1.000000	3.000000	38.000000	1.000000	0.000000	31.000000
max	891.000000	1.000000	3.000000	80.000000	8.000000	6.000000	512.329200

In [102...

```
train_data=train_data.drop(columns="PassengerId",axis=1)
```

In [110...

```
train_data['Sex']=train_data['Sex'].map({'male':1,'female':0})
train_data['Embarked']=train_data['Embarked'].map({'C':1,'S':2,'Q':3})
train_data['Cabin']=train_data['Cabin'].map(lambda x:0 if pd.isna(x) else 1)
```

In [111...

```
train_data.head()
```

Out[111]:

	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	0	3	Braund, Mr. Owen Harris	1	22.0	1	0	A/5 21171	7.2500	0	2.0
1	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	0	38.0	1	0	PC 17599	71.2833	1	1.0
2	1	3	Heikkinen, Miss. Laina	0	26.0	0	0	STON/O2. 3101282	7.9250	0	2.0
3	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	0	35.0	1	0	113803	53.1000	1	2.0
4	0	3	Allen, Mr. William Henry	1	35.0	0	0	373450	8.0500	0	2.0

In [112...

```
x=train_data.drop(columns=["Survived","Name","Ticket"],axis=1)
```

In [113...

```
y=train_data["Survived"]
```

In [114...

```
print(x.head())
```

```
   Pclass  Sex   Age  SibSp  Parch     Fare  Cabin  Embarked
0       3    1  22.0      1      0   7.2500      0       2.0
1       1    0  38.0      1      0  71.2833      1       1.0
2       3    0  26.0      0      0   7.9250      0       2.0
3       1    0  35.0      1      0  53.1000      1       2.0
4       3    1  35.0      0      0   8.0500      0       2.0
```

In [115…

```
x.isna().sum()
```

Out[115]:

```
Pclass        0
Sex           0
Age         177
SibSp         0
Parch         0
Fare          0
Cabin         0
Embarked      2
dtype: int64
```

In [118…

```
x['Age']=x['Age'].fillna(x['Age'].mean())
x['Embarked']=x["Embarked"].fillna(x["Embarked"].max())
```

In [119…

```
x.isna().sum()
```

Out[119]:

```
Pclass      0
Sex         0
Age         0
SibSp       0
Parch       0
Fare        0
Cabin       0
Embarked    0
dtype: int64
```

In [120…

```
y.value_counts()
```

Out[120]:

```
Survived
0    549
1    342
Name: count, dtype: int64
```

In [146…

```
y.head()
```

Out[146]:

```
0    0
1    1
2    1
3    1
4    0
Name: Survived, dtype: int64
```

In [121…

```
scaler=StandardScaler()
```

In [122…

```
standardized_train_data=scaler.fit_transform(x)
```

In [123…

```
print(standardized_train_data)
```

```
[[ 0.82737724  0.73769513 -0.5924806  ... -0.50244517 -0.54492498
   0.19322457]
 [-1.56610693 -1.35557354  0.63878901 ...  0.78684529  1.835115
  -1.74119218]
 [ 0.82737724 -1.35557354 -0.2846632  ... -0.48885426 -0.54492498
   0.19322457]
 ...
 [ 0.82737724 -1.35557354  0.         ... -0.17626324 -0.54492498
   0.19322457]
 [-1.56610693  0.73769513 -0.2846632  ... -0.04438104  1.835115
  -1.74119218]
 [ 0.82737724  0.73769513  0.17706291 ... -0.49237783 -0.54492498
   2.12764132]]
```

In [124…

```
x=standardized_train_data
```

In [125…

```
classifier=svm.SVC(kernel="linear")
```

In [126…

```
classifier.fit(x,y)
```

Out[126]:

```
▾          SVC
SVC(kernel='linear')
```

In [128…

```
prediction=classifier.predict(x)
```

In [129…

```
accuracy=accuracy_score(prediction,y)
```

In [130…

```
accuracy
```

Out[130]:

```
0.7867564534231201
```

In [184…

```
test_data=pd.read_csv("test.csv")
```

In [185…

```
test_data.head()
```

Out[185]:

	PassengerId	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	892	3	Kelly, Mr. James	male	34.5	0	0	330911	7.8292	NaN	Q
1	893	3	Wilkes, Mrs. James (Ellen Needs)	female	47.0	1	0	363272	7.0000	NaN	S
2	894	2	Myles, Mr. Thomas Francis	male	62.0	0	0	240276	9.6875	NaN	Q
3	895	3	Wirz, Mr. Albert	male	27.0	0	0	315154	8.6625	NaN	S
4	896	3	Hirvonen, Mrs. Alexander (Helga E Lindqvist)	female	22.0	1	1	3101298	12.2875	NaN	S

In [186…

```
test_data=test_data.drop(columns=["Name","Ticket","PassengerId"])
```

In [187…

```
test_data.head()
```

Out[187]:

	Pclass	Sex	Age	SibSp	Parch	Fare	Cabin	Embarked
0	3	male	34.5	0	0	7.8292	NaN	Q
1	3	female	47.0	1	0	7.0000	NaN	S
2	2	male	62.0	0	0	9.6875	NaN	Q
3	3	male	27.0	0	0	8.6625	NaN	S
4	3	female	22.0	1	1	12.2875	NaN	S

In [188…

```
test_data['Sex']=test_data['Sex'].map({'male':1,'female':0})
test_data['Embarked']=test_data['Embarked'].map({'C':1,'S':2,'Q':3})
test_data['Cabin']=test_data['Cabin'].map(lambda x:0 if pd.isna(x) else 1)
```

In [189…

```
test_data.isna().sum()
```

Out[189]:

```
Pclass       0
Sex          0
Age         86
SibSp        0
Parch        0
Fare         1
Cabin        0
Embarked     0
dtype: int64
```

In [190…

```
test_data['Age']=test_data['Age'].fillna(test_data['Age'].mean())
test_data['Fare']=test_data["Fare"].fillna(test_data["Fare"].mean())
```

In [191…

```
test_data.isna().sum()
```

Out[191]:

```
Pclass      0
Sex         0
Age         0
SibSp       0
Parch       0
Fare        0
Cabin       0
Embarked    0
dtype: int64
```

In [192…

```
test_data.head()
```

Out[192]:

	Pclass	Sex	Age	SibSp	Parch	Fare	Cabin	Embarked
0	3	1	34.5	0	0	7.8292	0	3
1	3	0	47.0	1	0	7.0000	0	2
2	2	1	62.0	0	0	9.6875	0	3
3	3	1	27.0	0	0	8.6625	0	2
4	3	0	22.0	1	1	12.2875	0	2

In [193…

```
test_data["Embarked"]=test_data["Embarked"].astype(float)
```

In [194…

```
test_data.head()
```

Out[194]:

	Pclass	Sex	Age	SibSp	Parch	Fare	Cabin	Embarked
0	3	1	34.5	0	0	7.8292	0	3.0
1	3	0	47.0	1	0	7.0000	0	2.0
2	2	1	62.0	0	0	9.6875	0	3.0
3	3	1	27.0	0	0	8.6625	0	2.0
4	3	0	22.0	1	1	12.2875	0	2.0

In [195…

```
standardized_test_data=scaler.fit_transform(test_data)
```

In [196…

```
print(standardized_test_data)
```

```
[[ 0.87348191  0.75592895  0.3349926  ... -0.49840706 -0.52752958
   1.95594094]
 [ 0.87348191 -1.32287566  1.32553003 ... -0.51327429 -0.52752958
   0.23108163]
 [-0.31581919  0.75592895  2.51417495 ... -0.46508846 -0.52752958
   1.95594094]
 ...
 [ 0.87348191  0.75592895  0.65196458 ... -0.50879189 -0.52752958
   0.23108163]
 [ 0.87348191  0.75592895  0.         ... -0.4944482  -0.52752958
   0.23108163]
 [ 0.87348191  0.75592895  0.         ... -0.23790598 -0.52752958
  -1.49377768]]
```

In [197…

```python
test_prediction=classifier.predict(standardized_test_data)
```

In [198…

```python
test_prediction
```

Out[198]:

```
array([0, 1, 0, 0, 1, 0, 1, 0, 1, 0, 0, 0, 1, 0, 1, 1, 0, 0, 1, 1, 0, 0,
       1, 0, 1, 0, 1, 0, 0, 0, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 0, 0, 0, 1,
       1, 0, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 1,
       1, 0, 0, 1, 1, 0, 1, 0, 1, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 1,
       1, 0, 1, 0, 1, 0, 0, 0, 1, 0, 1, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0,
       0, 1, 1, 1, 1, 0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 0, 0, 0, 0,
       1, 0, 0, 0, 0, 0, 1, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1,
       0, 0, 1, 1, 0, 1, 1, 0, 1, 0, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 0, 1,
       1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1,
       0, 1, 1, 0, 0, 1, 0, 0, 1, 0, 1, 0, 0, 0, 0, 1, 1, 0, 1, 0, 1, 0,
       1, 0, 1, 0, 1, 1, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1,
       0, 0, 0, 0, 1, 0, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 1,
       0, 0, 0, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 0, 0, 1, 0, 1, 1, 1, 0,
       0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0,
       0, 1, 0, 0, 0, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0,
       1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 1, 0, 1, 1, 0,
       0, 0, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 0, 0,
       1, 1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1,
       0, 1, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 0, 1, 0, 0, 0],
      dtype=int64)
```

In [199…

```python
test_output=pd.read_csv("gender_submission.csv")
```

In [200…

```python
test_output.head()
```

Out[200]:

	PassengerId	Survived
0	892	0
1	893	1
2	894	0
3	895	0
4	896	1

In [201…

```python
test_output=test_output["Survived"]
```

In [202…

```python
test_output.head()
```

Out[202]:

```
0    0
1    1
2    0
3    0
4    1
Name: Survived, dtype: int64
```

In [203…

```
accuracy_test=accuracy_score(test_prediction,test_output)
```

In [204…

```
print(accuracy_test)
```

```
1.0
```

In [ ]:

```
```