Assignment 1 Set A3

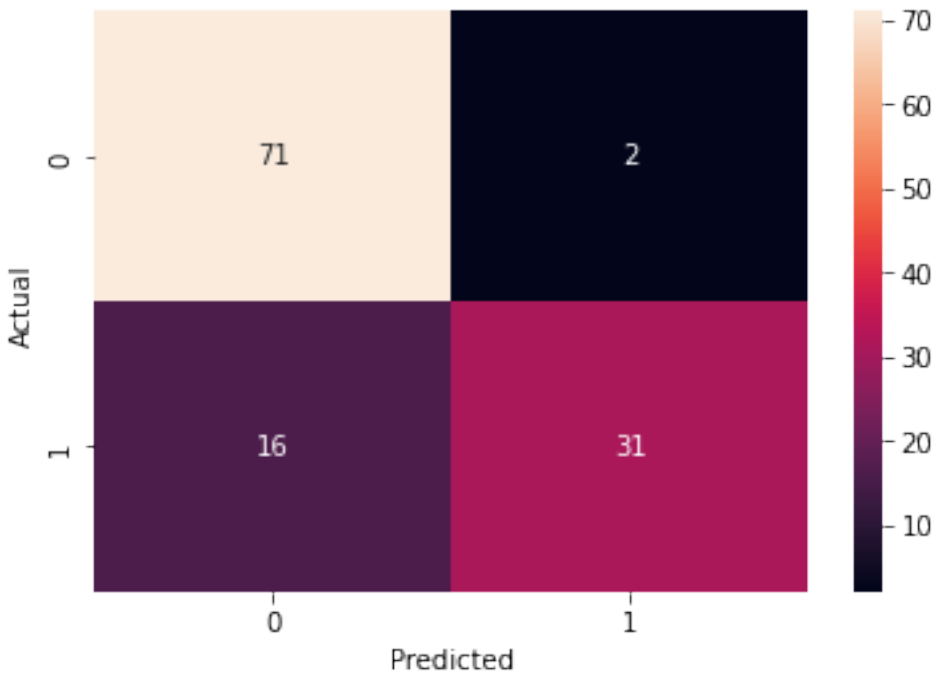
March 2, 2024

1 Create 'User' Data set having 5 columns namely: User ID, Gender, Age, EstimatedSalary and Purchased. Build a logistic regression model that can predict whether on the given parameter a person will buy a car or not.

[2]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sn
from sklearn.linear_model import LogisticRegression
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import accuracy_score,␣
↪classification_report,confusion_matrix

user_df = pd.read_csv('/home/mmcc/Desktop/DA Data Sets/User_Data.csv')
print(user_df)

X=user_df[['Age','EstimatedSalary']]
y=user_df['Purchased']
X_train, X_test, y_train, y_test =train_test_split(X, y, test_size=0.
↪3,random_state=42)
scaler =StandardScaler()
X_train_scaled =scaler.fit_transform(X_train)
X_test_scaled =scaler.transform(X_test)
model =LogisticRegression()
model.fit(X_train_scaled, y_train)
y_pred =model.predict(X_test_scaled)
accuracy =accuracy_score(y_test, y_pred)
print("\nAccuracy:", accuracy)
print("\nClassification Report:")
print(classification_report(y_test, y_pred))
# Confusion matrix
confusion_matrix=pd.
↪crosstab(y_test,y_pred,rownames=['Actual'],colnames=['Predicted'])
```

```
sn.heatmap(confusion_matrix,annot=True)
print ("\nConfusion Matrix : \n", confusion_matrix)
```

```
      User ID  Gender  Age  EstimatedSalary  Purchased
0    15624510    Male   19            19000          0
1    15810944    Male   35            20000          0
2    15668575  Female   26            43000          0
3    15603246  Female   27            57000          0
4    15804002    Male   19            76000          0
..        …     …  …              …        …
395  15691863  Female   46            41000          1
396  15706071    Male   51            23000          1
397  15654296  Female   50            20000          1
398  15755018    Male   36            33000          0
399  15594041  Female   49            36000          1

[400 rows x 5 columns]

Accuracy: 0.85

Classification Report:
              precision    recall  f1-score   support

           0       0.82      0.97      0.89        73
           1       0.94      0.66      0.78        47

    accuracy                           0.85       120
   macro avg       0.88      0.82      0.83       120
weighted avg       0.86      0.85      0.84       120


Confusion Matrix :
 Predicted   0   1
Actual
0          71   2
1          16  31
```


71
2
16
31
0
1
Actual
0
1
Predicted
70
60
50
40
30
20
10

[ ]: