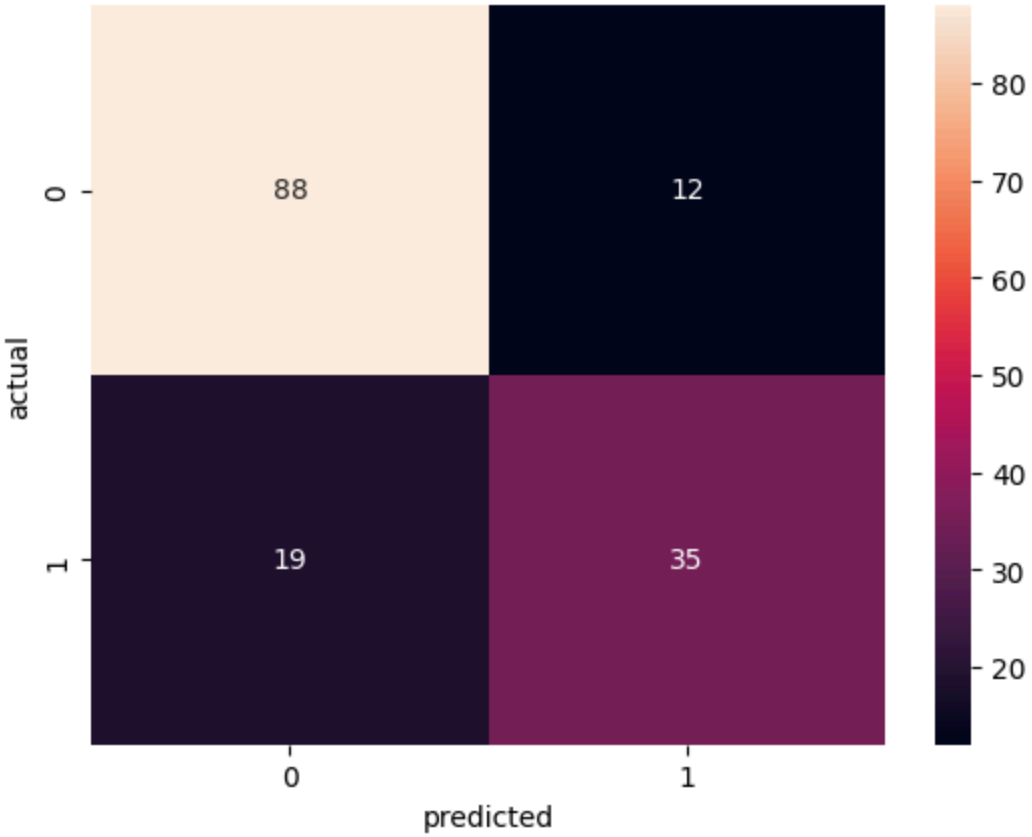
In [23]:
```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn import metrics
import seaborn as sn
from google.colab import files
uploaded = files.upload()  # This will open a file picker in Colab


col_names = ['Pregnant', 'glucose', 'bp', 'skin', 'insulin', 'bmi', 'pedigree'
data = pd.read_csv("diabetes.csv", header=0, names=col_names)
print(data.shape)
data.head()

data.isnull().sum()

feature_cols = ['Pregnant', 'insulin','bmi','age','glucose','bp','pedigree']
x = data[feature_cols]
y = data.label
x_train, x_test,y_train,y_test = train_test_split(x,y,test_size = 0.2,random_s
display(x_train.shape, y_train.shape,x_test.shape,y_test.shape)
model = LogisticRegression(solver = 'lbfgs', max_iter = 1000)
model.fit(x_train, y_train)
y_pred = model.predict(x_test)
conf_mat = metrics.confusion_matrix(y_test,y_pred)
print('confusion Matrix :', conf_mat)
accuracy_score = metrics.accuracy_score(y_test, y_pred)

print('accuracy score :', accuracy_score)
print('accuracy in  percentage:', int(accuracy_score*100),'%')

conf_mat = pd.crosstab(y_test, y_pred, rownames = ['actual'],colnames = ['pred
sn.heatmap(conf_mat, annot = True)
```

Browse...

Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.

```
Saving diabetes.csv to diabetes (9).csv
(768, 9)
(614, 7)
(614,)
(154, 7)
(154,)
confusion Matrix : [[88 12]
 [19 35]]
accuracy score : 0.7987012987012987
accuracy in  percentage: 79 %
```

Out[23]: <Axes: xlabel='predicted', ylabel='actual'>


88
12
19
35
0
1
actual
0
1
predicted
80
70
60
50
40
30
20