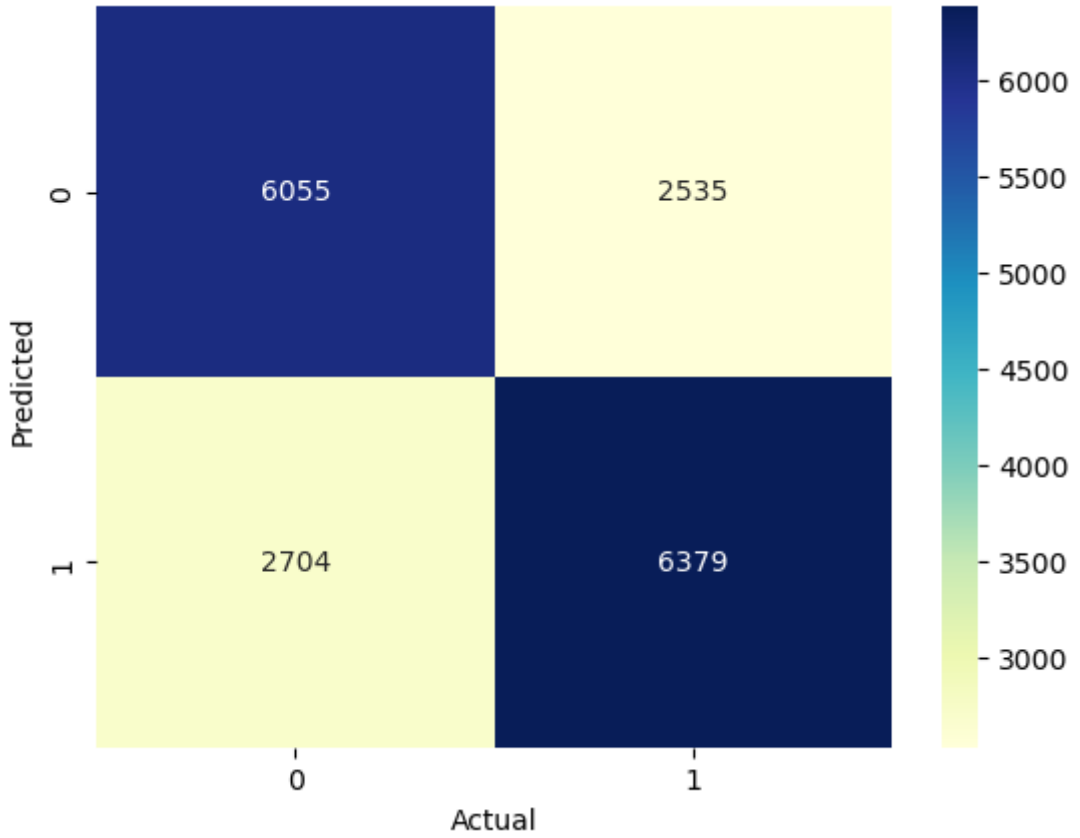
In [ ]:

In [50]:
```
import numpy as np
import pandas as pd
import pickle
import xgboost as xgb
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score
import seaborn as sn
import matplotlib.pyplot as plt

data = pd.read_csv("Data1.csv")
train, test = train_test_split(data)
feature_columns = ["HighBP","HighChol","CholCheck","BMI","Smoker","Stroke","HeartDi
target_column = "Diabetes_binary"
xgtrain = xgb.DMatrix(train[feature_columns].values, train[target_column].values)
xgtest = xgb.DMatrix(test[feature_columns].values, test[target_column].values)

param = {'max_depth': 20, 'eta': 1, 'objective': 'binary:logistic'}
param['nthread'] = 4
param['eval_metric'] = 'auc'

watchlist  = [(xgtest,'eval'), (xgtrain,'train')]
num_round = 10
bst = xgb.train(param, xgtrain, num_round, watchlist)
labels = xgtest.get_label()
pred = bst.predict(xgtest)
pred_copy = pred
result = list()
sum = 0
tp=0
tn=0
fp=0
fn=0
for i in range(len(pred)):
    if pred[i] > 0.5:
        pred_copy[i] = 1
    else:
        pred_copy[i] = 0
    if pred_copy[i] == labels[i]:
        sum += 1
    if pred_copy[i] == 0 and labels[i] == 0:
        tp += 1
    if pred_copy[i] == 1 and labels[i] == 0:
        fn += 1
    if pred_copy[i] == 0 and labels[i] == 1:
        fp += 1
    if pred_copy[i] == 1 and labels[i] == 1:
        tn += 1

auc = sum/len(pred)
print("Accuracy" + str(auc))
print("tp" + str(tp) + " tn" + str(tn) + " fp" + str(fp) +  " fn" + str(fn))


data = [[0,0],[0,0]]
data[0][0] = tp
data[0][1] = fp
data[1][0] = fn
data[1][1] = tn
hm = sn.heatmap(data = data, annot=True, cmap="YlGnBu", fmt='g')
plt.xlabel('Actual')
```

```
plt.ylabel('Predicted')
plt.show(hm)
```

```
[0]     eval-auc:0.79012        train-auc:0.89451
C:\Users\panyu\anaconda3\Lib\site-packages\xgboost\core.py:726: FutureWarning: Pas
s `evals` as keyword args.
  warnings.warn(msg, FutureWarning)
[1]     eval-auc:0.78087        train-auc:0.94120
[2]     eval-auc:0.77720        train-auc:0.96579
[3]     eval-auc:0.77623        train-auc:0.97910
[4]     eval-auc:0.77394        train-auc:0.98729
[5]     eval-auc:0.77302        train-auc:0.99269
[6]     eval-auc:0.77293        train-auc:0.99531
[7]     eval-auc:0.77308        train-auc:0.99674
[8]     eval-auc:0.77321        train-auc:0.99767
[9]     eval-auc:0.77328        train-auc:0.99827
Accuracy0.7035591014541956
tp6055 tn6379 fp2535 fn2704
```



In [ ]:

In [ ]: