

**University of North Carolina at Charlotte**  
**Department of Electrical and Computer Engineering**

*Linear Regression Report HW#2*

**Logistic Regression**

---

**Logistic Regression**

Author: Eric Lee

Student ID: 801165284

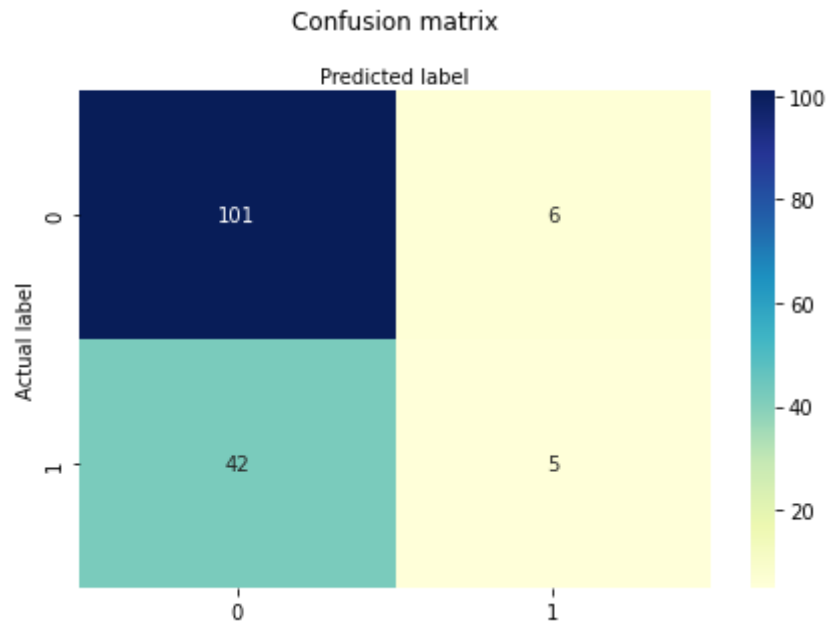
Date: 10/9/2022

This report was submitted in compliance with UNCC POLICY 407  
THE CODE OF STUDENT ACADEMIC INTEGRITY, Revised November 6, 2014  
(<http://legal.uncc.edu/policies/up-407>) \_\_\_\_PEL\_\_\_\_. (Student's Initials)

### Problem #1

After performing the logistic regression binary classifier for the diabetes data set the calculated accuracy, precision, and recall can be seen below. The confusion matrix is also seen below.

Accuracy: 0.6883116883116883  
Precision: 0.45454545454545453  
Recall: 0.10638297872340426



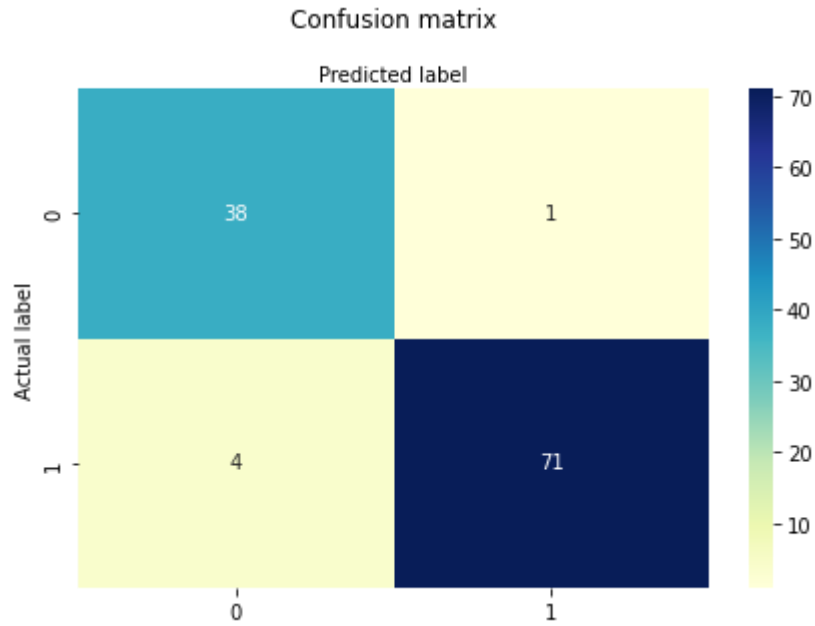
### Problem #2

When performing the K-fold cross-validation a recorded accuracy of 66.017% was done with 5 splits and an accuracy of 66.025% was done with 10 splits. When comparing the accuracy to the accuracy of problem 1 there can be seen a decrease in accuracy of about 2%.

### Problem #3.1

Using the given library to import the cancer data set a logistic regression model was built and the accuracy, precision, and recall was recorded below. The confusion matrix is given below as well.

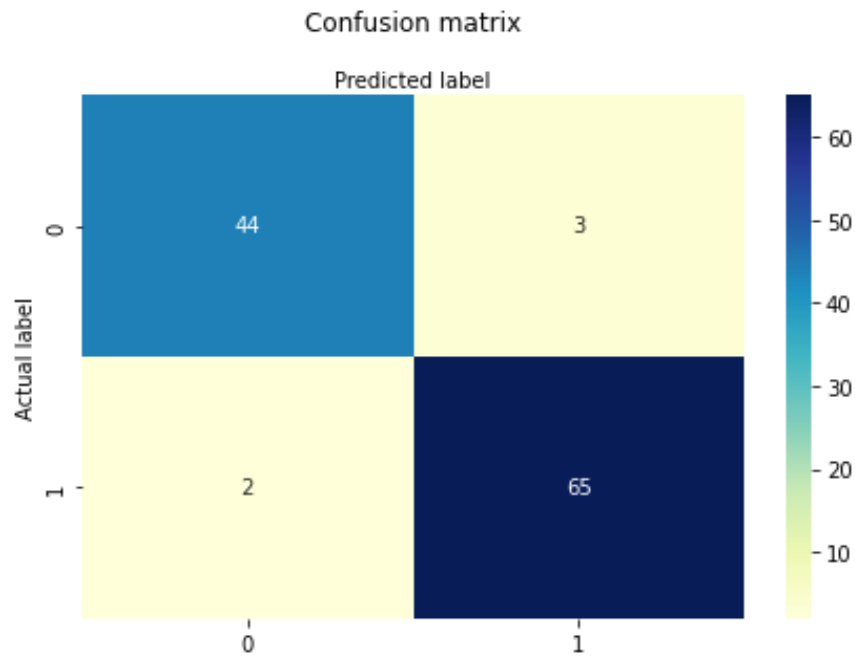
Accuracy: 0.956140350877193  
Precision: 0.9861111111111112  
Recall: 0.9466666666666667



### Problem #3.2

When adding in the penalty weight the accuracy of the model stayed the same, but the precision and recall changed which also changed the confusion matrix as well. The difference can be seen below.

Accuracy: 0.956140350877193  
Precision: 0.9558823529411765  
Recall: 0.9701492537313433



**Problem #4.1**

Using the cross-validation method, when  $k$  is equal to 5 it was reported that the accuracy was 95.434% and when  $k$  was equal to 10 it was 95.432%. When comparing the accuracy to the accuracy of problem 3 it is seen that the two accuracies are almost identical.

**Problem #4.2**

When adding in the penalty term there was an increase in accuracy of about 1%. So, this gave the a accuracy of about 96%.

**References**

<https://github.com/plee41/ECGR-4105>