

Homework 2

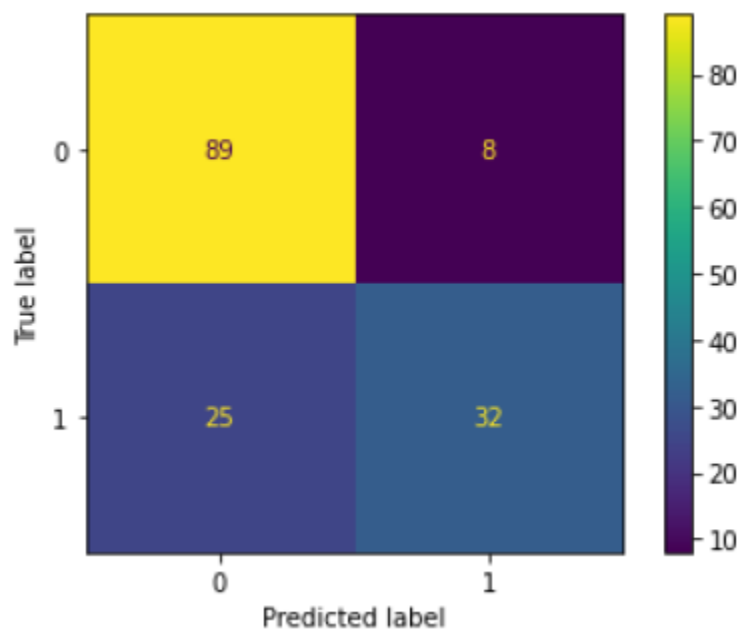
GitHub: https://github.com/pballou/ECGR_4105/tree/master/Homework/homework_2

1. Logistic regression for the diabetes dataset. Results:

Accuracy: 0.7857142857142857

Precision: 0.8

Recall: 0.5614035087719298



2. Repeated problem 1 but used k-fold validation. Used k=5 and k=10. Results:

Accuracy for K = 5: 77.341% (1.944%)

Accuracy for K = 10: 77.346% (4.689%)

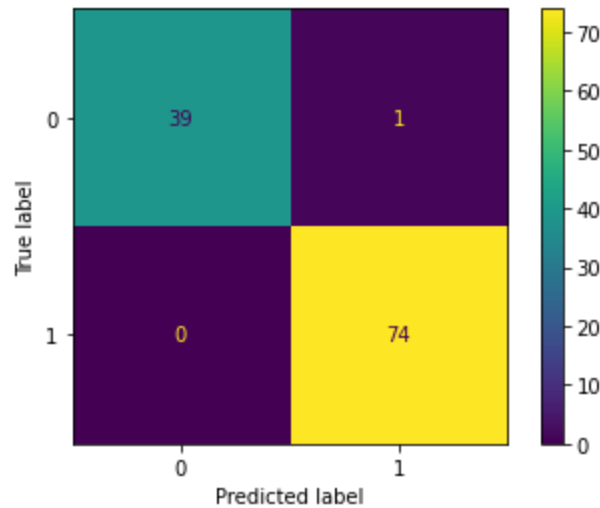
3. Logistic regression for the cancer dataset.

- a. Results without penalty:

Accuracy: 0.9912280701754386

Precision: 0.9866666666666667

Recall: 1.0

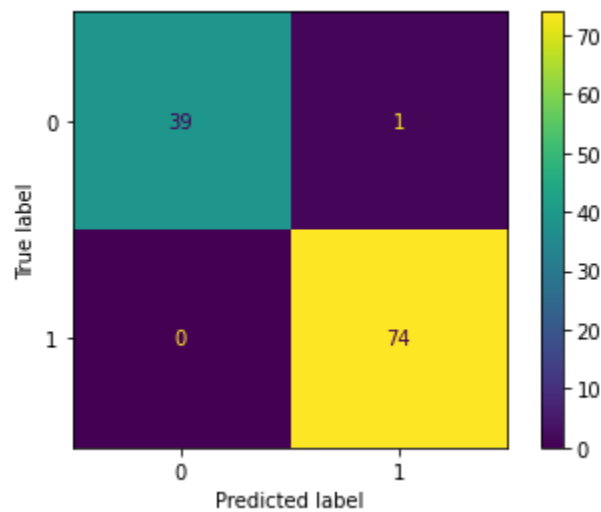


b. Results with penalty:

Accuracy: 0.9912280701754386

Precision: 0.9866666666666667

Recall: 1.0



4. Using k-fold validation on the cancer dataset. Used k=5 and k=10.

a. Results without penalty:

Accuracy for K = 5: 97.539% (1.024%)

Accuracy for K = 10: 97.716% (1.578%)

b. Results with penalty:

Accuracy for K = 5: 97.539% (1.024%)

Accuracy for K = 10: 97.716% (1.578%)