

# Lab 5 (due: Feb 28)

## MACHINE LEARNING - COSC 4360

Department of Computer Science and Electrical Engineering

Spring 2025

### Exercises

Create a **New Project** for every exercise. Take a screenshot of the source code along with its output and place the **source code** and the **screenshot** in a **zipped folder** named **LastNameFirstName\_Lab5**

#### Exercise 1

Given the following dataset: *Student-Pass-Fail.csv*, use **logistic regression** to find the **odds**, and predict whether a student will **Pass** or **Fail** given the following *three* data points: [7, 28], [10, 34], [2, 39]. Print the **probabilities** for each one of the *three* data points.

**Note:** You can compute the odds using:  $e^x$ , where  $x$  is the coefficient.

In python: `odds = np.exp(logisticRegression.coef_)`

#### Exercise 2

In continuation of Ex. 1, estimate the **accuracy score** of the model and print the **Confusion Matrix**. You will have to create your **own** functions to **split** the data (given by the user), estimate the model's **accuracy score**, and compute the **Confusion Matrix**.

**Note:** You can use *built-in* functions only for comparison.

#### Exercise 3

Given the following dataset: *Bank-data.csv*, use **logistic regression** to find the **odds**, and predict whether a client will subscribe a term deposit or not given the following *two* data points: [1.335, 0, 1, 0, 0, 109], [1.25, 0, 0, 1, 0, 279]. Print the **probabilities** for each one of the *two* data points.

**Note:** Drop the first column, the **target** (y) variable is the last column.

**Note:** Submit through **Canvas**