Homework 3 *(due: Feb 10)*
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_HW3***

Exercise 1

Using the following dataset: **Square Feet**: 100, 150, 185, 235, 310, 370, 420, 430, 440, 530, 600, 634, 718, 750, 850, 903, 978, 1010, 1050, 1990 and **Price** ($): 12300, 18150, 20100, 23500, 31005, 359000, 44359, 52000, 53853, 61328, 68000, 72300, 77000, 89379, 93200, 97150, 102750, 115358, 119330, 323989. Find the **inliers** and **outliers** in the data set using the **RANSAC** algorithm, based off slide 139. Print the line parameters and plot the regression lines *before* and *after* **RANSAC** implementation. Your output should be similar to Fig. 1 below. You can use any *built-in* functions.

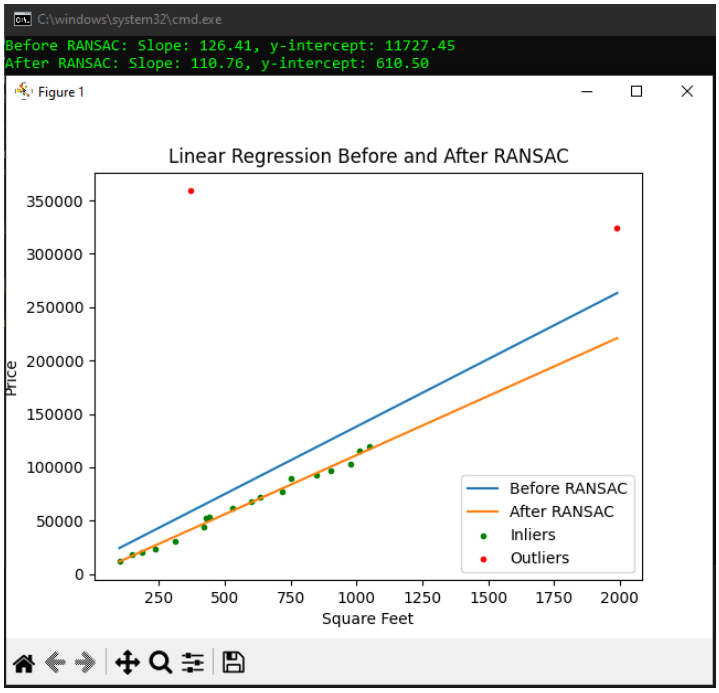

Figure 1: Regression lines before and after RANSAC

Exercise 2 *(Optional)*

Given the dataset: *iris.data.csv*, perform **5-fold Cross Validation** as shown on slide 108 (assume no randomness in selecting data points). Then, perform classification using **KNN** with **K=9**. Print the **accuracies** for all **5** test datasets, and finally, print the **average accuracy**.

Note: Use your **own** implementation of *Cross Validation* without using the *built-in* functions.

Note: Submit through **Canvas**