Assignment #4 Support Vector Machines

Problem# 1(5 pts):

Employ the Task_Data dataset Inside the compressed file, you will find two files:
1. data.txt: provides the examples with the last column for class category (1 or 2).
2. readme.data.txt: describes the information regarding this dataset

You are required to train and understand SVM classification model, by randomly select 60% of the examples from the data.txt file as training examples, and use the remaining examples for testing.

1. Build a model of SVM classifier of scikit-learn with a linear kernel, train the classifier for ten times and report the classification averaged accuracy of SVM on the test data over these ten trials averaged by ten trials. Note: each time the difference is in the data that is randomly chosen for testing and training.

2. Normalize the data and do any preprocessing step that you see that is needed and train the same model in (1) and rerun the experiment in (1). First Discuss:
   A. Report the Difference between the dataset used in (1) and those used in (2).
   B. Report the averaged accuracy over the ten trails. Note: each time the difference is in the data that is randomly chosen for testing and training.
   C. Discuss the difference in the averaged accuracy of (1) and (2).
   D. Report all the preprocessing steps you did to the data.

Problem# 2(3 pts):

Implement, from scratch, linear SVM model using Gradient descent as an optimization function as discussed in the section. Your model should have a function called fit for training the data and predict for the prediction and any other needed function. Use Iris Data for training and testing your binary classifier model.

Hint: Use only two features. and plot them together.

General Instructions
1. This is an individual task.
2. Please report al your findings in a neatly created pdf report.
3. The source code as well as the report should be submitted through Black board.
4. The due date for the submission of this phase is Tuesday, December, 14, 2021 at 11:59 pm.
5. Please Review the definition of cheating in the first presentation.