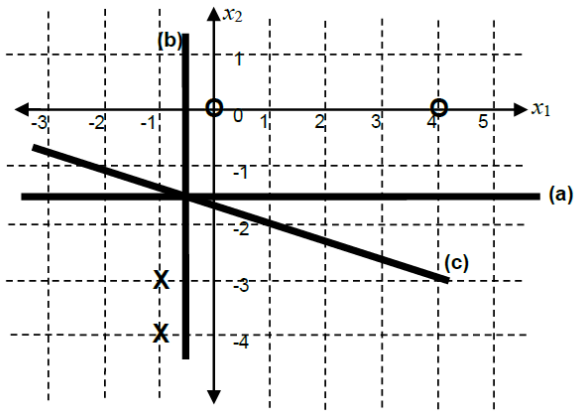
German International University of Applied Sciences
Informatics and Computer Science
Dr. Caroline Sabty
Eng. Nadeen Amr

Machine Learning, Winter 2022
Practice Assignment 12

Exercise 12-1

Consider the dataset shown in the figure below where a linear Support Vector Machine (SVM) without slack variables is supposed to be used:



a) **Which** of the decision boundaries (a), (b) or (c) shown on the figure would be the resulting decision boundary of linear SVM? **Show** your calculations. When answering this question, no need to solve by optimizing the SVM objective function.

b) **What** are the support vectors based on your answer in (a)?

c) **How** would adding a training point in location (1, 1) to the dataset that belongs to the (**O**) class change the decision boundary?

Exercise 12-2

Consider a Support Vector Machine (SVM) without slack variables that is supposed to be used to classify the 1-dimension data given below:

x	t
-2	-1
1	1
3	1

a) Given that the objective function of SVM takes the form
$J = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j t_i t_j \phi(x_i)^T \phi(x_j)$
Find J for the dataset given above in terms of α_i if the kernel used is $K(x_i, x_j) = |x_i||x_j|$. (Don't solve the optimization problem. Just find the expression of J).

b) **Would** solving this optimization problem for this data using the given kernel lead to a decision boundary that classifies the given data correctly? **Explain** your answer.

Exercise 12-3 Coding Question

Using the SVM model in sklearn do the following:

a) Split the Iris dataset into training and testing portions (80,20 split).

b) Train your model using the training portion

c) Test your model using the testing portion

d) Produce the accuracy of your model

Exercise 12-4 Coding Question

Using the SVM model in sklearn, train the model using this entire dataset using the linear kernal function.

x_1	x_2	y
-3	-1	0
0	-2	1
-2.5	2	0
-1	-1	1
3	0.5	1
0.5	3	0
-3	-3	1

Afterwards, test your model using this data point [2, 4]