UPRM CSE Department
Introduction to Deep Learning
Spring 2018

Practice-Homework 3

Submit the solution for problem 1. Solve problem 2 or problem 4 and submit your answer. If you choose to solve problem 4, submit the code and explain the procedure you used to obtain your results.

Due Date, what to submit and how to submit

Be aware that collaboration with classmates is allowed but be sure to understand and provide your own solution.

Practice Problems

Using your knowledge on SVM answer the following questions.

1. Consider data from two types of tumor, benign (-) and malignant (+) which is described by two features: cellular distributions and size. The scatter plot is provided in Fig. 1.
   a. What is the width of the margin?
   b. Compute $\vec{w}$ and b for the decision bounary $h(\vec{u}) = \vec{w}.\vec{u} + b \geq 0$
   c. Calculated the weights (alphas) of each data point.
   d. What would be the alpha of a new negative point placed at (0,6)?
   e. What would be the alpha of a new negative point placed at (0,0)?
   f. Supoose you move (replace) point (3,3) to (4,2), how would it modify the magnitud of its alpha: it will be increased, it will remain the same or it will be decreased. Explain your answer.
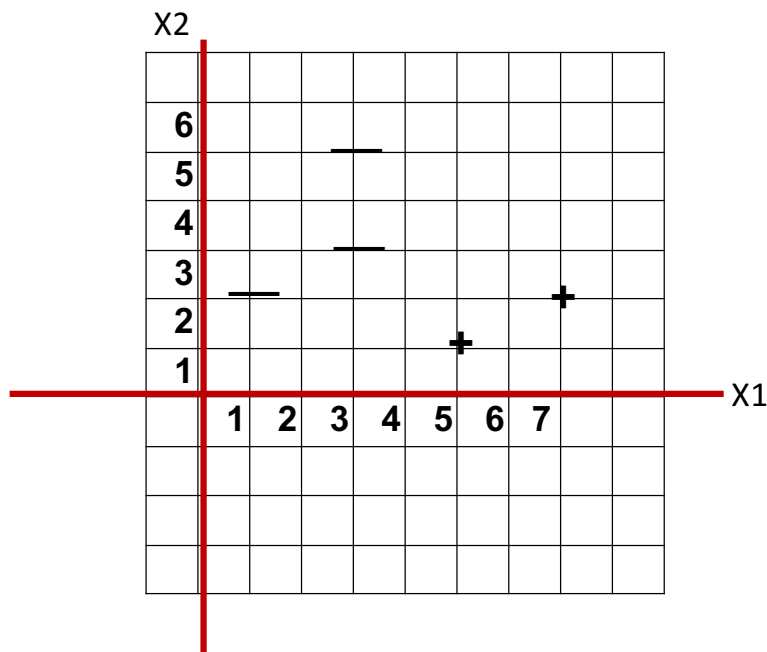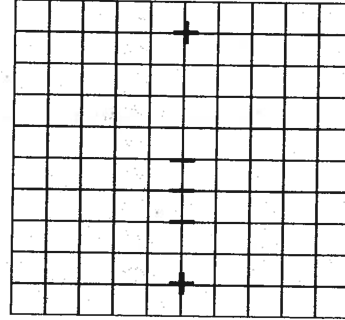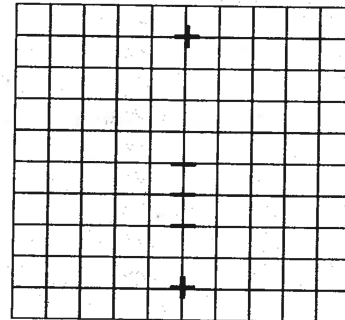

Figure 1.

2. For the following cases indicate Yes or Not wheater the kernels can be used to perfectly classify the test points, and if Yes, sketch the decision boundary and indicate which data points are the support vectors.
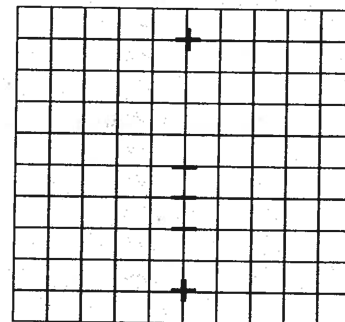   Note: more than one answer may apply for some of the cases.

   a. Case 1: $K(\vec{u}, \vec{v}) = \vec{u} \cdot \vec{v}$

   b. Case 2: $K(\vec{u}, \vec{v}) = (\vec{u} \cdot \vec{v} + 1)^2$

   c. Case 3: $K(\vec{u}, \vec{v}) = e^{-(|\vec{u} - \vec{v}|)^2 / 2}$

3. (Optional) Use the example provided by Matworks at https://www.mathworks.com/matlabcentral/fileexchange/63158-support-vector-machine as a practice for aplying SVM in linearly seperable data using a linear Kernel with Gradient ascent.

4. The file ***anascodata.mat*** provides samples of six clases of geografical areas from the region of Añasco-Puerto Rico captured with a multispectral sensor equiped with seven spectral channels. Analyze the performace of the SVM classifier for this data providing the classification accuracy. You can use the number of testing samples classified correctly to measure the accuaracy.

    Data Description: The ***anascodata.mat*** contains data for 6 clases corresponding to two types of water, two types of crops, montain areas and urban areas. The data for each class is stored in 2D data arrays. The last two letters of the file name are Tr for training samples and Te for testing samples.