

ECE 449 Machine Learning (ZJU-UIUC Institute)

Assignment 3

Due Date: May 18th, 2021.

1 Overview

Welcome to ECE449 assignment 3!

In this assignment, you will complete 2 Python programming tasks and answer 4 conceptual questions. Assignment 3 has 50 pts in total: 25 pts from programming tasks and 25 pts from conceptual questions.

2 Programming Questions

2.1 Programming environment update

In assignment 3, you need a new library plotly to visualize your final result. Even if missing plotly will not influence your answer to the question, I would still recommend you to install this to see the final result.

To do this, you may use: `pip install plotly`

2.2 Programming questions brief descriptions

There are 2 Python notebook files you need to work on in this assignment.

The first Python notebook file, `exercise1-template.ipynb`, is about support vector machines with 4 functions to implement.

The second Python notebook file, `exercise2-template.ipynb`, is about PCA and K-means with 7 functions to implement.

The programming part has 25 pts in total.

2.3 Autograder notice

We use autograder to grade your implementations, and here are some things we want to highlight:

1. Do NOT modify the names of any functions we grade. Autograder will give 0 for functions with names modified.
2. Do NOT modify the input interface of any functions, Autograder will give 0 for functions with input interfaces modified. However, in some cases you are allowed to modify the return values of some functions.
3. Do NOT import extra libraries. The Python libraries given in the `.py` scripts should be adequate for you to obtain a correct answer.

2.4 Collaboration policy

You should not look at python code solutions from other students, the Python implementations should be your own work.

3 Conceptual Questions

There are 4 questions you need to answer in the conceptual question part. You can write down your answer on a few sheets of paper and scan them to a single PDF file, or you can use a digital pen to work on these questions and export your handwritten answers to a single PDF file. We also accept answers in pdf format generated using L^AT_EX.

3.1 Problem 1 : SVM[6pts]

The “maximum margin classifier” (also called linear “hard margin” SVM) is a classifier that leaves the largest possible margin on either side of the decision boundary. The samples lying on the margin are called support vectors.

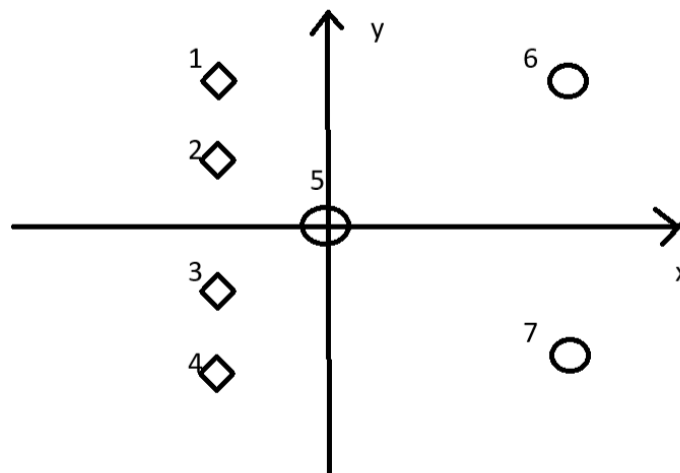


Figure 1

3.1.a Draw on Figure 1 the decision boundary obtained by the linear hard margin SVM method with a thick solid line. Draw the margins on either side with thinner dashed lines. Circle the support vectors.

3.2.b The removal of which sample will change the decision boundary?

3.2 Problem 2: k-means cluster[6pts]

Cluster the following eight points (with (x, y) representing locations) into three clusters:

$A_1(2, 10)$, $A_2(2, 5)$, $A_3(8, 4)$, $A_4(5, 8)$, $A_5(7, 5)$, $A_6(6, 4)$, $A_7(1, 2)$, $A_8(4, 9)$

Initial cluster centers are: $A_1(2, 10)$, $A_4(5, 8)$ and $A_7(1, 2)$.

The distance function between two points $a = (x_1, y_1)$ and $b = (x_2, y_2)$ is defined as-
 $P(a, b) = |x_2 - x_1| + |y_2 - y_1|$

Use K-Means Algorithm to find the three cluster centers after the second iteration.

3.3 Problem 3: PCA algorithm [6pts]

Consider the two dimensional patterns (2, 1), (3, 5), (4, 3), (5, 6), (6, 7), (7, 8).

Compute the principal component using PCA Algorithm.

3.4 Problem 4: SVM [7pts]

Assume we have 6 points showing 6 observations in a 2-D Euclidean space as below:

Observation	X_1	X_2	Y
1	0	2	class 1
2	0	3	class 1
3	1	3	class 1
4	1	1	class 2
5	2	1	class 2
6	2	2	class 2

Using a maximum margin classifier, determine the optimal separating hyperplane and give an equation for it, and determine which observations are the support vectors.

4 Submission

You only need to submit 3 files to BlackBoard: the 2 Python notebook files with your programming implementations and a PDF file with your answers to the conceptual questions.

For example, a student whose ID number is 3180100111 should submit the following 3 files:

exercise1_3180100111.ipynb, exercise2_3180100111.ipynb
conceptual_3180100111.pdf

If you want to submit a zip file, please name it as [your_studentID]_assignment3.zip. For example, a student whose ID number is 3170100111 should submit the zip file 3170100111_assignment3.zip to BlackBoard. Please make sure that you add the 3 files directly to your zip file, not a folder containing your 3 files. A quick verification for this is that if you unzip your zip file, the 3 files should appear in the same folder as your zip file. **Even if you submit a zip file, renaming the 3 files inside that zip file is required.** Otherwise our autograder will not be able to recognize the files you submit, which may even result in a 0 score for the file not properly renamed.