

ITU, Computer Engineering Dept.
BLG527E, Machine Learning HW2

Due: December 20th, 2024, 23:00 through Ninova.

NO LATE SUBMISSION WILL BE ACCEPTED. DO NOT SUBMIT THROUGH E-MAIL.

Instructors: Ayşe Tosun (tosunay@itu.edu.tr) Yusuf Yaslan (yyaslan@itu.edu.tr)

Grading: You must complete the table below according to what you expect to get out of each question. At the beginning of your report you should give the following table. You can use the given report template.

		Q1	Q2	Q3	Total
Grade	Max	4	3	4	10
	Expected				

Policy:

Please do your homeworks on your own. You are encouraged to discuss the questions with your class mates, but the code and the hw you submitted must be your own work. Cheating is highly discouraged for it could mean a zero or negative grade from the homework.

If a question is not clear, please let us know (via email or in class). Unless we indicate otherwise, do not use libraries for machine learning methods. When in doubt, email me.

There will be 3 homeworks this term. Each hw is worth 10 points and each question will be evaluated on a 0/1 basis.

Q1) **Gaussian Mixture Models and EM.** Generate a 2 dimensional dataset with 3 clusters. You will fit a Gaussian mixture with isotropic covariance each. You may use Python libraries and toolboxes for data generation. Make sure each cluster has 200-400 samples that are generated from each cluster's Gaussian. Initialize the cluster means and covariance matrices, and implement EM algorithm without using any toolbox/library. Plot the Gaussian contours after the EM.

Q2) **EM and Coin Flipping.** Write a program to solve the coin flipping problem that we covered in the last lecture. You may use the resources provided to you via Ninova. Your program should generate 10 sequences of data representing flipping two coins, and later estimate the parameters. Assume that there is an equal chance of choosing each coin (50%).

Your program should first print the 10 sequences of coin flips, and the estimated parameters.

Each time you run your program, you must generate a new sequence of data, and parameters should be estimated accordingly.

Q3) **EM and Updated Coin Flipping.**

a) Write down the EM steps to find the parameters of the coin flipping problem when the coins are NOT chosen with equal chance, i.e., the probability of choosing coin A is **different** than 50%.

b) Update your program in Q2 to handle this case and estimate the parameters. Please note that you have an additional parameter, probability of choosing coin one or two, that needs to be estimated in this scenario.

Your program should print the 10 sequences of coin flips, and the estimated parameters.