



MITx 6.86x

Machine Learning with Python-From Linear Models to Deep Learning[Course](#)[Progress](#)[Dates](#)[Discussion](#)[Resources](#)[Course](#) / [Unit 4. Unsupervised Learning...](#) / [Project 4: Collaborative Filtering vi](#)[< Previous](#)

2. K-means

[Bookmark this page](#)

Project due Apr 26, 2023 08:59 -03 Completed

K-means

1.0/1.0 point (graded)

For this part of the project you will compare clustering obtained via K-means to the (so-called) EM. In order to do so, our K-means algorithm will differ a bit from the one you learned. It will estimate the means of the component Gaussians exactly as before but the algorithm returns additional information. More specifically, it returns the resulting clusters of points to estimate a Gaussian model for each cluster. Thus, our K-means algorithm returns a mixture model where the means of the component Gaussians are the K centroids. This is to make it such that we can now directly plot and compare the results of the K-means algorithm. This is to make it such that we can now directly plot and compare the results of the two algorithms as if they were both estimating mixtures.

Read a 2D toy dataset using `X = np.loadtxt('toy_data.txt')`. Your task is to run this data using the implementation we have provided in `kmeans.py`. Initialize K-means using `common.init(X, K, seed)`, where K is the number of clusters and `seed` is the random seed. This will randomly initialize the parameters.

Note that `init(X, K)` returns a K-component mixture model with means, variances and weights. The K-means algorithm will only care about the means, however, and returns a mixture model based on the K-means solution.

Try $K = [1, 2, 3, 4]$ on this data, plotting each solution using our `common.plot` function. Since the initialization is random, please use seeds `0, 1, 2, 3, 4` to and select the one that minimizes the cost. Report the associated plots (best solution for each K). The code for this task can be written in

Report the lowest cost for each K : $\text{Cost}_{K=1} = 5462.297452340002$ [About](#)[Affiliates](#)[edX for Business](#)[Open edX](#)[Careers](#)[News](#) $\text{Cost}_{K=2} = 1684.9079502962372$  $\text{Cost}_{K=3} = 1329.59486715443$  $\text{Cost}_{K=4} = 1035.499826539466$ 

Legal

[Terms of Service & Honor Code](#)[Privacy Policy](#)[Accessibility Policy](#)[Trademark Policy](#)

You have used 2 of 50 attempts



[Previous](#)

[Next >](#)



© 2023 edX LLC. All rights reserved.

深圳市恒宇博科技有限公司 [粤ICP备17044299号-2](#)