Optional Programming Assignment

Help Center

Goals

- This assignment is optional and worth 5% extra credit.
- In this assignment you will implement the following clustering algorithms: K-means and Kernel K-means. In doing so, you will become more familiar with the details of k-means algorithm and its kernel variation.
- In addition, you will also implement the following supervised clustering evaluation metrics: purity and NMI, which will help you understand the performance of different clustering algorithms.

Instructions

- You do not need to implement all functions. You have the option of using the provided implementation for I/O and initialization instead of writing your own. The main framework for the assignment has been provided in the .zip file below, and the only thing you need to do is to fill in the designated key functions, all of which are highly related to our course materials.
- The instructions provide support for three programming languages, C++, Java, and Python. However, you can still write in your own favorite language. If you do so, please pay attention to the details in the initializations because failure to do so may lead to different results. In particular, we use the first two data points as initial cluster centers for all implementation languages. However, it is strongly suggested you use the framework provided in the instructions. It is worth noting that if you choose to use your favorite programming language, you do so at your own risk, since limited resources would be provided for debugging guidance.
- Please read the "README.md" files carefully.
- Click to download the .zip file for this programming assignment

Time Estimation

Students can expect to finish this homework in 1 – 2 hours by first reviewing the corresponding slides, then implementing the key functions, and finally, testing and debugging.

Self-evaluation

In Week 3, we provide a small dataset for local testing purposes. The expected outputs are given in each subfolder, corresponding to different programming languages. You can follow the instructions to verify whether or not your implementation is correct.

Quiz Evaluation

In Week 4, a new dataset will be released. We will ask you to choose the correct multiple-choice answers based on the screen output.

Q & A

Optional Programming Assignment Help Forum

Created Mon 11 May 2015 9:20 AM PDT

Last Modified Wed 13 May 2015 7:00 AM PDT