10th May 2022

Your Internship Practice and Schedule at OMU

- Task: k-Means type clustering and its application to personalized recommendation

Clustering is a basic method for revealing the intrinsic data structure in unsupervised learning. k-means is the most practical clustering algorithms and fuzzy c-means (FCM) is an extension for handling ambiguous cluster boundaries. In this practice, you should first learn how to implement k-means clustering and FCM followed by their application to collaborative filtering, which is a basic technique for personalized recommendation.

- k-means implementation (by the end of May)

First, learn how to implement k-means clustering through programing exercise. You can use any programming language but your tutor may use C language and Python only.

- Fuzzy c-means (FCM) implementation (by the end of June)

Second, learn how to implement fuzzy c-means (FCM) clustering through programing exercise.

- GroupLens recommendation system (by the end of July)

Third, learn the basic concept of GroupLens recommendation system referring to the following document:

J. L. Herlocker, J. A. Konstan, A. Borchers and J. Riedl: An algorithmic framework for performing collaborative filtering; Proc. of Conference on Research and Development in Information Retrieval (1999)

Then, reproduce the experimental results presented in the document.

- Application of k-means and FCM to collaborative filtering and comparison with GroupLens results (by the end of August)

Fourth, investigate the applicability of k-means and FCM in collaborative filtering tasks, where you should compare their performances with GroupLens.

- Your tutors

Mr. Koki Kitamori can support your internship days.