

Department of Computer Science

St. Francis Xavier University

Presents

Higher-order Affinity Propagation Algorithm Based on Multiple Similarity Measurements

by

Zihao Jiang

St. Francis Xavier University

M.Sc. Thesis Proposal Presentation

Time: Jul 14, 2023 01:00 PM Halifax

Join Zoom Meeting

<https://us02web.zoom.us/j/88194706378>

Affinity propagation (AP) clustering algorithm has been a popular research topic in data mining. AP is appealing because it is efficient, insensitive to initialization, and it produces clusters at a lower error rate than other clustering algorithms. Despite its significant progress in recent years, there still exists a challenging issue. Nowadays, more and more large scale data are provided from the services, applications, and platforms. When facing big data, most of recent AP clustering algorithms lack the abilities to represent and handle the rich relationships between objects at higher order. To address these massive data, and dig out more useful values in these data, this thesis presents a novel AP clustering algorithm called higher-order affinity propagation clustering algorithm based on multiple similarity measurements.

This proposed method involves multi-view clustering concept in many places. Similarities between the samples are calculated by using various similarity measurements, mainly three methods about structural similarity measurement, tensor distance similarity measurement and cosine similarity measurements are used in this thesis. And then, the similarities obtained by these different similarity measurements are constructed into a tensor. After the construction, the tensor is asked to be standardized and optimized, otherwise, the constructed tensor is difficult to analyze as a whole. That is because each similarity measurement characterizes the similarity in a different sense. Finally, this optimized tensor is applied to the higher-order AP algorithm, which can coordinate these different similarity measurements and make them cooperate with each other to generate higher quality clustering centers to get better accuracy. From the beginning of the higher-order AP algorithm based on multiple similarity measurements to the final clustering, there are many places where voting or weighted decision-making is required due to the presence of different similarity measurements. And finally, it was experimentally verified that this novel algorithm is highly efficient in this era of complex data.