



Annie Qu, PhD

Chancellor's Professor
Department of Statistics
University of California, Irvine

Friday, September 16, 2022

12:00-1:00 pm

Biotech Center Auditorium *or* via Zoom Link

<https://uwmadison.zoom.us/j/95990198725>

Query-augmented Active Metric Learning

Abstract: We propose an active metric learning method for clustering with pairwise constraints. The proposed method actively queries the label of informative instance pairs, while estimating underlying metrics by incorporating unlabeled instance pairs, which leads to a more accurate and efficient clustering process. In particular, we augment the queried constraints by generating more pairwise labels to provide additional information in learning a metric to enhance clustering performance. Furthermore, we increase the robustness of metric learning by updating the learned metric sequentially and penalizing the irrelevant features adaptively. Specifically, we propose a new active query strategy that evaluates the information gain of instance pairs more accurately by incorporating the neighborhood structure, which improves clustering efficiency without extra labeling cost. In theory, we provide a tighter error bound of the proposed metric learning method utilizing augmented queries compared with methods using existing constraints only. Furthermore, we also investigate the improvement using the active query strategy instead of random selection. Numerical studies on simulation settings and real datasets indicate that the proposed method is especially advantageous when the signal-to-noise ratio between significant features and irrelevant features is low.



**School of Medicine
and Public Health**

UNIVERSITY OF WISCONSIN-MADISON