THE UNIVERSITY OF CHICAGO

Computational and Applied Mathematics
&
Statistics Student Seminar

Erasmo Tani

Department of Computer Science
University of Chicago

Recent Progress on Learning Partitions with a Faulty Same-Cluster Oracle

Tuesday, January 16, 2024
12:30 PM
Searle 240A

ABSTRACT

Learning cluster structure is a fundamental task in Machine Learning. In this talk, we focus on partitioning a finite set. We consider the active setting in which one gains information about the clusters exclusively by asking potentially expensive questions of the form: "Are u and v part of the same cluster?" The main question of interest is: how many queries are needed to guarantee full recovery of the underlying partition if a fixed number of questions could return an erroneous answer? We will discuss recent progress on the problem, including an algorithm that achieves provably optimal query complexity up to constant factors.