# Skeleton Clustering and Regression

Zeyu (Jerry) Wei[1], Yen-Chi Chen[1]

[1]*Department of Statistics, University of Washington, U.S.A.*

## Abstract

In this talk, we introduce a novel method called Skeleton Clustering for detecting clusters in multivariate and high-dimensional data with irregular shapes. To bypass the curse of dimensionality, we propose surrogate density measures that are less dependent on the dimension but have intuitive geometric interpretations. Our clustering framework constructs a concise representation of the data using a combination of prototype methods, density-based clustering, and hierarchical clustering. We then discuss a new regression framework designed to handle large-scale, complex data that lies around a low-dimensional manifold. Our approach constructs a graph representation, referred to as the skeleton, to capture the underlying geometric structure, and defines metrics on the skeleton graph to estimate the regression function using nonparametric regression techniques. We provide statistical guarantees for the proposed method and demonstrate its effectiveness through simulations and real data examples. The proposed method is robust to additive noise and noisy observations and can handle the union of multiple manifolds.

Keywords: Manifold learning; high-dimensional clustering; density-based clustering; nonparametric regression.