

**“YOU CALL IT A MANIFOLD, I CALL IT A SUBSPACE”
- SELECTED EXAMPLES ON THE INTERFACE
BETWEEN COMPUTER SCIENCE AND STATISTICS IN
THE CONTEXT OF CLUSTERING AND MANIFOLD
LEARNING**

Daniyal Kazempour¹ and Peer Kröger¹

¹ Department of Computer Science,
Christian-Albrechts-Universität zu Kiel,
(e-mail: [dka, pkr]@informatik.uni-kiel.de)

ABSTRACT: Do prominent data mining methods in computer science have anything in common with well-established techniques in statistics? Are there any benefits in combining methods from statistics with those from computer science, and if yes, why do we gain such benefits? These and further aspects are approached at the interface between computer science and statistics.

This talk first provides a brief introduction to the clustering and dimensionality reduction tasks from a computer science perspective. Furthermore, a brief introduction to the manifold learning task is given. This foundation is followed by an elaboration on similarities and distinct properties between two seemingly different tasks from different domains (cs and statistics), more specifically the subspace clustering and the manifold learning task.

Pursuing this path on the interface between computer science and statistics, it is elaborated on endeavors of enhancing cluster analysis through manifold learning while investigating why the combination of two methods from different domains (clustering and manifold learning) results in a symbiotic relationship.

In conclusion, this talk aims to sketch selected examples of the potential synergies that can emerge on the interface of computer science and statistics in the context of data mining and machine learning including its challenges and benefits. The examples in this talk do not only target a formal level but also the interdisciplinary experiences gained in collaborations between statisticians and computer scientists encouraging future endeavors between both scientific domains.