

The advantages of our proposed Saturn coefficient over continuity and trustworthiness for UMAP dimensionality reduction evaluation

RESEARCH SEMINAR

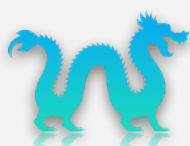
JULY 3 | 2026 11:00 AM - 12:00 PM



Davide Chicco

Dipartimento di Informatica Sistemistica e Comunicazione, Università di Milano-Bicocca, Milan, Italy
Institute for Health Policy Management and Evaluation, University of Toronto, Toronto, Ontario, Canada

Understanding the structure of a dataset is an easy task when the dimensions are two or three, but it can become extremely difficult when a dataset consists of tens, hundreds, or thousands of variables. Dimensionality reduction methods are computational techniques with solid mathematical foundations that allow for the projection of high-dimensional datasets into smaller data spaces. Uniform Manifold Approximation and Projection (UMAP) is one of the most effective and popular algorithms for dimensionality reduction, and has been proven effective on biomedical datasets, in particular. Even though UMAP is commonly utilized by thousands of researchers worldwide, no consensus has been reached on how to assess the output of dimensionality reduction informatively: to date, researchers often evaluate UMAP's outcomes by eyeballing its two-dimensional plots each time. To address this issue, we introduced the Saturn coefficient, a new simple statistical metric that expresses the conservation of local structure and the conservation of global structure in UMAP through a real value ranging from 0 (no preservation) to 1 (complete preservation).. Our results demonstrate the validity of our Saturn coefficient across all artificial datasets and in seven out of fifteen real-world biomedical datasets. We therefore recommend the use of our Saturn coefficient to anyone wishing to assess UMAP results: our statistic, for example, can be used to test several sets of UMAP hyperparameters and to select the best configuration among them.



Location: Room S1

Faculty of Electrical Engineering, University of Sarajevo