

How to assess my results? Evaluation metrics for supervised and unsupervised machine learning results

RESEARCH SEMINAR

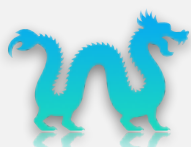
JULY 1 | 2026 11:00 AM - 12:00 PM



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

Machine learning has become a pervasive tool in any application of computer science nowadays, and new algorithms are invented on a regular basis. However, researchers often avoid innovation on the statistical metrics to utilize to assess the results, and use traditional scores that sometimes can lead to biased, misleading, or even wrong outcomes. In this talk, I will explain why some statistical metrics are better than others in assessing the quality of machine learning results. For binary classification, the Matthews correlation coefficient (MCC) should be employed; for regression analyses, the coefficient of determination (R^2) is the way to go; while for clustering, the Silhouette coefficient and Davies-Bouldin index are fine for convex-shaped clusters, while the DBCV index should be employed for concave-shaped nested clusters. Innovation should be pursued in all the scientific areas of machine learning, including the selection of the metrics for result assessment.



Location: Room S1

Faculty of Electrical Engineering, University of Sarajevo