



Contribution ID: 38

Type: **not specified**

Ivano Baronchelli - UMLAUT: A Data-Driven Machine Learning Approach with Automated Parameter Tuning and Outlier Detection (Invited)

Tuesday, 13 June 2023 17:30 (25 minutes)

UMLAUT (Unbiased Machine Learning Algorithm Using Topology) is a versatile machine learning algorithm initially developed as a variant of the highly efficient K-Nearest Neighbors (KNN) algorithm, known for its effectiveness in handling large datasets. While originally designed for the automatic classification of single emission lines, UMLAUT possesses remarkable potential that extends far beyond its initial application. The algorithm combines diverse sources of information, automatically identifying the most relevant ones, and enabling dimensionality reduction in problem-solving. With the inclusion of a modified DBSCAN-like module, UMLAUT facilitates the automatic identification of spurious sources and the discovery of new types of sources or data points not originally present in the training set. It can be used for regression, classification, and for the identification of different groups. UMLAUT is available in multiple programming languages and can be integrated into various algorithms, allowing users to combine it with different machine learning techniques. Its adaptability and versatility render it applicable to a wide range of contexts and problems, including sub-mm astronomy.

Session Classification: Machine learning, analysis tools