

NSF HDR ML Anomaly Detection Challenge

Monday 14 October 2024 11:00 (15 minutes)

Harnessing the Data Revolution (HDR), is an effort by the National Science Foundation (NSF) to promote the exploration of fundamental scientific questions using data-driven techniques. To raise interest in these approaches, and the HDR community, we have developed a Machine Learning (ML) challenge for anomaly detection, taking advantage of widespread data from several HDR institutes. This challenge seeks to connect these endeavors across several scientific disciplines, using a range of datasets spanning climate science, phylogenetics, materials science, and gravitational wave data from LIGO. The aim of participants is to design a novel ML model for their anomaly detection task. Using a single metric, their algorithm should detect anomalies across various datasets. We utilize the open-source benchmark ecosystem Codabench to host the challenge and ensure a Findable, Accessible, Interoperable, and Reusable (FAIR) dataset and workflow for participants from any community to contribute. In involving participating members from various communities, we promote collaboration and advance the broader goal of data-driven discovery.

Authors: GOVORKOVA, Katya (Massachusetts Inst. of Technology (US)); HARRIS, Philip Coleman (Massachusetts Inst. of Technology (US)); CHOU, Yuan-Tang (University of Washington (US))

Presenter: ANAND, Advaith (University of Washington (US))

Session Classification: Morning Plenary