

1146-68-415

Francis Charles Motta* (fmotta@fau.edu). *Automated Learning of Topological Features: Predicting Protein Stability.*

Considerable effort has been spent developing tools that are designed to strengthen the applicability of computational statistics and machine learning to persistent homology. For example, several methods that transform persistence diagrams into vectors, which aim to map the algebraic content of a persistence diagram to feature vectors to be subsequently used with machine learning paradigms, have been proposed. Often these transformations require choices, which may be regarded as model hyperparameters to be tuned to a given problem or determined by prior subject-matter expertise. We show how to use the Cover-tree Differencing via Entropy Reduction algorithm (CDER)—a general tool for performing supervised learning on labelled point clouds—to automatically learn discriminating and parsimonious sets of coordinates to vectorize persistence diagrams. We apply this technique to the problem of predicting synthetic protein stability and discuss its potential to reveal new biological insights via interpretation of the learned coordinates. This material is based upon work supported by the Defense Advanced Research Projects Agency (DARPA) and the Air Force Research Laboratory under Contract No. FA8750-17-C-0054 (and related contracts by SD2 Publication Consortium Members). (Received January 28, 2019)