THE UNIVERSITY OF CHICAGO | DEPARTMENT OF STATISTICS

MASTER'S THESIS PRESENTATION

JERRY WHITE
Department of Statistics
The University of Chicago

Exploring Denoising Autoencoder Architectures for Self-Supervised Learning

THURSDAY, February 16, 2023, at 1:00 PM
Zoom Meeting

ABSTRACT

Modern deep learning networks rely on massive amounts of labeled data for training, which can be prohibitively costly and time consuming to generate. The field of self-supervised learning aims to circumvent this issue by using unsupervised pretraining to generate lower-dimensional data representations. Cutting-edge methods utilize two "views" or transformations of each data point to create an embedding space that maintains the underlying structure of the data distribution. Out of these methods, "contrastive" methods-- which create positive pairs and negative pairs of each image in a batch-- reign supreme despite their dependence on hyper-parameters like batch size. In this work, we propose a denoising autoencoder-based training architecture as an alternative to contrastive methods. The novelty of this method comes from utilizing a denoising autoencoder in combination with various types of embedding losses to generate representations. We perform sensitivity analysis to obtain the optimal weighting of the embedding loss with respect to the denoising reconstruction loss and then compare the resulting models to contrastive methods. We evaluate the learned representations via a standard linear evaluation protocol as well as utilizing k-nearest neighbors. In linear evaluation, our method achieves 67.3\% classification accuracy, compared with 69.7\% for the contrastive SimCLR despite our method not requiring the use of negative pairs. The code repository used for all experiments can be found at https://github.com/artenyx/SSLProject.