Differentiable Unsupervised Feature Selection based on a Gated Laplacian

Ofir Lindenbaum [1*] Uri Shaham [1*] Jonathan Svirsky [2]
Erez Peterfreund[3]
Yuval Kluger[1†]

[1]Yale University, USA; [2] Technion, Israel; [3] Hebrew University, Israel

[†]Corresponding author. E-mail: yuval.kluger@yale.edu

Address: 333 Cedar St, New Haven, CT 06510, USA

[*] These authors contributed equally.

Abstract

Scientific observations may consist of a large number of variables (features). Identifying a subset of meaningful features is often ignored in unsupervised learning, despite its potential for unraveling clear patterns hidden in the ambient space. To address the problem of unsupervised feature selection, we propose a differentiable loss function that combines the Laplacian score, which favors low-frequency features, with a gating mechanism for feature selection. We improve the Laplacian score, by replacing it with a gated variant computed on a subset of features. This subset is obtained using a continuous approximation of Bernoulli variables whose parameters are trained to gate the full feature space. We mathematically motivate the proposed approach and demonstrate using extensive simulations its efficacy and advantage over current baselines.

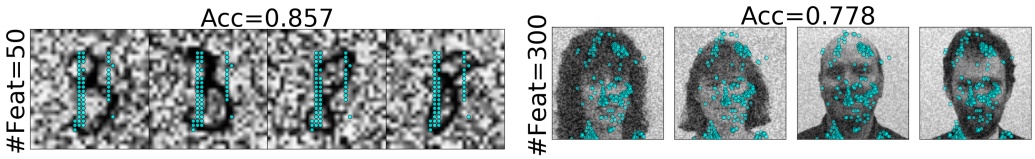

Figure 1: Examples from noisy MNIST and PIX10 along with the leading features selected by the proposed approach. We indicate the clustering accuracy obtained using these features above the images.