Homework 4, due September 23rd, 11:59pm

September 16, 2020

1. Implement the TISP variable selection method for classification (as described in pages 12 of the Regularized Loss course notes), with the hard-thresholding penalty, described in page 11 (with $\eta = 0$). Take special care to **normalize each column** of the X matrix to have zero mean and variance 1 and to use the same mean and standard deviation that you used for normalizing the train set also for normalizing the test set.

a) Using the `Gisette` data, train a TISP classifier on the training set, starting with $\mathbf{w}^{(\mathbf{0})} = \mathbf{0}$, with 100 iterations. Find appropriate λ-s to select approximately $10, 30, 100, 300, 500$ features. Plot the train misclassification error vs iteration number when selecting 300 features. Plot the final train and test misclassification error vs the number of selected features. Report in a table these misclassification errors on the training and test set, the corresponding numbers of selected features and the values of λ. (3 points)

b) Repeat point a) on the `dexter` dataset. (2 points)

c) Repeat point a) on the `madelon` dataset. (2 points)