

# Inductive Regularized Learning of Kernel Functions

Speaker: Prateek Jain

Joint work with: Brian Kulis, Inderjit S. Dhillon

Poster No: T55

# Kernel Learning

- **Goal:** Learn appropriate kernel function for a given task
  - Supervision should be flexible
    - Labels
    - Pair-wise distance constraints
    - Relative distance constraints
  - Generalization to new points with good error bounds
- Existing approaches:
  - Parametric approaches: Multiple Kernel Learning  
( $\kappa(x, y) = \sum_i \mu_i \kappa_i(x, y)$ ), Gaussian Kernel Learning  
( $\kappa(x, y) = \exp(-\|x - y\|^2 / \sigma)$ )
  - Non-parametric approaches: mostly limited to **Transductive** settings

# Our Approach

- Define a kernel *matrix* learning framework

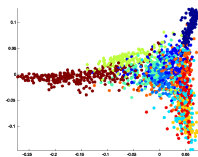
- $K = \Phi^T \Phi$ ,  $K \in \mathbb{R}^{n \times n}$
- Learn a new kernel matrix  $K_W$ :

$$\begin{aligned} \min_{K_W \succeq 0} f(K^{-1/2} K_W K^{-1/2}) \\ \text{s.t. } g_i(K_W) \leq b_i, \quad 1 \leq i \leq m. \end{aligned}$$

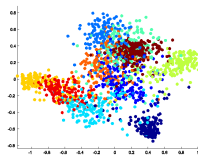
- $f: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ . *regularizer*,  $g_i$ : *constraints*.
- **Problem**: how to compute  $K_W$  for points outside  $[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]$ .
- Show that the learned kernel matrix parameterizes a general kernel *function*
  - Learns linear transformation kernel  $\kappa_W(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T W \mathbf{y}$  for *spectral* regularizer  $f$
  - Known to have good generalization bounds
- Use the learned kernel *function* to compute kernel value for unseen points

# Results

- Special Cases:
  - Kernelizes most existing convex formulations for Mahalanobis metric learning
  - Provides efficient method for inductive semi-supervised kernel dimensionality reduction
- Empirical Results:
  - UCI datasets: efficiently learns kernel for k-NN classification
  - USPS digits: maps digits dataset to 2-dimensions, better than baseline Kernel-PCA



(a) KPCA



(b) Our Method

- For more details: [T55](#)