New Theoretical Frameworks in Metric Learning

Irina Nicolae
PhD student

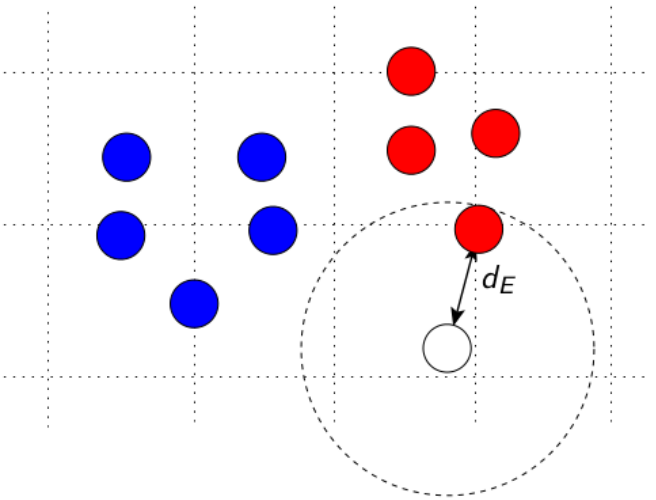
Supervisors:
Éric Gaussier, Joseph Fourier University
Marc Sebban, Jean Monnet University

October 15, 2014

Importance of Metrics

Some applications

- classification: k-Nearest Neighbors, SVMs;
- clustering: k-Means;
- information retrieval: distance between query and document.



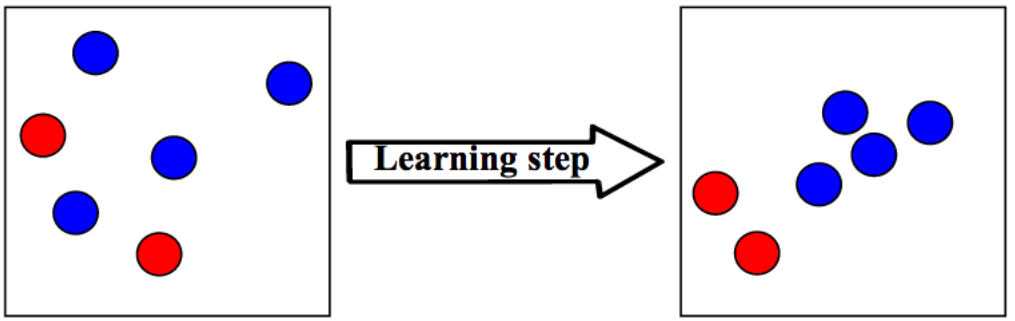
Importance of Metrics

Limitations

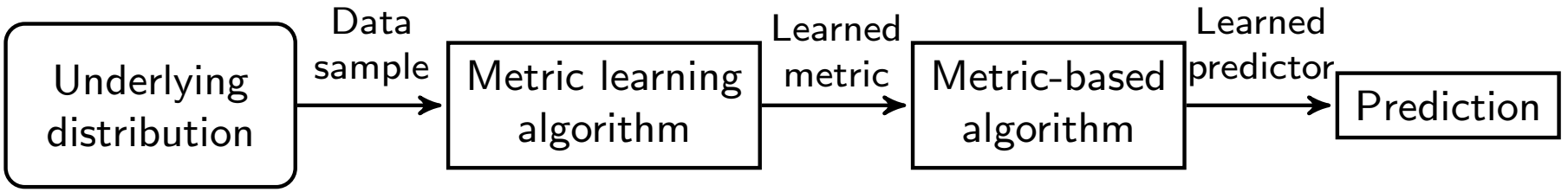
These distances are not **parameterizable** to take into account ground truth.

Solution

Learn the metric from data.



Metric Learning [Yan06, BHS13]



- Aims at optimizing parameterized distances/similarities.
- Leads to transformations of the input space before learning.
- Takes its constraints from side information of the input data.

Examples

- bilinear similarity $K_{\mathbf{A}}(\mathbf{x}, \mathbf{x}') = \mathbf{x}^T \mathbf{A} \mathbf{x}'$
- Mahalanobis distance $K_{\mathbf{A}}(\mathbf{x}, \mathbf{x}') = \sqrt{(\mathbf{x} - \mathbf{x}')^T \mathbf{A} (\mathbf{x} - \mathbf{x}')}$,

where $\mathbf{A} \in \mathbb{R}^{d \times d}$.

Semi-Supervised Metric Learning

Objective

- Provide a framework for **learning the metric and the separator** at the same time.
- Derive **consistency guarantees** for the learned similarity.

Setting

- use d_u unlabeled examples to contruct the feature space;
- use d_l labeled examples to learn a global linear separator α.

Method

Formulation

Solve w.r.t. $\alpha \in \mathbb{R}^{d_u}$ and $\mathbf{A} \in \mathbb{R}^{d \times d}$ (where $K_{\mathbf{A}}$ is the similarity):

$$\text{minimize } \frac{1}{d_l} \sum_{i=1}^{d_l} \left[1 - \sum_{j=1}^{d_u} \alpha_j l(x_i) K_{\mathbf{A}}(x_i, x_j) \right]_+ + \lambda \sum_{k=1}^{d} |\mathbf{A}_{kk}|$$

$$\text{s.t. } \sum_{j=1}^{d_u} |\alpha_j| \leq 1/\gamma$$

$$\mathbf{A} \text{ diagonal, } |\mathbf{A}_{kk}| \leq 1.$$

Prediction rule

$$l(x) = \text{sgn} \left(\sum_{j=1}^{d_u} \alpha_j K_{\mathbf{A}}(x, x'_j) \right)$$

Surveys

Aurélien Bellet, Amaury Habrard, and Marc Sebban.
A survey on metric learning for feature vectors and structured data.
Technical report, 2013.

Liu Yang.
Distance metric learning: A comprehensive survey, 2006.