Parent Item: Neighbourhood Components Analysis

Summary

1 Intro

- Traditional KNN suffers from computational drawback and modeling issues in how the distance metric should be defined
- This paper aims to learn quadratic distance metric to optimize and can force the learned metric to be low rank to reduce storage and search costs

2 Stochastic Nearest Neighbors for Distance Metric Learning

- Goal is to find a good distance metric
- Restrict themselves to only quadratic distance metrics
- introduces a stochastic neighbor assignment (each point gives a probability for it to be one of its neighbors)
- They have never observed an overtraining of this model
- Also learning a estimate of the optimal number of neighbors to have at the same time.

3 Low Rank Distance Metrics and Nonsquare Projection

- Reducing the dimensionality can vastly reduce storage and search time of KNN

4 Experiments in Metric Learning and Dimensionality Reduction

- Compared to classical methods of PCA and LDA, NCA is able to outperform them in class separation tasks

5 Extensions to Continuous Labels and Semi-Supervised Learning

- This can be extended further to semi-supervised tasks, continuous label types, and nonlinear methodology

6 Relationships to Other Methods and Conclusions

- NCA is a non-parametric learning method that can distance learn and dimension reduce in a unified manner