ScholarWorks@UMass Amherst

<u>OPEN ACCESS DISSERTATIONS</u>

Title

<u>A Geometric Framework for Transfer Learning Using Manifold Alignment</u>

Author

Chang Wang, *University of Massachusetts Amherst* Follow

Date of Award

9-2010

Document Type

Open Access Dissertation

Degree Name

Doctor of Philosophy (PhD)

Degree Program

Computer Science

First Advisor

Sridhar Mahadevan

Second Advisor

Andrew McCallum

Third Advisor

Erik Learned-Miller

Subject Categories

Computer Sciences

Abstract

Many machine learning problems involve dealing with a large amount of high-dimensional data across diverse domains. In addition, annotating or labeling the data is expensive as it involves significant human effort. This dissertation explores a joint solution to both these problems by exploiting the property that high-dimensional data in real-world application domains often lies on a lower-dimensional structure, whose geometry can be modeled as a graph or manifold. In particular, we propose a set of novel manifold-alignment based approaches for transfer learning. The proposed approaches transfer knowledge across different domains by finding low-dimensional embeddings of the datasets to a common latent space, which simultaneously match corresponding instances while preserving local or global geometry of each input dataset. We develop a novel two-step transfer learning method called Procrustes alignment. Procrustes alignment first maps the datasets to low-dimensional latent spaces reflecting their intrinsic geometries and then removes the translational, rotational and scaling components from one set so that the optimal alignment between the two sets can be achieved. This approach can preserve either global geometry or local geometry depending on the dimensionality reduction approach used in the first step. We propose a general one-step manifold alignment framework called manifold projections that can find alignments, both across instances as well as across features, while preserving local domain geometry. We develop and mathematically analyze several extensions of this framework to more challenging situations, including (1) when no correspondences across domains are given; (2) when the global geometry of each input domain needs to be respected; (3) when label information rather than correspondence information is available. A final contribution of this thesis is the study of multiscale methods for manifold alignment. Multiscale alignment automatically generates alignment results at different levels by discovering the shared intrinsic multilevel structures of the given datasets, providing a common representation across all input datasets.

Recommended Citation

Download

DOWNLOADS

Since December 01, 2010

Included in

Computer Sciences Commons

Share

COinS