

Metric Learning

Synthesis Lectures on Artificial Intelligence and Machine Learning

Editors

Ronald J. Brachman, *Yahoo! Labs*

William W. Cohen, *Carnegie Mellon University*

Peter Stone, *University of Texas at Austin*

Metric Learning

Aurélien Bellet, Amaury Habrard, and Marc Sebban

2015

Graph-Based Semi-Supervised Learning

Amarnag Subramanya and Partha Pratim Talukdar

2014

Robot Learning from Human Teachers

Sonia Chernova and Andrea L. Thomaz

2014

General Game Playing

Michael Genesereth and Michael Thielscher

2014

Judgment Aggregation: A Primer

Davide Grossi and Gabriella Pigozzi

2014

An Introduction to Constraint-Based Temporal Reasoning

Roman Barták, Robert A. Morris, and K. Brent Venable

2014

Reasoning with Probabilistic and Deterministic Graphical Models: Exact Algorithms

Rina Dechter

2013

Introduction to Intelligent Systems in Traffic and Transportation

Ana L.C. Bazzan and Franziska Klügl
2013

A Concise Introduction to Models and Methods for Automated Planning

Hector Geffner and Blai Bonet
2013

Essential Principles for Autonomous Robotics

Henry Hexmoor
2013

Case-Based Reasoning: A Concise Introduction

Beatriz López
2013

Answer Set Solving in Practice

Martin Gebser, Roland Kaminski, Benjamin Kaufmann, and Torsten Schaub
2012

Planning with Markov Decision Processes: An AI Perspective

Mausam and Andrey Kolobov
2012

Active Learning

Burr Settles
2012

Computational Aspects of Cooperative Game Theory

Georgios Chalkiadakis, Edith Elkind, and Michael Wooldridge
2011

Representations and Techniques for 3D Object Recognition and Scene Interpretation

Derek Hoiem and Silvio Savarese
2011

A Short Introduction to Preferences: Between Artificial Intelligence and Social Choice

Francesca Rossi, Kristen Brent Venable, and Toby Walsh
2011

Human Computation

Edith Law and Luis von Ahn
2011

Trading Agents

Michael P. Wellman
2011

Visual Object Recognition

Kristen Grauman and Bastian Leibe
2011

Learning with Support Vector Machines

Colin Campbell and Yiming Ying
2011

Algorithms for Reinforcement Learning

Csaba Szepesvári
2010

Data Integration: The Relational Logic Approach

Michael Genesereth
2010

Markov Logic: An Interface Layer for Artificial Intelligence

Pedro Domingos and Daniel Lowd
2009

Introduction to Semi-Supervised Learning

XiaojinZhu and Andrew B.Goldberg
2009

Action Programming Languages

Michael Thielscher
2008

Representation Discovery using Harmonic Analysis

Sridhar Mahadevan
2008

Essentials of Game Theory: A Concise Multidisciplinary Introduction

Kevin Leyton-Brown and Yoav Shoham
2008

A Concise Introduction to Multiagent Systems and Distributed Artificial Intelligence

Nikos Vlassis
2007

Intelligent Autonomous Robotics: A Robot Soccer Case Study

Peter Stone
2007

© Springer Nature Switzerland AG 2022

Reprint of original edition © Morgan & Claypool 2015

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means—electronic, mechanical, photocopy, recording, or any other except for brief quotations in printed reviews, without the prior permission of the publisher.

Metric Learning

Aurélien Bellet, Amaury Habrard, and Marc Sebban

ISBN: 978-3-031-00444-5 paperback

ISBN: 978-3-031-01572-4 ebook

DOI 10.1007/978-3-031-01572-4

A Publication in the Springer series

SYNTHESIS LECTURES ON ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Lecture #30

Series Editors: Ronald J. Brachman, *Yahoo! Labs*

William W. Cohen, *Carnegie Mellon University*

Peter Stone, *University of Texas at Austin*

Series ISSN

Print 1939-4608 Electronic 1939-4616

Metric Learning

Aurélien Bellet

Télécom ParisTech

Amaury Habrard

Université de Saint-Etienne

Marc Sebban

Université de Saint-Etienne

*SYNTHESIS LECTURES ON ARTIFICIAL INTELLIGENCE AND MACHINE
LEARNING #30*

ABSTRACT

Similarity between objects plays an important role in both human cognitive processes and artificial systems for recognition and categorization. How to appropriately measure such similarities for a given task is crucial to the performance of many machine learning, pattern recognition and data mining methods. This book is devoted to metric learning, a set of techniques to automatically learn similarity and distance functions from data that has attracted a lot of interest in machine learning and related fields in the past ten years. In this book, we provide a thorough review of the metric learning literature that covers algorithms, theory and applications for both numerical and structured data. We first introduce relevant definitions and classic metric functions, as well as examples of their use in machine learning and data mining. We then review a wide range of metric learning algorithms, starting with the simple setting of linear distance and similarity learning. We show how one may scale-up these methods to very large amounts of training data. To go beyond the linear case, we discuss methods that learn nonlinear metrics or multiple linear metrics throughout the feature space, and review methods for more complex settings such as multi-task and semi-supervised learning. Although most of the existing work has focused on numerical data, we cover the literature on metric learning for structured data like strings, trees, graphs and time series. In the more technical part of the book, we present some recent statistical frameworks for analyzing the generalization performance in metric learning and derive results for some of the algorithms presented earlier. Finally, we illustrate the relevance of metric learning in real-world problems through a series of successful applications to computer vision, bioinformatics and information retrieval.

KEYWORDS

metric learning, similarity learning, Mahalanobis distance, edit distance, structured data, learning theory

Contents

1	Introduction	1
1.1	Metric Learning in a Nutshell	2
1.2	Related Topics	2
1.3	Prerequisites and Notations	3
1.4	Outline	4
2	Metrics	7
2.1	General Definitions	7
2.2	Commonly Used Metrics	8
2.2.1	Metrics for Numerical Data	8
2.2.2	Metrics for Structured Data	11
2.3	Metrics in Machine Learning and Data Mining	14
3	Properties of Metric Learning Algorithms	17
4	Linear Metric Learning	19
4.1	Mahalanobis Distance Learning	19
4.1.1	Early Approaches	19
4.1.2	Regularized Approaches	22
4.2	Linear Similarity Learning	26
4.3	Large-Scale Metric Learning	28
4.3.1	Large n : Online, Stochastic and Distributed Optimization	28
4.3.2	Large d : Metric Learning in High Dimensions	31
4.3.3	Large n and large d	32
5	Nonlinear and Local Metric Learning	33
5.1	Nonlinear Methods	33
5.1.1	Kernelization of Linear Methods	33
5.1.2	Learning Nonlinear Forms of Metrics	35
5.2	Learning Multiple Local Metrics	37

6	Metric Learning for Special Settings	43
6.1	Multi-Task and Transfer Learning	43
6.2	Learning to Rank	46
6.3	Semi-Supervised Learning	47
6.3.1	Classic Setting	47
6.3.2	Domain Adaptation	48
6.4	Histogram Data	50
7	Metric Learning for Structured Data	53
7.1	String Edit Distance Learning	53
7.1.1	Probabilistic Methods	54
7.1.2	Gradient Descent Methods	56
7.2	Tree and Graph Edit Distance Learning	58
7.3	Metric Learning for Time Series	58
8	Generalization Guarantees for Metric Learning	61
8.1	Overview of Existing Work	61
8.2	Consistency Bounds for Metric Learning	63
8.2.1	Definitions	63
8.2.2	Bounds based on Uniform Stability	64
8.2.3	Bounds based on Algorithmic Robustness	68
8.3	Guarantees on Classification Performance	75
8.3.1	Good Similarity Learning for Linear Classification	75
8.3.2	Bounds based on Rademacher Complexity	80
9	Applications	85
9.1	Computer Vision	85
9.2	Bioinformatics	92
9.3	Information Retrieval	95
10	Conclusion	99
10.1	Summary	99
10.2	Outlook	99
A	Proofs of Chapter 8	103
A.1	Uniform Stability	103
A.2	Algorithmic Robustness	108
A.3	Similarity-based Linear Classifiers	112

Bibliography	115
Authors' Biographies	139