Automatic Speech and Speaker Recognition: Large Margin and Kernel Methods

edited by

Joseph Keshet
Samy Bengio

Contents

List of Contributors

Yasemin Altun
Dept. Schölkopf,
Max Planck Institute for Biological Cybernetics
yasemin.altun@tuebingen.mpg.de

Francis Bach
INRIA - Willow project,
Département d'Informatique,
Ecole Normale Supérieure
francis.bach@mines.org

Samy Bengio
Google Research Labs,
Google Inc.
bengio@google.com

Dan Chazan
Dept. of Electrical Engineering,
The Technion Institute of Technology
dan_chazan@yahoo.com

Koby Crammer
Dept. of Computer and Information Science,
University of Pennsylvania
crammer@cis.upenn.edu

Mark Gales
Dept. of Engineering,
University of Cambridge
mjfg@eng.cam.ac.uk

Yves Grandvalet
Heudiasyc,
Université de Technologie de Compiègne
yves.grandvalet@utc.fr

David Grangier
Dept. of Machine Learning,
NEC Laboratories America, Inc.
dgrangier@nec-labs.com

Michael I. Jordan
Computer Science Div. and Dept. of Statistics,
University of California at Berkeley
jordan@eecs.berkeley.edu

Joseph Keshet
Idiap Research Institute,
Martigny, Switzerland
jkeshet@idiap.ch

Johnny Mariéthoz
Idiap Research Institute,
Martigny, Switzerland
marietho@idiap.ch

Lawrence Saul
Dept. of Computer Science and Engineering,
University of California at San Diego
saul@cs.ucsd.edu

Brian Roark
Dept. of Computer Science and Electrical Eng.,
OGI School of Science and Engineering
roark@cslu.ogi.edu

Fei Sha
Computer Science Dept.,
University of Southern California
feisha@usc.edu

Shai Shalev-Shwartz
Toyota Technological Institute at Chicago
shai@tti-c.org

Yoram Singer
Google Research Labs,
Google Inc.
singer@google.com

Nathan Srebo
Toyota Technological Institute at Chicago
nati@uchicago.edu

Foreword

Preface

This is the first book dedicated to uniting research related to speech and speaker recognition based on the recent advances in large margin and kernel methods. The first part of the book presents theoretical and practical foundations of large margin and kernel methods, from support vector machines to large margin methods for structured learning. The second part of the book is dedicated to acoustic modeling of continuous speech recognizers, where the grounds for practical large margin sequence learning are set. The third part introduces large margin methods for discriminative language modeling. The last part of the book is dedicated to the application of keyword spotting, speaker verification and spectral clustering.

The book is an important reference to researchers and practitioners in the field of modern speech and speaker recognition. The purpose of the book is twofold; first, to set the theoretical foundation of large margin and kernel methods relevant to speech recognition domain; second, to propose a practical guide on implementation of these methods to the speech recognition domain. The reader is presumed to have basic knowledge of large margin and kernel methods and of basic algorithms in speech and speaker recognition.

August 2008

Joseph Keshet
Samy Bengio