

1. Record Nr.	UNINA9910146404403321
Titolo	Automatic speech and speaker recognition : large margin and kernel methods / / [edited by] Joseph Keshet, Samy Bengio
Pubbl/distr/stampa	Chichester, U.K. ; , : J. Wiley & Sons, , 2009 [Piscataqay, New Jersey] : , : IEEE Xplore, , [2009]
ISBN	1-282-34941-4 9786612349416 0-470-74204-6 0-470-74203-8
Descrizione fisica	1 online resource (271 p.)
Altri autori (Persone)	KeshetJoseph BengioSamy
Disciplina	006.4 006.4/54 006.454
Soggetti	Automatic speech recognition
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	Description based upon print version of record.
Nota di bibliografia	Includes bibliographical references and index.
Nota di contenuto	List of Contributors -- Preface -- I Foundations -- 1 Introduction (Samy Bengio and Joseph Keshet) -- 1.1 The Traditional Approach to Speech Processing -- 1.2 Potential Problems of the Probabilistic Approach -- 1.3 Support Vector Machines for Binary Classification -- 1.4 Outline -- References -- 2 Theory and Practice of Support Vector Machines Optimization (Shai Shalev-Shwartz and Nathan Srebro) -- 2.1 Introduction -- 2.2 SVM and L2-regularized Linear Prediction -- 2.3 Optimization Accuracy From a Machine Learning Perspective -- 2.4 Stochastic Gradient Descent -- 2.5 Dual Decomposition Methods -- 2.6 Summary -- References -- 3 From Binary Classification to Categorical Prediction (Koby Crammer) -- 3.1 Multi-category Problems -- 3.2 Hypothesis Class -- 3.3 Loss Functions -- 3.4 Hinge Loss Functions -- 3.5 A Generalized Perceptron Algorithm -- 3.6 A Generalized Passive / Aggressive Algorithm -- 3.7 A Batch Formulation -- 3.8 Concluding Remarks -- 3.9 Appendix. Derivations of the Duals of the Passive / Aggressive Algorithm and the Batch Formulation -- References -- II

Acoustic Modeling -- 4 A Large Margin Algorithm for Forced Alignment (Joseph Keshet, Shai Shalev-Shwartz, Yoram Singer and Dan Chazan) -- 4.1 Introduction -- 4.2 Problem Setting -- 4.3 Cost and Risk -- 4.4 A Large Margin Approach for Forced Alignment -- 4.5 An Iterative Algorithm -- 4.6 Efficient Evaluation of the Alignment Function -- 4.7 Base Alignment Functions -- 4.8 Experimental Results -- 4.9 Discussion -- References -- 5 A Kernel Wrapper for Phoneme Sequence Recognition (Joseph Keshet and Dan Chazan) -- 5.1 Introduction -- 5.2 Problem Setting -- 5.3 Frame-based Phoneme Classifier -- 5.4 Kernel-based Iterative Algorithm for Phoneme Recognition -- 5.5 Nonlinear Feature Functions -- 5.6 Preliminary Experimental Results -- 5.7 Discussion: Can we Hope for Better Results? -- References -- 6 Augmented Statistical Models: Using Dynamic Kernels for Acoustic Models (Mark J. F. Gales) -- 6.1 Introduction -- 6.2 Temporal Correlation Modeling. 6.3 Dynamic Kernels -- 6.4 Augmented Statistical Models -- 6.5 Experimental Results -- 6.6 Conclusions -- Acknowledgements -- References -- 7 Large Margin Training of Continuous Density Hidden Markov Models (Fei Sha and Lawrence K. Saul) -- 7.1 Introduction -- 7.2 Background -- 7.3 Large Margin Training -- 7.4 Experimental Results -- 7.5 Conclusion -- References -- III Language Modeling -- 8 A Survey of Discriminative Language Modeling Approaches for Large Vocabulary Continuous Speech Recognition (Brian Roark) -- 8.1 Introduction -- 8.2 General Framework -- 8.3 Further Developments -- 8.4 Summary and Discussion -- References -- 9 Large Margin Methods for Part-of-Speech Tagging (Yasemin Altun) -- 9.1 Introduction -- 9.2 Modeling Sequence Labeling -- 9.3 Sequence Boosting -- 9.4 Hidden Markov Support Vector Machines -- 9.5 Experiments -- 9.6 Discussion -- References -- 10 A Proposal for a Kernel Based Algorithm for Large Vocabulary Continuous Speech Recognition (Joseph Keshet) -- 10.1 Introduction -- 10.2 Segment Models and Hidden Markov Models -- 10.3 Kernel Based Model -- 10.4 Large Margin Training -- 10.5 Implementation Details -- 10.6 Discussion -- Acknowledgements -- References -- IV Applications -- 11 Discriminative Keyword Spotting (David Grangier, Joseph Keshet and Samy Bengio) -- 11.1 Introduction -- 11.2 Previous Work -- 11.3 Discriminative Keyword Spotting -- 11.4 Experiments and Results -- 11.5 Conclusions -- Acknowledgements -- References -- 12 Kernel-based Text-independent Speaker Verification (Johnny Mariethoz, Samy Bengio and Yves Grandvalet) -- 12.1 Introduction -- 12.2 Generative Approaches -- 12.3 Discriminative Approaches -- 12.4 Benchmarking Methodology -- 12.5 Kernels for Speaker Verification -- 12.6 Parameter Sharing -- 12.7 Is the Margin Useful for This Problem? -- 12.8 Comparing all Methods -- 12.9 Conclusion -- References -- 13 Spectral Clustering for Speech Separation (Francis R. Bach and Michael I. Jordan) -- 13.1 Introduction -- 13.2 Spectral Clustering and Normalized Cuts. 13.3 Cost Functions for Learning the Similarity Matrix -- 13.4 Algorithms for Learning the Similarity Matrix -- 13.5 Speech Separation as Spectrogram Segmentation -- 13.6 Spectral Clustering for Speech Separation -- 13.7 Conclusions -- References -- Index.

---

## Sommario/riassunto

This book discusses large margin and kernel methods for speech and speaker recognition. *Speech and Speaker Recognition: Large Margin and Kernel Methods* is a collation of research in the recent advances in large margin and kernel methods, as applied to the field of speech and speaker recognition. It presents theoretical and practical foundations of these methods, from support vector machines to large margin methods for structured learning. It also provides examples of large margin based acoustic modelling for continuous speech recognizers, where the

grounds for practical large margin sequence learning are set. Large margin methods for discriminative language modelling and text independent speaker verification are also addressed in this book. Key Features: . Provides an up-to-date snapshot of the current state of research in this field . Covers important aspects of extending the binary support vector machine to speech and speaker recognition applications . Discusses large margin and kernel method algorithms for sequence prediction required for acoustic modeling . Reviews past and present work on discriminative training of language models, and describes different large margin algorithms for the application of part-of-speech tagging . Surveys recent work on the use of kernel approaches to text-independent speaker verification, and introduces the main concepts and algorithms . Surveys recent work on kernel approaches to learning a similarity matrix from data This book will be of interest to researchers, practitioners, engineers, and scientists in speech processing and machine learning fields.

---