Lecture Notes in Artificial Intelligence 9925

Subseries of Lecture Notes in Computer Science

More information about this series at http://www.springer.com/series/1244

Ronald Ortner · Hans Ulrich Simon
Sandra Zilles (Eds.)

Algorithmic Learning Theory

27th International Conference, ALT 2016
Bari, Italy, October 19–21, 2016
Proceedings

Springer

Editors
Ronald Ortner
Montanuniversität Leoben
Leoben
Austria

Hans Ulrich Simon
Ruhr-Universität Bochum
Bochum
Germany

Sandra Zilles
University of Regina
Regina, SK
Canada

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Artificial Intelligence
ISBN 978-3-319-46378-0 ISBN 978-3-319-46379-7 (eBook)
DOI 10.1007/978-3-319-46379-7

Library of Congress Control Number: 2016950899

LNCS Sublibrary: SL7 – Artificial Intelligence

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Preface

This volume contains the papers presented at the 27th International Conference on Algorithmic Learning Theory (ALT 2016). ALT 2016 was co-located with the 19th International Conference on Discovery Science (DS 2016). Both conferences were held during October 19–21 in the beautiful city of Bari, Italy.

The technical program of ALT 2016 had five invited talks (presented jointly to both ALT 2016 and DS 2016) and 24 papers selected from 45 submissions by the ALT Program Committee. ALT is dedicated to the theoretical foundations of machine learning and provides a forum for high-quality talks and scientific interaction in areas such as statistical learning theory, online learning, inductive inference, query models, unsupervised learning, clustering, semi-supervised and active learning, stochastic optimization, high-dimensional and non-parametric inference, exploration–exploitation trade-off, bandit theory, reinforcement learning, planning, control, and learning with additional constraints. ALT is furthermore concerned with the analysis of the theoretical properties of existing algorithms such as boosting, kernel-based methods, SVM, Bayesian methods, graph- or manifold-based methods, methods for latent-variable estimation or clustering, decision tree methods, and information-based methods.

The present volume of LNAI contains the text of the 24 papers presented at ALT 2016 as well as the abstracts of the invited talks:

- Avrim Blum (Carnegie Mellon University, Pittsburgh):
  "Learning About Agents and Mechanisms from Opaque Transactions"
  (Invited talk for ALT 2016)
- Gianluca Bontempi (Interuniversity Institute of Bioinformatics, Brussels):
  "Perspectives of Feature Selection in Bioinformatics: From Relevance to Causal Inference"
  (Invited tutorial for DS 2016)
- Kristian Kersting (Technische Universität Dortmund):
  "Collective Attention on the Web"
  (Invited talk for DS 2016)
- Gábor Lugosi (Pompeu Fabra University, Barcelona):
  "How to Estimate the Mean of a Random Variable"
  (Invited tutorial for ALT 2016)
- John Shawe-Taylor (University College London):
  "Margin-Based Structured Output Learning"
  (Invited talk for ALT 2016 and DS 2016)

Since 1999, ALT has been awarding the E.M. Gold Award for the most outstanding student contribution. This year, the award was given to Areej Costa for her paper "Exact Learning of Juntas from Membership Queries" co-authored with Nader Bshouty.

ALT 2016 was the 27th meeting in the ALT conference series, established in Japan in 1990. The ALT series is supervised by its Steering Committee: Shai Ben-David (University of Waterloo, Canada), Marcus Hutter (Australian National University, Canberra, Australia), Sanjay Jain (National University of Singapore, Republic of Singapore), Ronald Ortner (Montanuniversität Leoben, Austria), Hans U. Simon (Ruhr-Universität Bochum, Germany), Frank Stephan (National University of Singapore, Republic of Singapore), Csaba Szepesvári (University of Alberta, Edmonton, Canada), Eiji Takimoto (Kyushu University, Fukuoka, Japan), Akihiro Yamamoto (Kyoto University, Japan), and Sandra Zilles (Chair, University of Regina, Canada).

We thank the following people and institutions who contributed to the success of the conference. Most importantly, we would like to thank the authors for contributing and presenting their work at the conference. Without their contribution this conference would not have been possible. We are very grateful to the Fondazione Puglia and to the Consorzio Interuniversitario Nazionale per l'Informatica (National Interuniversity Consortium for Informatics, CINI) for their financial support. We would also like to acknowledge the support of the European Commission through the project MAESTRA — Learning from Massive, Incompletely Annotated, and Structured Data (grant number ICT-2013-612944).

ALT 2016 and DS 2016 were organized by the University of Bari A. Moro. We thank the local arrangements chairs, Annalisa Appice, Corrado Loglisci, Gianvito Pio, Roberto Corizzo, and their team for their efforts in organizing the two conferences.

We are grateful for the collaboration with the conference series Discovery Science. In particular, we would like to thank the general chair of DS 2016 and ALT 2016, Donato Malerba, and the DS 2016 Program Committee chairs, Toon Calders and Michelangelo Ceci.

We are also grateful to EasyChair, the excellent conference management system, which was used for putting together the program for ALT 2016. EasyChair was developed mainly by Andrei Voronkov and is hosted at the University of Manchester. The system is free of charge.

We are grateful to the members of the Program Committee for ALT 2016 and the additional reviewers for their hard work in selecting a good program for ALT 2016. Special thanks go to Frank Stephan from the National University of Singapore for maintaining the ALT website. Last but not the least, we thank Springer for their support in preparing and publishing this volume in the *Lecture Notes in Artificial Intelligence* series.

July 2016

Ronald Ortner
Hans U. Simon
Sandra Zilles

Organization

Conference Chair

Donato Malerbo	University of Bari A. Moro, Italy

Program Committee

Nir Ailon	Israel Institute of Technology, Haifa, Israel
Dana Angluin	Yale University, USA
Peter Bartlett	UC Berkeley, USA and Queensland University of Technology, Brisbane, Australia
Shai Ben-David	University of Waterloo, Canada
Alina Beygelzimer	Yahoo! Labs, New York, USA
Corinna Cortes	Google Research, New York, USA
Malte Darnstädt	Continentale Insurance Group, Dortmund, Germany
Sanjoy Dasgupta	UC San Diego, USA
Rong Ge	Duke University, Durham, USA
Steve Hanneke	Princeton, USA
Kohei Hatano	Kyushu University, Fukuoka, Japan
Jyrki Kivinen	University of Helsinki, Finland
Wouter M. Koolen	Centrum Wiskunde & Informatica, Amsterdam, The Netherlands
Mehrdad Mahdavi	Toyota Technical Institute, Chicago, USA
Eric Martin	University of New South Wales, Sydney, Australia
Hanna Mazzawi	Google Research, Mountain View, California, USA
Mehryar Mohri	Courant Institute of Mathematical Sciences, New York, USA
Ronald Ortner (Chair)	Montanuniversität Leoben, Austria
Hans U. Simon (Chair)	Ruhr-Universität Bochum, Germany
Frank Stephan	National University of Singapore, Republic of Singapore
Gilles Stoltz	GREGHEC: HEC Paris - CNRS, France
Csaba Szepesvári	University of Alberta, Edmonton, Canada
Balázs Szörényi	Israel Institute of Technology, Haifa, Israel
György Turán	University of Illinois at Chicago, USA
Liwei Wang	Peking University, China
Akihiro Yamamoto	Kyoto University, Japan
Sandra Zilles	University of Regina, Canada

Local Arrangements

Annalisa Appice	University of Bari A. Moro, Italy
Roberto Corizzo	University of Bari A. Moro, Italy
Corrado Loglisci	University of Bari A. Moro, Italy
Gianvito Pio	University of Bari A. Moro, Italy

Additional Reviewers

Giulia Desalvo
Christos Dimitrakakis
András György
Elad Hoffer
Prateek Jain
Pooria Joulani
Sumeet Katariya
Vitaly Kuznetsov
Tor Lattimore
Guillaume Lecué
Alan Malek
David Pal
Bernardo Ávila Pires
Lev Reyzin
Afshin Rostamizadeh
Vikas Sindhwani
Marta Soare
Daiki Suehiro
Ambuj Tewari
Scott Yang
Felix Yu
Huizhen Yu

Sponsoring Institutions

Fondazione Puglia
Consorzio Interuniversitario Nazionale per l'Informatica, CINI
European Commission through the project MAESTRA (*Learning from Massive, Incompletely Annotated, and Structured Data*), Grant number ICT-2013-612944

Abstract of Invited Talks

Learning about Agents and Mechanisms from Opaque Transactions

Avrim Blum

School of Computer Science, Carnegie Mellon University,
Pittsburgh, PA 15213–3891, USA
avrim@cs.cmu.edu

In this talk I will discuss some learning problems coming from the area of algorithmic economics. I will focus in particular on settings known as combinatorial auctions in which agents have preferences over items or sets of items, and interact with an auction or allocation mechanism that determines what items are given to which agents. We consider the perspective of an outside observer who each day can only see which agents show up and what they get, or perhaps just which agents' needs are satisfied and which are not. Our goal will be from observing a series of such interactions to try to learn the agent preferences and perhaps also the rules of the allocation mechanism.

As an example, consider observing web pages where the agents are advertisers and the winners are those whose ads show up on the given page. Or consider observing the input-output behavior of a cloud computing service, where the input consists of a set of agents requesting service, and the output is a partition of them into some whose requests are actually fulfilled and the rest that are not—due to overlap of their resource needs with higher-priority requests. From such input-output behavior, we would like to learn the underlying structure. We also consider a classic Myerson single-item auction setting, where from observing who wins and also being able to participate ourselves we would like to learn the agents' valuation distributions.

In examining these problems we will see connections to decision-list learning and to Kaplan-Meier estimators from medical statistics.

This talk is based on work joint with Yishay Mansour and Jamie Morgenstern.

Perspectives of Feature Selection in Bioinformatics: From Relevance to Causal Inference

Gianluca Bontempi

Machine Learning Group, Interuniversity
Institute of Bioinformatics in Brussels (IB)^2
Université libre de Bruxelles, Bld de Triomphe, 1050 Brussels, Belgium
mlg.ulb.ac.be

A major goal of the scientific activity is to model real phenomena by studying the dependency between entities, objects or more in general variables. Sometimes the goal of the modeling activity is simply predicting future behaviors. Sometimes the goal is to understand the causes of a phenomenon (e.g. a disease). Finding causes from data is particular challenging in bioinformatics where often the number of features (e.g. number of probes) is huge with respect to the number of samples [5]. In this context, even when experimental interventions are possible, performing thousands of experiments to discover causal relationships between thousands of variables is not practical. Dimensionality reduction techniques have been largely discussed and used in bioinformatics to deal with the curse of dimensionality. However, most of the time these techniques focus on improving prediction accuracy, neglecting causal aspects. This tutorial will introduce some basics of causal inference and will discuss some open issues: may feature selection techniques be useful also for causal feature selection? Is prediction accuracy compatible with causal discovery [2]? How to deal with Markov indistinguishable settings [1]? Recent results based on information theory [3], and some learned lessons from a recent Kaggle competition [4] will be used to illustrate the issue.

References

1. Bontempi, G., Flauder, M.: From dependency to causality: a machine learning approach. JMLR **15**(16), 2437–2457 (2015)
2. Bontempi, G., Haibe-Kains, B., Desmedt, C., Sotiriou, C., Quackenbush, J.: Multiple-input multiple-output causal strategies for gene selection. BMC Bioinformatics **12**(1), 458 (2011)
3. Bontempi, G., Meyer, P.E.: Causal filter selection in microarray data. In: Proceedings of ICML (2010)
4. Bontempi, G., Olsen, C., Flauder, M.: D2C: predicting causal direction from dependency features (2014). R package version 1.1
5. Meyer, P.E., Bontempi, G.: Information-theoretic gene selection in expression data. In: Biological Knowledge Discovery Handbook. IEEE Computer Society (2014)

Margin Based Structured Output Learning

John Shawe-Taylor

Department of Computer Science, CSML,
University College London, London WC1E 6EA, UK
j.shawe-taylor@cs.ucl.ac.uk

Structured output learning has been developed to borrow strength across multi-dimensional classifications. There have been approaches to bounding the performance of these classifiers based on different measures such as microlabel errors with a fixed simple output structure. We present a different approach and analysis starting from the assumption that there is a margin attainable in some unknown or fully connected output structure. The analysis and algorithms flow from this assumption but in a way that the associated inference becomes tractable while the bounds match those attained were we to use the full structure. There are two variants depending on how the margin is estimated. Experimental results show the relative strengths of these variants, both algorithmically and statistically.

Collective Attention on the Web

Kristian Kersting

Computer Science Department, TU Dortmund University,
44221 Dortmund, Germany
Kristian.Kersting@cs.tu-dortmund.de

It's one of the most popular YouTube videos ever produced, having been viewed more than 840 million times. Its hard to understand why this clip is so famous and actually went viral, since nothing much happens. Two little boys, Charlie and Harry, are sitting in a chair when Charlie, the younger brother, mischievously bites Harrys finger. There's a shriek and then a laugh. The clip is called "Charlie Bit My Finger–Again!"

Generally, understanding the dynamics of collective attention is central to an information age where millions of people leave digital footprints everyday. So, can we capture the dynamics of collective attention mathematically? Can we even gain insights into the underlying physical resp. social processes? Is it for instance fair to call the video "viral" in an epidemiological sense?

In this talk I shall argue that computational methods of collective attention are not insurmountable. I shall review the methods we have developed to characterize, analyze, and even predict the dynamics of collective attention among millions of users to and within social media services. For instance, we found that collective attention to memes and social media grows and subsides in a highly regular manner, well explained by economic diffusion models [2, 4]. Using mathematical epidemiology, we find that so-called viral videos show very high infection rates and, hence, should indeed be called viral [1]. Moreover, the spreading processes may also be related to the underlying network structures, suggesting for instance a physically plausible model of the distance distributions of undirected networks [3]. All this favors machine learning and discovery science approaches that produce physically plausible models.

This work was partly supported by the Fraunhofer ICON project SoFWIReD and by the DFG Collaborative Research Center SFB 876 project A6.

References

1. Bauckhage, C., Hadiji, F., Kersting, K.: How viral are viral videos? In: Proceedings of the Ninth International Conference on Web and Social Media (ICWSM 2015), pp. 22–30 (2015)
2. Bauckhage, C., Kersting, K., Hadiji, F.: Mathematical models of fads explain the temporal dynamics of internet memes. In: Proceedings of the Seventh International Conference on Weblogs and Social Media (ICWSM 2013) (2013)
3. Bauckhage, C., Kersting, K., Hadiji, F.: Parameterizing the distance distribution of undirected networks. In: Proceedings of the Thirty-First Conference on Uncertainty in Artificial Intelligence (UAI 2015), pp. 121–130 (2015)

4. Bauckhage, C., Kersting, K., Rastegarpanah, B.: Collective attention to social media evolves according to diffusion models. In: 23rd International World Wide Web Conference (WWW 2014), pp. 223–224 (2014)

How to Estimate the Mean of a Random Variable?

Gabor Lugosi

Department of Economics, Pompeu Fabra University,
Ramon Trias Fargas 25-27, 08005, Barcelona, Spain
gabor.lugosi@gmail.com

Given n independent, identically distributed copies of a random variable, one is interested in estimating the expected value. Perhaps surprisingly, there are still open questions concerning this very basic problem in statistics. In this talk we are primarily interested in non-asymptotic sub-Gaussian estimates for potentially heavy-tailed random variables. We discuss various estimates and extensions to high dimensions. We apply the estimates for statistical learning and regression function estimation problems. The methods improve on classical empirical minimization techniques.

This talk is based on joint work with Emilien Joly, Luc Devroye, Matthieu Lerasle, Roberto Imbuzeiro Oliveira, and Shahar Mendelson.

Contents