Lecture Notes in Artificial Intelligence **9047**

Subseries of Lecture Notes in Computer Science

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

Alexander Gammerman · Vladimir Vovk
Harris Papadopoulos (Eds.)

Statistical Learning and Data Sciences

Third International Symposium, SLDS 2015
Egham, UK, April 20–23, 2015
Proceedings

Springer

Editors
Alexander Gammerman
University of London
Egham, Surrey
UK

Harris Papadopoulos
Frederick University
Nicosia
Cyprus

Vladimir Vovk
University of London
Egham, Surrey
UK

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Artificial Intelligence
ISBN 978-3-319-17090-9 ISBN 978-3-319-17091-6 (eBook)
DOI 10.1007/978-3-319-17091-6

Library of Congress Control Number: 2015935220

LNCS Sublibrary: SL7 – Artificial Intelligence

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

In memory of Alexey Chervonenkis

Preface

This volume contains the Proceedings of the Third Symposium on Statistical Learning and Data Sciences, which was held at Royal Holloway, University of London, UK, during April 20–23, 2015. The original idea of the Symposium on Statistical Learning and Data Sciences is due to two French academics – Professors Mireille Gettler Summa and Myriam Touati – from Paris Dauphine University. Back in 2009 they thought that a "bridge" was required between various academic groups that were involved in research on Machine Learning, Statistical Inference, Pattern Recognition, Data Mining, Data Analysis, and so on; a sort of multilayer bridge to connect those fields. This is reflected in the symposium logo with the Passerelle Simone-de-Beauvoir bridge. The idea was implemented and the First Symposium on Statistical Learning and Data Sciences was held in Paris in 2009. The event was indeed a great "bridge" between various communities with interesting talks by J.-P. Benzecri, V. Vapnik, A. Chervonenkis, D. Hand, L. Bottou, and many others. Papers based on those talks were later presented in a volume of the *Modulad* journal and separately in a post-symposium book entitled *Statistical Learning and Data Sciences*, published by Chapman & Hall, CRC Press. The second symposium, which was equally successful, was held in Florence, Italy, in 2012.

Over the last 6 years since the first symposium, the progress in the theory and applications of learning and data mining has been very impressive. In particular, the arrival of technologies for collecting huge amounts of data has raised many new questions about how to store it and what type of analytics are able to handle it – what is now known as Big Data. Indeed, the sheer scale of the data is very impressive – for example, the Large Hadron Collider computers have to store 15 petabytes a year (1 petabyte = 10^{15} bytes). Obviously, handling this requires the usage of distributed clusters of computers, streaming, parallel processing, and other technologies. This volume is concerned with various modern techniques, some of which could be very useful for handling Big Data.

The volume is divided into five parts. The first part is devoted to two invited papers by Vladimir Vapnik. The first paper, "Learning with Intelligent Teacher: Similarity Control and Knowledge Transfer," is a further development of his research on learning with privileged information, with a special attention to the knowledge representation problem. The second, "Statistical Inference Problems and their Rigorous Solutions," suggests a novel approach to pattern recognition and regression estimation. Both papers promise to become milestones in the developing field of statistical learning.

The second part consists of 16 papers that were accepted for presentation at the main event, while the other three parts reflect new research in important areas of statistical learning to which the symposium devoted special sessions. Specifically the special sessions included in the symposium's program were:

– Special Session on Conformal Prediction and its Applications (CoPA 2015), organized by Harris Papadopoulos (Frederick University, Cyprus), Alexander Gammerman (Royal Holloway, University of London, UK), and Vladimir Vovk (Royal Holloway, University of London, UK).

– Special Session on New Frontiers in Data Analysis for Nuclear Fusion, organized by Jesus Vega (Asociacion EURATOM/CIEMAT para Fusion, Spain).
– Special Session on Geometric Data Analysis, organized by Fionn Murtagh (Goldsmith College London, UK).

Overall, 36 papers were accepted for presentation at the symposium after being reviewed by at least two independent academic referees. The authors of these papers come from 17 different countries, namely: Brazil, Canada, Chile, China, Cyprus, Finland, France, Germany, Greece, Hungary, India, Italy, Russia, Spain, Sweden, UK, and USA.

A special session at the symposium was devoted to the life and work of Alexey Chervonenkis, who tragically died in September 2014. He was one of the founders of modern Machine Learning, a beloved colleague and friend. All his life he was connected with the Institute of Control Problems in Moscow, over the last 15 years he worked at Royal Holloway, University of London, while over the last 7 years he also worked for the Yandex Internet company in Moscow. This special session included talks in memory of Alexey by Vladimir Vapnik – his long standing colleague and friend – and by Alexey's former students and colleagues.

We are very grateful to the Program and Organizing Committees, the success of the symposium would have been impossible without their hard work. We are indebted to the sponsors: the Royal Statistical Society, the British Computer Society, the British Classification Society, Royal Holloway, University of London, and Paris Dauphine University. Our special thanks to Yandex for their help and support in organizing the symposium and the special session in memory of Alexey Chervonenkis. This volume of the proceedings of the symposium is also dedicated to his memory. Rest in peace, dear friend.

February 2015

Alexander Gammerman
Vladimir Vovk
Harris Papadopoulos

Organization

General Chairs

Alexander Gammerman, UK
Vladimir Vovk, UK

Organizing Committee

Zhiyuan Luo, UK
Mireille Summa, France
Yuri Kalnishkan, UK
Myriam Touati, France
Janet Hales, UK

Program Committee Chairs

Harris Papadopoulos, Cyprus
Xiaohui Liu, UK
Fionn Murtagh, UK

Program Committee Members

Vineeth Balasubramanian, India
Giacomo Boracchi, Italy
Paula Brito, Portugal
Léon Bottou, USA
Lars Carlsson, Sweden
Jane Chang, UK
Frank Coolen, UK
Gert de Cooman, Belgium
Jesus Manuel de la Cruz, Spain
Jose-Carlos Gonzalez-Cristobal, Spain
Anna Fukshansky, Germany
Barbara Hammer, Germany
Shenshyang Ho, Singapore
Carlo Lauro, Italy
Guang Li, China
David Lindsay, UK
Henrik Linusson, Sweden
Hans-J. Lenz, Germany
Ilia Nouretdinov, UK
Matilde Santos, Spain
Victor Solovyev, Saudi Arabia
Jesus Vega, Spain
Rosanna Verde, Italy

Contents