

The series *Studies in Computational Intelligence* (SCI) publishes new developments and advances in the various areas of computational intelligence – quickly and with a high quality. The intent is to cover the theory, applications, and design methods of computational intelligence, as embedded in the fields of engineering, computer science, physics and life science, as well as the methodologies behind them. The series contains monographs, lecture notes and edited volumes in computational intelligence spanning the areas of neural networks, connectionist systems, genetic algorithms, evolutionary computation, artificial intelligence, cellular automata, self-organizing systems, soft computing, fuzzy systems and hybrid intelligent systems. Critical to both contributors and readers are the short publication time and world-wide distribution – this permits a rapid and broad dissemination of research results.

Statistical implicative analysis is a data analysis method created by Régis Gras almost thirty years ago which has a significant impact on a variety of areas ranging from pedagogical and psychological research to data mining. Statistical implicative analysis (SIA) provides a framework for evaluating the strength of implications; such implications are formed through common knowledge acquisition techniques in any learning process, human or artificial. This new concept has developed into a unifying methodology, and has generated a powerful convergence of thought between mathematicians, statisticians, psychologists, specialists in pedagogy and last, but not least, computer scientists specialized in data mining.

This volume collects significant research contributions of several rather distinct disciplines that benefit from SIA. Contributions range from psychological and pedagogical research, bioinformatics, knowledge management, and data mining.

Gras · Suzuki · Guillet
Spagnolo (Eds.)



Statistical Implicative Analysis

Régis Gras · Einoshin Suzuki
Fabrice Guillet · Filippo Spagnolo
(Eds.)

Statistical Implicative Analysis

Theory and Applications

 Springer

ISSN 1860-949X

ISBN 978-3-540-78982-6



9 783540 789826

Available
online

springerlink.com

Preface

Statistical implicative analysis is a data analysis method created by Régis Gras almost thirty years ago which has a significant impact on a variety of areas ranging from pedagogical and psychological research to data mining. This new concept has developed into a unifying methodology, and has generated a powerful convergence of thought between mathematicians, statisticians, psychologists, specialists in pedagogy and last, but not least, computer scientists specialized in data mining.

Statistical implicative analysis (SIA) provides a framework for evaluating the strength of implications; such implications are formed through common knowledge acquisition techniques in any learning process, human or artificial. Therefore, the epistemological interest of SIA is, in my opinion, of universal interest for researchers. In many applications implications appear as “rules” and, as it is often the case, rules have exceptions. SIA provides a powerful instrument for quantifying the quality of a rule taking into account the reality of these exceptions. Many applications, especially in data mining, extract large sets of rules that are impossible to assimilate by humans and used efficiently in decision processes. Therefore, it is important to develop measures of interestingness for these rules and the success of SIA-based techniques in this direction is indisputable.

This volume collects significant research contributions of several rather distinct disciplines that benefit from SIA. Contributions range from psychological and pedagogical research, bioinformatics, knowledge management, and data mining.

The first applications of SIA were in the realm of didactics and this field is richly represented here by several contributions that focus on such diverse problems as didactics of algebra and geometry, the teaching of functions representations and graphing, Bayesian inference, and student representations of physical activities.

Interesting data mining applications authored by leading researchers in the field range from applying SIA in the study of rules produced by decision trees, association rules generated by the analysis of transactional data, tempo-

VI Preface

ral rules, measures of interestingness for various types of rules, and hierarchical organization of rules. A novel method for analyzing DNA microarrays is formulated using SIA concepts. Furthermore, applications of SIA to the study of ontologies and textual taxonomies, as well as applications to fuzzy knowledge discovery are also included.

We have here a new volume that confirms the validity of a novel and powerful statistical methodology, though many convincing applications. The contributors have done a masterful job of exposition.

After reading this book, I have in mind a few applications of SIA in my own research. I am convinced that the readers will find this volume as stimulating as I did.

Boston,
September, 2007

Prof. Dan A. Simovici
Department of Computer Science
University of Massachusetts at Boston University

Review Committee

All published chapters have been reviewed by at least 2 referees.

- Saddo Ag Almouloud (University of Sao Paulo, Brazil)
- Carmen Batanero (University of Grenada)
- Hans Bock (Aachen University, Germany)
- Henri Briand (LINA, University of Nantes, France)
- Guy Brousseau (University of Bordeaux 3, France)
- Alex Freitas (University of Kent, UK)
- Athanasios Gagatsis (University of Chyprius)
- Robin Gras (University of Windsor, Canada)
- Howard Hamilton (University of Regina, Canada)
- Jiawei Han (University of Illinois, USA)
- David J. Hand (Imperial College, London, UK)
- André Hardy (University of Namur, Belgium)
- Robert Hilderman (University of Regina, Canada)
- Yves Kodratoff (LRI, University of Paris-Sud, France)
- Pascale Kuntz (LINA, University of Nantes, France)
- Ludovic Lebart (ENST, Paris, France)
- Amédéo Napoli (LORIA, University of Nancy, France)
- Maria-Gabriella Ottaviani (University of Roma, Italy)
- Balaaji Padmanabhan (University of Pennsylvania, USA)
- Jean-Paul Rassin (University of Namur, Belgium)
- Jean-Claude Régnier (University of Lyon 2, France)
- Gilbert Ritschard (Geneve University, Switzerland)
- Lorenza Saitta (University of Piemont, Italy)
- Gilbert Saporta (CNAM, Paris, France)
- Dan Simovici (University of Massachusetts Boston, USA)
- Djamel Zighed (ERIC, University of Lyon 2, France)

Associated Reviewers

Nadja Maria Acioly- Régnier,	Carmen Diaz, Pablo Gregori,	Philippe Lenca, Elsa Malisani,
Angela Alibrandi,	Alain Kuzniak,	Rajesh Natajara,
Jérôme Azé,	Eduardo Lacasta,	Pilar Orús,
Maurice Bernadet,	Dominique Lahanier-	Gérard Ramstein,
Julien Blanchard,	Reuter,	Ansaf Salleb,
Catherine-Marie Chiocca,	Stéphane Lallich,	Aldo Scimone,
Raphaël Couturier,	Letizia La Tona,	Benoît Vaillant,
Stéphane Daviet,	Patrick Leconte,	Ingrid Verscheure
Jérôme David,	Rémi Lehn,	

Manuscript coordinator

Bruno Pinaud (LINA, University of Nantes, France)

Acknowledgments

The editors would like to thank the chapter authors for their insights and contributions to this book.

The editors would also like to acknowledge the members of the review committee and the associated referees for their involvement in the review process of the book, and without whose support the book would not have been satisfactorily completed.

A special thank goes to H. Briand for his encouragements.

Thanks also to J. Blanchard who has managed the cyberchair web site.

Finally, we thank Springer and the publishing team, and especially T. Ditzinger and J. Kacprzyk, for their confidence in our project.

Nantes,
December 2007

Régis Gras
Eimoshin Suzuki
Fabrice Guillet
Filippo Spagnolo

Contents

Introduction

Régis Gras, Einoshin Suzuki, Fabrice Guillet, Filippo Spagnolo 1

Part I Methodology and concepts for SIA

An overview of the Statistical Implicative Analysis (SIA) development

Régis Gras, Pascale Kuntz 11

CHIC: Cohesive Hierarchical Implicative Classification

Raphaël Couturier 41

Assessing the interestingness of temporal rules with Sequential Implication Intensity

Julien Blanchard, Fabrice Guillet, Régis Gras 55

Part II Application to concept learning in education, teaching, and didactics

Student's Algebraic Knowledge Modelling: Algebraic Context as Cause of Student's Actions

Marie-Caroline Croset, Jana Trgalova, Jean-François Nicaud 75

The graphic illusion of high school students

Eduardo Lacasta, Miguel R. Wilhelmi 99

Implicative networks of student's *representations* of Physical Activities

Catherine-Marie Chiocca, Ingrid Verscheure 119

A comparison between the hierarchical clustering of variables, implicative statistical analysis and confirmatory factor analysis <i>Iliada Elia, Athanasios Gagatsis</i>	131
Implications between learning outcomes in elementary bayesian inference <i>Carmen Díaz, Inmaculada de la Fuente, Carmen Batanero</i>	163
Personal Geometrical Working Space: a Didactic and Statistical Approach <i>Alain Kuzniak</i>	185
<hr/>	
Part III A methodological answer in various application frameworks	
<hr/>	
Statistical Implicative Analysis of DNA microarrays <i>Gerard Ramstein</i>	205
On the use of Implication Intensity for matching ontologies and textual taxonomies <i>Jérôme David, Fabrice Guillet, Henri Briand, Régis Gras</i>	227
Modelling by Statistic in Research of Mathematics Education <i>Elsa Malisani and Aldo Scimone and Filippo Spagnolo</i>	247
Didactics of Mathematics and Implicative Statistical Analysis <i>Dominique Lahanier-Reuter</i>	277
Using the Statistical Implicative Analysis for Elaborating Behavioral Referentials <i>Stéphane Daviet, Fabrice Guillet, Henri Briand, Serge Baquedano, Vincent Philippé, Régis Gras</i>	299
Fictitious Pupils and Implicative Analysis: a Case Study <i>Pilar Orús, Pablo Gregori</i>	321
Identifying didactic and sociocultural obstacles to conceptualization through Statistical Implicative Analysis <i>Nadja Maria Acioly-Régnier, Jean-Claude Régnier</i>	347
<hr/>	
Part IV Extensions to rule interestingness in data mining	
<hr/>	
Pitfalls for Categorizations of Objective Interestingness Measures for Rule Discovery <i>Einoshin Suzuki</i>	383

Inducing and Evaluating Classification Trees with Statistical Implicative Criteria
Gilbert Ritschard, Vincent Pisetta, Djamel A. Zighed 397

On the behavior of the generalizations of the intensity of implication: A data-driven comparative study
Benoît Vaillant, Stéphane Lallich, Philippe Lenca 421

The TVpercent principle for the counterexamples statistic
Ricco Rakotomalala, Alain Morineau 449

User-System Interaction for Redundancy-Free Knowledge Discovery in Data
Rémi Lehn, Henri Briand, Fabrice Guillet 463

Fuzzy Knowledge Discovery Based on Statistical Implication Indexes
Maurice Bernadet 481

About the editors 507

Index 509