Game Theory for Data Science

Eliciting Truthful Information

Synthesis Lectures on Artificial Intelligence and Machine Learning

Robot Learning from Human Teachers
Sonia Chernova and Andrea L. Thomaz
2014

General Game Playing
Michael Genesereth and Michael Thielscher
2014

Judgment Aggregation: A Primer
Davide Grossi and Gabriella Pigozzi
2014

An Introduction to Constraint-Based Temporal Reasoning
Roman Barták, Robert A. Morris, and K. Brent Venable
2014

Reasoning with Probabilistic and Deterministic Graphical Models: Exact Algorithms
Rina Dechter
2013

Introduction to Intelligent Systems in Traffic and Transportation
Ana L.C. Bazzan and Franziska Klügl
2013

A Concise Introduction to Models and Methods for Automated Planning
Hector Geffner and Blai Bonet
2013

Essential Principles for Autonomous Robotics
Henry Hexmoor
2013

Case-Based Reasoning: A Concise Introduction
Beatriz López
2013

Answer Set Solving in Practice
Martin Gebser, Roland Kaminski, Benjamin Kaufmann, and Torsten Schaub
2012

Planning with Markov Decision Processes: An AI Perspective
Mausam and Andrey Kolobov
2012

Active Learning
Burr Settles
2012

Computational Aspects of Cooperative Game Theory
Georgios Chalkiadakis, Edith Elkind, and Michael Wooldridge
2011

Representations and Techniques for 3D Object Recognition and Scene Interpretation
Derek Hoiem and Silvio Savarese
2011

A Short Introduction to Preferences: Between Artificial Intelligence and Social Choice
Francesca Rossi, Kristen Brent Venable, and Toby Walsh
2011

Human Computation
Edith Law and Luis von Ahn
2011

Trading Agents
Michael P. Wellman
2011

Visual Object Recognition
Kristen Grauman and Bastian Leibe
2011

Learning with Support Vector Machines
Colin Campbell and Yiming Ying
2011

Algorithms for Reinforcement Learning
Csaba Szepesvári
2010

Data Integration: The Relational Logic Approach
Michael Genesereth
2010

Markov Logic: An Interface Layer for Artificial Intelligence
Pedro Domingos and Daniel Lowd
2009

Introduction to Semi-Supervised Learning
XiaojinZhu and Andrew B.Goldberg
2009

Action Programming Languages
Michael Thielscher
2008

Representation Discovery using Harmonic Analysis
Sridhar Mahadevan
2008

Essentials of Game Theory: A Concise Multidisciplinary Introduction
Kevin Leyton-Brown and Yoav Shoham
2008

A Concise Introduction to Multiagent Systems and Distributed Artificial Intelligence
Nikos Vlassis
2007

Intelligent Autonomous Robotics: A Robot Soccer Case Study
Peter Stone
2007

Game Theory for Data Science: Eliciting Truthful Information

Boi Faltings and Goran Radanovic

ISBN: 978-3-031-00449-0 paperback
ISBN: 978-3-031-01577-9 ebook

DOI 10.1007/978-3-031-01577-9

A Publication in the Springer series
SYNTHESIS LECTURES ON ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Lecture #35
Series Editors: Ronald J. Brachman, *Jacobs Technion-Cornell Institute at Cornell Tech*
Peter Stone, *University of Texas at Austin*
Series ISSN
Print 1939-4608 Electronic 1939-4616

Game Theory for Data Science

Eliciting Truthful Information

Boi Faltings
École Polytechnique Fédérale de Lausanne (EPFL)

Goran Radanovic
Harvard University

SYNTHESIS LECTURES ON ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING #35

ABSTRACT

Intelligent systems often depend on data provided by information agents, for example, sensor data or crowdsourced human computation. Providing accurate and relevant data requires costly effort that agents may not always be willing to provide. Thus, it becomes important not only to verify the correctness of data, but also to provide incentives so that agents that provide high-quality data are rewarded while those that do not are discouraged by low rewards.

We cover different settings and the assumptions they admit, including sensing, human computation, peer grading, reviews, and predictions. We survey different incentive mechanisms, including proper scoring rules, prediction markets and peer prediction, Bayesian Truth Serum, Peer Truth Serum, Correlated Agreement, and the settings where each of them would be suitable. As an alternative, we also consider reputation mechanisms. We complement the game-theoretic analysis with practical examples of applications in prediction platforms, community sensing, and peer grading.

KEYWORDS

data science, information elicitation, multi-agent systems, computational game theory, machine learning

Contents

Preface

Data has very different characteristics from material objects: its value is crucially dependent on novelty and accuracy, which are determined only from the context where it is generated. On the other hand, it can be freely copied at no extra cost. Thus, it cannot be treated as a resource with an intrinsic value, as is the focus in most of game theory.

Instead, we believe that game theory for data has to focus on incentives for *generating* novel and accurate data, and we bring together a body of recent work that takes this perspective.

We describe a variety of mechanisms that can be used to provide such incentives. We start by showing incentive mechanisms for verifiable information, where a ground truth can be used as a basis for incentives. Most of this book is about the much harder problem of incentives for unverifiable information, where the ground truth is never known. It turns out that even in this case, game-theoretic schemes can provide incentives that make providing accurate and truthful information the best interest of contributors.

We also consider scenarios where agents are mainly interested in influencing the result of learning algorithms through the data they provide, including malicious agents that do not respond to monetary rewards. We show how the negative influence of any individual data provider on learning outcomes can be limited and thus how to thwart malicious reports.

While our main goal is to make the reader understand the principles for constructing incentive mechanisms, we finish by addressing several other aspects that have to be considered for their integration in a practical distributed machine learning system.

This book is a snapshot of the state of the art in this evolving field at the time of this writing. We hope that it will stimulate interest for further research, and make it itself obsolete soon!

Boi Faltings and Goran Radanovic
July 2017

Acknowledgments

Our interest in this topic goes back to 2003 and much of the early work was carried out in collaboration with Radu Jurca, who has developed several of the mechanisms described in this book and is responsible for many important insights. We also thank numerous researchers for discussions and comments over the years, in particular Yiling Chen, Vincent Conitzer, Chris Dellarocas, Arpita Ghosh, Kate Larson, David Parkes, David Pennock, Paul Resnick, Tuomas Sandholm, Mike Wellmann, and Jens Witkowski.

Boi Faltings and Goran Radanovic
July 2017