Mach Learn
DOI 10.1007/s10994-013-5416-x

Spectral learning of weighted automata

A forward-backward perspective

Borja Balle · Xavier Carreras · Franco M. Luque · Ariadna Quattoni

Received: 8 December 2012 / Accepted: 3 September 2013

Abstract In recent years we have seen the development of efficient provably correct algorithms for learning Weighted Finite Automata (WFA). Most of these algorithms avoid the known hardness results by defining parameters beyond the number of states that can be used to quantify the complexity of learning automata under a particular distribution. One such class of methods are the so-called spectral algorithms that measure learning complexity in terms of the smallest singular value of some Hankel matrix. However, despite their simplicity and wide applicability to real problems, their impact in application domains remains marginal to this date. One of the goals of this paper is to remedy this situation by presenting a derivation of the spectral method for learning WFA that—without sacrificing rigor and mathematical elegance—puts emphasis on providing intuitions on the inner workings of the method and does not assume a strong background in formal algebraic methods. In addition, our algorithm overcomes some of the shortcomings of previous work and is able to learn from statistics of substrings. To illustrate the approach we present experiments on a real application of the method to natural language parsing.

Keywords Spectral learning · Weighted finite automata · Dependency parsing

Editors: Jeffrey Heinz, Colin de la Higuera, and Tim Oates.

B. Balle · X. Carreras (✉) · A. Quattoni
Universitat Politècnica de Catalunya, Barcelona 08034, Spain
e-mail: carreras@lsi.upc.edu

B. Balle
e-mail: bballe@lsi.upc.edu

A. Quattoni
e-mail: aquattoni@lsi.upc.edu

F.M. Luque
Universidad Nacional de Córdoba and CONICET, X5000HUA Córdoba, Argentina
e-mail: francolq@famaf.unc.edu.ar

Published online: 03 October 2013