Pattern Mining with Evolutionary Algorithms

Sebastián Ventura • José María Luna

Pattern Mining with Evolutionary Algorithms

Springer

Sebastián Ventura
Department of Computer Science
and Numerical Analysis
University of Cordoba
Cordoba, Spain

José María Luna
Department of Computer Science
and Numerical Analysis
University of Cordoba
Cordoba, Spain

ISBN 978-3-319-33857-6 ISBN 978-3-319-33858-3 (eBook)
DOI 10.1007/978-3-319-33858-3

Library of Congress Control Number: 2016939025

Printed on acid-free paper

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

–Success is a journey, not a destination. The doing is often more important than the outcome–

Arthur Ashe.

To our families.

Preface

This book is intended to provide a general and comprehensible overview of the field of pattern mining with evolutionary algorithms. To do so, the book provides formal definitions about patterns, pattern mining, type of patterns, and the usefulness of patterns in the knowledge discovery process. As it is described within the book, the discovery process suffers from both high runtime and memory requirements, especially when high-dimensional datasets are analyzed. To solve this issue, many pruning strategies have been developed. Nevertheless, with the growing interest in the storage of information, more and more datasets comprise such a dimensionality that the discovery of interesting patterns becomes a hard process. In this regard, the use of evolutionary algorithms for mining pattern enables the computation capacity to be reduced, providing sufficiently good solutions.

The book also provides a survey on evolutionary computation with particular emphasis on genetic algorithms and genetic programming. Additionally, this book carries out an analysis of the set of quality measures most widely used in the field of pattern mining with evolutionary algorithms. This book serves as a good review on the most important evolutionary algorithms for pattern mining. In this sense, it considers the analysis of different algorithms for mining different types of patterns and relationships between patterns, such as frequent patterns, infrequent patterns, patterns defined in a continuous domain, or even positive and negative patterns.

The book also introduces a completely new problem in the pattern mining field, which is known by the name of the mining of exceptional relationships between patterns. In this problem, the goal is to identify patterns where distribution is exceptionally different from the distribution in the complete set of data records. Finally, this book deals with the subgroup discovery task, a method to identify a subgroup of interesting patterns that is related to a dependent variable or target attribute. This subgroup of patterns satisfies two essential conditions: interpretability and interestingness.

Cordoba, Spain
February 2016

Sebastián Ventura
José Marí a Luna

Acknowledgments

We would like to thank the Springer editorial team for giving us the opportunity to publish this book, for their great support toward the preparation and completion of this work, and for their valuable editing suggestions to improve the organization and readability of the manuscript. We also want to thank our colleagues for their valuable help during the preparation of the book, whose comments were very helpful for improving its quality.

This work was supported by the Spanish Ministry of Economy and Competitiveness under the project TIN2014-55252-P and FEDER funds.

Contents