

# Preface

This book is dedicated to Stan Matwin in recognition of the numerous contributions he has made to the fields of machine learning, data mining, and big data analysis to date. With the opening of the Institute for Big Data Analytics at Dalhousie University, of which he is the founder and the current Director, we expect many more important contributions in the future.

Stan Matwin was born in Poland. He received his Master's degree in 1972 and his Ph.D. in 1977, both from the Faculty of Mathematics, Informatics and Mechanics at Warsaw University, Poland. From 1975 to 1979, he worked in the Institute of Computer Science at that Faculty as an Assistant Professor. Upon immigrating to Canada in 1979, he held a number of lecturing positions at Canadian universities, including the University of Guelph, York University, and Acadia University. In 1981, he joined the Department of Computer Science (now part of the School of Electrical Engineering and Computer Science) at the University of Ottawa, where he carved out a name for the department in the field of machine learning over his 30+ year career there (he became a Full Professor in 1992, and a Distinguished University Professor in 2011). He simultaneously received the State Professorship from the Republic of Poland in 2012.

He founded the Text Analysis and Machine Learning (TAMALE) lab at the University of Ottawa, which he led until 2013. In 2004, he also started cooperating as a “foreign” professor with the Institute of Computer Science, Polish Academy of Sciences (IPI PAN) in Warsaw. Furthermore, he was invited as a visiting researcher or professor in many other universities in Canada, USA, Europe, and Latin America, where in 1997 he received the UNESCO Distinguished Chair in Science and Sustainable Development (Universidade de Sao Paulo, ICMSC, Brazil).

In addition to his position as professor and researcher, he served in a number of organizational capacities: former president of the Canadian Society for the Computational Studies of Intelligence (CSCSI), now the Canadian Artificial Intelligence Society (CAIAC), and of the IFIP Working Group 12.2 (Machine Learning), Founding Director of the Information Technology Cluster of the Ontario Research Centre for Electronic Commerce, Chair of the NSERC Grant Selection

Committee for Computer Science, and member of the Board of Directors of Communications and Information Technology Ontario (CITO).

Stan Matwin is the 2010 recipient of the Distinguished Service Award of the Canadian Artificial Intelligence Society (CAIAC). He is Fellow of the European Coordinating Committee for Artificial Intelligence and Fellow of the Canadian Artificial Intelligence Society.

His research spans the fields of machine learning, data mining, big data analysis and their applications, natural language processing and text mining, as well as technological aspects of e-commerce. He is the author and co-author of over 250 research papers.

In 2013, he received the Canada Research Chair (Tier 1) in Visual Text Analytics. This prestigious distinction and a special program funded by the federal government allowed him to establish a new research initiative. He moved to Dalhousie University in Halifax, Canada, where he founded, and now directs, the Institute for Big Data Analytics.

The principal aim of this Institute is to become an international hub of excellence in Big Data research. Its second goal is to be relevant to local industries in Nova Scotia, and in Canada (with respect to applications relating to marine biology, fisheries and shipping). Its third goal is to develop a focused and advanced training program that covers all aspects of big data, preparing the next generation of researchers and practitioners for research in this field of study.

On the web page of his Institute, he presents his vision on Big Data Analytics. He stresses, “Big data is not a single breakthrough invention, but rather a coming together and maturing of several technologies: huge, inexpensive data harvesting tools and databases, efficient, fast data analytics and data mining algorithms, the proliferation of user-friendly data visualization methods and the availability of affordable, massive and non-proprietary computing. Using these technologies in a knowledgeable way allows us to turn masses of data that get created daily by businesses and the government into a big asset that will result in better, more informed decisions.”

He also recognizes the potential transformative role of big data analysis, in that it could support new solutions for many social and economic issues in health, cities, the environment, oceans, education access, personalized medicine, etc. These opinions are reflected in the speech he gave at the launch of his institute, where his recurring theme was “Make life better.” His idea is to use big data (i.e., large and constantly growing data collections) to learn how to do things better. For example, he proposes to turn data into an asset by, for instance, improving motorized traffic in a big city or ship traffic in a big port, creating personalized medical treatments based on a patient's genome and medical history, and so on.

Notwithstanding the advantages of big data, he also recognizes its risks for society, especially in the area of privacy. As a result, since 2002, he has been engaged in research on privacy preserving data mining.

Other promising research directions, in his opinion, include data stream mining, the development of new data access methods that incorporate sharing ownership mechanisms, and data fusion (e.g., geospatial applications).

We believe that this book reflects Stan Matwin's call for careful research on both the opportunities and the risks of Big Data Analytics, as well as its impact on society.

Nathalie Japkowicz  
Jerzy Stefanowski

Big Data Analysis: New Algorithms for a New Society

Japkowicz, N.; Stefanowski, J. (Eds.)

2016, XII, 329 p. 63 illus., 35 illus. in color., Hardcover

ISBN: 978-3-319-26987-0