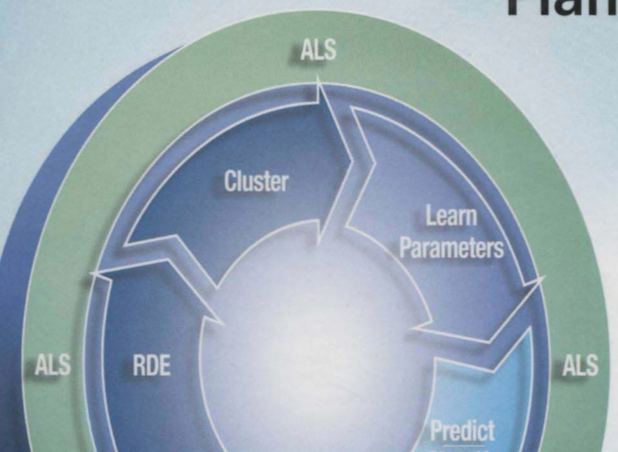
Autonomous Learning Systems

From Data Streams to Knowledge in Real-time

Plamen Angelov



WILEY

Autonomous Learning Systems

From Data Streams to Knowledge in Real-time

Plamen Angelov

Lancaster University, UK

WILEY
A John Wiley & Sons, Ltd., Publication

This edition first published 2013

Registered office
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com.

Library of Congress Cataloging-in-Publication Data:

Angelov, Plamen P.
Autonomous learning systems : from data streams to knowledge in real-time / Plamen P. Angelov.
pages cm
Includes bibliographical references and index.
ISBN 978-1-119-95152-0 (cloth)
1. Self-organizing systems. 2. Machine learning. I. Title.
Q325.A54 2013
006.3′1–dc23

2012025907

A catalogue record for this book is available from the British Library.

ISBN: 978-1-119-95152-0

Set in 10/12.5pt Palatino by Aptara Inc., New Delhi, India

Printed and bound in Malaysia by Vivar Printing Sdn Bhd

Autonomous Learning Systems

Forewords

Adrian Stoica

Efficient and robust performance in imperfectly known, nonstationary, environments – and this characterizes the vast majority of real-world applications – requires systems that can improve themselves, transcending their initial design, continuously optimizing their parameters, models, and methods. These improvements come predominantly from learning – about the environment, about the ageing self, about the interactions with, and within, the environment, and from the ability to put this learning to use. Batch learning – or at least repeated updating learning from most recent batches, is sufficient only for a limited number of applications. For other applications learning needs to be incremental, to sample level, a learn-or-perish, or at least learn-or-pay (a hefty price) situation. In particular, real-time learning is most critical for bots, virtual or real, agents of the cyberphysical systems that need the agility to swiftly react to virus attacks, or physical robots exposed to hazards while performing search and rescue in disaster areas, or dealing with what is, for now, a largely unpredictable partner: the human. The fast advancement in autonomous systems makes the subject of real-time autonomous learning critically important, and yet the literature addressing this topic is extremely scarce.

Dr Angelov's pioneering book addresses this problem at its core, focusing on real-time, online learning from streaming data on a sample-by-sample basis. It offers a basic framework for understanding and for designing such systems. It importantly contributes to a more powerful learning, not only of the parameters but of a better structure as well. Conventional approaches are characterized by the fact that the system structure (model) is determined in the beginning, by human designers of the system, and only parameters are learned from the interaction. The entire model identification–learning process can, however, be posed as an optimization problem, as the author points out, and this includes automatically determining the optimal structure in conjunction with the optimal parameters for it. This is done automatically in the methods described in the book, and constitutes a significant and valuable contribution. The system is continuously evolving, not in the evolutionary (genetic) sense of improvement over generations, but continuously perfecting its development.

A valuable contribution of the book is that it offers a high-level, holistic perspective of the field, which helps both students and expert practitioners better comprehend the interplay of various disciplines involved in learning autonomous systems, as diverse as adaptive control and evolutionary algorithms, offering analogies between different disciplines and referring to the equivalency of the concepts characterized by different terminology in different disciplines. It is not meant to be a comprehensive reference of concepts and methods in the field, the author instead paints the landscape with selected brush strokes that allows the viewer to see the forest without getting lost in seeing the trees. It is a work that charters a new field, innovates in methods to advance into it, and outlines new challenges to be addressed by future explorers.

The selected concepts and methods, a good number of which come from the author's own prior work, are used in the second part of the book to illustrate the implementation of learning in autonomous systems. Concepts such as that of evolving clusters, 'age' of an (evolving) local submodel, and methods such as recursive density estimation (RDE) introduced by the author, showing significant improvement over the state-of-the art, are important additions to the arsenal of tools for real-time learning. In particular, I believe that adaptive, self-learning (evolving) classifiers will play a fundamental role in future autonomous learning systems.

The book's last part is a review of three applications: autonomous learning sensors for chemical and petrochemical applications, autonomous learning in mobile robots, and autonomous novelty detection and object tracking in video systems. These diverse domains illustrate the general applicability of the set of methodology presented in the book and focused on the main theme of this work: the real-time autonomous online learning from data streams.

The field of autonomous learning systems is destined to play an increasingly important role in most systems that will surround us in cyberphysical space. Converting information in data streams, larger and larger, to actionable knowledge, in real time: this is the great challenge ahead, and this book is an important step towards addressing it.

Adrian Stoica
Jet Propulsion Laboratory
California Institute of Technology
June 2012

Vladik Kreinovich

In many practical situations, we have experts who are skilled in doing certain tasks: expert medical doctors are skilled in diagnosing and curing diseases, professional drivers are skilled in driving cars – in particular, driving them in difficult traffic and/or weather conditions, etc. It is desirable to incorporate the knowledge of these top experts in an automatic system that would help other users perform the corresponding tasks – and, ideally, perform these tasks automatically.

Experts are usually willing to share their knowledge, but the difficulty is that in many situations, experts describe their knowledge by using imprecise ("fuzzy") words from natural language, like "small". For example, an expert driver rarely describes his or her experience in precise terms like "if the car in front slows down by 10 km/h and it is at a distance of 10 meters, you should press the brake for 0.6 seconds with a force of 2.7 Newtons"; most probably, the rule described by an expert driver is "if the car in front of you is close, and it suddenly slows down some, then you should brake right away". In this rule, "close" and "some" are imprecise terms: while everyone would agree that, say 100 meters is not close while 5 meters is close, there will not be a precise threshold so that before this threshold the distance is close, and a 1 cm larger distance is not close.

To describe such imprecise (fuzzy) knowledge in computer-understandable precise terms, Professor Lotfi A. Zadeh invented, in the 1960s, a new approach called *fuzzy logic*. Zadeh's ideas led to *revolutionary* changes in many control situations: from the first successful control applications in the 1970s through the fuzzy control boom in the 1990s – when fuzzy-controlled washing machines, camcorders, elevators, trains were heavily promoted and advertised – to the current ubiquity of fuzzy controllers. Just like nowadays computers are ubiquitous – companies no longer brag about computer chips in their cars, since all the cars have such chips – similarly, fuzzy control is ubiquitous: for example, in many cars, automatic transmission systems use fuzzy control.

The existing fuzzy controllers are very successful, but they have a serious limitation: they do not learn. Once the original expert rules are implemented, these same rules are used again and again, even when it becomes clear that the rules need to be updated. We still need an expert to update these rules.

There are, of course, numerous intelligent systems that *can* learn, such as artificial neural networks, but from the viewpoint of the user, these systems are "black boxes": we may trust them, but we cannot easily understand the recommendations. In contrast, fuzzy rules, by definition, are formulated in terms of understandable rules. If we could make fuzzy systems themselves learn, make them automatically update the rules – this would combine the clarity of fuzzy rules with the autonomous learning ability of neural networks. This would make learning fuzzy controllers even more efficient – and therefore, even more widely used. This would lead to a *second revolution* in intelligent control.

And this revolution is starting. This book, by Dr. Plamen Angelov, one of the world's leading specialists in learning fuzzy systems, is the first book that summarizes the current techniques and successes of autonomously learning fuzzy (and other) systems – techniques mostly developed by Dr. Angelov himself, often in collaboration with other renowned fuzzy researchers (like Dr. Ronald Yager). Some of these techniques have previously appeared in technical journals and proceedings of international conferences, some appear here for the first time.

Ideas are many, it is difficult to describe them all in a short preface, so let us just give a few examples. The first example is an interesting AnYa algorithm invented by *An*gelov and *Ya*ger (Anya is also a Russian short form of Anna (Anne)). In fuzzy logic, each fuzzy term like "small" is described by a *membership function*, i.e. a function that assigns, to each possible value x, the degree $\mu(x)$ from the interval [0, 1] to which this value is small. The value $\mu(x) = 1$ means that x is absolutely small, every expert would agree to this; $\mu(x) = 0$ means that x is definitely not small, while values between 0 and 1 represent the expert's uncertainty.

In the traditional fuzzy control algorithms, we select a finite-parametric family of membership functions – e.g., functions that are of triangular, trapezoid, or Gaussian shapes – and adjust parameters of these functions based on the expert opinions. As a result, sometimes, the resulting membership functions provide a rather crude and not very accurate description of the expert knowledge. To improve the situation, AnYa does not limit the shape of the membership function. Instead, it uses all the value $x_1, \ldots, x_n$, that the expert believes to be satisfying the property (like "small"), and defines the desired membership function by formalizing the statement "x is close to x_1 or x is close to x_2, …". Now, all we need to do is describe what experts mean by "close" (and by "or"), and we will no only have a well-shaped membership function, we will also have a way to update its shape when new observations appear.

A similar idea can be implemented in probabilistic terms, when we use probability density functions (pdf) $\rho(x)$ instead of membership functions, but the authors show a clear computational advantage of their fuzzy approach: A pdf is normalized by the condition that the total probability is 1: $\int \rho(x)dx = 1$, so we need to go through a computationally intensive process of renormalize all its values every time we update one value of $\rho(x)$. In contrast, a membership function is usually normalized by the condition that $\max_x \mu(x) = 1$. Thus, if we change a value of the membership function, we only need to renormalizing other values when the changed value is $\mu(x) = 1$ – and this happens rarely.

Similar ideas are used to automatically decide how to adjust the rule's conclusions, when to subdivide the original rule into two subrules – that would provide a more subtle description of actions, when to dismiss the old data that is no longer representative of the system's inputs, etc.

Researchers and practitioners who have been using fuzzy techniques will definitely benefit from learning how to make fuzzy systems learn automatically (pun intended :-). But this book is not only for them. Readers who are not familiar with the current fuzzy techniques will also greatly benefit: the book starts with a nice introduction that explains, in popular terms, what is fuzzy, and why and how we can use fuzzy techniques. (Some math is needed – but math taught to engineers is quite enough.)

This book is not just for the academics, practitioners will surely benefit. In the last part of the book, numerous applications are described in detail, providing the reader with an understanding of how these new methods can be used in practical situations. It may be a good idea to glance through these exciting applications first, this will give the readers an excellent motivation to grind through all the formulas and algorithms in the main part of the book.

Applications include learning sensors for chemical and petrochemical industries – industries where the chemical contents of the input (such as oil) changes all the time, and intelligent adjustments need to be constantly made. Another successful application example is mobile robotics, where the robot's ability to learn how to navigate in a new environment – and learn fast – is often crucial for the robot's mission. The new methods have also been applied to novelty detection and object tracking in video streams, to wireless sensor networks, and to many other challenging application areas.

The second revolution – of making intelligent control systems fast learners – has started. Its preliminary results are already exciting. This book will definitely help promote the ideas of this second revolution – and thus, further improve its methods and use these improved methods to solve numerous challenging problems of today.

Vladik Kreinovich
President
North American Fuzzy Information Processing Society (NAFIPS), El Paso, Texas
August 2012

Arthur Kordon

One of the most significant changes during the twenty-first century is the fast dynamics in almost all components of life. Economic, social, and financial systems, to name a few, are moving more of their activities to a real-time mode of operation. Extracting knowledge from data streams becomes as important as was information retrieval from data bases several decades ago. The need for fast adaptation to unknown conditions, due to the new complex nature of the global economy, is another big challenge in operating the systems in the twenty-first century.

Unfortunately, the existing classical modeling techniques, based on first principles, statistics, system identification, etc., cannot deliver satisfying solutions adequate to the new fast dynamics. Adaptive systems are limited to models with a fixed structure and linear relationships. Some recent computational intelligence methods with nonlinear and adaptive capabilities, such as evolutionary computation and swarm intelligence, are too slow for real-time operation.

A potential solution to the new needs is the fast-growing research area of evolving intelligent systems. They offer a system that simultaneously learns its structure and calculates its parameters "on the fly" from data streams. An advantage of this approach is the simplicity and very low memory requirement of the used algorithms, which makes them appropriate for real time. In addition, the algorithms are universal (i.e. can be applied in various areas with no or minor changes), with minimal number of tuning parameters, and in the case of evolving fuzzy systems – the suggested models are interpretable by the users. Autonomous learning systems (ALS) is the ultimate solution of an evolved intelligent system that integrates the broadest possible range of algorithms and requires minimal human intervention. The research area of ALS has grown significantly in terms of publications, conference presence, and funding support. An impressive feature of this emerging technology is its fast applicability in different areas, such as inferential sensors, mobile robotics, video streams processing, defence applications, etc.

Unfortunately, the literature for ALS is mostly available in journal or conference papers. Some recently published books on evolving intelligent systems and evolving fuzzy systems give a generic overview and some guidelines on the different components of the technology. However, researchers and practitioners need a book that describes in sufficient details the foundation of ALS, offers software for investigating the key algorithms in a popular environment, such as Matlab, and gives appropriate application examples. *Autonomous Learning Systems* is the first book on the market that fills this need.

Purpose of the Book

The purpose of the book is to give the reader a comprehensive view of the current state of the art of ALS. The key topics of the book are:

1. *What are the fundamentals of ALS?* The first main topic of the book focuses on the ambitious task of describing the diverse research foundation of ALS. The key methods, such as probability theory, machine learning and pattern recognition, clustering, and fuzzy system theory are presented at a level of detail sufficient for understanding the ALS mechanisms.
2. *How to develop an ALS?* The second key topic of the book is the description of the methodology of ALS. Its main focus is on presenting the key steps in developing of an ALS from streaming data, such as dynamic data space partitioning, normalization and standardisation, autonomous monitoring of the structure quality, and autonomous learning parameters of the local submodels.
3. *How and where to apply ALS in practice?* The third key topic of the book covers the implementation issues and application areas of ALS. It includes an overview of the potential application areas, such as autonomous predictors, estimators, filters, and inferential sensors; autonomous learning classifiers; autonomous learning controllers and collaborative autonomous learning systems. Of special interest are the results from several real-world applications of ALS, such as inferential sensors in the chemical and petrochemical industries, ALS in mobile robotics, and autonomous novelty detection and object tracking in video streams.

Who this Book is for?

Due to the wide potential application areas, the targeted audience is much broader than the traditional scientific communities in computer science, data mining, and engineering. The readers who can benefit from this book are described below:

- *Academics* – This group includes a large class of academics in the fields of computer science, data mining, and engineering who are not familiar with the research and technical details of this new field. They will benefit from the book by using it as an introduction to the field, exploring the described algorithms, and understanding its broad application potential.
- *Students* – Undergraduate and graduate students can benefit from the book by understanding this new field. The book could be a basis for a graduate course on this topic.
- *Industrial researchers* – This group includes scientists in industrial labs who create new solutions for the busines. They will benefit from the book by understanding the value of this new emerging technology in delivering novel solutions in the area of real-time modelling based on data streams.
- *Governmental agencies* – ALS have almost unlimited potential in various military applications and space exploration projects. This book can be used by the governmental decision makers as an introduction to the technology and even may open new application areas.
- *Software vendors* – This group includes vendors of process monitoring and control systems, data-mining software, robotic systems, etc. They will benefit from the book by understanding a new emerging technology, exploring the described algorithms, and defining new application areas, related to data streams.

Features of the Book

The key features that differentiate this book from other titles on learning and evolving systems are:

1. *A systematic description of autonomous learning systems* – One of the most valuable features of the book is the systematic and comprehensive way of presenting related methods. It gives the reader a solid basis for understanding the research foundation of ALS, which is of critical importance for introducing this emerging technology to a broad audience.
2. *A detailed methodology for development of autonomous learning systems* – Another important topic of the book is the step-by-step description of the key algorithms, used in the proposed technology. This allows the reader to easily implement and explore the large potential of ALS on a simulation level.
3. *A broad range of autonomous learning systems applications* – The third key feature of the book is the impressive list of described real-world applications across several application areas. It illustrates one of the unique advantages of ALS – the fast transition from theory to practice.

Arthur Kordon
Lake Jackson, Texas
March, 2012

Lawrence O. Hall

Today's data-driven world has many large streams of data, for example a day's worth of images posted to the internet. Imagine the task of finding all instances, perturbed or not, of an image. It would require learning and adaptation as the way images are placed and modified evolve. This book provides an informative snapshot of how to build autonomous learning systems for data streams. It covers the math you need, probability, normalization, fuzzy systems, the basics of machine learning and pattern recognition, clustering, feature selection and more. It shows how to do learning in an autonomous setting and how to tune out noise/outliers.

Sensor learning, autonomous classifier learning, collaborative learning, learning controls systems in an autonomous way are all covered in an interesting way. There are case studies given to show how it all fits together. These include object tracking in videos and autonomous learning robots. It is clear that we need adaptable robots capable of modifying their behaviour. This is an ambitious, early book that covers robot learning and the basics for much more autonomous learning. It is well worth perusing for those interested in this broad subject.

Lawrence O. Hall
Dept. of Computer Science and Engineering, University of South Florida, USA
May 16, 2012

Preface

This book comes as a result of focused research and studies for over a decade in the emerging area that is on the crossroads of a number of well-known and well-established disciplines, such as (Figure 1):

- machine learning (ML);
- system engineering (specifically, system identification), SI;
- data mining, DM;
- statistical analysis, SA;
- pattern recognition including clustering, classification, PR;
- fuzzy logic and fuzzy systems, including neurofuzzy systems, FL;
- and so on.

On the one hand, there is a very strong trend of innovation of all of the above well-established branches of research that is linked to their *online* and *real-time* application; their adaptability, flexibility and so on (Liu and Meng, 2004; Pang, Ozawa and Kasabov, 2005; Leng, McGuinty and Prasad, 2005). On the other hand, a very strong driver for the emergence of *autonomous learning systems* (ALS) is industry, especially defence and security, but also aerospace and advanced process industries,

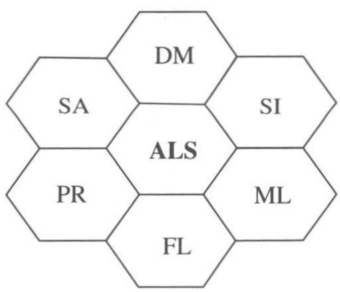


Figure 1 Autonomous learning systems theory is build upon other well-established areas of research (the list is, of course, not exhaustive)

the Internet, eHealth (assisted living), intelligent transport, and so on. The demand in defence was underpinned recently by a range of multimillion research and development projects funded by DARPA, USA (notably, two Grand Challenge competitions (Buehler, Iagnemma and Singh, 2010)); by MoD and BIS, UK (Defence Technology Centre on Systems Engineering and Autonomous Systems; ASTRAEA and GAMMA multimillion programmes, in which the author has played the research provider role, being the technical lead for several projects) and similar programmes in other European countries (France, Sweden, Spain, Czech Republic, Russia), and Israel. Major global companies have established their own programmes, such as IBM's autonomous computing initiative (IBM, 2009) and BT's intelligent network of autonomous elements (Detyenecki and Tateson, 2005). The International Neural Network Society (INNS) has established in 2011 a section on Autonomous Machine Learning of which the author is a founding member, together with scientists such as Bernard Widrow – the father of the famous least mean squares (LMS) algorithm.

This book attempts to address these challenges with a systematic approach that can be seen as laying the foundations of what can become a fast-growing area of research that can underpin a range of technological applications so needed by industry and society. The author does not claim that this represents a finished and monolithic theory; this is rather a catalyser for future developments, an inoculum, a vector pointing the direction rather than a full solution of the problems.

An important aim of preparing this book was also to make it a one-stop shop for students, researchers, practicing engineers, computer specialists, defence and industry experts and so on that starts with the motivation, presents the concept of the approach, describes details of the theoretical methodology based on a rigorous mathematical foundation, presents a wide range of applications, and more importantly, provides illustrations and algorithms that can be used for further research. The software (subject to a license) can be downloaded from the author's web site: http://www.lancs.ac.uk/staff/angelov/Downloads.htm. It is covered by USA patent # 2010-0036780, granted 21 August 2012 (priority date 1 Nov. 2006) and two pending patent applications and distributed by the spin-out company of Lancaster University called EntelSenSys Ltd. (www.entelsensys.com). From the same web site there will also be available for the readers for this book a set of lecture notes that will be a useful tool for delivering specialised short courses or an advanced Master level module as a part of various related programmes that cover the topics of machine learning, pattern recognition, control systems, computational intelligence, data mining, systems engineering, and so on.

The book was initially planned at the end of 2006 during the very successful IEEE Symposium on Evolving Fuzzy Systems held in Ambleside in the Lake District, UK but the turn of events (as usually happens) postponed its appearance by more than five years, which gave an opportunity for the concepts to mature and evolve further and new results and applications to be added.

It would not have become a reality without the support of the colleagues and collaborators, students, associates and visitors of the author. In the hope not to miss someone this includes Prof. Ronald Yager (Iona College, NY, USA), Dr. Dimitar Filev

(Ford, MI, USA), Prof. Nikola Kasabov (Auckland University, New Zealand), Prof. Fernando Gomide (University of Campinas, Brazil), Dr. Xiaowei Zhou (HW Communications, UK), Dr. Jose Antonio Iglesias (University Carlos III, Madrid, Spain), Dr. Jose Macias Hernandez (CEPSA, Tenerife, Spain), Dr. Arthur Kordon (The Dow Chemical, TX, USA), Dr. Edwin Lughofer (Johan Kepler University, Linz, Austria), Prof. Igor Skrjanc (University of Ljubljana, Slovenia), Prof. Frank Klawonn (Ostfalia University of Applied Sciences, Germany), Mr. Jose Victor Ramos (University of Coimbra, Portugal), Dr. Ana Cara Belen (University of Granada), Mr. Javier Andreu (Lancaster University), Mr. Pouria Sadeghi-Tehran (Lancaster University), Mr. Denis Kolev (Rinicom Ltd.), Mrs. Rashmi Dutta Baruah (Lancaster University), Mr. Ramin Ramezani (Imperial College, London), Mr. Julio Trevisan (Lancaster University), and many more.

The feedback on the manuscript by Professor Vladik Kreinovich (University of Texas, USA) who is also President of the North American Fuzzy Information Processing Society (NAFIPS); Dr. Adrian Stoica, Senior Research Scientist at the Autonomous Systems Division, NASA Jet Propulsion Laboratory, Pasadena, CA, USA, Dr. Arthur Kordon, Team Leader at Dow Chemical, TX, USA; as well as from Dr. Larry Hall, Distinguished Professor and Chair at the Department of Computer Science and Engineering, University of South Florida, USA was also instrumental to improve and smooth out the presentation and remove some omissions.

Plamen Angelov
November 2009–May 2012
Lancaster, UK