Preface to the book "Empirical Approach to Machine Learning"

We are all fascinated by the recent successes of deep learning. Deep learning has performed wonders in image processing, in natural language processing, in bioinformatics, in many other application areas. Deep learning has fulfilled a long-time dream of Artificial Intelligence community: to design a computer program that excels in Go. These successes are well known, but it is also well known that sometimes deep learning makes mistakes. For example, small changes in a picture -- changes invisible to a human eye -- can lead to a gross misclassification, like interpreting a coffee machine for a cobra. Because of these mistakes, we cannot always trust the results of these exciting machine learning algorithms.

The problem is not so much the complexity of these algorithms: we often trust the results of very complex algorithms. For example, we trust the planes when they are controlled by very complex algorithm implementing the autopilot, we trust the X-ray and MRI machines where algorithms are also very complex. We trust them because in all these situations, while the algorithms are complex, we -- or at least experts whom we trust -- have an intuitive understanding of these algorithms. Not only we have the results of the complex algorithms, we also have a clear qualitative explanation of these results.

In contrast, many modern machine learning algorithms are, to a large extent, black boxes -- they provide predictions and recommendations, but they do not explain in clear terms why. This is a known serious problem of machine learning. Many researchers are working on it, and some progress have been achieved -- but the challenge largely remains. This book describes a new innovative approach to making machine learning intuitively interpretable -- an approach whose success has been proved by many successful applications. To fully understand this approach, one needs to read this very interesting book.

What do we mean by intuitive explanation? In a nutshell, it means that a simple IF...THEN set of rules using prototypes (actual data, which may be images or else) is automatically derived. Thus we have commonsense explanations described by using words from natural language such as IF...is like....OR...is like...THEN.

To solve this new task, Dr. Angelov comes up with a systematic way to generate powerful models (classifiers, predictors, anomaly detectors, even controllers) which can be expressed through fuzzy sets based on data only.

Expert knowledge can optionally be used (if available and convenient) and, equally, experts can interrogate and analyse the automatically extracted form data models.

Summarizing: this book is an important step in Computational Intelligence. It provides a new way to make machine learning results interpretable, and it also provides a new innovative application of fuzzy techniques. This book will definitely inspire new ideas, new techniques, and new applications.

Vladik Kreinovich, El Paso, Texas

Vice President, International Fuzzy Systems Association