Journal of Intelligent & Fuzzy Systems 16 (2005) 151–155
IOS Press

Book reviews

Data Mining: Multimedia, Soft Computing, and Bioinformatics, Sushmita Mitra and Tinku Acharya, Wiley, 2003, ISBN 0 471 460540.

The main objective of this book is to provide the reader with an introduction to data mining, especially data mining that uses soft computing techniques, with an emphasis on bioinformatics applications. Chapter 1 is a general introduction to data mining. Chapter 2 is an overview of different techniques that form soft computing: fuzzy, neural, genetic algorithms, rough sets, with wavelets thrown in. Chapter 3 describes different techniques for data compression: an important but often overlooked first stage for data mining. This chapter contains the general notions of compression rate, and examples of compression algorithms, from Huffman and Lempel-Ziv algorithms for text compression to JPEG2000 image compression techniques. Chapter 4 describes data matching algorithms, Chapters 5 and 6 describe different methods for classification and clustering, Chapters 7 and 8 describe possible rules and how to mine them from the data. Chapter 9 describes image mining techniques, and Chapter 10 overview bioinformatics.

Overall, the book is an impressive and broad overview of various subjects. At the end, a reader gets a good general picture of different techniques that form data mining and its applications to bioinformatics. Of course, with such a broadness comes a negative side: so many topics and techniques are covered that it is impossible to provide the reader with the detailed algorithmic ability of implementing all these techniques. While several simple algorithms (such as Huffman code and Princial Component Analysis) are described in all the necessary detail, most descriptions (e.g., of wavelets) are rather high-level. The reader who is really interested in applying the corresponding techniques must, in most cases, go to the references papers and books to learn the details.

What this book provides is a general roadmap of what methods are available and where to look.

Minor drawback: while the bibliography to each chapter is good, it is listed in the order of referencing; in my opinion, the more traditional alphabetic order is more convenient for the readers: it allows them to easily find, e.g., a paper by Zadeh that attracted their attention when reading one of the previous chapters.

It is also somewhat disappointing (but understandable) that, in spite of the promising word "bioinformatics" in the title, only a 24-page Chapter 10 provides the reader with an introduction to Bioinformatics, and this chapter is even more high-level than most other chapters.

Vladik Kreinovich
Book Review Editor
Journal of Intelligent & Fuzzy Systems

Combining Pattern Classifiers: Methods and Algorithms, Ludmila I. Kuncheva, Wiley, 2004, ISBN 0 471 210781.

Many approaches exist for classification (clustering):

- there are approaches based on discriminant analysis in which we try to find a linear (or quadratic or polynomial) surface that best separates the classes;
- there are non-parametric approaches like Parzen classifiers where we reconstruct a probability distribution $\rho_i(x)$ characterizing each class i, and then assign a new object x to the class i for which the probability $\rho_i(x)$ is the largest;
- there is a nearest neighbor approach, in which a new object x is assigned to the class that contains the most of its k closest neighbors from the training sample;
- there are methods based on neural networks training.

For situations where the number of classes is huge, there are special methods in which, instead of trying to come up with all the classes at once, we try to describe the classification hierarchically (usually, as a tree).

The very fact that there exist so many different classification techniques means that each of these methods has its advantages and disadvantages. For each of these methods, there are real-life examples when this method produces a much better classification than all other techniques, and there are also situations when this particular method does not work as well as some other techniques. It is therefore reasonable to try to combine ("fuse") the existing methods so that the resulting combined classifiers will combine the advantages of all the different techniques. Different successful combination rules have been proposed, and the overview of such rules is the main objective of this book.

There are two main approaches to the classifier fusion:

- fusing the output labels, and
- fusing the continuous values that has lead to these output labels.

Techniques for combining the output labels range:

- from a simple majority rule, according to which we assign an object x to a class to which the largest number of combined methods assign it,
- to weighted majority rules, in which we take into considerations the relative weight of different methods,
- to complex fusion rules like Singular Value Decomposition, in which, crudely speaking, the weights of different combination rules are determined by comparing their outcomes.

When we combine the output labels, we ignore the fact that some output labels may have resulted from a solid conclusion, while other outcomes might have been on the edge of indecision. For example, for linear discrimination techniques, we are much more confident about the classification of the objects x that are far away from the separating (hyper)plane than about the objects that are close to this inter-class plane. Most classification methods generate such "degrees of certainty" before they make a classification decision. It makes sense to produce, as an outcome, not only the classification itself, but also the degree of confidence with which we classify the given object x to the corresponding class. It is therefore reasonable, when combining the outcomes, to take into consideration these degree of certainty, i.e., to fuse such continuous-valued outputs as well. In this case, in addition to combining the corresponding degrees, we also have an option of simply selecting a classifier methods for which the degree of confidence is the largest.

When combining different methods, it is important to make sure that the combined methods are diverse: combining very similar methods, whose results are very similar, does not add much to out ability to classify.

The book under review is the world's first monograph on combining different classification techniques. This is an extremely well written and well designed book. It starts with a clear and detailed overview of the classification methods themselves, and then goes ahead to describe and analyze different ways of combining these techniques. Dr. Kuncheva tried her best in incorporating as much theoretical justifications as possible; at the same time, she made this book as accessible as possible by including Matlab codes and examples for all the described algorithms.

Let me also briefly mention what this book is not. It is not a handbook. A practitioner who is interested in applying these methods will find a good exposition of the basic techniques that will cover the basic applications, but he or she may have to go deeper in more complex practical cases.

Also, as the author herself mentions in the introduction, this book is not an encyclopedia. There are so many different classification techniques and so many ways to combine them, that it is impossible to include all of them in a single readable book. To make the material manageable, the author had to make a (inevitably subjective) selection. Her selection is very coherent, based on the probabilistic techniques. It does lead out many successfully classification techniques that other authors might have included, such as fuzzy techniques. Maybe this book will inspire others to write their own survey monographs emphasizing these other techniques.

In the present form, this book is a unique and perfect book for researchers and practitioners who want to get a grasp of this new exciting area.

Vladik Kreinovich
Book Review Editor
Journal of Intelligent & Fuzzy Systems

Practical Genetic Algorithms, 2nd Edition, Randy L. Haupt and Sue Ellen Haupt, Wiley, 2004, ISBN 0 471 188735.

This book is a nice step-by-step introduction to genetic algorithms (GA) which is specifically designed for practitioners who are interested in how genetic algorithms work, how to program them, and how to use them.

It starts by explaining the basics: what is optimization, why it is practically important, what is the gradient method, etc., and goes all the way to explaining complex modern genetic algorithms.

It is really amazing how the authors go from the basic calculus-type material to real sophisticated techniques. Numerous exercises build on one another and help the readers grasp the main concepts.

A lot of new and exciting material has been added since the 1997 first edition, including:

– the detailed information on how to parallelize the genetic algorithms (and parallelizability is one of the main advantages of GA), together with an explicit High Performance Fortran codes, and
– new sections explaining such techniques as simulated annealing, swamp optimization, and ant colony techniques.

From the GA viewpoint, this is the best introduction to GAs for practitioners I have ever seen.

A word of caution: the author also overview more traditional numerical (non-GA) optimization techniques in Chapter 1. While the main ideas behind numerical optimization are represented well, the author's references are outdated (latest is 1992), and the authors' statement (in Section 1.5) that traditional methods are efficient only for convex function of few variables is not true.

There is even a serious mathematical mistake in this chapter: on p. 8, the authors claim that the point where the gradient ∇f is 0 is a minimum if $\nabla^2 f > 0$; this is false, the correct criterion is that the Hessian matrix is positive definite.

It may sound grave, but, in my opinion, it can all be forgiven as a natural consequence of the authors' enthusiasm. I suggest the readers follow the book in GA optimization techniques and take the author's statements about the traditional optimization methods with a grain of salt – and hopefully the authors will correct their false statements in the next (3rd) edition.

Vladik Kreinovich
Book Review Editor
Journal of Intelligent & Fuzzy Systems

Foundations of Soft Case-Based Reasoning, Sankar K. Pal and Simon C.K. Chiu, Wiley, 2004, ISBN 0 471 08635 0.

Traditional approach to reasoning is based on logical *rules* of the type "if A then B". The main idea behind case-based reasoning is that when we need to make a decision in a new situation, we do not look for decision-making rules. Instead, we look for similar previous situations (*cases*) in which someone has already made a decision, and that decision worked well. If such cases are found, it is reasonable to make a similar decision in our new situation.

This idea sounds very reasonable, so, at first glance, this reasoning should make decision-making systems more efficient. However, in practice, very few successful decision-making systems use case-based reasoning. The main reason: it is very difficult to describe the intuitive notion of similarity in a way that a computer can understand. Similarity is a very informal, very fuzzy notion; so, to describe it adequately, it is reasonable to use the formalisms like fuzzy logic (and soft computing in general) whose main emphasis is on representing and processing such "soft", informal knowledge.

The book shows that soft computing can indeed lead to an efficient use of case-based reasoning in automated decision-making systems.

Chapter 1 introduces the basic ideas behind case-based reasoning and behind soft computing techniques (such as fuzzy logic) that the authors use to formalize the corresponding notion of similarity.

The rest of the book describes the techniques that are needed to design an efficient soft case-based reasoning system.

By definition, case-based reasoning means finding similar cases. So, to make this reasoning efficient, we must use efficient algorithms that search for similar cases; to make efficient search possible, we must develop efficient schemes for storing cases. Efficient case-storing schemes are described in Chapter 2, and the resulting efficient algorithms for case selection and retrieval are described in Chapter 3. In both problems, many soft computing techniques turn out to be very efficient:

– techniques from fuzzy logic, like fuzzy clustering, that formalize the expert's informal clustering rules;
– techniques from (Bayesian) statistics that use frequency of different situations;
– neural networks techniques that enable us to learn, i.e., to adjust our schemes as new cases are added, and
– genetic algorithms that help us optimize our choices.

Once we selected similar cases, we must now adjust the corresponding solutions to the new situation. This adjustment – called *case adaptation* – is described in Chapter 4.

After we applied the prescribed solution and got good results, we must add this situation to the list of cases on which we base our reasoning. We have already mentioned that the cases are not simply listed, they are organized so as to make search for similar cases most efficient. As a result, adding a new case to the corresponding efficient structure can be very time-consuming; often, it is more efficient to wait until we have several cases before restructuring the case base. The corresponding case-based maintenance algorithms are described in Chapter 5.

Finally, Chapter 6 describes example of working soft case-base reasoning systems. The domains of these systems range from web mining to medical diagnostics to legal reasoning (of course) to shoe design.

The book is well-written and self-contained. For readers who are not familiar with the basic soft computing techniques, fuzzy, neural, general, and rough set techniques are described in special appendices. Every chapter (and each appendix) has a list of helpful references.

This book can be used by students who want to learn the new techniques, it can be used by practitioners who want to design new decision-making systems, and it can be used by researchers who want to study the existing techniques – so that they will be able to come up with better ones.

Vladik Kreinovich
Book Review Editor
Journal of Intelligent & Fuzzy Systems