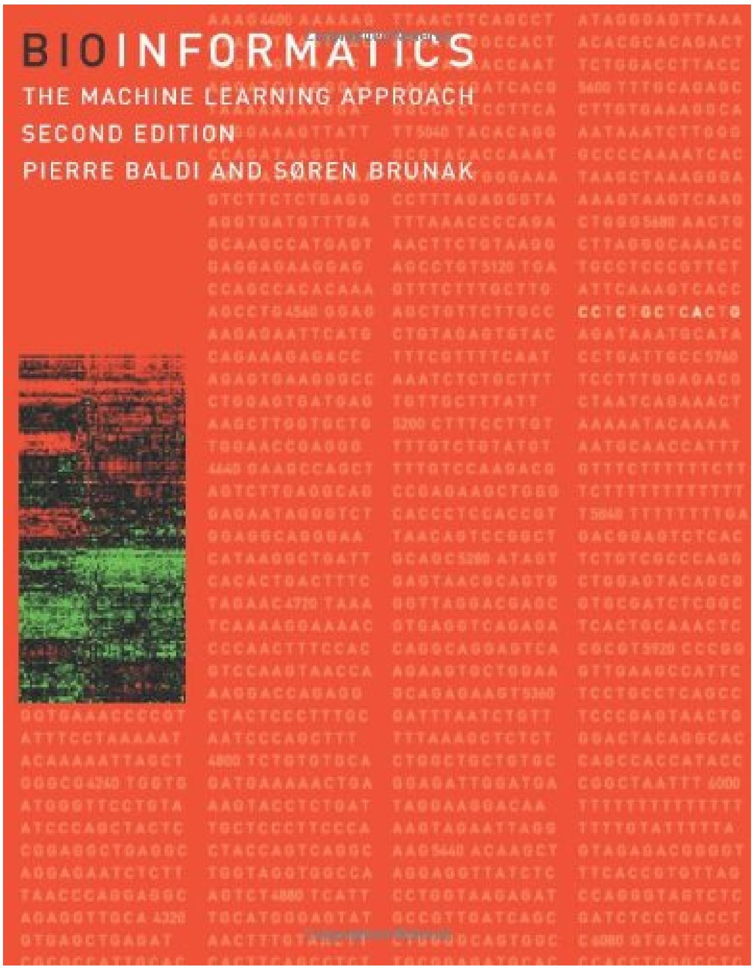
Bioinformatics: The Machine Learning Approach, Second Edition (Adaptive Computation and Machine Learning)

Pierre Baldi, Sren Brunak



#1353895 in Books 2001-08-01Ingredients: Example IngredientsOriginal language:EnglishPDF # 1 9.00 x 1.25 x 7.00l, 2.38 #File Name: 026202506X476 pages | File size: 22.Mb

Pierre Baldi, Sren Brunak : Bioinformatics: The Machine Learning Approach, Second Edition (Adaptive Computation and Machine Learning) before purchasing it in order to gage whether or not it would be worth my time, and all praised Bioinformatics: The Machine Learning Approach, Second Edition (Adaptive Computation and Machine Learning):

21 of 21 people found the following review helpful. Could have been a great one.By wiredweirdThis book is decidedly a mix: some very good information, combined with some very puzzling omissions and uneven editing.First, the good. The description of stochastic context free grammars is the best I've seen. I don't know any other reference that even hint at how to use generative grammars to evaluate likelihoods. Once they caught my interest, though, the authors did not carry through with training and evaluation algorithms I could really use. I suspect that parts of the information are

there, but I'll have to go back over their opaque notation again to work out just what they've given and just what's been left out.This same pattern - an interesting introduction with missing or mysterious development - recurs throughout the book. The discussion on clustering and phylogeny goes the same way: a number of techniques are mentioned but not developed. The authors mention a tree drawing problem, not just building the tree's topology, but ordering the branches for the most informative rendering. Again, a critical topic and one that most authors miss - in the end, these authors miss it, too, by mentioning but not filling in the idea.Their discussion of neural nets suffers badly from the authors' partial presentation. Evaluation of network output for a given input is relatively straightforward, and they present it in some detail. Training the net is the real problem, though, and is given less than a page.Baldi and Brunak give more of the fundamentals than most authors. For example, they explain the maximum entropy principle well enough that I'll use it in lots of other areas. They give some coverage to topics of intermediate complexity, such as the forward and backward algorithms for HMM training. Finally, they fizzle out at the higher levels of complexity - the Baum-Welch algorithm could have followed from the forward and backward methods, but is left as a reference to another book.There is some good here, especially in the fundamentals behind important techniques. The discussions I wanted - the more avanced topics, in forms I can use - are often weak, missing, or impenetrable. Just a bit more work, clearly within the authors' capability, would have made this a landmark reference.7 of 14 people found the following review helpful. Great bookBy A CustomerThe book of P.Baldi and S.Brunak presents a clear and exhaustive review of the main topics concerning Machine Learning techniques, as well as a broad discussion on the most significant problems that have faced Bioinformatics in recent years together with many hints on the future directions for the ML approach in BI. In the book the description of ML tools (Probabilistic Models, ANNs, HMMs, Hybrid Systems, etc.) unified under the Bayesian framework, is always clear and rigorous. Most of the theoretical materials that are unnecessary for an immediate comprehension -but that some readers may require for a deeper foundation of the ML approach- are presented in the rich appendices, a fair choice to keep the text clear. In any case the specific techniques are described in enough detail, so that any smart reader should be able to implement the models presented without further information. The biological aspects are described at a similar level of detail. As a result the book is very useful both for CS researchers interested in Computational Biology and for Biologists who want to acquire a deeper knowledge of the ML algorithmic tools used for biological data processing. It is obvious that ML plays a broad role in Bioinformatics and that sometimes some of its different aspects seem to be so weakly related that it seems a hard task to systematically review the state of the art of this approach. Anyway, the book of P.Baldi and S.Brunak performs the task successfully and actually represents both the first comprehensive book on ML in Bioinformatics and an incredibly rich pointer to all the resources (books, papers, servers and biological databases on the web) concerning this very promising discipline.10 of 17 people found the following review helpful. A must-haveBy A CustomerThis book is an excellent source of information for beginning the study of machine learning algorithms applied to biology. Reading the book you get a clear feeling that bioinformatics is not just one of the many application fields of computer science and artificial intelligence, it is perhaps the most challenging set of problems for intelligent algorithms not primarily focused on replicating human intelligence. There is an amazing wealth of open problems, some of which apparently very difficult. No doubt that unless you are already an expert you need an accurate map of this complex territory and the book by Baldi and Brunak is an excellent and up-to-date map that may suggest new exciting ideas for research.As a computer scientist I can say that the book is sometimes difficult to read if you have no previous knowledge of biology. This is because the authors didn't take the simplificative approach of reducing biological problems to abstract mathematics. Rather, they preserved the full biological flavor of the problems. Although this approach costs you more at the beginning, you can eventually get a more accurate and nontrivial picture of the problems.My conclusion: it is perhaps unlikely that you can learn about bioinformatics using only this book. However, if you want to learn about bioinformatics, this book is a must-have reference.

A guide to machine learning approaches and their application to the analysis of biological data.An unprecedented wealth of data is being generated by genome sequencing projects and other experimental efforts to determine the structure and function of biological molecules. The demands and opportunities for interpreting these data are expanding rapidly. Bioinformatics is the development and application of computer methods for management, analysis, interpretation, and prediction, as well as for the design of experiments. Machine learning approaches (e.g., neural networks, hidden Markov models, and belief networks) are ideally suited for areas where there is a lot of data but little theory, which is the situation in molecular biology. The goal in machine learning is to extract useful information from a body of data by building good probabilistic models -- and to automate the process as much as possible.In this book Pierre Baldi and Sren Brunak present the key machine learning approaches and apply them to the computational problems encountered in the analysis of biological data. The book is aimed both at biologists and biochemists who need to understand new data-driven algorithms and at those with a primary background in physics, mathematics, statistics, or computer science who need to know more about applications in molecular biology.This new second edition contains expanded coverage of probabilistic graphical models and of the applications of neural networks, as well as a new chapter on microarrays and gene expression. The entire text has been extensively revised.

This is a very good book, written with a high level of erudition and insight. (Gustavo A. Stolovitzky Physics Today)From the PublisherAn unprecedented wealth of data is being generated by genome sequencing projects and other experimental efforts to determine the structure and function of biological molecules. The demands and opportunities for interpreting these data are expanding rapidly. Bioinformatics is the development and application of computer methods for management, analysis, interpretation, and prediction, as well as for the design of experiments. Machine learning approaches (e.g., neural networks, hidden Markov models, and belief networks) are ideally suited for areas where there is a lot of data but little theory, which is the situation in molecular biology. The goal in machine learning is to extract useful information from a body of data by building good probabilistic models--and to automate the process as much as possible. In this book, Pierre Baldi and Sren Brunak present the key machine learning approaches and apply them to the computational problems encountered in the analysis of biological data. The book is aimed both at biologists and biochemists who need to understand new data-driven algorithms and at those with a primary background in physics, mathematics, statistics, or computer science who need to know more about applications in molecular biology. This edition contains expanded coverage of probabilistic graphical models and of the applications of neural networks, as well as a new chapter on microarrays and gene expression. The entire text has been extensively revised.About the AuthorPierre Baldi is Professor of Information and Computer Science and of Biological Chemistry (College of Medicine) and Director of the Institute for Genomics and Bioinformatics at the University of California, Irvine.Sren Brunak is Professor and Director of the Center for Biological Sequence Analysis at the Technical University of Denmark.