

Preface

This book is aimed at both data miners/complex systems analysts wanting to analyse postgenomic biodata and at computational biologists interested in new complex data analysis methods from other disciplines. Novel and innovative data mining and visualization techniques from artificial intelligence and knowledge engineering are covered.

In the first two chapters, the complexity of biodata and the available large and exciting datasets are reviewed, followed by local pattern analysis methods for short DNA, RNA and protein sequences. The fourth chapter explores large-scale genome comparison methods and the fifth chapter covers small scale molecular structure searches. The next chapter discusses large-scale metadata analysis and ontology-based searching and classification, including latent semantic indexing. The seventh chapter explores genomic data analysis by Support Vector Machines, microarray gene selection methods and analysis of genomes as long time series, among other techniques. The penultimate chapter deals with the integration of multimodal systems biology data. The ninth and last chapter then assesses the future challenges and includes an extensive up-to-date review on biocomputing using Graphical Processing Units and virtual machines.

Strengths and weaknesses of various approaches are reviewed, promising research trends are identified and copious up-to-date references are given for further reading. After perusal of the book,

the reader can select the most promising algorithms and datasets for their interests and is prepared to tackle the future challenges of bioinformatics in the postgenomics era.

Ilkka Havukkala
May 2010