

1. Record Nr.	UNISA996466406003316
Autore	Bao Feng
Titolo	Computational reconstruction of missing data in biological research // Feng Bao
Pubbl/distr/stampa	Gateway East, Singapore : , : Tsinghua University Press : , : Springer, , [2021] ©2021
ISBN	981-16-3064-X
Edizione	[1st ed. 2021.]
Descrizione fisica	1 online resource (XVII, 105 p. 43 illus., 41 illus. in color.)
Collana	Springer theses
Disciplina	570.285
Soggetti	Biology - Data processing Biologia Processament de dades Aprentatge automàtic Estructures de dades (Informàtica) Estadística matemàtica Llibres electrònics
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di bibliografia	Includes bibliographical references.
Nota di contenuto	Chapter 1 Introduction -- Chapter 2 Fast computational recovery of missing features for large-scale biological data -- Chapter 3 Computational recovery of information from low-quality and missing labels -- Chapter 4 Computational recovery of sample missings -- Chapter 5 Summary and outlook.
Sommario/riassunto	The emerging biotechnologies have significantly advanced the study of biological mechanisms. However, biological data usually contain a great amount of missing information, e.g. missing features, missing labels or missing samples, which greatly limits the extensive usage of the data. In this book, we introduce different types of biological data missing scenarios and propose machine learning models to improve the data analysis, including deep recurrent neural network recovery for feature missings, robust information theoretic learning for label missings and structure-aware rebalancing for minor sample missings. Models in the book cover the fields of imbalance learning, deep learning, recurrent

neural network and statistical inference, providing a wide range of references of the integration between artificial intelligence and biology. With simulated and biological datasets, we apply approaches to a variety of biological tasks, including single-cell characterization, genome-wide association studies, medical image segmentations, and quantify the performances in a number of successful metrics. The outline of this book is as follows. In Chapter 2, we introduce the statistical recovery of missing data features; in Chapter 3, we introduce the statistical recovery of missing labels; in Chapter 4, we introduce the statistical recovery of missing data sample information; finally, in Chapter 5, we summarize the full text and outlook future directions. This book can be used as references for researchers in computational biology, bioinformatics and biostatistics. Readers are expected to have basic knowledge of statistics and machine learning.

---