

1. Record Nr.	UNINA9910847581003321
Autore	Shimura Shuichi
Titolo	The First Discriminant Theory of Linearly Separable Data : From Exams and Medical Diagnoses with Misclassifications to 169 Microarrays for Cancer Gene Diagnosis / / by Shuichi Shimura
Pubbl/distr/stampa	Singapore : , : Springer Nature Singapore : , : Imprint : Springer, , 2024
ISBN	9789819994205 9819994209
Edizione	[1st ed. 2024.]
Descrizione fisica	1 online resource (373 pages)
Disciplina	164
Soggetti	Statistics Biometry Diagnosis Cancer - Genetic aspects Quantitative research Mathematical optimization Statistical Theory and Methods Biostatistics Cancer Genetics and Genomics Data Analysis and Big Data Discrete Optimization Aprendentatge automàtic Xarxes neuronals (Informàtica) Diagnòstic Genètica mèdica Càncer Llibres electrònics
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di bibliografia	Includes bibliographical references and index.
Nota di contenuto	The most important knowledge by 27 Revolutionary Findings and the Outlook of this book -- LINGO Programs Usage and New Facts by Iris Data -- Swiss banknote data and CPD data -- Test Pass/Fail Judgment

and Japanese Compact Cars and Regular Cars -- First Theory of Cancer Gene Data Analysis by 169 Microarrays: Four Universal Data Structures of Discriminant Data -- Three Important Studies for Cancer Gene Diagnosis -- Two-Step Practical Screening Method for Cancer Gene Diagnoses: Multivariate Oncogenes among 169 Microarrays.

Sommario/riassunto

This book deals with the first discriminant theory of linearly separable data (LSD), Theory3, based on the four ordinary LSD of Theory1 and 169 microarrays (LSD) of Theory2. Furthermore, you can quickly analyze the medical data with the misclassified patients which is the true purpose of diagnoses. Author developed RIP (Optimal-linear discriminant function finding the combinatorial optimal solution) as Theory1 in decades ago, that found the minimum misclassifications. RIP discriminated 63 (=261) models of Swiss banknote (200*6) and found the minimum LSD: basic gene set (BGS). In Theory2, RIP discriminated Shipp microarray (77*7129) which was LSD and had only 32 nonzero coefficients (first Small Matryoshka; SM1). Because RIP discriminated another 7,097 genes and found SM2, the author developed the Matryoshka feature selection Method2 (Program3), that splits microarray into many SMs. Program4 can split microarray into many BGSs. Then, the wide column LSD (Revolution-0), such as microarray ($n < p$), is found to have several Matryoshka dolls, including SM up to BGS. Theory3 shows the surprising results of six ordinary data re-analyzed by Theory1 and Theory2 knowledge. Essence of Theory3 is described by using cephalopelvic disproportion (CPD) data. RIP discriminates CPD data (240*19) and finds two misclassifications unique for cesarean and natural-born groups. CPD238 omitting two patients becomes LSD, which is the first case selection method. Program4 finds BGS (14 vars.) the only variable selection method for Theory3. 32 (=25) models, including BGS, become LSD among (2191) models. Because Program2 confirms BGS has the minimum average error rate, BGS is the most compact and best model satisfying Occam's Razor. With this book, physicians obtain complete diagnostic results for disease, and engineers can become a true data scientist, by obtaining integral knowledge of statistics and mathematical programming with simple programs.