

Preface

The volume “Granular Neural Networks, Pattern Recognition and Bioinformatics” is an outcome of the granular computing research initiated in 2005 at the Center for Soft Computing Research: A National Facility, Indian Statistical Institute (ISI), Kolkata. The center was established in 2005 by the Department of Science and Technology, Govt. of India under its prestigious IRHPA (Intensification of Research in High Priority Area) program. Now it is an Affiliated Institute of ISI.

Granulation is a process like self-production, self-organization, functioning of brain, Darwinian evolution, group behavior and morphogenesis—which are abstracted from natural phenomena. Accordingly, it has become a component of natural computing. Granulation is inherent in human thinking and reasoning process, and plays an essential role in human cognition. Granular computing (GrC) is a problem-solving paradigm dealing with the basic elements, called granules. A granule may be defined as the clump of indistinguishable elements that are drawn together, for example, by indiscernibility, similarly, proximity or functionality. Granules with different levels of granularity, as determined by its size and shape, may represent a system differently. Since in GrC, computations are performed on granules, rather than on individual data points, computation time is greatly reduced. This made GrC a very useful framework for designing scalable pattern recognition and data mining algorithms for handling large data sets.

The theory of rough sets that deals with a set (concept) defined over a granulated domain provides an effective tool for extracting knowledge from databases. Two of the important characteristics of this theory that drew the attention of researchers in pattern recognition and decision science are its capability of uncertainty handling and granular computing. While the concept of granular computing is inherent in this theory where the granules are defined by equivalence relations, uncertainty arising from the indiscernibility in the universe of discourse can be handled using the concept of lower and upper approximations of the set. Lower and upper approximate regions respectively denote the granules which definitely, and definitely and possibly belong to the set. In real-life problems the set and granules, either or both, could be fuzzy; thereby resulting in fuzzy-lower and fuzzy-upper approximate regions, characterized by membership functions.

Granular neural networks described in the present book are pivoted on the characteristics of lower approximate regions of classes demonstrating its significance. The basic principle of design is—detect lower approximations of classes (regions where the class belonging of samples is certain); find class information granules, called knowledge; form basic networks based on those information, i.e., by knowledge encoding; and then grow the network with samples belonging to upper approximate regions (i.e., samples of possible as well as definite belonging). Information granules considered are fuzzy to deal with real-life problems. The class boundaries generated in this way provide optimum error rate. The networks thus developed are capable of efficient and speedy learning with enhanced performance. These systems have a strong promise to Big data analysis.

The volume, consisting of seven chapters, provides a treatise in a unified framework in this regard, and describes how fuzzy rough granular neural network technologies can be judiciously formulated and used in building efficient pattern recognition and mining models. Formation of granules in the notion of both fuzzy and rough sets is stated. Judicious integration in forming fuzzy-rough information granules based on lower approximate regions enables the network in determining the exactness in class shape as well as handling the uncertainties arising from overlapping regions. Layered network and self-organizing map are considered as basic networks.

Based on the existing as well as new results, the book is structured according to the major phases of a pattern recognition system (e.g., classification, clustering, and feature selection) with a balanced mixture of theory, algorithm and application. Chapter 1 introduces granular computing, pattern recognition and data mining for the convenience of readers. Beginning with the concept of natural computing, the chapter describes in detail the various characteristics and facets of granular computing, granular information processing aspects of natural computing, its different components such as fuzzy sets, rough sets and artificial networks, relevance of granular neural networks, different integrated granular information processing systems, and finally the basic components of pattern recognition and data mining, and big data issues. Chapter 2 deals with classification task, Chaps. 3 and 5 address clustering problems, and Chap. 4 describes feature selection methodologies, all from the point of designing fuzzy rough granular neural network models. Special emphasis has been given to dealing with problems in bioinformatics, e.g., gene analysis and RNA secondary structure prediction, with a possible use of the granular computing paradigm. These are described in Chaps. 6 and 7 respectively. New indices for cluster evaluation and gene ranking are defined. Extensive experimental results have been provided to demonstrate the salient characteristics of the models.

Most of the texts presented in this book are from our published research work. The related and relevant existing approaches or techniques are included wherever necessary. Directions for future research in the concerned topic are provided. A comprehensive bibliography on the subject is appended in each chapter, for the convenience of readers. References to some of the studies in the related areas might have been omitted because of oversight or ignorance.

The book, which is unique in its character, will be useful to graduate students and researchers in computer science, electrical engineering, system science, data science, medical science, bioinformatics and information technology both as a textbook and a reference book for some parts of the curriculum. The researchers and practitioners in industry and R&D laboratories working in the fields of system design, pattern recognition, big data analytics, image analysis, data mining, social network analysis, computational biology, and soft computing or computational intelligence will also be benefited.

Thanks to the co-authors, Dr. Avatharam Ganivada for generating various new ideas in designing granular network models and Dr. Shubhra S. Ray for his valuable contributions to bioinformatics. It is the untiring hard work and dedication of Avatharam during the last ten days that made it possible to complete the manuscript and submit to Springer in time.

We take this opportunity to acknowledge the appreciation of Prof. Janusz Kacprzyk in accepting the book to publish under the SCI (Studies in Computational Intelligence) series of Springer, and Prof. Andrzej Skowron, Warsaw University, Poland for his encouragement and support in the endeavour. We owe a vote of thanks to Dr. Thomas Ditzinger and Dr. Lavanya Diaz of Springer for coordinating the project, as well as the office staff of our Soft Computing Research Center for their support. The book was written when Prof. S.K. Pal held J.C. Bose Fellowship and Raja Ramanna Fellowship of the Govt. of India.

Kolkata, India
January 2017

Sankar K. Pal
Principal Investigator
Center for Soft Computing Research
Indian Statistical Institute

Granular Neural Networks, Pattern Recognition and
Bioinformatics

Pal, S.K.; Ray, S.S.; Ganivada, A.

2017, XIX, 227 p. 54 illus., 31 illus. in color., Hardcover

ISBN: 978-3-319-57113-3