Data Mining With Computational Intelligence—Lipo Wang and Xiuju Fu (Berlin: Springer-Verlag, 2005, pp. 276, ISBN 3-540-24 522-7). *Reviewed by Jacek M. Zurada*

In recent years, data mining has been one of the hotter topics in computing due to the accumulation of large volumes of data carrying valuable but nonetheless hidden information. Short of being a new discipline equipped with a new core of knowledge, data mining often builds on an interdisciplinary bundle of specialized techniques from fields such as statistics, artificial intelligence, machine learning, data bases, pattern recognition, computer-based visualization, and, more recently, also from the computational intelligence (CI) field that encompasses neural, fuzzy, and genetic approaches to data analysis.

Data Mining with Computational Intelligence by L. Wang and X. Fu enables readers, including both intermediate and more experienced data analysts, to understand data mining processes and improve related practices at many steps of model design and its implementation using CI methods. As shown in this book, the aim of data mining techniques is to discover hidden information from data sets and to present it in an understandable way. This book reviews main data mining stages, such as data preprocessing, data dimensionality reduction, data modeling, classification, elements of knowledge discovery, and linguistic rule extraction.

Data Mining with Computational Intelligence features ten chapters and systematically presents how to utilize fuzzy neural networks, multilayer perceptron (MLP) neural networks, radial basis function (RBF) neural networks, wavelet transformation, genetic algorithms, and support vector machines (SVMs) in data mining tasks.

In the introductory chapter, data mining tasks including data dimensionality reduction, classification and clustering, and rule extraction are introduced. In the later part, CI methods for data mining are described for different tasks. Chapter 2 describes applications of MLP neural works in time-series prediction and data classification. A wavelet MLP, a wavelet packet MLP, and a cost-sensitive MLP are introduced to show how to modify the standard MLP in order to achieve specific data mining goals. In Chapter 3, fuzzy neural networks are discussed and experiments are described in applications to micro-array cancer classification and the problem of small disjuncts. The advantages of fuzzy neural networks are highlighted by comparisons with C4.5 rules.

Chapters 4–8 present applications of RBF networks to various data mining tasks, including classification, attribute importance ranking for data dimensionality reduction, class-dependent feature selection, linguistic rule extraction, and secondary protein structure prediction with a hybrid network combining RBF and MLP. RBF networks combined with genetic algorithms and the MLP are also illustrated as basic tools of data mining. These chapters illustrate the differences between class-dependent and class-independent feature selection, and discuss comparison of linguistic rules and complex classification models. The classification accuracy of linguistic rules is also shown to improve with more appropriate feature selection. Two final chapters offer insights to applications of SVMs in data mining, i.e., classification and linguistic rule extraction.

The book demonstrates that a particular data mining task can be approached with different CI methods or tools based on their individual characteristics. An added value of this work is that it provides detailed experimental results for applications of different tools. This will help readers to better explore and compare different data mining techniques regarding their fitness for a particular task. As the authors have shown, a thorough understanding of data mining processes and further progress will be needed for successful application of CI methods toward solutions of other real-world problems. By timely publishing of their book, L. Wang and X. Fu have provided the technical community with an informative and readable volume in the fast growing area of data mining.

The reviewer is with the Electrical and Computer Engineering Department, University of Louisville, Louisville, KY 40292 USA (e-mail: j.zurada@ieee.org).