Editorial Preface

Special Issue on Recent Advances in Data Analysis with Computational Intelligence Techniques

H. S. Behera, Veer Surendra Sai University of Technology, Department of Computer Science Engineering and Information Technology, Odisha, India

D. P. Mohapatra, National Institute of Technology, Odisha, India

In this decade, with advancements in data mining, many new models and tools in discovering knowledge and extracting intelligence brought forth revolutionary developments with the help of Computational Intelligence techniques. The present scenario of storage of the amount of data is quite huge in the modern database due to the availability and popularity of internet. Thus, the information needs to be summarized and structured in order to maintain effective decision making. When the quantity of data, dimensionality, and complexity of the relations in the database are beyond of human capacities, there is a requirement for intelligent data analysis techniques, which could discover useful knowledge from data. While Data Mining evolves with innovative learning algorithms and knowledge discovery techniques, Computational Intelligence harness the results of data mining for becoming more intelligent than ever. In the present scenario of computing, Computational Intelligence is playing a major role in solving the real world complex problems. Intelligence involves both reasoning and decision making and goal of the computational intelligence is to understand the principles that make intelligent behavior possible in both natural and artificial systems. Although, Computational Intelligence is a developed area of research, many new and novel findings are being evolved by the researchers in several applications. This special issue covers both the theory and applications of various computational intelligence techniques embedded to the diversified spanning fields of evolutionary computation, neural networks, connectionist system, artificial intelligence, fuzzy systems, wireless networks, fault tolerant systems, etc.

In the first article, the concepts of functional dependency and closure of an attribute of database technology are used by Das and Das for finding the most important set of genes for cancer detection. Their Proposed method computes similarity factor between each pair of genes. Based on the similarity factors a set of gene dependency is formed from which closure set is obtained. Subsequently, conditional probability based interestingness measurements have been used to determine the most informative gene for disease classification. They applied the proposed method on some publicly available cancerous gene expression dataset. The simulation results show the effectiveness and robustness of the algorithm. From the outcomes of the proposed method, the authors claim that the method will be useful for biologists and clinicians.

Since last decade, nature inspired optimization algorithms has always been a hot topic of research for the optimization community. In the earlier days of computing, many real life complex problems were solved by using some heat and trail methods. Unfortunately, these methods do not work for solving complex real valued problems. Naik et al. have developed a hybrid black hole optimization based FLANN model to solve data classification problem. They have thoroughly discussed about some of the earlier developed metaheuristic based FLANN model with their pros and cons. In their proposed method, Black hole optimization is used for selection of optimized weights for FLANN model for solving classification problem. They have conducted a vast literature survey on the related works. Their proposed BHO-FLANN is found to be free from complicated operations, complex parameter

setting process and mathematical operations. The proposed work endows with a framework for solving a variety of demanding classification problems in science and engineering such as bio-medical data classification. The performance in terms of classification accuracy of the proposed classifier is compared with some other competent models such as GA-FLANN and PSO-FLANN. Experimental results divulge that the proposed method is efficient to solve real world data classification problem which are non-linear in nature as compared to other techniques described in the literature.

Realizing the importance of social network analysis, the third article elaborates about the detection of communities in dynamic social networks using modularity ensembles SOM. Enugala et al. have proposed a community detection method in dynamic social networks, which explores Self Organizing Maps (SOM) for cluster selection and modularity measure for community strength identification. The method also emphasizes on the usage of modularity definition to identify strengths of communities. Modularity and subsequent eigenvector usage help to operate on the data with reduced dimensionality. They have considered the synthetic data sets for the validation of the proposed method for detection of communities.

With the integration of Adaptive Cuckoo Search algorithm and Intrinsic Discriminant Analysis (IDA), in the fourth article Sahoo and Choudhury have designed an intelligent object classifier for inspection of defective object like bottle in a manufacturing unit. Authentication of the proposed scheme is achieved by using different bench mark test functions along with an effective inspection procedure for identification of bottle by using Adaptive Cuckoo Search algorithm, PCA and IDA. The projected procedure remains used for defective bottle inspection in a manufacturing unit is recognized by means of adaptive cuckoo search intrinsic discriminant investigation. By the use of proposed method, the response time becomes faster as compared to other techniques.

Data in real life are mostly imprecise in nature and so the conventional tools for formal modeling, reasoning and computing, which are crisp, deterministic and precise in characteristics, are inadequate to handle them. After analyzing the importance of rough set based approximate equalities, Tripathy and Mohanty have discussed a new notion called Covering based pessimistic multigranular approximate rough equalities and studied their properties. These properties will be very much helpful in real life applications where equality of concepts is being studied. They thoroughly discussed about both the covering based pessimistic multigranular lower and upper approximation with the help of real life example. They claim that approximate reasoning can be carried out by using the properties established by them, which are approximate in character instead of being certain in their form.

In the sixth article is concerned with an EOQ (Economic Order Quantity) model for deteriorating items developed by Singh et. al. The proposed model has some important features such as Deteriorating items follow a three-parameter Weibull distribution deterioration rate, Shortages are allowed and are partially backlogged, Salvage value of items is incorporated, Demand is deterministic and a time-dependent quadratic function of time etc. The principal objective of the introduced model is to minimize the average total inventory cost by finding an optimal replenishment policy. The effectiveness of the model is validated with a numerical example and the sensitivity analysis of the optimal solutions to changes in the values of the various parameters associated with the model has been performed.

The seventh article is based on Fuzzy Knowledge based Artificial Neural Network Routing (ANNR) fault tolerance mechanism for Wireless Sensor Networks developed by Acharya and Tripathy. The proposed method uses an exponential Bi-directional Associative Memory (eBAM) for the encoding and decoding of data packets and application of Intelligent Sleeping Mechanism (ISM) to conserve energy. A combination of fuzzy rules is used to identify the faulty nodes in the network. The Cluster Head (CH) acts as the data aggregator in the network. The performance of the proposed ANNR is compared with that of Low-Energy Adaptive Clustering Hierarchy (LEACH), Dual Homed Routing (DHR) and Informer Homed Routing (IHR) and found to be superior to others. The proposed ANNR mechanism is observed to provide better performance than the existing three considered mechanisms in terms of network lifetime, throughput and reliability. Thus, the authors claim that ANNR is a hybrid algorithm that merges the powerful data recall capability of eBAMs with fuzzy

knowledge based fault detection mechanism for efficient data aggregation to design a more reliable and fault-tolerant WSN with better performance characteristics and a prolonged network lifetime as validated through the simulation results.

The final article based on the Enhancement of TOPSIS for Evaluating the Web-Sources to Select as External Source for Web-Warehousing discusses the main concerns of evaluating web-sources, which are selected as an external source for web-warehousing. In order to identify the web sources, they are evaluated on the basis of their multiple features. Jeffrey Divergence has been proposed instead of Euclidean Distance to compute the similarity measure which includes asymmetric and symmetric distances during computation. Experimental analysis of both the variations of TOPSIS approach have been conducted and the results show the enhancement in the selection of web sources.

To sum, since the data are often vague in nature and complexity in handling their distributions and the hidden relations, there is a need for the tools that can cope with the lack of information, complexity, and imprecision. Among such tools, the computational intelligence techniques are proved to be very effective. The methods of computational intelligence are quite powerful in handling the problems of data analysis. The wider use and successful applications in various diversified problem domains discussed in this special issue show the efficiency of the computational intelligence methods. As guest editors, we hope that spectrum of research works covered under this special issue will be of value for whole host of readers/researchers working in the domain of data analysis and related areas. It is important to have a good balance of different article type within the special issue.

We are grateful to our authors who have contributed their valued research to this special issue and always supported us during the reviewing of the articles. The technical standards and quality of published articles in this special issue is based on the strength and expertise of the reviewer board members who have been grossly involved in providing high quality reviews for the submitted papers. Our special thanks go to the Editor-in-Chief of the International Journal of Rough Sets and Data Analysis (IJRSDA), Dr. Nilanjan Dey for all his continued guidance and input on the policies of the journal as well as for his volunteered significant time despite of his busy schedules.

H. S. Behera
D. P. Mohapatra
Guest Editors