

Editorial

Dear Colleague:

Welcome to volume 11(4) of *Intelligent Data Analysis* – An international Journal.

Following is a brief summary of all six articles in this issue of the IDA journal. The first three articles are about clustering and classification where some new approaches are proposed and a number of existing methods are enhanced and evaluated. The last three articles, mostly on applied research, are on applications of graph grammars, data streams and analysis of sensor data.

The first article by Su and Dy focus on performing better partitions in clustering and compare the application of Principal Component Analysis (PCA) and Variance partitioning (Val-Part). They introduce Val-Part a computationally less expensive approach than PCA to provide a better alternative for minimizing the sum-squared-error in clustering. Their experiments showed better results in K-Means clustering. Chakraborty and Chakraborty in the second article explain how to efficiently construct a fuzzy rule base, as a knowledge base from a set of fuzzy rules that could be used as a knowledge base of an expert system. They introduce an interactive fuzzy rule classification approach that includes a fuzzy rule similarity index and decision maker's response criteria. The approach is evaluated for an online purchasing application. In the third article of this issue, Chehreghani et al. discuss how clustering can be applied for better data storage, retrieval and rule extraction and introduce a heuristic algorithm for clustering tree structured data. The approach is based on using a representative for each cluster where the algorithm only compares each input tree with the representative trees of clusters and as a result allows a significant reduction in processing time. The article includes the evaluation of their approach and its comparison with other methods.

Graph grammars are tools for representing relational aspects of graphs and could play a key role in discovering knowledge from data. Kukluk et al. in the next article describe an approach for learning node replacement grammars. Their approach is based on frequent isomorphic subgraph discovery where one can infer a hierarchy of productions by compressing portions of a graph described by a production. The approach is validated using some experiments where the authors generate graphs from known grammars and measure its performance. Similar work is also discussed in this article. Huang and Dong in the fifth article of this issue, address the problem of mining time changing data and how to effectively identify changes in the training data and react appropriately. For this, they propose an active learning system. Their approach is based on identifying changes in the data and choosing the most informative instances to label. While this approach is sensitive to the most significant changes in the data it is robust with respect to noise and can quickly adapt to concept-drift. Their article include evaluation of the proposed approach using synthetic and real-world data, where they show the advantages of their method. The last article of this issue, by Hu and Huang, introduces a modeling method to handle inverse problems with missing data. The introduced modeling method, which explores the input space using particle swarm optimization, is applied a complex problem with some interesting results.

And finally, the IDA-2007 (<http://www.ida2007.org/>) conference that will be held in Slovenia this year is approaching fast. This bi-annual event will be held from September 6–8, in the city of Ljubljana. We look forward to your attendance to this event and contributions to the IDA journal.

With our best wishes,

Dr. A. Famili
Editor-in-Chief