Intelligent Data Analysis 8 (2004) 323–324
IOS Press

Editorial

Dear Colleague:

Welcome to volume 8(4) of the journal *Intelligent Data Analysis*!

The contents of this issue of the IDA journal represent a variety of topics, all related to the theoretical or applied research in Intelligent Data Analysis. The topics consist of clustering with some applications to real-world problems, genetic algorithms, evaluation of machine learning algorithms in KDD, and association rules.

In the first article of this issue, Liu at al introduce an unsupervised learning algorithm that is based on a two stage process and they apply it for intrusion detection in networks. In stage 1, the algorithm takes the nearest neighbor clustering to group objects, (network data, in this case). This results in reducing the search space for better efficiency. The second stage of operation of this algorithm is a tabu search optimization that allows obtaining the near optimal normal cluster. Results reported in the paper show that the algorithm is feasible and effective for intrusion detection. Savaresi and Boley, in the second article, report a comparative analysis of two types of clustering algorithms. The two particular algorithms are: Bisecting K-Means and Principal Direction Divisive Partitioning (PDDP). The pros and cons of the two algorithms are given by the authors. The best results in terms of computational efforts and cluster quality are apparently obtained when K-means is initilized with PDDP results.

The third and fourth articles in this issue are about genetic algorithms. Stejic, et al propose a modified hierarchical genetic algorithm for relevance feedback in image retrieval. Two characteristics of this algorithm are: its generality in terms of existing relevance feedback methods and its automatic capability to switch between different combinations of adaptation targets. The results reported in the article show that the algorithm is an effective approach to the relevance feedback. In the next genetic algorithm article, Yang introduces a new operator, called "immigration operator", and show its effectiveness to improve the evolutionary performance of genetic algorithms. Two benchmarking problems are used in this study. The results show a number of advantages, such as better performance in convergence and its flexibility for different kinds of applications.

The fifth article of this issue discusses about the failure of machine learning algorithms in a real-world applications, namely intrusion detection. The article describes the deficiencies and limitations of the KDD data sets and particularly that the KDD training and testing data subsets represent dissimilar target hypotheses. The results showed that no pattern classification or machine learning algorithm can be trained successfully with the KDD data sets to perform misuse detection. The last article, by Tseng and Lin, investigate the problem of maintaining the discovered multi-supported, generalized association rules when new transactions are added into the original database. This research introduces two new algorithms which can incrementally update the discovered generalized association rules with non-uniform support specification and are capable of effectively reducing the number of candidate sets. Their results showed the two algorithms are substantially faster.

And finally, the next issue of the IDA journal, Volume 8(5) will contain six of the best papers from the fifth Intelligent Data Analysis symposium that was held in Berlin, Germany from August 28–30,

2003. In addition, planning to hold the 6th International Symposium of Intelligent Data Analysis (IDA-2005) in Madrid is complete. Detailed information is available at the IDA Society home page at: http://www.ida-society.org or http://www.ida-2005.org.

With our best wishes,

Dr. A. Famili
Editor-in-Chief