

## Editorial

---

Dear Colleague:

Welcome to volume 18(3) of Intelligent Data Analysis Journal.

This issue of the IDA journal consists of eleven articles, all related to various aspects of theoretical and applied research and development related to the field of Intelligent Data Analysis.

In the first article of this issue Pichara and Soto highlight some shortcomings of feature selection methods and suggest to adaptively find a local subset of features for each data instance so that classification of each instance is performed according to its own selective subset. Their proposed approach is based on Gaussian processes which improves classification performance by learning a set of features that quantify the discriminative power of each feature. Their experimental results show that by using local discriminative subspaces one can achieve a higher level of classification accuracy. Do, Song and Jia in the second article of this issue discuss challenges in detecting concept drift and propose a method that applies entropy over an adaptive sliding window. In their approach the sliding window is not fixed but dynamically determined. Their approach also integrates an algorithm to find the exact timestamp for retraining the classifier whenever a concept drift is detected. Their experimental results show that comparing to four benchmarks, their proposed method is better than or comparable to others. The next article by Padrol and Muntés-Mulero is about graph anonymization. They emphasize the importance of this topic and suggest that classical anonymization techniques could be used in this case. They propose to embed a graph into a multidimensional vector space that approximately preserves the distance between any two vertices in a typical graph. Their experiments show that they can successfully anonymize graphs using most common techniques.

The next two articles are about clustering. Alizadeh *et al.* discuss the issue of cluster stability and explain the drawbacks of some of the most common approaches used in cluster quality. They propose a new asymptotic criterion to assess the association between a cluster and a set of partitionings. They also propose a clustering ensemble framework that incorporates their criterion in order to find the best performing clusters. Their empirical studies show that their proposed approach outperforms several other ensemble approaches. Ben Hariz and Elouedi in the fifth article of this issue explain dynamic clustering and propose a new dynamic clustering approach that is based on k-means. In this approach, using cluster cohesion and separation concepts, the algorithm updates cluster partitions. Their experiments on several benchmark data sets show that their approach outperforms static approaches.

Nearest neighbor method is the subject of the next two articles. Zhai *et al.* discuss the limitations of K-NN and Fuzzy K-NN and propose two condensed fuzzy nearest neighbor methods that each consist of three steps. The proposed steps in their methods are: obtaining a fuzzy attribute reduct, finding sets of prototypes, and extracting fuzzy classification rules. Their experimental results and the statistical analysis of the results show that their proposed methods outperform other methods compared in their study. Zhang and Song highlight the main limitation of K-NN where there is no symmetrical solution to determine the specific value of K. In order to address this problem the authors propose a novel method of using back propagation neural networks to explore the relationship between data set characteristics and the optimal values of K. Their experimental results reported in the article show that compared with the recommended K values, the required time for determining the K values is substantially decreased.

The last four articles of this issue are mostly on applied research. Nickaein *et al.* introduce a novel approach to apply support vector regression for rate prediction in distributed video coding. The authors introduce a new method for bitrate estimation using  $\nu$ -SVM regression with the aid of a novel set of features. Their approach also includes a hybrid coding mode which reduces the computational complexity in a conventional Stanford codec. They evaluate their method using three different video sequences. Their simulation results for the feedback-free method show that the average decrease in the decoded frames is reasonably low. Burton *et al.* in the next article discuss a novel mining technique for analysis of questionnaire data. They argue that traditional association rule mining algorithms only produce obvious rules from questions data. The authors propose an enhancement to association rule mining that uses clustering to identify related questions to pre-prune rules involving similar questions. Their proposed approach reduces the search space of rules which in turn results in improving the algorithm's efficiency. They demonstrate this improvement through analyzing two real-world data sets. Guopeng *et al.* in the tenth article of this issue discuss the importance of community discovery in social network data and propose extensions to nonnegative matrix factorization method where they use diagonally dominant matrix as constraint condition to obtain the community membership as well as the interaction information among communities. Their experiments demonstrate the meaningful results produced which also contain additional information about node characteristics. And finally, Rahie and Beigy in the last article of this issue discuss the importance of credit assignment in artificial intelligence and consider use of reinforcement learning as a possible solution. They define a framework called expertness framework and use it in a multi-agent credit assignment problem. In their proposed approach the critic agent, responsible for distributing credit among agents, is equipped with learning capability, and the proposed credit assignment solution is based on the critic to learn to assign a proportion of the credit to each agent. Their experimental results show the superiority of their method over several methods of credit assignment used in other domains.

In conclusion, with this issue of the IDA journal, which is Volume 18(3), we are glad to report continuous increase in submission of manuscripts to our journal. Preparation of two special issues of the IDA journal is in progress. In addition, this year's IDA symposium will be held from October 30<sup>th</sup> to November 1<sup>st</sup>, in Leuven, Belgium (<http://www.ida2014.org/>). We look forward to receiving your feedback along with more and more quality articles in both applied and theoretical research related to the field of IDA.

With our best wishes,  
*Dr. A. Famili*  
*Editor-in-Chief*